MICHIGAN INTER-UNIVERSITY
COMMUNITY OF MATHEMATICAL GEOGRAPHERS

DISCUSSION PAPER

NUMBER 10

Effects of Boundary Shape and the Concept of Local Convexity

by

John D. Nystuen

On the Length of Empirical Curves

and

An Attempt at Objective Generalization

by

Julian Perkal

DECEMBER, 1966

Dr. John D. Nystuen, *Editor*
Department of Geography
University of Michigan
Ann Arbor, Michigan

IMaGe REPRINTS
VOLUME I

MICMG DISCUSSION PAPERS

1. Arthur Getis, "Temporal land use pattern analysis with the use of nearest neighbor and quadrat methods." July, 1963.

2. Marc Anderson, "A working bibliography of mathematical geography." September, 1963.

3. William Bunge, "Patterns of location." February, 1964.

4. Michael F. Dacey, "Imperfections in the uniform plane." June, 1964.

5. Robert S. Yuill, "A simulation study of barrier effects in spatial diffusion problems." April, 1965.

6. William Warntz, "A note on surfaces and paths and applications to geographical problems." May, 1965.

7. Stig Nordbeck, "The law of allometric growth." June, 1965.

8. Waldo R. Tobler, "Numerical map generalization;" and Waldo R. Tobler, "Notes on the analysis of geographical distributions." January, 1966.

9. Peter R. Gould, "On mental maps." September, 1966.

10. John D. Nystuen, "Effects of boundary shape and the concept of local convexity;" Julian Perkal, "On the length of empirical curves;" and Julian Perkal, "An attempt at objective generalization." December, 1966.

11. E. Casetti and R. K. Semple, "A method for the stepwise separation of spatial trends." April, 1968.

12. W. Bunge, R. Guyot, A. Karlin, R. Martin, W. Pattison, W. Tobler, S. Toulmin, and W. Warntz, "The philosophy of maps." June, 1968

ACKNOWLEDGMENT

The reprinting of these papers is done with the permission of John D. Nystuen. He has asked that the profits from the sales of this first volume in the IMaGe Reprint Series be used as a fund to promote projects in mathematical geography. This fund will be administered by IMaGe and will be disposed of by Nystuen in collaboration with IMaGe.

Sandra L. Arlinghaus (Ph.D.)
Director, IMaGe
1441 Wisteria Drive
Ann Arbor, Michigan 48104

October, 1986.

TEMPORAL LAND USE PATTERN ANALYSIS WITH THE USE OF NEAREST NEIGHBOR AND QUADRAT METHODS [1]

ARTHUR GETIS

MICHIGAN STATE UNIVERSITY

Nearest neighbor analysis as developed by Clark and Evans [2] and Thompson [3] has been successfully used by plant ecologists, but with less success and frequency by geographers. Geographers, such as Dacey [4] and Berry [5], who have attempted to use this technique to measure map patterns have criticized it for a number of reasons, the most prominent being insufficient rationale for objectively defining the study area. Nevertheless, we hope to show that the nearest neighbor method is a useful geographic tool if study conclusions "make sense" in light of expectations.

Berry [6], Bachi [7], and others have suggested the use of standard distance measures, after Mahalanobis, in order to describe spatial patterns. However, these measures, geared mainly toward describing dispersion characteristics in a population, do not permit analysis of map patterns in terms of their departures from randomness. Therefore, standard distance measures might best be thought of as a complementary tool for map pattern analysis.

In this paper, hypotheses based on land use patterns within cities are tested. It is felt that the character of an urban transportation system has a great influence on land use patterns. Our analysis was aimed at bringing to light the importance of the transportation technology variable. Although no attempt was made to explore this variable in depth, we have tried to show, quantitatively, how grocery store patterns reflect innovations in transportation technology.

Grocery store locations in the city of Lansing, Michigan, for the time periods 1900, 1910, 1920, 1930, 1940, 1950, and 1960 were used to indicate land use patterns. Patterns described in one period of time are tested for significant changes with patterns in other time periods. As a check on some of this work, a quadrat method, based on the Poisson distribution (as is the nearest neighbor technique) was used. Grocery stores were selected as indicators of commercial land use because of certain theoretical and empirical considerations. A discussion of these considerations follows.

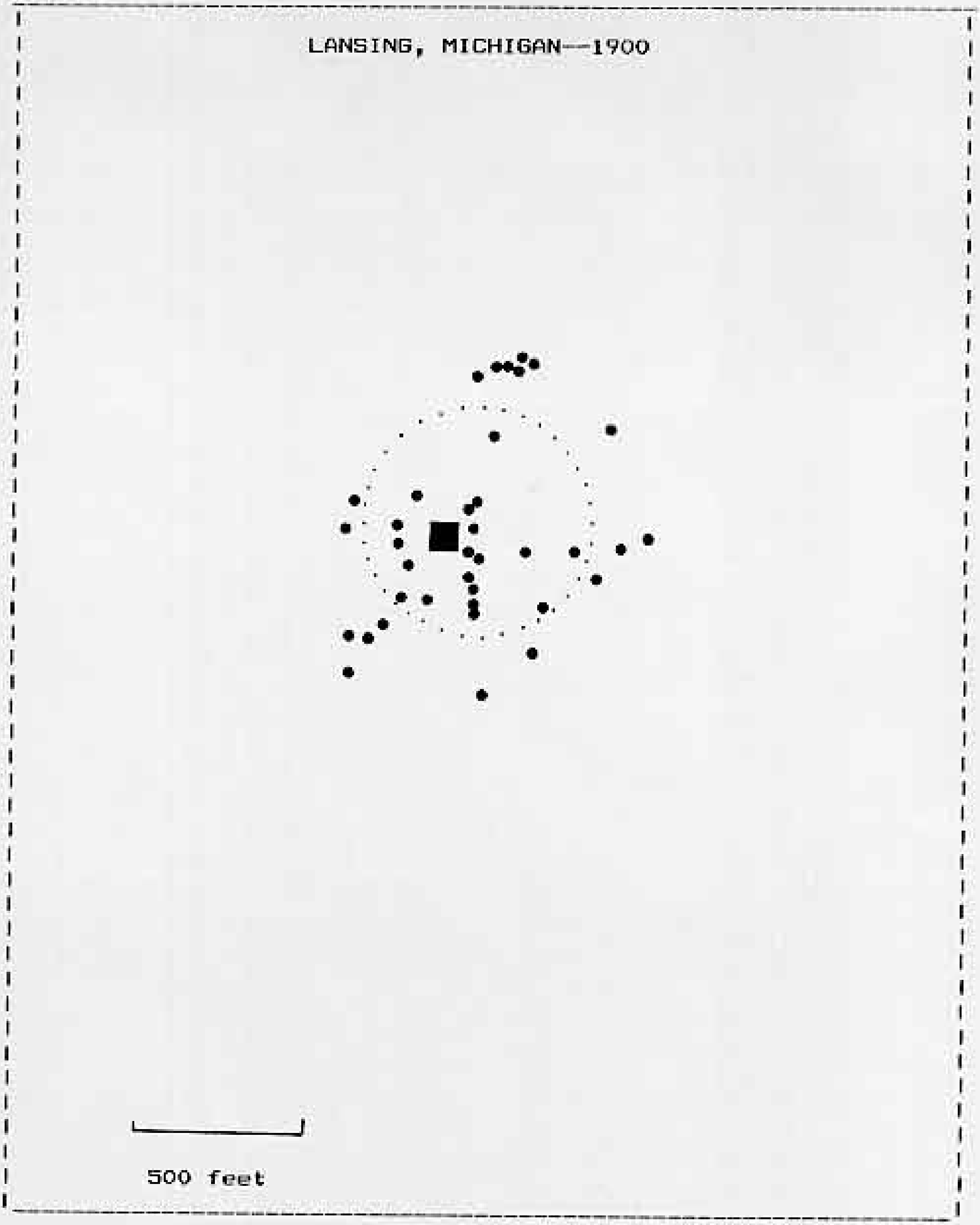
Assuming no market overlap, a central good will locate in the center of the area it serves. Central place theory asserts this proposition for all central goods, but, unlike high order goods, the frequently visited, convenience type firms of low order reflect this proposition rather well. The validity of the market overlap assumption for low order goods appears to make the difference. Goods supplied at grocery stores are of low order, and it can be shown that such stores appear in both large and small shopping centers. Only in well developed central business districts and along interurban arterials do commercial areas exist without grocery stores. In other words, for all practical purposes, all commercial districts have at least one grocery store.

If grocery store organizational characteristics were constant, then we would expect the grocery store location pattern to approximate the population density characteristics. In the pre-supermarket period, the grocery store pattern duplicated the population pattern. Furthermore, in 1961 it was shown that this was approximately the case in Tacoma, Washington, even though organizational characteristics did vary from store to store [8]. It follows that any change in population density characteristics will be reflected in grocery store location patterns, and therefore in all commercial land use patterns.

Knowledge of the urban environment enables us to hypothesize about expected store location patterns. One would expect the density of stores to fall off as distance from the center of the city increases, as is the case with population. This is obvious, but, density gradients vary greatly from city to city. Generally, within cities there is more than one area of dense settlement. Within a city we might expect to find a number of different kinds of patterns, depending on our level of analysis—for example, grouped neighborhood patterns but dispersed city patterns. We would also expect that as a city's population density increases unevenly, as is generally the case, the number of groups of stores would dominate the pattern. However, in the early days of settlement this groupedness would be less evident. As the automobile becomes more important, the population densities would decrease and therefore we would expect a return to the dispersed patterns found in the early days of settlement. Between a period having a grouped and a period having a dispersed, or uniform, pattern one would expect to find a random pattern--that is, a pattern reflecting neither extreme population densities nor very moderate densities. This allows us to test hypotheses based on randomness.

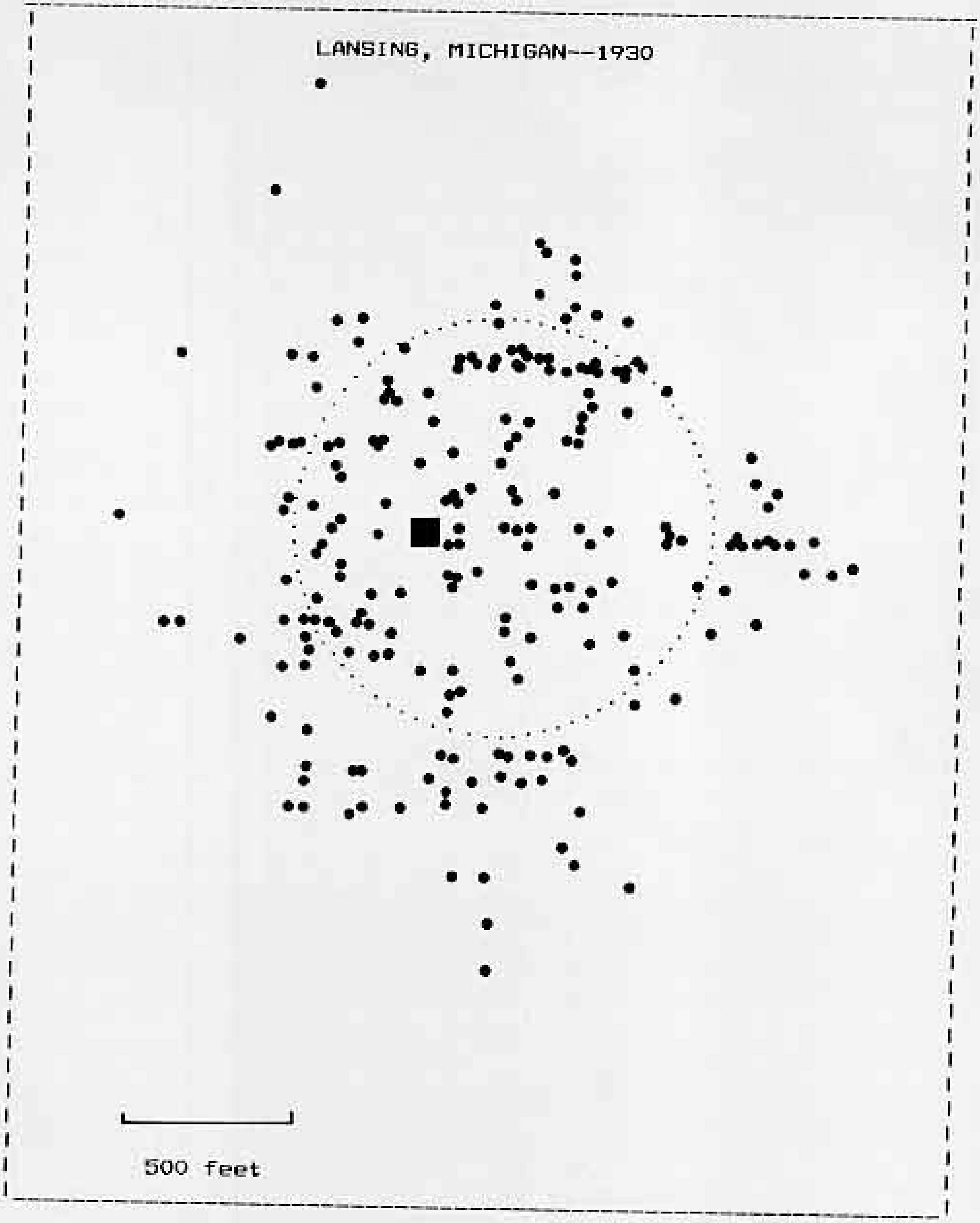
Grocery store location information was collected from Lansing City Directories for each of seven years at ten year intervals, starting in 1900. The data were mapped and the patterns analyzed. It should be noted that nearest neighbor analysis does not allow us to select the entire urbanized region as the study area (nor a larger area than the urbanized region). Rather, the analysis permits us to make statements about the character of the distribution well within the limits of the total pattern. A discussion of this point will be made in the following section.

Maps 1-3 show the location of grocery stores in three of the seven time periods. The nearest neighbor method is explained in terms of two dimensional spacing. The dot distribution of grocery stores in part of Lansing for each of the designated time periods was measured, and the manner and degree to which the distribution of points departs from random expectation was noted. The distance from an individual point to its nearest neighbor provided the basis for the initial analysis. Using a randomly selected sample of points, a series of these measures was made. The mean distance to the nearest neighbor is compared statistically to the expected distance, that is, the average distance the points would be in a random pattern of the same



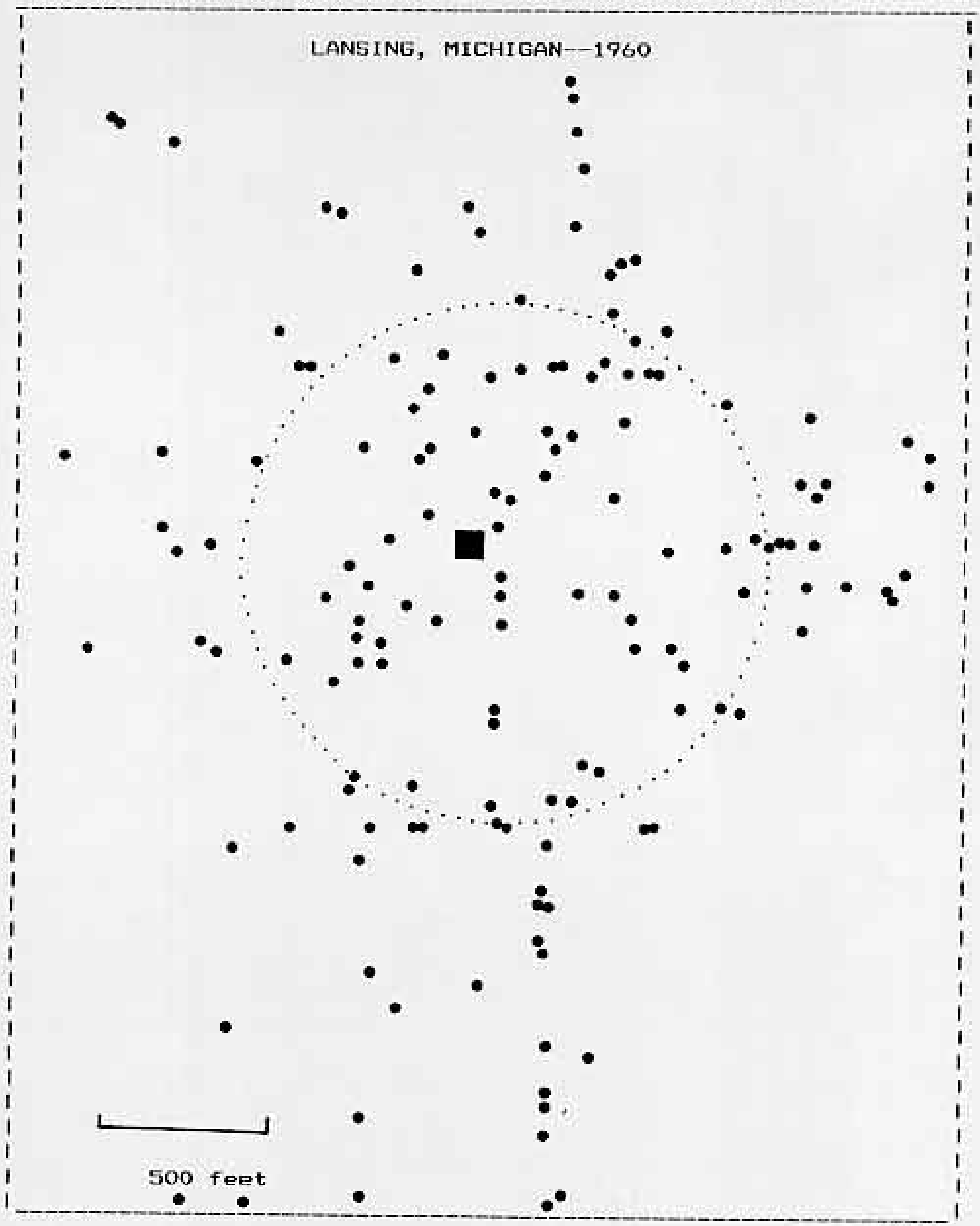
A square represents the state capitol

MAP 1



A square represents the state capitol

MAP 2



A square represents the state capitol

MAP 3

density as that of the random sample. The ratio of the observed mean distance to the expected mean distance serves as the measure of departure from randomness, and these ratios are compared from one time period to another. The analysis also included measures to three nearest neighbors; a discussion of this is found in the section "An Extension to the Method."

Tests of significance were made about the hypothesis of randomness. Table 1 summarizes the results of the initial nearest neighbor analysis. N is equal to the number of dots used as centers of measurement; ρ is the density of the observed distribution expressed as the number of individuals per unit of area; $\bar{r}_A$ is the mean of the series of distances to nearest neighbors; $\bar{r}_E$ is the mean distance to the nearest neighbor expected in an infinitely large random distribution of density ρ; R is the measure of the degree to which the observed distribution departs from random expectation, $\sigma_{\bar{r}_E}$ is the standard error of the mean distance to the nearest neighbor in a randomly distributed population of density ρ; and Z is the standard variate of the normal curve.

Before we discuss the results of this part of the study, it is important to mention some of the problems of procedure. For tests of significance to be meaningful, a sample rather than the whole population must be used. Because the population is very small in the early periods, small samples had to be selected. For 1900, for instance, only ten randomly selected observations were used. The problem of sample size is most acutely felt when it is realized that the size of the area selected must be well within the total dot pattern. This eliminates many more possibilities. However, in the period from 1920 through 1950 adequate samples of 47 or 50 points were used, and this adds confidence to conclusions drawn about those years.

A second problem of procedure is related to the size of the study area. If the sample area were the size of the State of Michigan, our study would, of course, show that departures from random expectations were extreme in the direction of groupedness. It is evident, then, that the size of area selected will have a great bearing on the results of the analysis. In order to logically consider the spacing of stores, it was necessary to exclude that area of the city where the general dot pattern was altered by low population densities. Nearest neighbor analysis is based on the Poisson distribution, where, supposedly, each location has an equal chance of containing a point. Therefore, the rims of the settled portions of the city were eliminated from the analysis and a circular area within the total dot pattern was selected. In order to choose a realistic area, the location of the center of the circle varies from period to period. Of course, the radius of the circle also varies. One of the problems facing geographers is that of selecting meaningful study areas where spatial bias is minimized.

For further discussion on the mechanics of the analysis, readers are referred to Clark and Evans [9], Clark [10], Thompson [11], Berry [12], and Dacey [13]. There are a few unpublished papers in the geographical literature which demonstrate the empirical use of the technique. However, to the author's knowledge only Dacey has attempted an empirical analysis that has been published, and this work is applied to a special case and is therefore of limited value [14]. If the ratio of the observed distribution to the random expectation is equal to R=1, then this would signify a random observed distribution. If the value is R=0, this would signify maximum aggregation, that is, all observations are at the same place. Clark and Evans have shown that under conditions of maximum spacing, that is, when the points are distributed in an hexagonal pattern, each point being equidistant from six other points, the value of R is 2.1491 [15]. The range of R, then, allows us to determine the degree to which an observed distribution departs from randomness. Statistical statements can be made with regard to significant departures from randomness. Also, it is possible to conclude that an R value of 0.5, for example, indicates that nearest neighbors are, on the average, half as far as expected under conditions of randomness. One unfortunate characteristic of the nearest neighbor technique, when only the first nearest neighbor of a point is used, is that a repeated pattern of two or more closely spaced points occurring far from one another would yield a low value of R, even when the pattern may appear dispersed. This problem can be alleviated to some degree by using measures to more that the first nearest neighbor and obtaining new values of R. In the section "An Extension to the Method" this type of analysis is discussed.

Table 1 shows that the R values for the different time periods are all less than 1. This means that all of the patterns show tendencies toward groupedness. However, in the case of 1900 the R value of 0.721 was shown not to be significant at the .15 level. All other values were significant at less than the .01 level. Figure 8 shows the trend in R values over time. It appears that in recent years the value is increasing and from this it might be predicted that in the not too distant future the R value will not be significantly different from a random expectation. The greatest tendency for groupedness appears in the year 1920, and again in 1940 the value is very low. These values indicate the very dense arrangements of grocery stores in those years.

Table 2 summarizes the results of a series of "student's t" tests made on all combinations of study periods. These tests are designed to show whether or not significant differences exist between R values of the various time periods [16]. In eleven of the 21 tests, significant differences were found. The 1900 pattern differs significantly from the 1920, 1930, and 1940 patterns. The 1960 pattern differs significantly from all patterns but that of 1900. Since the 1900 and 1960 patterns were rather similar, patterns between these two periods might be thought of as coming from a different population. Tests on variations in the period from 1910 to 1950 show that in three cases these patterns were significantly different; 1920 different from 1930 and 1950, and 1940 different from 1950.

It appears that the most radically different pattern is that of 1960--more significant departures with other patterns were found. The thesis that present-day patterns are similar to the

TABLE 1

Summary of Nearest Neighbor Measures [a]

YEAR	N	ρ	$\bar{r}_A$	$\bar{r}_E$	R	$\sigma_{\bar{r}_E}$	Z
1900	10	.000393	18.20	26.25	0.721	4.18	-1.69
1910	16	.000577	10.81	20.82	0.519*	2.72	-3.68
1920	47	.000607	7.72	20.29	0.380*	1.55	-8.11
1930	50	.000577	10.86	20.69	0.525*	1.53	-6.42
1940	50	.000475	9.08	23.49	0.387*	1.70	-8.48
1950	50	.000451	12.80	23.55	0.543*	1.74	-6.19
1960	34	.000248	18.41	31.75	0.580*	2.85	-4.68

[a] One map distance unit equals 36.36 feet. For example, the $\bar{r}_A$ value of 18.20 is equivalent to 666.175 feet.

*Significant at the .01 level.

TABLE 2
Student's t Values for Difference Between Means Test on All Combinations of Study Periods

	1900	1910	1920	1930	1940	1950	1960
1900	--	1.58	3.56**	2.14	2.94*	1.61	0.04
1910		--	1.55	0.02	0.80	0.84	2.22*
1920			--	2.05	0.99	3.41*	5.11**
1930				--	1.13	1.15	3.33*
1940					--	2.42*	4.38**
1950						--	2.50*

*Significant at the .01 level.
**Significant at the .001 level.

TABLE 3

R Values for One, Two, and Three Nearest Neighbors by Study Time Periods

Year	1st	2nd	3rd
1900	0.72	0.82	0.83
1910	0.52	0.74	0.77
1920	0.38	0.58	0.78
1930	0.53	0.63	0.63
1940	0.39	0.53	0.63
1950	0.54	0.57	0.62
1960	0.58	0.67	0.69

early developmental patterns, and much more dispersed than those of the middle period, is borne out in these tests. Although the R value of the 1960 pattern shows significant departures from randomness in the grouped direction, still the pattern is significantly less grouped than those of earlier periods. Since the 1950 pattern is statistically significantly different from that of 1940, one might conceive of the automobile revolution beginning between 1940 and 1950. Data on the number of stores operating in Lansing show a downward trend in the depression years, a rise, and then a new decrease starting about 1940.

An Extension To the Method

A more incisive analysis would consider distance to more than the first nearest neighbor. Clark and Evans [17] and Dacey [18] suggest techniques where measures to two to six nearest neighbors can be made and R values obtained. Dacey provides a table for expected mean distances in a random distribution when one to six nearest neighbors are used. In our analysis three nearest neighbors were used. The technique provided by Dacey involves dividing a circle into three equal sectors, placing the center of the circle over each randomly selected point, and measuring the distance to the nearest neighbor in each of the sectors. The closest neighbor is considered the first nearest neighbor; the closest neighbor in either of the two sectors not containing the first nearest neighbor is considered the second nearest neighbor; and the closest neighbor in the remaining sector is the third nearest neighbor. It was hoped that this analysis would give some idea of the more general pattern, rather than the limited idea available using only first nearest neighbors.

Although significance tests were not made, nor tests between various time periods carried out, still the R values were obtained and differences from the first nearest neighbor R values noted (see Table 3). In all cases the R value increased; all of the patterns, if not random to start with, were approaching randomness. These results shed light on grouping characteristics of grocery stores. As second and third nearest neighbors are included in the analysis, the R values show less of a tendency toward grouping, and a more dispersed pattern emerges. In the case of the 1920 R values, the tendency to group locally is much in evidence, but the more general pattern appears to be similar to the others.

Preliminary Conclusions

The results of the nearest neighbor analysis allow us to make statements regarding the evolution of land use patterns in the Lansing area. With increased accessibility in an urban area, the urban landscape approaches the isotropic surface assumed by Christaller. All areas in the city are becoming equally accessible. The advent of new and faster means of transportation, better roads, and the ability of nearly all of our gainfully employed people to own an automobile have caused a transportation revolution. With increased accessibility, the functional land use structure spreads itself out until it approaches an even pattern. However, there appears to be a lag effect in the consummation of that scheme. Further research will be aimed at discovering the characteristics of this lag effect.

Although one might argue that the pattern is simply returning to its original configuration, the pattern of 1900, and although statistically this is a sound conclusion, I would like to propose that the pattern in the past was never even in its arrangement, while the pattern of the future will approach evenness. In the formative stages of a community, there is a period when the city fills its empty areas with homes, and, appropriately, a rather haphazard, random pattern of commercial land use develops, only to become more grouped as population densities increase. The original random pattern is not caused by the same phenomena as the random pattern expected in the near future. The first is an initial city form evolving from the single shopping area of the village, while the expected pattern is a transitional form, standing between the grouped pattern of the pre-automobile period and the dispersed or even pattern of the mature period of adjustment to the automobile's influence. The size of grocery stores is indicative of this trend. Until 1940, most stores, including those of chain organizations, were of the service type--small and dependent on a limited number of customers. There are still many of these small stores in existence, but they are declining rapidly in number. The large supermarket is an answer to the automobile revolution, and it is only a matter of time before the smaller stores cease to exist in their present form. Other studies have provided us with information. In one study it was found that the threshold level, in terms of dollars of sales, for supermarkets is nearly twice that of the service type of store [19]. In another it was found that only about 25% of the stores under a particular proprietor are still under the same proprietorship ten years later [20]. This latter fact indicates that patterns can possibly change very rapidly. While it is true that some large regional shopping centers contain more than one supermarket, we feel that present diseconomies of scale are short lived.

The Quadrat Method

The Poisson distribution was the foundation for nearest neighbor analysis. The same probability distribution can be used to determine the expected occurrence of points in a cell of a plane divided into many cells. At this time a short discussion of the Poisson distribution is in order.

Following the notation of Dacey, the Poisson distribution is given by

$$(m^x e^{-m})/(x!)$$

where m is the average number of points per cell and x is the number of points expected in a randomly chosen area within the designated plane [21]. e is the base for natural logarithms. The basis for the development of this distribution comes from a probability situation where the number of observations (cells) is large and the probability of occurrence (points) is small.

In the quadrat method, the plane is divided into a number of equal sized cells, and the number of points occurring in each cell or in randomly selected cells is noted. Tests are made about the hypothesis of randomness. If the observed frequency of points occurring in cells does not approximate the expected distribution which is derived from the observed density function, then the distribution can be thought of as either grouped or dispersed, depending on the direction in which the observed values differ from the expected values.

A major shortcoming of the technique is the effect of various quadrat sizes on the results. In fact, the probable reason for the method's disuse is this shortcoming. A related problem facing geographers is finding appropriate cell sizes for quadrat analysis. However, the technique does not depend on a random sample--the entire population may be used, thereby adding more information and permitting more meaningful conclusions. In our analysis we used three different cell sizes and compared results with each other and with the nearest neighbor results. In this way the method can be thought of as being supplementary to the nearest neighbor technique; and possibly as a test of the significance of nearest neighbor results.

Table 4 gives the results of the quadrat analysis for cell sizes $(545)^2$ feet, $(909)^2$ feet, and $(1{,}818)^2$ feet. One useful characteristic of the Poisson distribution is that the variance equals the mean. Tests of significance can be made using the chi-square statistic because the distribution of chi-square approximates the ratio of the sum of squares to the variance. Therefore, the equation

(1) $$\chi^2 = (\Sigma (x - \bar{x})^2) / \sigma^2$$

but,

(2) $$\chi^2 = (\Sigma (x - \bar{x})^2) / \bar{x}$$

The ratio of the mean to the variance would equal 1 in a random distribution, and values higher or lower than 1 would be equivalent to the R values found in the nearest neighbor work.

(3) $$R = (N\bar{x}) / (\Sigma (x - \bar{x})^2)$$

The values for R and chi-square were obtained and are summarized in Table 4. In all but one case the R value is below 1, signifying a tendency toward groupedness. A test was made for significance based on percentiles of chi-square as a ratio to the number of degrees of freedom. In all three cell size analyses, the 1900, 1910 and 1960 R values were not significantly different from random, but in the 1960 case there appears conclusive evidence of randomness. In fact, the results of these tests more closely follow our pre-study expectations than the nearest neighbor tests. Let us look at one of the analyses in more detail. When the cell size is $(1{,}818)^2$ feet the R values start at a value of 0.6562 in 1900. This value did not differ

TABLE 4

Summary of Quadrat Method Data and Measures

Frequency of Points Per Cell	0	1	2	3	4	5	6	7	8	9	10	11	al Points	Total Cells	Mean Points Per Cell	R	Probability of a greater X^2 than that observed
(a) Size: $(545)^2$ Feet Per Cell																	
1900	89	13	3										19	105	0.181	0.882	.16
1910	112	20	4	3									37	139	0.266	0.696	.0005
1920	261	43	12	2	3	1							110	322	0.342	0.762	.0005
1930	315	66	18	5	3								129	407	0.317	0.679	.0005
1940	389	82	15	5	2	2							145	495	0.293	0.640	.0005
1950	433	76	15	5	1								125	530	0.236	0.746	.0005
1960	589	64	4										72	657	0.110	0.999	.50
(b) Size: $(909)^2$ Feet Per Cell																	
1900	24	9	2		1								17	36	0.472	0.681	.027
1910	32	12	7	2	1	1							41	55	0.764	0.614	.002
1920	69	20	12	8	1		2						84	112	0.750	0.509	.0005
1930	83	41	15	12	4	1		1					135	157	0.860	0.585	.0005
1940	103	47	23	8	2		2		1				145	186	0.780	0.545	.0005
1950	115	59	11	11	2		1						129	199	0.643	0.682	.0005
1960	231	45	14										73	290	0.252	0.884	.064
(c) Size: $(1,818)^2$ Feet Per Cell																	
1900	1	2	2	1			1						15	7	2.143	0.656	.10
1910	1	5	1	5	1	1		1	1				46	16	2.875	0.607	.035
1920	7		3	8	3	3		1	1	2			90	28	3.214	0.467	.0005
1930	6	6	7	8	3	6		4		2	1		142	43	3.302	0.484	.0005
1940	7	7	15	8	1	5	5	2		2		1	163	53	3.075	0.480	.0005
1950	9	13	8	9	7	2	1	2	2				130	53	2.453	0.562	.0005
1960	23	20	17	8	1								82	69	1.188	1.044	.54

significantly from random at the 10% level. With succeeding time periods the R values decrease to 0.4669 in 1920 and remain close to that level until 1950 when a less grouped pattern begins to assert itself. By 1960 the value has actually passed the 1 level. Although the degree of groupedness differs between the various cell sizes, still this same pattern of high then low then high R values is evident in all. This same pattern was borne out in the nearest neighbor analysis.

Conclusions

In a recent study Muth carried out a multiple regression analysis using an independent variable based on certain characteristics of the housing market's density gradient. He concluded that, "in line with central place theory, the spatial distribution of retail sales appears to be a result rather than a cause of urban population distribution" [22]. This lends greater credibility to the assertion that the location of grocery stores corresponds to population density patterns. Wingo has stated that cities which grew the most from 1920-1950 tended to have substantially lower gross densities than those whose primary growth took place before the impact of the automobile [23]. Wingo and others have rightly concluded that various aspects of the changing urban transportation system are responsible for these differences. Measuring a transportation variable is a difficult task; usually, indirect methods are used. We feel that, with the use of nearest neighbor analysis and the associated quadrat analysis, a rather simple, straightforward transportation variable might be derived. Of course, the derivation would have to be made in terms of the expected effect of transportation changes on urban land use changes.

Our results give hope to geographers and planners who are interested in measuring land use change in light of transportation innovation. The numerous shortcomings of the two techniques, it is felt, should not seriously inhibit their use. Common sense modifications might add much to their potential value. We were able to show quantitatively how land use patterns in Lansing, Michigan changed over time. Values for the relative spacing of grocery stores in one time period were tested against the values in another. Significant differences in the land use patterns were evident between certain time periods. These differences, for the most part, followed our previous notions on the evolution of land use patterns in light of population density and transportation changes. The nearest neighbor and the quadrat methods provide useful tools for such measurements and tests.

Footnotes

1. The author would like to thank Professor P. J. Clark, Michigan State University, for his comments and suggestions.

2. P. J. Clark and F.C. Evans, "Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations," Ecology. Vol. 35 (1954), pp. 445-453.

3. H.R. Thompson, "Distribution of Distance to n-th Neighbor in a Population of Randomly Distributed Individuals," Ecology. Vol. 37 (1956), pp. 391-394. For a review of nearest neighbor procedures with regard to patterns of points see P. Greig-Smith, Quantitative Plant Ecology, New York: The Academic Press, 1957.

4. Brian J. L. Berry, "Methods and Problems of Taxonomy." Unpublished working paper. University of Chicago. Portions of this work were discussed in a paper entitled "Statistical Tests of Value in Grouping Geographic Phenomena" given at the Annual Meeting of the Association of American Geographers, Pittsburgh, 1959.

5. Michael F. Dacey, "Analysis of Map Distributions by Nearest Neighbor Methods," Department of Geography, University of Washington, unpublished Discussion Paper Number 1, March 8, 1958, and Idem. "Analysis of Central Place Patterns by Nearest Neighbor Method," Department of Geography, University of Washington, unpublished Discussion Paper Number 20, May 15, 1959.

6. Brian J. L. Berry, op. cit.

7. Roberto Bachi, "Statistical Methods for Spatial Analysis," unpublished paper submitted to the Second European Congress of the Regional Science Association, Zurich, 1962.

8. Arthur Getis, "The Determination of the Location of Retail Activities with the Use of a Map Transformation," Economic Geography, Vol. 39, No. 1, January, 1963, pages 14-22.

9. P. J. Clark and F. C. Evans, op. cit.

10. P. J. Clark, "Grouping in Spatial Distributions," Science, Vol. 123 (1956), pp. 123-125.

11. H. R. Thompson, op. cit.

12. Brian J. L. Berry, op. cit.

13. Michael F. Dacey, op. cit., and idem. "Comments on the Experimental Design of the Nearest Neighbor Statistic," Department of Geography, University of Washington, unpublished Discussion Paper Number 13, December 12, 1958, and idem. "A Note on the Derivation of Nearest Neighbor

Distances," Journal of Regional Science, Vol. 2, No. 2, Fall, 1960, pp. 81-87.

14. Michael F. Dacey, "The Spacing of River Towns," Annals of the Association of American Geographers, Vol. 50, No. 1, March 1960, pp. 59-61.

15. P. J. Clark and F. C. Evans, op. cit., p. 447.

16. Strictly speaking, multiple t tests are illogical. The probability of more than one t value to yield spurious results is greater than .05. However, the results are indicative of the comparable character of the data.

17. P. J. Clark and F. C. Evans, op. cit., p. 450.

18. M. F. Dacey, 1960, op. cit.

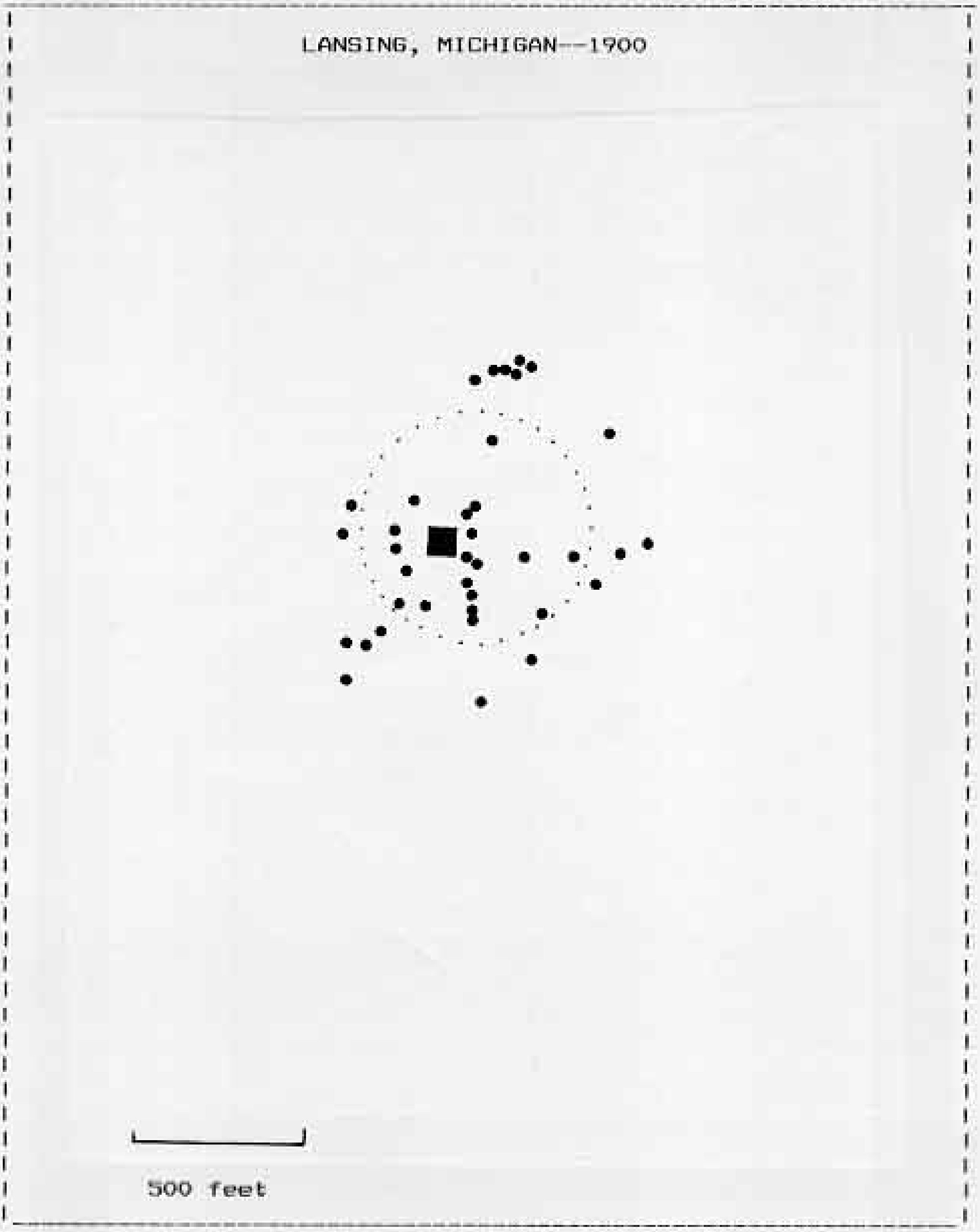
19. A. Getis, "The Service Function of Cities," abstract of paper in Annals of the Association of American Geographers, Vol. 50, No. 4, December 1960.

20. This conclusion was based on examination of the data used in this study.

21. M. F. Dacey, 1960, op. cit.

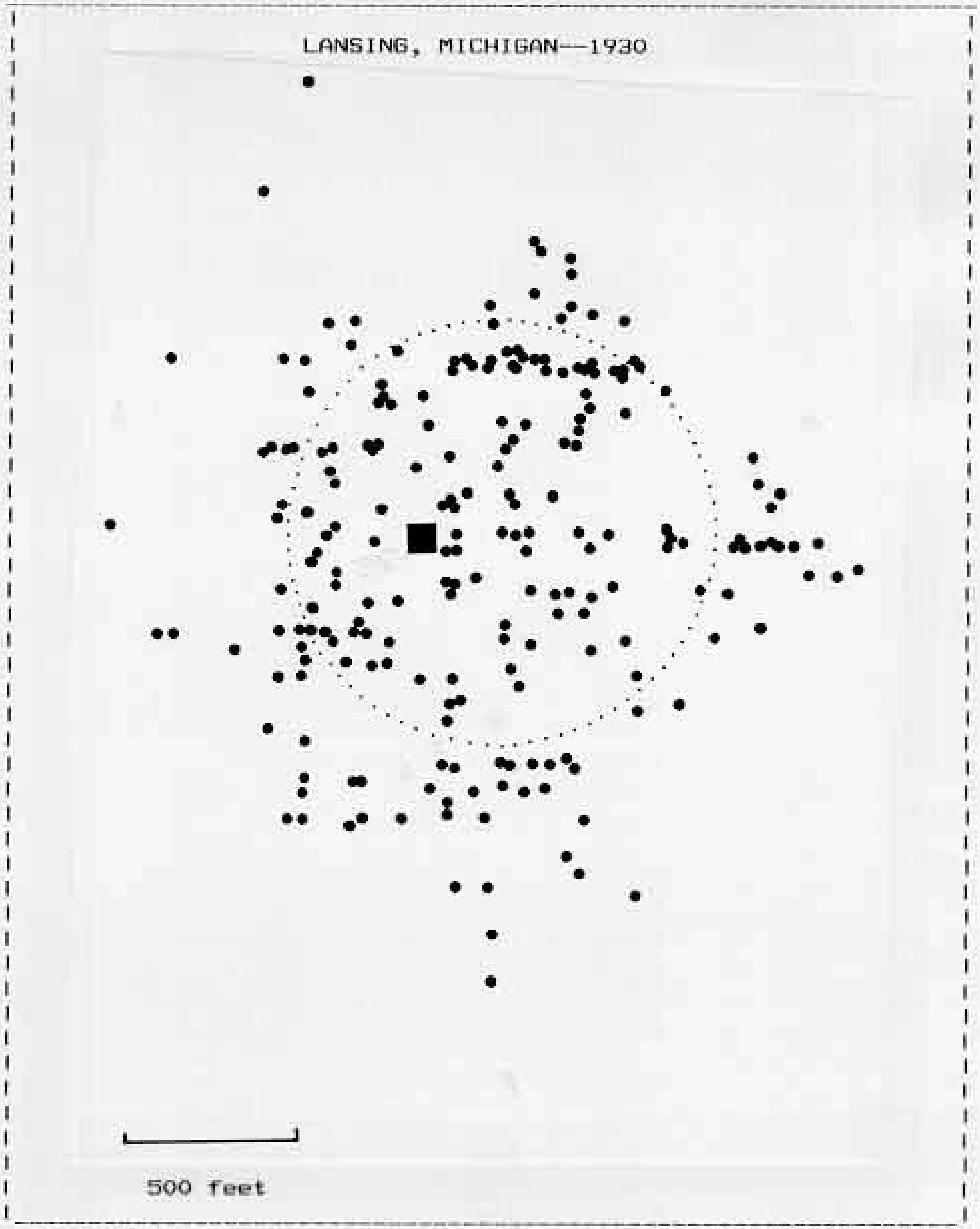
22. Richard F. Muth, "The Spatial Structure of the Housing Market," Papers and Proceedings of The Regional Science Association, Vol. 7, 1961, p. 218.

23. Lowdon Wingo, Jr., Transportation and Urban Land, Resources for the Future Inc., 1961, pp. 23, 25.



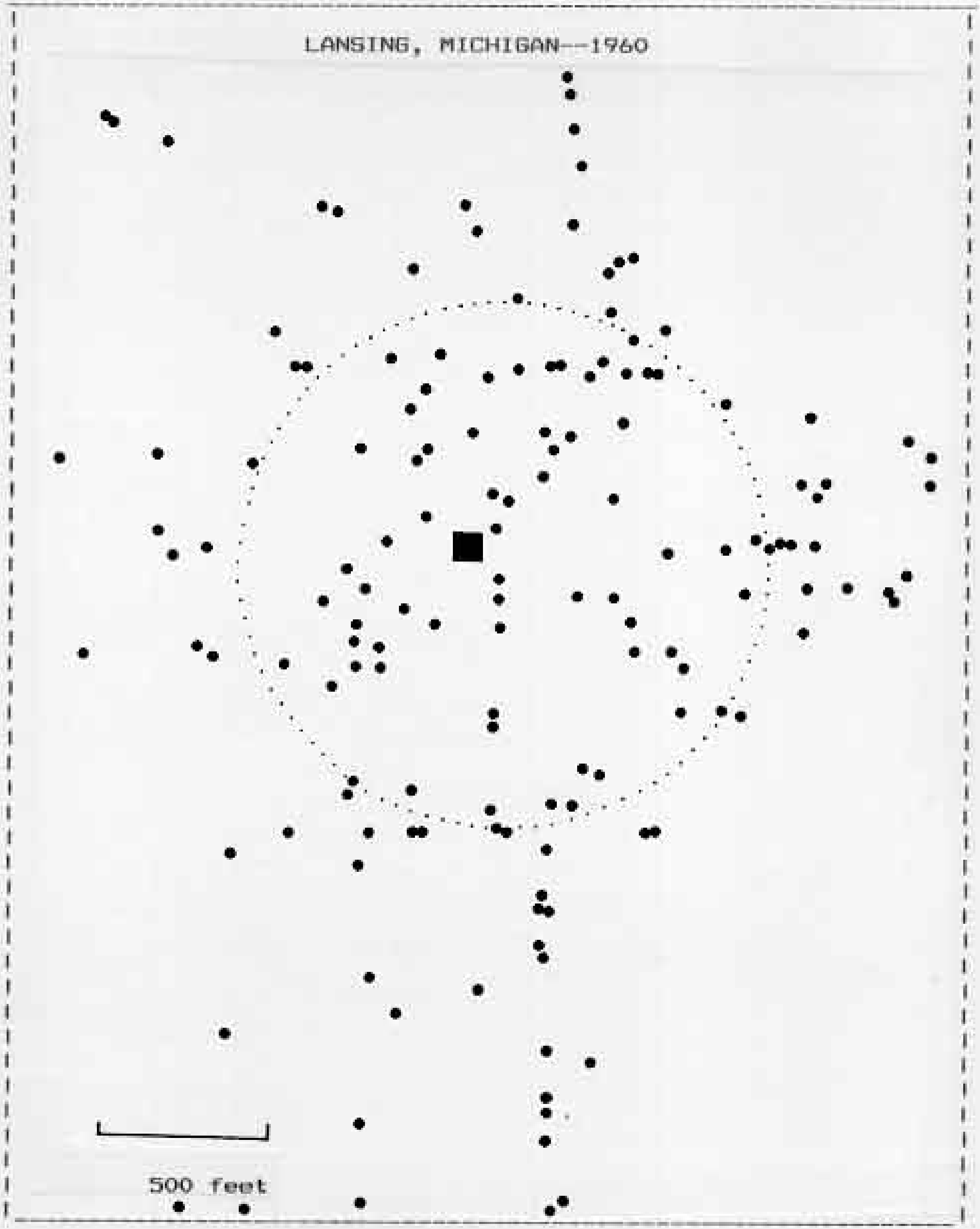
A square represents the state capitol

MAP 1



A square represents the state capitol

MAP 2



A square represents the state capitol

MAP 3

"A Working Bibliography

of

Mathematical Geography"

(Discussion Paper No. 2)

The "Working Bibliography of Mathematical Geography" is precisely that. There are no annotations, no subject groupings, and even the usual categorization into "book," "article," and "other" is missing. The accuracy of the entrees have not been checked in all cases. As are all works of this nature, the "Working Bibliography" is only partially complete. The virtues are that the working mathematical geographer has access to the bibliographies of his fellow research workers. The references are usually accurate enough so that the research worker can at least find them, which is taken here to be the primary function of a bibliography, not bibliographic perfection.

The bibliography has been compiled from several sources, primarly reading lists contributed by many mathematical geographers for this purpose and references cited in current literature. If you find such a bibliography of use please send in your course reading lists and other references and if and when we have accumulated enough material we will issue a second edition of this clearinghouse for working mathematical geographers.

Please send all additions, corrections, and comments to:

Marc Anderson
Department of Geography
Wayne State University
Detroit 2, Michigan

A Working Bibliography
of
Mathematical Geography

by Marc Anderson

ABBOTT, E. A. Flatlands: A Romance of Many Dimensions. New York: Dover, 1952 (orig. pub. 1884).

Abrahamson, Hans. "Jordbrukets Mekanisering en Detaljundersokining (The Mechanization of Agriculture - A Detailed Study)." Svensk Geografisk Arsbok, Vol. 33 (1957), 7-28.

Ackerman, Edward A. Geography as a Fundamental Research Discipline. Chicago: University of Chicago, Department of Geography, Research Paper No. 53, June, 1958.

Adams, O. C. General Theory of Equivalent Projections. Washington: Government Printing Office, Coast and Geodetic Survey Special Publication 236, 1945.

Ajo, Reino. Contributions to "Social Physics." Lund, Sweden: Royal University of Lund, Lund Studies in Geography, Series B., Human Geography, No. 11, 1953.

__________. An Analysis of Automobile Frequencies in a Human Geographic Continuum. Lund, Sweden: Royal University of Lund, Lund Studies in Geography, Series B., Human Geography, No. 15, 1955.

__________. "An Approach to Demographical System Analysis." Economic Geography, Vol. 38, No. 4 (1962), 359-71.

__________. "Telephone Call Markets." Fennia, Vol. 86 (1962), 1-34.

__________. "Fields of Population Change Around the 60th Parallel Capitals and Maritime Cities - Oslo - Stockholm - Helsinki about 1960." Acta Geographica, Vol. 17 (1963), 1-19.

Alampiyev, P. M., et. al. "Letter of Protest Regarding Yu. G. Saushkin's Article in Economic Geography, 1962, No. 1." Soviet Geography: Review and Translation, Vol. 4, No. 1 (January, 1963), 60-62.

Alexander, John W. "An Isarithmic-dot Population Map." Economic Geography, Vol. 19 (1943), 431-432.

__________. Economic Geography. Englewood Cliffs, New Jersey: Prentice-Hall, 1963.

__________ and Lindberg, James B. "Measurements of Manufacturing: Coefficients of Correlation." Journal of Regional Science, Summer, 1961, 71-81.

__________ and Zahorchak, G. A. "Population-density Maps of the United States: Techniques and Patterns." The Geographical Review, Vol. 33 (1943) 457-466.

Alexandersson, Gunnar. The Industrial Structure of American Cities. Lincoln, Nebraska: University of Nebraska Press, 1956.

Allen, P. and Krumbein, W. C. "Secondary Trend Components in the Top Ashdown Pebble Bed, A Case History." Journal of Geology, Vol. 70 (1962), 507-538.

Alonso, William. "A Theory of the Urban Land Market." Papers and Proceedings, Regional Science Association, Vol. 6 (1960), 149-157.

Andersen, Nels. "Intermetropolitan Migration: A Correlation Analysis." American Journal of Sociology, Vol. 61, No. 5 (1956), 459-462.

Anderson, Marc, Compiler. "A Working Bibliography of Mathematical Geography," Michigan Inter-University Community of Mathematical Geography, Discussion Paper No. 2 (1963).

Anderson, Theodore R. "Intermetropolitan Migration: A Comparison of the Hypotheses of Zipf and Stouffer." American Sociological Review, Vol. 20, No. 3 (1955), 287-291.

Anerson, R. L. and Bancroft, T. A. Statistical Theory in Research. New York: McGraw-Hill, 1952.

Anonymous. Geography. Vol. 1, No. 1 (April, 1962) (Unpublished).

Anuchin, V. A. "On the Criticism of the Unity of Geography." Soviet Geography: Review and Translation, Vol. 3, No. 7 (1962), 22-39.

Argunov, B. I. and Skornyakov, L. A. Configuration Theorems. In the series: Topics in Mathematics, Boston: D. C. Health and Company, (to be published).

Arnold, B. H. Intuitive Concepts in Elementary Topology. Englewood Cliffs, New Jersey: Prentice-Hall, 1962.

Aron, Raymond, et. al. "The Diffusion of Ideologies, Part I." Confluence, Vol. 2, No. 1 (March, 1953), 3-65.

Artle, R. Studies of the Structure of the Stockholm Economy. Stockholm: Stockholm School of Economics, 1959.

Austin, T. L., Fagen, R. E., Penney, W. F., and Riordan, John. "The Number of Components in Random Linear Graphs." Ann. Math. Statist., Vol. 30 (1959), 747-754.

Avery, G. S. "Growth of the Tobacco Leaf." American Journal of Botany, Vol. 20 (1933).

Azzi, Girolamo. Geografia Teorica. Bologna: Casa Editrice R. Patron, 1961.

BACHI, Roberto. "Statistical Analysis of Geographic Series." Bulletin de l'Institut International de Statistique, Vol. 36 (1958), 229-240.

Bailey, Norman T. J. The Mathematical Theory of Epidemics. New York: Hafner Publishing Co., 1957.

Bain, Joe S. "Economies of Scale, Concentration, and the Condition of Entry in Twenty Manufacturing Industries," The American Economic Review, Vol. 44 (1954), 15-39.

Bakker, J. P. and Le Heux, J. W. N. "Theory on Central Rectilinear Recession of Slopes, I and II," Koninklijke Nederlandsche Akademie Van Wetenschappen Proceedings, Vol. 50, No. 8 and 9 (1947).

____________. and ________________. "Theory on Central Rectilinear Recession of Slopes, III and IV," Koninklijke Nederlandsche Akademie Van Wetenschappen Proceedings, Vol. 50, No. 7 and 9 (1950).

Balinski, M. L. "Fixed-cost Transportation Problems," Naval Research Logistics Quarterly, Vol. 8 (1961), 41-54.

Ball, W. W. Rouse. Mathematical Recreations and Essays. London: Macmillan and Co, Ltd., 1940.

Barger, Harold. The Transportation Industries, 1889-1946: A Study of Output, Employment and Productivity. New York: National Bureau of Economic Research, 1951.

Barlow, W. "Probable Nature of the Internal Symmetry of Crystals," Nature (1883), 186-188.

Barnes, J. A. and Robinson, A. H. "A New Method for the Representation of Dispersed Rural Population," The Geographical Review, Vol. 30 (1940), 134-137.

Barzel, Yoram, Grunfeld, Yehuda, Morrill, R. L., Taaffe, E. J., and Warner, S. L. Transportation Geography Research, a report prepared for the U.S. Army Transportation Research Command, Fort Eustis, Virginia by the Transportation Center at Northwestern University, Evanston, Illinois, July 1, 1960 under Contract DA 44-177-TC-574.

Baskin, Carlisle W. "A Critique and Translation of Walter Christaller's Die zentralen Orte in Süddeutschland," Charlottesville: University of Virginia, 1957 (unpublished Doctoral dissertation).

Baulig, H. "Morphometrie," Ann. de Geographie, No. 369 (1959), 385-408.

Baumol, William J. and Ide, Edward A. "Variety in Retailing," Management Science Vol 3 (1956), 93-101.

Beckmann, Martin J. "A Continuous Model of Transportation," Econometrica, Vol. 20 (1952), 643-660.

________________. "The Partial Equilibrium of a Continuous Space Market," Weltwirtschaftliches Archiv, Vol. 71, No. 1 (1953), 73-89.

________________. "The Economics of Location," Kyklos, Vol. 8, No. 4 (1955), 416-421.

________________. "Some Reflections on Lösch's Theory of Location," Papers and Proceedings, Regional Science Association, Vol. 1 (1955), N1-N9.

Beckmann, Martin J. "On the Equilibrium Distribution of Population in Space," Bulletin of Mathematical Biophysics, Vol. 19 (1957), 81-90.

________________. "City Hierarchies and the Distribution of City Size," Economic Development and Cultural Change, Vol. 6 (1958), 243-248.

________________. "Principles of Optimum Location of Transportation Networks," presented at the Symposium on Quantitative Problems in Geography, sponsored by the Office of Naval Research, Chicago, May 5-6, 1960.

________________ and Koopmans, T. C. "Assignment Problems and the Location of Economic Activities," Econometrica, Vol. 25, No. 1 (1957), 53-76 (Cowles Foundation Paper No. 108.).

________________, McGuire, C. B., and Winsten, C. B. Studies in the Economics of Transportation. New Haven: Yale University Press, 1956. (Part I: A Study of Highway Transportation, 1-110; Part II: A Study of Railroad Transportation, 111-218).

________________ and Marschak, T. "An Activity Analysis Approach to Location Theory," Second Symposium, Linear Programming 1955, 331-380 (Cowles Foundation Paper No. 99).

Behrens, Svend E. Morfometriska, morfogenetiska och tektonoska studier av de nordvastskanska urbergsasarna, sarskilt Kullaberg. Lund, Sweden: Lund Studies in Geography, Series A., Physical Geography, No. 5, 1953.

Bellman, Richard E. and Dreyfus, Stuart E. Applied Dynamic Programming. Princeton, New Jersey: Princeton University Press, 1962.

Benvenuti, B. Farming in Cultural Change. Wageningen, Netherlands: Agricultural University, 1960 (Doctoral dissertation).

Berge, Calude. "Two Theorems in Graph Theory," Proceedings, Vol. 43 (1957), 842-844.

Bergson, Henri. An Introduction to Metaphysics. New York: The Liberal Arts Press, 1950.

Berry, Brian J. L. "A Note Concerning Methods of Classification," Annals, Association of American Geographers, Vol. 48 (1958), 300-303.

________________. "On the Relations of Statistics, Scientific Method, and the Science of Geography," 1958 (unpublished).

________________. "Further Comments Concerning 'Geographic' and 'Economic' Economic Geography," The Professional Geographer, Vol. 11, No. 1, Part 1 11-12.

________________. "Ribbon Developments in the Urban Business Pattern," Annals, Association of American Geographers, Vol. 49 (1959), 145-155.

________________. "Statistical Tests of Value in Grouping Geographic Phenomena," 1959 (unpublished).

Berry, Brian J. L. "The Impact of Expanding Metropolitan Communities Upon the Central Place Hierarchy," Annals, Association of American Geographers, Vol. 50 (1960), 112-116.

________________. "An Inductive Approach to the Regionalization of Economic Development," in Norton Ginsburg, Essays on Geography and Economic Development, Chicago, Illinois: University of Chicago, Department of Geography, Research Paper No. 62, 1960, 78-107.

________________. "A Method for Deriving Multi-Factor Uniform Regions," Przeglad Geograficzny, Vol. 33, No. 2 (1961).

________________. "On Research Frontiers in Urban Geography," paper prepared for the Committee on Urbanization of the Social Science Research Council, Spring, 1961 (unpublished).

________________. "Sampling, Coding, and Storing Flood Plain Data," Agriculture Handbook No. 237, U. S. Department of Agriculture, Economic Research Service, Farm Economics Division, August, 1962.

________________. A Geography of Retail and Service Business. Englewood Cliffs; New Jersey: Prentice Hall, Inc. (forthcoming).

________________. "Grouping and Regionalizing; An Approach to the Problem Using Multivariate Analysis," in William L. Garrison, ed., Quantitative Geography. Glencoe: The Free Press (forthcoming).

________________ and Garrison, William L. "Alternate Explanations of Urban Rank-Size Relationships," Annals, Association of American Geographers, Vol. 48 (1958), 83-91.

________________ and ________________. "A Note on Central Place Theory and the Range of a Good," Economic Geography, Vol. 34 (1958), 304-311.

________________ and ________________. "Recent Developments of Central Place Theory," Papers and Proceedings, Regional Science Association, Vol. 4 (1958), 107-120.

________________ and ________________. "Cities and Freeways," Landscape, May, 1961.

________________ and Mayer, Harold M. "Design and Preliminary Findings of the University of Chicago's Studies on the Central Place Hierarchy," in Knut Norborg, ed. Proceedings of the I.G.U. Symposium in Urban Geography, Lund, 1960. Lund Studies in Geography, Series B, Human Geography No. 24, 247-252.

________________ and ________________. Comparative Studies of C. P. Studies. Chicago: University of Chicago, 1962.

________________ and Pred, Allen. Central Place Studies: A Bibliography of Theory and Applications. Philadelphia: Regional Science Research Institute, 1961.

________________ et. al. "Retail Location and Consumer Behavior," Paper and Proceedings, Regional Science Association (forthcoming).

Biekerbach, L. Conformal Mapping (Translated by Steinhardt). New York: Chelsea, 1953.

Bing, R. H. "Elementary Point Set Topology," American Mathematical Monthly, Vol. 67, No. 7, Part 2.

Blackith, R. E. "Nearest-Neighbor Distance Measurements for the Estimation of Animal Population," Ecology 39 (1958), 150-157.

Blalock, Jr., H. M. and Blalock, Ann B. "Toward a Clarification of System Analysis in the Social Sciences," Philosophy of Science, Vol. 26 (1959), 84-92.

Blant, J. M. "Microgeographic Sampling: A Quantitative Approach to Regional Agricultural Geography," Economic Geography, Vol. 35 (1959), 79-88.

Blome, Donald A. A Map Transformation of the Time-Distance Relationships in the Lansing Tri-County Area. Lansing: Michigan State University, Institute for Community Development and Services, 1963.

Blumenfeld, Hans. "The Tidal Wave of Metropolitan Expansion," Journal, American Institute of Planners, Vol. 20, No. 1 (1954), 3-14.

______________. "Are Land Use Patterns Predictable?," Journal, American Institute of Planners, Vol. 25, No. 2 (May, 1959), 61-66.

Blumenstock, D. I. "The Reliability Factor in the Drawing of Isarithms," Annals, Association of American Geographers, Vol. 43 (1953), 289-304.

Blumenthal, L. M. Theory and Applications of Distance Geometry. Oxford: Clarendon Press, 1953.

Boggs, Whittemore. "Mapping the Changing World," Annals, Association of American Geographers, Vol. 31 (1941), 119-128.

Bogue, Donald J. The Structure of the Metropolitan Community. Ann Arbor: University of Michigan, 1949.

______________ and Harris, Dorothy L. Comparative Population and Urban Research via Multiple Regression and Covariance Analysis. Chicago: University of Chicago, Population Research and Training Center, 1954.

Boltyanskii, V. G. Equivalent and Equidecomposable Figures. In the Series: Topics in Mathematics, Boston: D. C. Health and Company (forthcoming).

Borchert, John R. Belt-line Industrial Commercial Development. Minneapolis: University of Minnesota, Highway Studies, 1960.

______________. "The Twin Cities Urbanized Area: Past, Present, and Future," Geographical Review, Vol. 51, No. 1 (Jan., 1961), 47-70.

______________. The Urbanization of the Upper Midwest: 1930-1960. Minneapolis 14, Minnesota: University of Minnesota, Upper Midwest Economic Study, Urban Study No. 2, February, 1963.

Borchert, John R. and Carroll, Donald D. "Time Series Maps for the Projection of Land-Use Patterns," Presented to the 40th Annual Meeting of the Highway Research Board, 1961.

Borel, E. Space and Time. New York: Dover, 1960 (first pub. 1922).

Bormann, F. H. "The Statistical Efficiency of Sample Plot Size and Shape in Forest Economy," Ecology, Vol. 34 (1953), 474-487.

Bos, H. C. and Koyck, L. M. "The Appraisal of Road Construction Projects: A Practical Example," Review of Economics and Statistics, Vol. 43 (1961), 13-20.

Boukidis, N. A., Boyce, David, Garrison, W. L. and Tobler, Waldo. "The Location of Transportation Routes: Connections between Three Points," A portion of a report submitted October 31, 1962, to U. S. Army Transportation Research Command, Fort Eustis, Virginia, by the Transportation Center at Northwestern University under Contract: DA-44-177-TC-685.

Box, G. E. P. "The Exploration and Exploitation of Response Surfaces: Some General Considerations and Examples," Biometrics, Vol. 10 (1954), 16-61.

___________ and Youle, P. V. "The Exploration and Exploitation of Response Surfaces II: An Example of the Link Between the Fitted Surface and the Basic Mechanism of the System," Biometrics, Vol 11 (1955), 287-323.

Boyce, David E. "A Bibliography of Compilations of National Aggregates on Transportation Stocks and Flows," A portion of a report submitted October 31, 1962, to U. S. Army Transportation Research Command, Fort Eustis, Virginia, by the Transportation Center at Northwestern University under Contract: DA-44-177-TC-685.

Boyce, Ronald. "Spatial Variable and CBD Sales," Papers and Proceedings, Regional Science Association (forthcoming).

___________ and Clark, W. A. V. The Concept of Shape in Geography with an Example of its Importance in Urban Geography. Chicago: University of Illinois, Department of Geography, Bureau of Community Planning (in progress, 1963).

Bradley, R. A. "Determination of Optimum Operating Conditions by Experimental Methods," Industrial Quality Control, Vol. 15, No. 1 (1958), 16-20.

Branch, Melville C., Jr. "Planning and Operations Research," Journal, American Institute of Planners, Vol. 23, No. 4 (1957), 168-175.

Braodbent, S. R. and Hammersley, J. M. "Percolation Process, Crystals and Mazes," Proc. Cambridge Philos. Soc., Vol. 53 (1957), 629-641.

Bright, Margaret L. and Thomas, Dorothy S. "Interstate Migration and Intervening Opportunities," American Sociological Review, Vol. 6 (1941), 773-783.

Bronowski, J. "The Logic of Experiment," The Advancement of Science, Vol. 9 (1952), 289-296.

Brooks, S. H. "A Comparison of Maximum-Seeking Methods," Operations Research, Vol. 7, No. 4 (July/August, 1959), 430-457.

Broscoe, A. J. "Quantitative Analysis of Longitudinal Stream Profiles of Small Watersheds," Columbia University Technical Report, No. 18 (1959).

Brumfiel, Charles F., Eicholz, R. E. and Shanks, M. E. Geometry. Reading, Mass: Addison-Wesley Pub. Co., 1960.

Brush, John E. "The Hierarchy of Central Places in Southwestern Wisconsin," The Geographical Review, Vol. 43 (1953), 380-402.

Bryson, R. A. "Fourier Analysis of Spatial Series," in W. L. Garrison, ed., Quantitative Geography, Glenco, Ill: Free Press (forthcoming).

____________ and Kuhn, P. Half-Hemispheric 500 MB Topography Description by Means of Orthogonal Polynomials, Madison: University of Wisconsin, Department of Meteorology, Rept. No. 4, 1956.

Bunge, William, The Economic Base of the Puget Sound Region: Present and Future. Washington: State Dept. of Commerce and Economic Development, 1960.

______________. "A Theory of Neighborhood Location," Iowa City: State University of Iowa, Department of Geography, Nov., 1960.

______________. Theoretical Geography, Lund, Sweden: The Royal University of Lund, Studies in Geography, Series C, General and Mathematical Geography, No. 1, 1962.

______________. "Patterns of Location," draft appearing in Michigan Inter-University Community of Mathematical Geographers, Discussion Paper No. 3, 1963.

______________. "Spatial Relations: The Subject of Theoretical Geography," article in special volume of Voprosy Geografii, edited by Julian Saushkin, University of Moscow, forthcoming.

______________. "Toward a Geography for Humans," New University Thought (forthcoming)

Burros, R. H. "A Mathematical Theory of Static Hierarchy," Paper read before Operations Research Society of America, Annual Meeting, May, 1957.

Burton, Ian and Adkinson, William. Accessibility in Northern Ontario: An Application of Geography Theory to a Regional Highway Network. Ontario, Canada: Department of Highways (unpublished).

Busemann, Herbert. The Geometry of Geodesics. New York: Academic Press, 1955.

Byerly, W. E. Fourier Series and Spherical, Cylindrical and Ellipsoidal Harmonics. Boston: Ginn, 1893.

Bylund, Erik. "Theoretical Considerations Regarding the Distribution of Settlement in Inner North Sweden," Geografiska Annaler, Vol. 17, No. 4 (1960), 225-231.

Byron, William Glenn. Methods of Mapping Population Destribution with Dots and Densitometer Derived Isopleths. Syracuse: Syracuse University, Doctoral Dissertation, 1954.

CALEF, W. and Newcomb, R. "An Average Slope Map of Illinois," Annals, Association of American Geographers, Vol. 43 (1953), 305-313.

Carroll, J. Douglas, Jr. "The Relation of Homes to Work Places and the Spatial Patterns of Cities," Social Forces, Vol. 30 (March, 1952), 271-282.

________________ and Bevis, Howard W. "Predicting Local Travel in Urban Regions," Papers and Proceedings, Regional Science Association, Vol. 3 (1957), 183-197.

Carrothers, Gerald A. P. "An Historic Review of the Gravity and Potential Concepts of Human Interaction," Journal, American Institute of Planners, Vol. 22 (1956), 94-102.

________________. "Population Projection by Means of Income Potential Models," Papers and Proceedings, Regional Science Association, Vol. 4,(1958), 121-160.

Carter, Everett C. and Stowers, Joseph R. "A Model for the Allocation of Funds for Capacity Improvements to Urban Highway Systems," paper presented to the 42nd Annual Meeting of the Highway Research Board, 1963 (unpublished).

Cattell, R. B. "A Note on Correlation Clusters and Cluster Search Methods," Psychometrika, Vol. 9 (1944), 169-184.

Cavalli-Sforzu, L. L. "The Distribution of Migration Distances: Models and Application to Genetics," Entretiens de Monaco, The Measure of Human Displacement, May, 1962.

________________ and Conterio, F. "Analisi dello Fluttuaziune di Frequenze Geniche Mella Popolaziune della Val Parma, A. G. I., Vol. 5 (1960).

Chamberlain, W. The Round Earth on Flat Paper. Washington, D. C.: National Geographical Society, 1947.

Chamberlin, Edward. The Theory of Monopolistic Competition. Cambridge: Harvard University Press, 1933.

Chapman, C. A. "A New Quantitative Method of Topographic Analysis," American Journal of Science, Vol. 250 (1952), 428-452.

Chayes, Fleix. Petrographic Modal Analysis: An Elementary Statistical Appraisal. New York: John Wiley & Sons, Inc., 1956.

__________. "On Correlation Between Variables of Constant Sum," Journal of Geophysical Research, Vol. 65 (1960), 4185-4193.

__________ and Suzuki, Y. "Geological Contours and Trend Surfaces," Journal of Petrology, Vol. 4 (1963) (in press).

Chicago Area Transportation Study. Final Report. Chicago: 1959 and 1960 (2 vols.).

Chorley, R. J. "Climate and Morphometry," Journal of Geology, Vol. 65, No. 6 (1957), 628-635.

__________. "Aspects of the Morphometry of Poly-Cyclic Drainage Basins," Geographical Journal, Sept. 1958, 370-374.

__________. Geomorphology and General Systems Theory. Washington, D. C.: United States Geological Survey, Professional Paper, 500-B, 1962.

__________, Malm, D. E. G. and Pogorzelski, H. A. "A New Standard for Estimating Drainage Basin Shapes," American Journal of Science, 1957, 138-141.

Choynowski, M. "Maps Based on Probabilities," Journal of the American Statistical Association, Vol. 54, No. 286 (June, 1959), 385-388.

Christaller, Walter. Die zentralen Orte in Süddeutschland. Jena: G. Fischer, 1933.

__________. "Some Considerations of Tourism Location in Europe," paper to be presented at the Third European Congress, Regional Science Association, Department of Geography, University of Lund, Lund, Sweden, August 26-29, 1963.

Christensen, David E. "A Simplified Traffic Flow Map," Professional Geographer, Vol. 13, No. 2 (March, 1961), 21-22.

Clark, C. "Urban Population Densities," Journal of the Royal Statistical Society, Vol. 114 (1951), 110-116.

Clark, P. J. "Grouping in Spatial Distributions," Science, Vol. 123 (1956), 373-374.

__________ and Evans, F. C. "Distance to Nearest Neighbour as a Measure of Spatial Relationships in Populations," Ecology, Vol. 35 (1954), 445-453.

__________ and __________. "On Some Aspects of Spatial Pattern in Biological Populations, Science, Vol. 121 (1954), 397-398.

Clarke, John I. "Statistical Map Reading," Geography, Vol. 44, No. 204, Pt. 2 (April, 1959), 96-104.

Clarke, J. I. and Orell, K. "An Assessment of Some Morphometric Methods," England, University of Durham, Department of Geography, 1958.

Clarke, R. D. "An Application of the Poisson Distribution," Journal of the Institute of Actuaries, Vol. 72 (1946), 481.

Clawson, Marion, Held, R. Burnell and Stoddard, C. H. Land for the Future. Baltimore John Hopkins, 1960.

Cline, Marlin G. "Basic Principles of Soil Classifications," Soil Science, Vol. 67 (1949), 81-91.

Coates, D. R. "Quantitative Geomorphology of Small Drainage Basins of Southern Indiana," Columbia University Technical Report, No. 10, 1958.

Cochran, William G. Sampling Techniques. New York: John Wiley & Sons, Inc., 1953.

Cochran, William G. Mosteller, F. and Tukey, J. W. "Principles of Sampling," *Journal of the American Statistical Association*, Vol. 49 (1954), 13-35.

Cohen, Morris R. and Nagel, Ernest. *An Introduction to Logic and Scientific Method.* New York: Harcourt, Brace and Co., 1934.

Coleman, James and James, John. "The Equilibrium Size Distribution of Freely-forming Groups," *Sociometry*, Vol. 24, No. 1 (March 1961), 36-45.

Committee of Geographers. *Introductory Predictive Geography*. New York: John Wiley (in preparation).

Conway, R. W., Johnson, B. M. and Maxwell, W. L. "Some Problems of Digital Systems Simulation," *Management Science*, Vol. 6 (1959), 92-110.

Courant, Richard and Robbins, Herbert. *What is Mathematics?* London: Oxford University Press, 1941.

Coxeter, Harold S. M. *Regular Polytopes*. New York: Pitman Publishing Co., 1947.

__________________. *Introduction to Geometry*. New York: John Wiley and Sons, 1961.

__________________. "The Problem of Packing a Number of Equal Nonoverlapping Circles on a Sphere," *Transactions*, the New York Academy of Sciences, Ser. II, Vol. 24 (1962), 320-331.

__________________. *Projective Geometry*. (forthcoming).

__________________ and Moser, W. O. J. *Generators and Relations for Discrete Groups*. Berlin: Springer-Verlag, 1957.

Cramer, Harold. *Mathematical Methods of Statistics*. Princeton, N. J.: Princeton University Press, 1946.

Creamer, M. C. "Isolines in Population Density Mapping," *The Professional Geographer*, Vol. 10 (1958), 14-15.

Crowe, P. R. "On Progress in Geography," *The Scottish Geographical Magazine*, Vol. 54, No. 1 (1938), 4.

Curry, Leslie. "The Climatic Resources of Intensive Grassland Farming: The Wikato, New Zealand," *The Geographical Review*, Vol. 52 (1962), 174-194.

____________. "The Geography of Service Centres Within Towns: The Elements of an Operational Approach," in Knut Norborg, ed, *Proceedings of the I.G.U. Symposium in Urban Geography, Lund, 1960*. Lund, Sweden: Royal University of Lund, Studies in Geography, Series B, Human Geography, No. 24, 1962, 31-54.

____________. "Climatic Change as a Random Series," *Annals*, the Association of American Geographers, Vol. 52 (1962), 21-31.

Curry, Leslie. "Explorations in Settlement Theory: The Random Spatial Economy: Part I,: Baltimore: University of Maryland (in progress).

____________. "Irrigated Taro in New Caledonia," Etudes Melanesiennes (forthcoming).

Czuber, Emanuel. Geometrische wahrscheinliehkerten und mitteluerte. (Geometric Mean and Geometric Probabilities.)(Translated by Herman Schuerhausel, 1902.

DACEY, Michael F. "Analysis of Map Distributions by Nearest Neighbor Methods," Seattle: University of Washington, Department of Geography, Discussion Paper No. 1, March 8, 1958.

____________. "Comments on the Experimental Design of the Nearest Neighbor Statistic," Seattle: University of Washington, Department of Geography, Discussion Paper Number 13, December 12, 1958.

____________. "Selection of an Initial Solution for the Traveling Salesman Problem," Seattle: University of Washington, Department of Geography, Discussion Paper No. 16, March 5, 1959.

____________. "Analysis of Central Place Pattern by Nearest Neighbor Methods," Seattle: University of Washington, Department of Geography, Discussion Paper No. 20, May 15, 1959.

____________. "The Spacing of River Towns," Annals, Association of American Geographers, Vol. 50 (1960), 59-61.

____________. "Analysis of Central Place and Point Patterns by a Nearest Neighbor Method," in Knut Norborg, ed., Proceedings of the I.G.U. Symposium on Urban Studies. Lund,Sweden: Royal University of Lund, Studies in Geography, Series B. Human Geography No. 24, 1962, 55-76.

____________. "A Note on Some Number Properties of Hexagonal Hierarchical Plane Lattice," Wharton School of Finance and Commerce, University of Pennsylvania, 1962 (unpublished).

____________. "Order Neighbor Relations for a Class of Random Patterns in Multidimensional Space," University of Pennsylvania, Wharton School of Finance and Commerce, January 14, 1962 (mimeographed).

____________. "Another Explanation for Rank-Size Regularity," University of Pennsylvania, Wharton School of Finance and Commerce, December 21, 1962 (mimeographed).

____________. "Certain Properties of Edges on a Polyhedron in a Two Dimensional Aggregate of Polyhedra Having Randomly Distributed Nuclei," University of Pennsylvania, Wharton School of Finance and Commerce, February 5, 1963 (mimeographed).

____________. "Experimental Results of Pattern Analysis," 1963 (unpublished).

____________. "Description of Linear Patterns," in William L. Garrison, ed., Quantitative Geography. Glencoe, Ill: Free Press (forthcoming).

____________ and Tung, T. "The Identification of Randomness in Point Patterns,I," Journal of Regional Science, (1962) (forthcoming).

Dalenius, Tore. Sampling in Sweden: Contributions to the Methods and Theories of Sample Survey Practice. Stockholm: Alurquist and Wihsell, 1957.

Dantzig, G., Fulkerson, R. and Johnson, S. "Solution of a Large-Scale Traveling-Salesman Problem," Journal of the Operations Research Society of America, Vol. 2 (1954), 215-221.

Das, A. C. "Two Dimensional Systematic Sampling and the Associated Stratified and Random Sampling," Sankhya, Vol. 10 (1950), 95-108.

Davis, Harold Thayer. The Analysis of Economic Time Series. Bloomington, Ind.: Principia Press Inc., 1941.

Dawson, K. R. and Whitten, E. H. T. "The Quantitative Mineralogical Composition and Variation of the Lacorne, La Motte, and Preissac Granitic Complex, Quebec, Canada," Journal of Petrology, Vol. 3 (1962), 1-37.

De Meirleir, M. J. "Industry Location Factors," paper to be presented at the Third European Congress, Regional Science Association, Department of Geography, University of Lund, Lund, Sweden, August 26-29, 1963.

Dempsey, Bernard W. The Frontier Wage. Chicago: Loyola University Press, 1960.

Dennis, Jack Bonnell. Mathematical Programming and Electrical Networks. New York: John Wiley & Sons, Inc., 1959.

Dickie, H. F. "The Use of Logarithmic Paper for Plotting Geographical Statistics," Geography, Vol. 24 (1939), 126-130.

Dodd, Stuart C. "The Interactance Hypothesis; A Gravity Model Fitting Physical Masses and Human Behavior," American Sociological Review, Vol. 15 (1950), 245-256.

____________. "Diffusion is Predictable: Testing Probability Models for Laws of Interaction," American Sociological Review, Vol. 20 (August, 1955).

____________ and Pitts, Forrest R. "Proposals to Develop Statistical Laws of Human Geography," Proceedings I.G.U. Regional in Meetings in Japan, 1957 (1959), 302-309.

Dohrs, Fred E. "Measurement of Location in Political Geography," paper prepared as a consultant in Political Geography, Naval War College (in progress).

Dorfman, Robert, Samuelson, Paul A. and Solow, Robert M. Linear Programming and Economic Analysis. New York: McGraw-Hill, Inc., 1958.

Douglas, E. M. Boundaries, Areas, Geographic Centres, and Altitudes of the United States and of the Several States. Washington, D. C.: U. S. Geological Survey, Bulletin 817, 1930.

Duguid, A. M. "Structural Properties of Switching Networks," Progress Report No. 7, issued by the Division of Applied Mathematics of Brown University, April 2, 1959.

____________. "Flows in Constricted Networks," Progress Report No. 8, issued by the Division of Applied Mathematics of Brown University, October 13, 1959.

Duncan, Beverly. "Factors in Work-Residence Separation: Wage and Salary Workers, Chicago, 1951," *American Sociological Review*, Vol. 21, No. 1 (1956), 48-56.

_______________. "Intra-Urban Population Movement," in P. K. Hatt and J. Reiss, Jr., eds., *Cities and Society*. Glencoe, Ill.: The Free Press, 1957, 297-308.

Duncan, O. D. "Urbanization and Retail Specialization," *Social Forces*, Vol. 30 (1951-52), 267-271.

___________. "Community Size and the Rural-Urban Continuum," in P. K. Hatt and A. J. Reiss, Jr. eds., *City and Society*. Glencoe, Ill.: The Free Press, 1957, 35-45.

___________. "Optimum Size of Cities," in P. K. Hatt and A. J. Reiss, Jr. eds., *City and Society*. Glencoe, Ill.: The Free Press, 1957, 759-772.

___________. "The Measurement of Population Distribution," *Population Studies*, Vol. 11 (1957-58), 27-45.

___________, Cuzzort, Ray P. and Duncan, Beverly. *Statistical Geography: Problems in Analyzing Areal Data*. Glencoe, Ill.: The Free Press, 1961.

___________ and Duncan, Beverly. "A Methodological Analysis of Segregation Indexes," *American Sociological Review*, Vol. 20 (1955), 210-217.

Dunn, Edgar S. *The Location of Agricultural Production.* Gainesville: University of Florida Press, 1954.

Durden, C. D. *Some Geographic Aspects of Motor Vehicle Travel in Rural Areas*. Seattle: University of Washington, 1955 (Doctoral Dissertation).

Durden, Dennis and Marble, Duane F. "The Rise of Theory in CBD Planning," *Journal* American Institute of Planners, Vol. 27, No. 1 (1961), 10-16.

"Editor's Note on the Center of Population and Point of Mininum Travel," (N.S.) *Journal*, American Statistical Association, Vol. 25 (1930), 447-452.

Edmonson, Monro S. "Hybrid Corn and the Economics of Innovation," *Science*, Vol. 132, No. 3422 (July 29, 1960), 275-280.

_______________. "Neolithic Diffusion Rates," *Current Anthropology*, Vol. 2, No. 2 (April, 1961), 71-102.

Eells, W. C. "The Center of Population - A Prophecy and Its Fulfillment," *Scientific Monthly*, Vol. 20 (1925), 78-84.

Eisenhart, L. P. *A Treatise on the Differential Geometry of Curves and Surfaces*. New York: Dover, 1960 (orig. pub. 1909).

Enke, Stephen. "Equilibrium Among Spatially Separated Markets: Solution by Electric Analogue," *Econometrica*, Vol. 19 (1951), 40-47.

Erdos, P. and Renyi, A. "On the Evolution of Random Graphs," *Magyar Tudomanyos Akademia Matematikai Kutato Intezetenek Kozlemenyeir* (*Publ. of the Math. Institute of Hungarian Acad. of Sci.*) Vol. 5, Series A, Fasc. 1-2 (1960), 17-61.

Ezekiel, Mordacai. "Factors Affecting Farmers' Earnings in Southeastern Pennsylvania Washington, D. C.: U. S. Dept. of Agric. Bull. No. 1400, 1926.

FELLER, William. An Introduction to Probability Theory and its Applications. New York: John Wiley and Sons, Inc., 1950.

Fellman, Jerome D. "Pre-Building Growth Patterns of Chicago," Annals, Association of American Geographers, Vol. 47 (1957), 59-82.

Few, L. "The Shortest Path and the Shortest Road Through n Points," Mathematics, Vol. 2 (1955), 141-144.

Finch, Vernor C. "Geographic Surveying" and "Montfort--A Study in Landscape Types in Southwestern Wisconsin," in Charles C. Colby, ed., Geographic Surveys, Parts I & II, Chicago.: The Geographical Society of Chicago, Bulletin No. 9., The University of Chicago Press, 1933.

Finney, D. J. "Random and Systematic Sampling in Timber Surveys," Forestry, Vol. 22 (1948), 64-99.

Fisher, Ronald A. "Mathematics of a Lady Tasting Tea," in James R. Newman, ed., The World of Mathematics. New York: Simon & Schuster, 1956, Vol. 3.

_______________ and Hoblyn, T. N. "Maximum and Minimum-Correlation Tables in Comparative Climatology," Geografiska Annaler, Vol. 10 (1928), 267-281.

Fisher, W. D. "Economic Aggregation as a Minimum Distance Problem," Kansas State College, 1957 (unpublished paper).

___________. "On Grouping for Maximum Homogeneiry," Journal of the American Statistical Association, Vol. 53 (1958), 789-98.

Fisk, Harold N. Geological Investigation of the Alluvial Valley of the Lower Mississippi River, for the U. S. Army, Corps of Engineers, Mississippi River Commission, 1944.

Flood, Merrill M. "The Traveling Salesman Problem," Journal of the Operations Research Society of America, Vol. 4 (1956), 61-75.

Florence, P. S., Fritz, W. G. and Gilles, R. C. "Measures of Industrial Distribution in Industrial Location and National Resources. Washington, D. C.: National Resources Planning Board, 1943.

Foley, Donald L. "The Daily Movement of Population into Central Business Districts," American Sociological Review, Vol. 17 (Oct. 1952), 538-543. Reprinted in Readings in Urban Geography, ed. by Mayer and Kohn, pp. 447-453.

______________. "Urban Daytime Population: A Field for Demographic-Ecological Analysis," Social Forces, Vol. 32 (May, 1954), 323-330.

Francheschini, G. A. "A Method for Determining the Average Precipitation Between Curved Isohyets," Transactions, American Geophysical Union, Vol. 39, No. 2 (April, 1958), 273-277.

Frank, Philipp. Philosophy of Science. Englewood Cliffs, N. J.: Prentice-Hall Inc., 1957.

Fratar, Thomas J. "Vehicular Trip Distribution by Successive Approximations," Traffic Quarterly, Vol. 8, No. 1 (Jan., 1954), 53-65.

Friedrich, Peter. Die Variationsrechnung als Planungsverfahren der Stadt-und Landesplanung. In series: Veröffentlichungen der Akademie für Raumforschung und Landesplanung, Raumforschung und Landesplanung Abhandlungen, Vol. 32, Bremen-Horn, Germany: W. Dorn, 1956.

Fuchs, R. J. "Intraurban Variations of Residential Quality," Economic Geography, Vol. 36, No. 4 (October, 1960), 313-325.

Furfey, P. H. "Note on Lefever's Standard Deviational Ellipse," American Journal of Sociology, Vol. 33 (1927), 94-98.

Furman, A. Ye. "On Interrelationships between Natural and Social Laws," Soviet Geography: Review and Translation, Vol. 3, No. ? (1962), 49-54.

Fwifey, P. H. "Note on Lefever's Standard Deviational Ellipse," American Journal of Sociology, Vol. 33 (1927), 94-98

GARDNER, Martin. "The Celebrated Four-Color Map Problem of Topology," Mathematical Games, Scientific American, Vol. 203 (September, 1960), 218-226.

_______________. "More About the Shapes that can be Made with Complex Dominoes," Mathematical Games, Scientific American, Vol. 203 (November, 1960), 186-194.

_______________. "Fiction about Life in Two Dimensions," Mathematical Games, Scientific American, Vol. 207 (July, 1962), 144-152.

_______________. "Curves of Constant Width, One of Which Makes It Possible to Drill Square Holes," Mathematical Games, Scientific American, Vol. 208 (February, 1963), 148-156.

Garrison, William L. "Rural Roads," Chapter in The Nature of Highway Benefits. Seattle: Washington State Council for Highway Research, 1954.

_______________. "The Spatial Impact of Transport Media: Studies of Rural Roads," Papers and Proceedings, Regional Science Association, Vol. I (1955).

_______________. "Allocation of Road and Street Costs," Part 4 of The Benefits of Rural Roads to Rural Property. Seattle: Washington State Council for Highway Research, 1956.

_______________. "Applicability of Statistical Inference to Geographical Research," The Geographical Review, Vol. 46 (1956), 427-429.

_______________. "Estimates of the Parameters of Spatial Interaction," Papers and Proceedings, Regional Science Association, Vol. 2 (1956), 280-288.

_______________. "Verification of a Location Model," in Malcolm J. Proudfoot Memorial Volume. Evanston, Illinois: Northwestern University, Studies in Geography, No. 2 (1956).

Garrison, William L. "Spatial Structure of the Economy," Annals, Association of American Geographers, Vol. 49,(1959); Part II, Ibid., Vol. 49 (1959); Part III, Ibid., Vol. 50 (1960).

__________. "Connectivity of the Interstate Highway System," Papers and Proceedings, Regional Science Association, Vol. 6 (1960), 121-137.

__________. "Notes on the Simulation of Urban Growth and Development," Occasional Papers, Canadian Association of Geographers, British Columbia Dominion, No. 1 (1960), 1-11.

__________. On the Flow of Water in Rivers. Seattle: University of Washington, Department of Geography, 1960 (unpublished manuscript available from Department).

__________. "Toward a Determination Model of Land Development at Freeway Interchanges," in Land Uses at Freeway Interchanges. Seattle: University of Washington, 1960, 53-73 (mimeographed).

__________. "Supply and Demand for Land at Highway Interchanges," presented at the 40th annual meeting, Highway Research Board, January 9-13, 1961 published in Land Use and Development at Highway Interchanges: A Symposium. Highway Research Board, Bulletin 288. Washington, D. C.: National Academy of Sciences, National Research Council, Highway Research Board, 1961, 61-66.

__________. "Transportation Forecast and Prediction Study," Paper read at Fort Eustis, Virginia, October 18, 1961, published in Proceedings, U. S. Army Panel on Environmental Research, 1961.

__________. Land Uses in the Vicinity of Freeway Interchanges - Models of Land Use Development and Related Traffic Flows. Seattle, Washington: University of Washington, Transportation Research Group, 1961 (mimeographed).

__________. "Intra and Interurban Transportation Networks," in Forest R. Pitts, ed., Urban Systems and Economic Development: Papers and Proceedings. Eugene, Oregon: University of Oregon, School of Business Administration, 1962.

__________. "Needed Additions to Central Place Theory," paper read at the Association of American Geographers meeting, April 22, 1962, Miami Beach, Florida.

__________. "Toward Simulation Models of Urban Growth and Development," in Knut Norborg, ed., Proceedings of the I.G.U. Symposium in Urban Geography, 1960. Lund, Sweden: Royal University of Lund, Studies in Geography, Series B, Human Geography No. 24, 1962, 91-108.

__________. "The Location of Transportation Routes," paper to be presented at the Third European Congress, Regional Science Association, Department of Geography, University of Lund, August 26-29, 1963.

__________, Berry, Brian J. L, Marble, Duane F., Nystuen, John D. and Morrill, Richard L. Studies of Highway Development and Geographic Change. Seattle: University of Washington Press, 1959.

Garrison, William L., Horwood, Edgar M. and Marble, Duane F. Progress Report on a Study of Land Development Problems at Freeway Interchanges. Seattle, Washington University of Washington, Highway Economic Studies, 1960.

_______________ and Marble, Duane F. "The Spatial Structure of Agricultural Activities," Annals, Association of American Geographers, Vol. 47 (1957), 137-144.

_______________ and _______________. "Analysis of Highway Networks: A Linear Programming Formulation," Highway Research Board Proceedings, Vol. 37 (1958), 1-17.

_______________ and _______________. "Approaches to the Development of a Forecasting Capability for National and Regional Transportation Systems," A portion of a report to be submitted October 31, 1962, to U. S. Army Transportation Research Command, Fort Eustis, Virginia, by the Transportation Center at Northwestern University under Contract DA-44-177-TC-685, Transportation Geography Study (unpublished).

_______________ and _______________. The Structure of Transportation Networks, a report to the U.S. Army Research Command by the Transportation Center at Northwestern University, Evanston, Illinois, under contract DA-44-177-TC-685, (October 31, 1961), published by the Army Transportation Corps and available from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D. C., listed under "Transportation" in Key Words Index to U.S. Government Technical Reports, Vol. 1, No. 12, 1962, page A 98.

_______________ and Marts, Marion E. Geographic Impact of Highway Improvement. Seattle: University of Washington, Department of Geography and Department of Civil Engineering, 1958.

_______________ and _______________. Influence of Highway Improvements and Urban Land: A Graphic Summary. Seattle: University of Washington, Department of Geography and Department of Civil Engineering, 1958.

_______________ and Pitts, Forrest R. "The Analysis of Geographic Problems," Proceedings of IGU, Regional Conference in Japan 1957, 1959, 320-323.

Gauss, L. F. General Investigations on Curved Surfaces. Princeton, 1902, (first publication 1827).

Gauthier, Howard. "Sampling Techniques in Land Use Analysis," Northwestern Universit Department of Geography, Spring, 1960 (unpublished paper).

Geary, R. C. "The Contiguity Ratio and Statistical Mapping," The Incorporated Statistician, Vol. 5 (1954), 115-146.

Gerlach, Arch C. "Marketing Maps, Their Sources and Uses," Advances in Chemistry Series, No. 10 (1954), 100-106.

Ghosh, B. "Topographic Variation in Statistical Fields," Calcutta Statistical Association Bulletin, Vol. 2 (1949), 11-28.

_______. "Random Distances Within a Rectangle and Between Two Rectangles," Calcutta Mathematical Society, Bulletin, Vol. 43 (1951), 17-24.

Gilbert, E. N. "Random Plane Networks," Bell Telephone System, Technical Publication Monograph 4121. Published in Journal of Society for Industrial and Applied Mathematics, Vol. 9 (1961) 533-543.

Gilbert, G. K. "The Convexity of Hilltops," Journal of Geology, Vol. 17 (1909), 344-350.

__________. "The Transportation of Debris by Running Water," U.S.G.S. Prof. Paper 86, 1914.

Gini, C. and Galvani, L. "Di talune estensioni dei concetti di media ai coratteri qualitatini," Metran, Vol. 8 (1929), 136-138. (Translated by A. J. Lotka, Journal, American Statistical Association, Vol. 25 (1930), 448-450.)

Ginsburg, Norton, ed. Essays on Geography and Economic Development. Chicago: University of Chicago, Department of Geography, Research Paper No. 62, 1960.

Glock, W. S. "Available Relief as a Factor of Control in the Profile of a Land Form," Journal of Geology, Vol. 40 (1932), 74-83.

Gluss, Brian. "An Alternative Solution to the "Lost at Sea" Problem," Naval Research Logistics Quarterly, Office of Naval Research, Vol. 8, No. 1 (March, 1961), 117-121.

__________. "The Minimax Path in a Search for a Circle in a Plane," Naval Research Logistics Quarterly, Office of Naval Research, Vol. 8, No. 4 (December, 1961), 357-360.

Godlund, Sven. Bus Service in Sweden. Lund, Sweden: Royal University of Lund, Lund Studies in Geography, Series B, Human Geography, No. 17, 1956.

__________. The Function and Growth of Bus Traffic Within the Sphere of Urban Influence. Lund, Sweden: Royal University of Lund, Lund Studies in Geography, Series B, Human Geography, No. 18, 1956.

Goldberg, Michael. "Rotors in Polygons and Polyhedra," Mathematics of Computation, Vol. 14 (1960), 229-239.

Goldberg, Samuel. Introduction to Difference Equations. New York: John Wiley and Sons, Inc., 1958.

Goldman, T. A. "Efficient Transportation and Industrial Location," Papers and Proceedings, Regional Science Association, Vol. 4 (1958), 91-106.

Golovina, L. I. and Yaglom, I. M. Induction in Geometry. In series: Topics in Mathematics, Boston: D. C. Heath and Company, (forthcoming).

Goodman, Leo A. and Kruskal, William H. "Measures of Association for Cross Classifications," Journal, American Statistical Association, Vol. 49 (1954), 732-764.

Gould, Peter R. "The Geographical Application of Nearest Neighbor Theory," Evanston, Illinois: Northwestern University, Department of Geography, 1957 (unpublished paper).

Grant, Fraser. "A Problem in the Analysis of Geophysical Data," Geophysics, Vol. 22 (1957), 309-343.

Greegg, J. R. The Language of Taxonomy: An Application of Symbolic Logic to the Study of Classifactory Systems. New York: Columbia University Press, 1954.

Gregory, S. Statistical Methods and the Geographer. Cambridge, England: W. Heffer & Sons, Ltd., 1963.

Grieg-Smith, P. Quantitative Plant Ecology. London: Butterworths Scientific Publications, 1957.

Griffin, Eldon. China's Railways as a Market for Pacific Northwest Products: A Study of a Phase of the External Relations of a Region. Seattle: University of Washington, Bureau of Business Research, College of Economics and Business, 1946.

Griffin, F. L. The Center of Population for Various Continuous Distributions of Population Over Areas of Various Shapes," Metron, Vol. 11, No. 1 (1933), 11-15.

Griliches, Zvi. "Hybrid Corn: An Exploration in the Economics of Technological Change," Econometrica, Vol. 25 (October, 1957), 501-522.

Gross, Oliver. "Some Special Search Problems," Talk presented at October meeting of the Institute of Radio Engineers Professional Group on Information Theory, October 12, 1959, in Lost Angeles. Reproduced by the Rand Corporation, Santa Monica, California.

Guilford, J. P. Psychometric Methods (second edition). New York: McGraw-Hill, 1954

Guillemin, Ernst A. Introductory Circuit Theory. New York: John Wiley & Sons, Inc. 1953.

Gunnarsson, Olof. "Optimal Linjestrackning av Trafikleder (Optimal Route Location of Main Routes," Teknish Tidskrift, Vol. 40 (1960), 689-693.

Gupta, S. D. "Methods of Isopleth Mapping," Geographical Review of India, Vol. 19 (September, 1957), 10-14.

HAGERSTRAND, Torsten. The Propagation of Innovation Waves. Lund, Sweden: The Royal University of Lund, Studies in Geography, Series B, Human Geography, No. 4, 1952.

__________, Innovationsförloppet Ur Korologisk Synpunkt. Lund, Sweden: Gleerupska University, Bokhanden, 1953.

__________. Migration in Sweden. Lund, Sweden: Royal University of Lund, Studies in Geography, Series B, Human Geography No. 13, 1957.

__________. "On Monte Carlo Simulation of Diffusion," in William L. Garrison, ed., Quantitative Geography,. New York: Atherton Press (or Glencoe Free Press), 1963 (forthcoming).

Hagood, M. J. et. al. "An Examination of the Use of Factor Analysis in the Problem of Sub-Regional Delineation," Rural Sociology, Vol. 6 (1941) 216-233.

__________. ______. "Statistical Methods for the Delineation of Regions Applied to Data on Agriculture and Population," Social Forces, Vol. 21 (1943), 287-297.

__________. and Price, Daniel O. Statistics for Sociologists. New York: Henry Holt, 1952.

Hammersley, J. M. "The Distribution of Distance in Hypersphere," Annals Mathematical Statistics, Vol. 21 (1950), 447-452.

Hannerberg, D., et. al., eds. Migration in Sweden: A Symposium. Lund, Sweden: Royal University of Lund, Studies in Geography, Series B, Human Geography, No. 13, 1957.

Hansen, Walter G. "How Accessibility Shapes Land Use," Journal, American Institute of Planners, Vol. 15 (1959), 73-76.

Hansen, Willard B. "An Approach to the Analysis of Metropolitan Residential Extension, Journal, Regional Science Association, Vol. 3, No. 1 (1961), 37-55.

Hararay, Frank and Norman, Robert Z. Graph Theory as a Mathematical Model in Social Science. Ann Arbor: University of Michigan, Institute for Social Research, 1953.

____________ and ______________. The Theory of Graphs. Reading, Mass.: Addison-Wesley (forthcoming).

Harling, John. "Simulation Techniques in Operations Research - A Review," Operations Research, Vol. 6, No. 3 (May-June, 1958), 307-319.

Harris, Britton. "Some Problems in the Theory of Inter-Urban Location," Operations Research, Vol. 9 (1961), 695-721.

_____________. "Experiments in Projections of Transportation and Land Use," Traffic Quarterly, April, 1962, 305-319.

Harris, Chauncey D. "A Manufacturing View of the United States," Chicago: University of Chicago, Department of Geography, 1954.

Hart, John Fraser. "Central Tendency in Areal Distributions," Economic Geography, Vol. 30, No. 1 (January, 1954), 48-59).

Hartman, G. W. and Hook, J. C. "Substandard Urban Housing in the United States: A Quantitative Analysis," Economic Geography, Vol. 32 (1956), 95-114.

Hartshorne, Richard. The Nature of Geography: A Critical Survey of Current Thought in the Light of the Past. Lancaster, Pennsylvania: Association of American Geographers, 1939.

_______________. Perspective on the Nature of Geography. Chicago: Rand McNally & Co., 1959.

Hasegawa, Norio. "Spatial Variation of Land Value and Land Use: Case Study of Sendai and Hirosaki," The Science Reports of the Tohoku University, Seventh Series (Geography), No. 12, 1963, 145-158.

Haurwitz, B. and Craig, R. Atmospheric Flow Patterns and their Representation by Spherical Surface Harmonics. (Geophysical Research Paper No. 14). Cambridge, Mass.: Cambridge Research Center, 1952.

Hauser, Philip M. "The Changing Population Pattern of the Modern City," in P. K. Hatt and A. J. Reiss, Jr., eds., Cities and Society. Glencoe, Ill.: The Free Press, 1957, 157-174.

Hawely, A. H. Human Ecology. New York: Ronald Press, 1950.

___________, and Duncan, O. D. "Social Area Analysis: A Critical Appraisal," Land Economics, Vol. 33 (1957), 337-344.

Hayford, John F. "What is the Center of an Area, or the Center of a Population?" Journal, American Statistical Association, Vol. 8, No. 58 (N.S.) (1902), 47-58.

Heald, K. L. Discussion of the Iowa Gravity Model. Traffic Distribution Program, 1960.

Henderson, James M. "The Utilization of Agricultural Land: A Regional Approach," Papers and Proceedings, Regional Science Association, Vol. 3 (1957) 99-117.

_______________. The Efficiency of the Coal Industry; An Application of Linear Programming. Cambridge, Mass.: Harvard University Press, 1958.

Herbert, J. D. and Stevens, B. "A Model for the Distribution of Residential Activity in Urban Areas," Journal of Regional Science, Vol. 2, No. 2 (1960) 21-36.

Hilbert, David. The Foundations of Geometry. LaSalle, Illinois: Open Court, 1902.

_____________ and Cohn-Vossen, Stephen. Geometry and the Imagination. Translated by P. Nemenyi. New York: Chelsea Publishing Co., 1952.

Hisle, George. "Program of Areas and Centroid," City of Detroit, Department of Public Works, Office of City Engineer (forthcoming 1963).

Hitchcock, Charles B. "The American Geographical Society - Annual Report of the Council," Geographical Review, Vol. 51 (1961), 290-303.

Hoel, Paul G. Introduction to Mathematical Statistics. New York: John Wiley & Sons, Inc., 1954.

Holloway, J. L., Jr. "Smothing and Filtering of Time Series and Space Fields," in H. E. Landsberg and J. Van Mieghem, eds., Advances in Geophysics, Vol. 4. New York: Academic Press, 1958.

Holm, P. "Simple Models for Community Expansion," paper to be read Third European Congress, Regional Science Association, Department of Geography, University of Lund, Lund, Sweden, August 26-29, 1963.

Honkala, Kauko. "Social Class and Visiting Patterns in Two Finish Villages," Acta Sociologica, Vol. 5, fasc. 1 (1960), 42-49.

Hooson, David J. M. "Methodological Clashes in Moscow," Annals, Association of American Geographers, Vol. 52 (1962), 469-475.

Hoover, Edgar M. "The Measurement of Industrial Localization," Review of Economic Statistics, Vol. 18 (1936), 162-171.

______________. The Location of Economic Activity. New York: McGraw-Hill Book Co., Inc., 1948.

Horton, R. E. "Erosional Development of Streams and their Drainage Basins: Hydrophysical Approach to Quantitative Morphology," Bull. Geol. Soc. Am., Vol. 56 , (1945), 272-370.

Horwood, Edgar M. and Boyce, Ronald R. Studies of the Central Business District and Urban Freeway Development. Seattle: University of Washington Press, 1959.

Hotelling, Harold. "A Mathematical Theory of Migration," Seattle: University of Washington, 1921, (unpublished Master of Arts thesis).

______________. "Stability in Competition," The Economic Journal, Vol. 39 (1929), 41-57.

______________. "Analysis of a Complex of Statistical Variables into Principal Components," Journal of Educational Psychology, Vol. 24 (1933), 417-441, 496-520.

Howells, William. "The Distribution of Man," Scientific American, Vol. 203 (Sept., 1960), 113-127.

Huddelston, H. F. "An Inverted Matrix Approach for Determining Crop-Weather Relationships," Biometrics, Vol. 11 (1955), 231-236.

Huff, David L. "A Topographical Model of Consumer Space Preferences," Papers and Proceedings, Regional Science Association, Vol. 6 (1960), 159-175.

Hughes, R., Jr. "Volume Estimates from Contours," Economic Geology, Vol. 54, No. 4 (July, 1959), 730-737.

Humlum, Johannes. Nogle Aktuelle Kanske Trafik-og Befolkningsproblemer. Aarhus, Denmark: Aarhus University, Series in Cultural Geography, No. 12, 1961.

______________. Problemer I Danmarks Landsplanlaegning. Aarhus, Denmark: Aarhus University, Series in Cultural Geography, No. 68.

Hunter, J. S. "Determination of Optimum Operating Conditions by Experimental Methods Models and Methods," Industrial Quality Control, Vol. 15, No. 5 (Nov., 1958), 16-24; No. 7 (Jan., 1959), 7-15; No. 8 (Feb., 1958), 6-14.

Huntington, Ellsworth. "The Water Barriers of New York City," Geographical Review, Vol. 2 (1916).

Hurley, P. M., et. al. "Radiogenic Strontium-87 Model of Continent Formation," Journal of Geophysical Research, Vol. 67 (1962), 5315-5334.

IMBRIE, John. "Factor Analysis Program Package," New York: Columbia University, Department of Geology (forthcoming).

Ipsen, G., Christaller, W., Kollmann, W. and Makensen, R. Standort und Wohnort: Okologische Studien. Dologne: Westdeutcher Verlag, 1957. Forschungsberichte des Wirtschafts -- und Verkehrsministeriums Nordheim-Westfalen.

Isard, Walter. Location and Space-Economy. Cambridge: The Technology Press of Massachusetts Institute of Technology, 1956.

__________. Methods of Regional Analysis: An Introduction to Regional Science. Cambridge: The Technology Press of M. I. T., 1960.

Isbell, J. R. "An Optimal Search Pattern," Naval Research Logistics Quarterly, Office of Naval Research, Vol. 4, No. 4 (December, 1957), 357-359.

JEFFERSON, Mark. "The Law of the Primate City," Geographical Review, 1939, 226-232.

Jenks, G. F. and Knos, O. S. "The Use of Shading Patterns in Graded Series," Annals, Association of American Geographers, Vol. 51, No. 3 (Sept., 1961), 316-334.

Joerg, W. L. G. "The Geographic Center: It's Definition and Determination (abstract) Annals, Association of American Geographers, Vol. 6 (1915), 127-128.

Johnson, F. A. and Hixon, H. J. "The Most Efficient Size and Shape of Plot to Use for Crusing in Old-Growth Douglas-Fir Timber," Journal of Forestry, Vol. 50 (1952), 17-20.

Johnston, G. H. "Minimum Path Route Selection Program," Washington, D. C., Departmen' of Highways, Traffic Section, May, 1962.

Jolly, G. M. "The Theory of Sampling," in Dorothy Brown, ed., Methods of Surveying and Measuring Vegetation. Farnham Royal, Bucks, England: Commonwealth Agricultural Bureaux, 1954, Ch. 2.

Jones, W. D. "Ratios and Isopleth Maps in Regional Investigation of Agricultural Land Occupance," Annals, Associatión of American Geogmaphers, Vol. 20 (1930), 177-195.

KADAS, C. "The Impact of Development of Transportation on the Optimal Size of Plants and on Regional Location," Faculty of Transport Engineering, Budapest, paper to be presented at the Third European Congress, Regional Science Associatic Department of Geography University of Lund, Lund, Sweden, August 26-29, 1963.

Kahan, B. C. "Precis of The Calculus of Variations as a Method for Town and Country Planning' by P. Friedrich," Road Research Laboratory, Research Note RN/4087/ BCK, 1961.

Kalaba, R. E. and Juncosa, M. L. "Optimal Design and Utilization of Communication Networks," Management Science, Vol. 3 (1956), 33-44.

Kalesnik, S. V. "About 'Monism' and 'Dualism' in Soviet Geography," Soviet Geography Review and Translation, Vol. 3, No. 7 (1962), 3-16.

Kansky, Karel J. "International and Interregional Comparative Studies," A portion of a report to be submitted October 31, 1962, to U. S. Army Transportation Research Command, Fort Eustis, Virginia, by the Transportation Center at Northwestern University under Contract: DA-44-177-TC-685, Transportation Geography Study.

Kant, Edgar. "Suburbanization, Urban Sprawl and Commutation," in David Hannerberg, et. al., eds., Migration in Sweden. Lund, Sweden: Royal University of Lund, Studies in Geography, Series B, Human Geography, No. 13, 1956, 244-309.

Kao, R. C. "Geometric Projections of the Sphere and Spheroid," Canadian Geographer, Vol. 5, No. 3 (Autumn, 1961), 12-21.

Kazarinoff, Nicholas D. Geometric Inequalities. New York: Random House, 1961.

Kemeny, John G. A Philosopher Looks at Science. Princeton, N. J.: D. Van Nostrand Co., 1959.

______________, Laurie, J. and Thompson, Gerald L. Introduction to Finite Mathematics. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1957.

______________ and Snell, J. L. Mathematical Models in the Social Sciences. New York: Ginn and Co., 1962.

Kendall, M. G. "The Geographical Distribution of Crop Productivity in England," Journal of the Royal Statistical Society, Vol. 102 (1939), 21-62.

Kensit, H. E. M. "The Centre of Population Moves West," Canadian Geographical Journal, Vol. 9 (1934), 262-269.

Keyser, Cassius J. "The Group Concept," in James R. Newman, ed., The World of Mathematics, New York: Simon & Schuster, 1956, Vol. 3.

King, Leslie J. "Discriminatory Analysis as a Tool in Geographic Research," Paper read before the Western Lakes Section of the Association of American Geographers, Minneapolis, 1959 (unpublished).

______________. "Central Place Theory and the Spacing of Towns in the United States, Land and Livelihood, Geographical Essays in Honour of George Jobberns, 1961, 238-254.

______________. "A Multivariate Analysis of the Spacing of Urban Settlements in the United States," Annals, Association of American Geographers, Vol. 51 (1961), 222-233.

______________. "A Quantitative Expression of the Pattern of Urban Settlements in Selected Areas of the United States," Tijdschrift Voor Econ. En. Soc. Geografie, Vol. 53 (1962), 1-7.

Kitagawa, E. M. and Bogue, D. J. Suburbanization of Manufacturing Activity within Standard Metropolitan Areas. Chicago: University of Chicago, P.R.T.C., 1955.

Klimm, L. E. "The D-Line Method of Analyzing a Distribution," Technical Report 3, Contract NONR 551 (01), University of Pennsylvania, 1959.

Kolb, J. H. and Brunner, Edmund de S. A. In William F. Ogburn, ed., Study of Rural Society. Boston: Houghton Mifflin Co., 1946.

Kolosovskiy, N. N. "On the Concept of the Unity of Geography," (Letter to V. M. Chetyrkin), Soviet Geography: Review and Translation, Vol. 3, No. 7 (1962), 39-44.

Konovalenko, V. G. "The Concept of the Unity of Geography in the Solution of Basic Problems of Geography," Soviet Geography: Review and Translation, Vol. 3, No. 7 (1962), 45-49.

Koopmans, Tjalling C. and Beckmann, Martin J. "Assignment Problems and the Location of Economic Activities," Econometrica, Vol. 25 (1957), 53-76.

Kratchick, M. Alignment Charts: Construction and Use. New York: Van Nostrand, 1944.

Krebs, N. "Mass und Zahl in der Physischen Geographie. Ergebnisse und Aufgaben geographischer Forschung," Peterm. Mitt. Erganzungsheft 209, 1930.

Kreyszig, E. Differential Geometry. Toronto: University of Toronto, 1959.

Krumbein, W. C. "Experimental Design in the Earth Sciences," Transactions of the American Geophysical Union, Vol. 36 (1959), 1-11.

____________. "Statistical Analysis of Facies Maps," Journal of Geology, Vol. 63 (1955), 452-70.

____________. "Regional and Local Components in Facies Maps," Bulletin, of the American Association of Petroleum Geologists, Vol. 40, No. 9 (Sept., 1956), 2163-2194.

____________. "Trend Surface Analysis of Contour-Type Maps with Irregular Control-Point Spacing," Journal of Geophysical Research, Vol. 64, No. 7 (July, 1959), 823-834.

____________. "The Geological 'Population' as a Framework for Analyzing Numerical Data in Geology," The Liverpool and Manchester Geological Journal, Vol. 2, Part 3 (1960), 341-368.

____________. "Stratigraphic Maps from Data Observed at Outcrop," Yorkshire Geological Society Proceedings, Vol. 32, (1960), 353-366.

____________. "The Computer in Geology," Science, Vol. 136 (1962), 1087-1092.

____________. "Open and Closed Number Systems in Stratigraphic Mapping," Bulletin, American Association of Petroleum Geologists, Vol. 46 (1962), 2229-2245.

____________. "Confidence Intervals on Low-order Trend Surfaces," Journal of Geophysical Research, Vol. 68 (1963) (in press).

Krumbein, W. C. and Miller, R. L. "Design of Experiments for Statistical Analysis of Geological Data," Journal of Geology, Vol. 61 (1953), 510-532.

____________, and ___________. "A Note on Transformation of Data for Analysis of Variance," Journal of Geology, Vol. 62 (1954), 192-193.

____________ and Slack, H. A. "Statistical Analysis of Low-level Radioactivity of Pennsylvania Black Fissile Shale in Illinois," Bulletin of the Geological Society of America, Vol. 67 (1956), 739-762.

Kulldorff, Gunnar. Migration Probabilities. Lund, Sweden: Royal University of Lund, Studies in Geography, Series B., Human Geography, No. 14, 1955.

Kulldorff, Gunnar, Contributions to the Theory of Estimation from Grouped and Partially Grouped Samples. Uppsala, Sweden: 1961 (Doctoral Dissertation, Royal University of Lund, Lund, Sweden).

LA LANNE, Leon. "An Essay on the Theory of Railway Systems, Based on Observation of Facts and the Basic Laws Governing Population Distribution," Comptes Rendus Des Seances De L'Academie Des Sciences, Vol. 57 (July-December, 1863), 206-210.

Lane, E. P. Metric Differential Geometry of Curves and Surfaces. Chicago: Universit of Chicago Press, 1940.

Latham, J. P. "The Distance Relations and some Other Characteristics of Cropland Areas in Pennsylvania," Technical Report 6, Contract NONR 551 (01), University of Pennsylvania, 1959.

Lathrop, John. Compendious Treatise on the Use of Globes and Maps. Boston: Wells and Lilly and J. W. Burditt, 1821.

Lazardfeld, Paul F. Mathematical Thinking in the Social Sciences. Glencoe, Ill.: Free Press, 1954.

Learmouth, A. T. A. "Sample Survey and National Planning in India." Paper read at meetings of the Institute of British Geographers, 1960.

Leech, J. "The Problem of the Thirteen Spheres," Math. Gazette, Vol. 40 (1956), 22-23.

Leible, Otto. "Verteilung der Radioaktivitat, der Thorium-und Urangehalte im Malsburggranit Sud Schwarzwald," Zeitschrift fur Erzbergbau und Metallhuttenwesen (Stuttgart), Vol. 12 (1959), 1-4.

Leighly, J. "Meandering Arroyas of the Dry Southwest," Geographical Review, Vol. 26,(1936), 270-282.

Leopold, L. G. "Downstream Change of Velocity Rivers," Am. J. Sci., No. 251 (1953), 606-624.

____________ and Langbein, W. B. "The Concept of Entropy in Landscape Evolution," U.S.G.S. Professional Paper, 1962.

Leopold, L. B. and Maddock, T. "The Hydraulic Geometry of Stream Channels and some Physiographic Implications," U.S.G.S. Prof. Paper 252, (1953), 1-57.

__________ and Miller, J. P. "Ephemeral Streams - Hydraulic Factors and their Relation to the Drainage Net," U.S.G.S. Prof. Paper 282-A, 1956.

__________ and Wolman, M. G. "River Channel Patterns: Braided, Meandering, and Straight," U.S.G.S. Prof. Paper 282-B, 1957.

Lighthill, M. J. and Whitham, G. B. "On Kinematic Waves, II, A Theory of Traffic Flow on Long Crowded Roads," Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences, Vol. 229, No. 1178 (1955), 317-345.

Lionberger, Herbert F. Adoption of New Ideas and Practices. Ames, Iowa: Iowa State University Press, 1960.

Loeben, A. F. "Analysis of Stead Distributions by the D-Line Method," Technical Report 6, Contract NONR 551 (01), University of Pennsylvania, 1959.

Lopshits, A. M. Computation of Areas of Oriented Figures. In Series: Topics in Mathematics. Boston: D. C. Heath and Company (forthcoming).

Lorenz, M. O. "Methods of Measuring the Concentration of Wealth," American Statistic Association, New Series, Vol 9 (1904-5), 209-219.

Lösch, August. The Economics of Location (translated by William H. Woglom). New Haven: Yale University Press, 1954.

Lovgren, Ease. "The Geographical Mobility of Labour: A Study of Migrations," Geografiska Annaler, Vol. 38, No. 4 (1956), 344-394.

Lukermann, F. "Toward a More Geographic Economic Geography," The Professional Geographer, Vol. 10, No. 4 (1958), 2-10.

Lyusternik, L. A. Convex Figures and Polyhedra. In series: Topics in Mathematics. Boston: D. C. Heath and Company (forthcoming).

MAC ARTHUR, Robert H. "Oh the Relative Abundance of Bird Species," Proceedings, National Academy of Sciences, Vol. 43 (1957), 293-295.

McCarty, Harolds H., Hook, John C. and Knos, Duane S. The Measurement of Associatio in Industrial Geography. Iowa City: 1956.

__________ and Salisbury, Neil E. "Visual Comparison of Isopleth Maps as a Means of Determining Correlations between Spatially Distributed Phenomena," State University of Iowa, Department of Geography, Series No. 3, 1961.

McCracken, D. D. A Guide to FORTRAN Programming. New York: John Wiley and Sons, 1961.

McGuire, C. B. and Winsten, C. B. Studies in the Economics of Transportation. New Haven, Conn.: Yale University Press, 1956.

Mach, Ernest. The Science of Mechanics (translated by Thomas J. McCormack). LaSalle, Illinois: Open Court Publishing Co., 1960 (6th ed.).

McIntyre, D. B. Program for Computation of Trend Surfaces and Residuals of Degree 1 through 8. Claremont, California: Pomona College, Department of Geology, Seauer Laboratory, 1963.

Mackay, J. Ross. "Dotting the Dot Map: An Analysis of Dot Size, Number, and Visual Tone Density," Surveying and Mapping, Vol. 9 (1949), 3-10.

______________. "Some Problems and Techniques in Isopleth Mapping," Economic Geography, Vol. 27 (1951), 1-9.

______________. "The Alternative Choice is Isopleth Interpolation," The Professional Geographer, Vol. 5, No. 4 (1953), 2-4.

______________. "Percentage Dot Maps," Economic Geography, Vol. 29 (1953), 263-266.

______________. "Experiments with Some Symbols and Map Projections," Annals, Association of American Geographers, Vol. 44, No. 2 (1954), 225-226.

______________. "Geographic Cartography," The Canadian Geographer, Vol. 4 (1954), 1-14.

______________. "An Analysis of Isopleth and Choropleth Class Intervals," Economic Geography, Vol. 31 (1955), 71-81.

______________. "Percentage Isopleth Maps," The Professional Geographer, Vol. 7 (1955), 10-12.

______________. "Chi Square as a Tool for Regional Studies," Annals, Association of American Geographers, Vol. 48 (1958), 164.

______________. "Regional Geography: A Quantitative Approach," in Melanges geographiques canadiens offerts a Raoul Blanchard, published under the auspices of the Institute of Geography of the University of Laval, University of Laval Press, Quebec, 1959, 57-63.

______________ and Berry, Brian, J. L. "Comments on the Use of Chi Square," Annals, Association of American Geographers, Vol. 49 (1959), 89.

McKean, Roland N. Efficiency in Government Through Systems Analysis, With Special Emphasis on Water Resources Development. New York: John Wiley & Co., 1959.

McKinney, William M. "Squire Hoffman's Comet," Idaho Yesterdays, Summer, 1959.

______________. "Experimental Proofs of the Earth's Rotation," Journal of Geography, Vol. 61, No. 4 (April, 1962) 171-174.

______________. Geography Via Use of the Globe. Booklet No. 5, "Do It This Way," Series, National Council for Geographic Education, 1963.

______________. "Laboratory Demonstrations in Mathematical Geography," The Professional Geographer, Vol. 40 (1963), 7-10.

McQuitty, Louis L. "Elementary Linkage Analysis for Isolating Orthogonal and Oblique Types and Typal Relevancies," Educational and Psychological Measurement, Vol. 17, No. 2 (Summer, 1957), 207-229.

McVoy, Edgar C. "Patterns of Diffusion in the United States," American Sociological Review, Vol. 5, No. 2 (April, 1940), 219-227.

Maisel, S. J. "Factors Influencing Suburban Land Development," paper to be read at the Third European Congress, Regional Science Association, Department of Geograph University of Lund, Lund, Sweden, August 26-29, 1963.

Malcolm, D. C. "Bibliography on the Use of Simulation in Management Analysis," Operations Research, Vol. 8 (1960), 169-177.

Mandelbaum, H. "Statistical and Geological Implications of Trend Mapping with Non Orthogonal Polynomials," Journal of Geophysical Research, Vol. 68 (1963),

Manheim, M. L. The Interdependence of Transportation and Land Use Planning (forthcoming).

Mansfield, Edwin. "Technological Change and the Rate of Imitation," Paper read at the 1959 meeting of the Econometric Society in Washington, D.C.: noted on p. 672 of Econometrica, Vol. 28, No. 3 (July, 1960).

Marble, Duane F. "Transport Imputs at Urban Residential Sites," Seattle: University of Washington, Department of Geography, Discussion Paper No. 15, January 9, 1959 (Micro film print available from University Microfilms, Ann Arbor, Michigan).

______________. "The Transportation Problem," Seattle; University of Washington, Department of Geography, Discussion Paper No. 23, September 1, 1959.

______________. "Some Models of Individual Travel Behavior," in W. L. Garrison, ed., Quantitative Geography (forthcoming).

______________ and Durden, C. D. "The Role of Theory in CBD Planning," Journal, American Institute of Planners, Vol. 27 (1961), 10=16.

______________ and Nystuen, John.. "An Approach to the Direct Measurement of Community Mean Information Fields," Papers and Proceedings, Regional Science Association.(forthcoming).

Marfunin, A. S. "Some Petrological Aspects of Order-Disorder in Feldspars," Minerological Magazine, Vol. 33 (1962), 298-314.

Massey, F. J. "The Distribution of the Maximum Deviation between Two Sample Cumulative Step Functions," Annals of Mathematical Statistics, Vol. 22 (1951), 125-128.

__________. "The Kolmogorov-Smirnov Test for Goodness of Fit," Journal of the American Statistical Association, Vol. 46 (1951), 68-78.

Matern, Bertil. "Some Applications of the Theory of Geometric Probabilities," Svenska Skogsvård-sföreningens Tidskrift, Vol. 57 (1959), 453-458.

______________. "Forest Surveys and Statistical Theory Sampling: Some Recent Developments," Fifth World Forestry Congress Proceedings. Seattle, Washington: 1960, 276-281.

______________. Spatial Variation: Stochastic Models and Their Application to Some Problems in Forest Surveys and Their Sampling Investigations. Meddelanden Fran Statens Skogsforskningsinstitut, Vol. 49 (1960).

______________. "Maximum Distance Between the Seedings in Scarified Spots," in Stig Hagner, ed., Natural Regeneration Under Shelterwood Stands. Meddelanden Fran Statens Skogsforskningsinstitut, Vol. 52, No. 4 (1962), 237-239, 259-261.

Mather, E. C. "A Linear-distance Map of Farm Population in the United States," Annals, Association of American Geographers, Vol. 34 (1944), 173-180.

Mathes, G. H. "Basic Aspects of Stream Meanders," Transactions of the American Geophysical Union, Vol. 22 (1941), 632.

Matui, I. "Statistical Study of the Distribution of Scattered Villages in Two Regions of the Tonami Plain, Tpyama Prefecture," Japanese Journal of Geology and Geography, Vol. 9 (1932), 251-256.

Maxwell, J. Clerk. "On Hills and Dales," The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Vol. 40, Fourth Series (1870), 421-427.

Mayfield, Robert C. "Line Sampling and Land Use Estimation in a Complex Rural Area: An Example from a North Indian Valley," Durant, Oklahoma: Southeastern State College, Department of Geography (unpublished).

________________. "Conformations of Service and Retail Activities: An Example in the Lower Orders of an Urban Hierarchy in a Lesser Developed Area," in Knut Norborg, ed., Proceedings of the Symposium in Urban Geography, Lund, Sweden Royal University of Lund, Studies in Geography, Series B, Human Geography, No. 2 1962.

________________. "The Range of a Central Good in the Indian Punjab," Annals, Association of American Geographers, Vol. 53 (1963), 38-49.

Medawar, P. B. and Clark, W. E. Legros. Essays on Growth and Form. Oxford, England: Clarendon, 1945.

Mehnert, K. R. "Zur Geochemie der Alkalien im Tiefen Grundgebirge," Beiträge zur Mineralogie und Petrographie (Berlin), Vol. 7 (1960), 318-339.

____________, and Willgallis, A. "Die Alkaliverteilung im Malsburger Granit (Sudschwarzwald)," Jahreshefte des Geologische Landesamt im Baden-Württemberg, Vol. 5 (1961), 117-139.

Melton, M. A. "An Analysis of the Relations among Elements of Climate, Surface Properties, and Geomorphology," New York: Columbia University, Department of Geology, Vol. 8 (1957).

__________. "Correlation Structure of Morphometric Properties of Drainage Systems and their Controlling Agents," Journal of Geology, Vol. 66 (1958), 442-460.

Michle, William. "Link-Length Minimization in Networks," Journal of the Operations Research Society of America, Vol. 6 (1958), 232-243.

Mikheyeva, V. S. "An Economic-Mathematical Model of the Location of Farm Production by Regions of the Soviet Union," Soviet Geography: Review and Translation, Vol. 4, No. 3 (March, 1963), 24-29.

Miller, O. M. and Fisher, I. World Maps and Globes. New York: Essential Books, 1944

__________ and Summerson, C. H. "Slope Zone Maps," Geographical Review, Vol. 50 (1960), 194-202.

Miller, R. L. "Trend Surfaces: Their Application to Analysis and Description of Environments of Sedimentation," Journal of Geology, Vol. 64 (1956), 425-446.

__________ and Ziegler, J. M. "A Model Relating Dynamics and Sediment Pattern in Equilibrium in the Region of Shoaling Waves, Breaker Zone, and Foreshore," Journal of Geology, Vol. 66 (1958), 417-441.

Miller, Victor C. A Quantitative Geomorphic Study of Drainage Basin Characteristics in the Clinch Mountain Area Virginia and Tennessee. Columbia University, Department of Geology, Technical Report No. 3, 1953.

Mohring, Herbert D. and Harwitz, Mitchell. Highway Benefits: An Analytical Framework Evanston, Illinois: Northwestern University Press, 1962.

Monkhouse, F. J. and Wilkinson, H. R. Maps and Diagrams. London: Methuen, 1952.

Moore, Edward F. "The Shortest Path Through a Maze," (International Symposium on the Theory of Switching, April 2-5, 1957, Part II. pp. 285-292), Annals of the Computation Lab. of Harvard University, Harvard University Press, Vol. 30 (1959),

Moore, Frederick, T. "A Note On City Size Distributions," Economic Development and Cultural Change, Vol. 7, No. 4 (July, 1959), 465-466.

Moran, P. A. P. "The Interpretation of Statistical Maps," Journal of the Royal Statistical Society, Series B (Methodological), Vol. 10 (1948), 245-251.

Morisawa, M. E. "Relation of Morphometric Properties to Runoff in the Little Mill Creek, Ohio, Drainage Basin," New York: Columbia University, Department of Geology, Report No. 17, 1959.

__________. "Relation of Quantitative Geomorphology to Stream Flow in Representat Watersheds of the Appalachian Plateau Province," Columbia University Technical Report, No. 20, 1959.

Morisita, M. "Estimation of Population Density by Spacing Method," Memoirs of the Faculty of Science of Kyushu University, Series E, Vol. I (1954), 187-197.

__________. "A New Method for the Estimation of Density by the Spacing Method Applicable to Non-Randomly Distributed Populations," Physiology and Ecology, Vol. 7, No. 2 (November, 1957), 134-144.

__________. "Measuring the Dispersion of Individuals and Analysis of the Distributional Pattern," Memoirs of the Faculty of Science of Kyushu University, Series E, Biology, Vol. 2 (1959), 215-235.

Morrill, Richard L. "An Experimental Study of Trade in Wheat and Flour in the Flour Milling Industry," Seattle: University of Washington, Department of Geography, 1957 (unpublished Master's Thesis).

__________. "A Normative Model of Trade Areas and Transportation: With Special Reference to Highways and Physicians' Services," Seattle: University of Washington, Department of Geography, 1959 (unpublished Doctoral Dissertation).

__________. "The Development of Spatial Distributions of Towns in Sweden: An Historical-Predictive Approach," Annals of the Association of American Geographers, Vol. 53 (1963), 1-14.

__________. "The Distribution of Migration Distances," Papers and Proceedings Regional Science Association (forthcoming).

__________. "A Model of Interregional Movement: An Adaptation to Realistic Conditions of Imperfect Information and/or Non-Economic Human Response," in W. L. Garrison, ed., Quantitative Geography (forthcoming).

__________ and Garrison, W. L. "Projections of Interregional Patterns of Trade in Wheat and Flour," Economic Geography, Vol. 36, No. 2 (April, 1960), 116-126.

__________ and Pitts, Forest R. "Marriage, Migration and the Mean Information Field: A Study in Uniqueness and Generality" (Hectographed).

Morrissett, Irving. "The Economic Structure of American Cities," Papers and Proceedin Regional Science Association, Vol. 4 (1958), 239-259.

Morse, M. "Fields of Geodesics Issuing from a Point," Proceedings, National Academy of Sciences, 1960, 105-111.

Moses, Leon N. "A General Equilibrium Model of Production, Interregional Trade, and Location of Industry," Rev. Econ. Stat., Vol. 42 (1960), 376-397.

Mosher, Walter W. Jr. "A Capacity Restrained Algorithm for Assigning Flow to a Transport Network," ITTE, University of California, Berkeley, (in progress).

Mowrer, E. W. "The Isometric Map as a Technique of Social Research," American Journal of Sociology, Vol. 44 (July, 1938), 86-96.

Murphy, Raymond E. "The Relative Change Map," The Professional Geographer, Vol. 5, No. 2 (1953), 4-5.

________________ and Spittal, Hugh E. "Movements of the Center of Coal Mining in the Appalachian Plateaus," Geographical Review, Vol. 35 (1945), 624-633.

Murray, Cecil D. "The Physiological Principle of Minimum Work Applied to the Angel of Branching Arteries," Journal of General Physiology, Vol. 9 (1926), 835-841.

______________. "A Relationship Between Circumference and Weight in Trees and Its Bearing on Branching Angels," Journal of General Physiology, Vol. 10 (1927), 725-729.

Muth, Richard F. "Economic Change and Rural-Urban Land Use Conversions," Econometrica Vol. 29 (1961), 1-23.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Insights Into Modern Mathematics. Washington: National Council of Teachers of Mathematics, 1957 [23rd Yearbook].

Nelson, Howard J. "A Service Classification of American Cities," Economic Geography, Vol. 31 (1955), 189-210.

Newell, G. F. "Mathematical Models for Freely-Flowing Highway Traffic," Journal of the Operations Research Society of America, Vol. 3 (1955), 176-186.

Neyman, Jerzy and Scott, Elizabeth L. "On a Mathematical Theory of Populations Conceived as a Conglomeration of Clusters," Cold Spring Harbor Symposia on Quantitative Biology, Vol. 22 (1957), 109-120.

________________________________ and Shane, C. D. "Statistics of Images of Galaxies with Particular Reference to Clustering," Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. 3 (1956), 75-111.

Nir, D. "The Ratio of Relative and Absolute Silitudes of Mt. Carmel: A Contribution to the Problem of Relief Analysis and Relief Classification," Geographical Review, Vol. 47 (October, 1957), 564-569.

Nordbeck, Stig. Location of Areal Data for Computer Processing. Lund, Sweden: Royal University of Lund, Studies in Geography, Series C, General and Mathematical Geography, No. 2, 1962.

Nystuen, John D. "Location Theory and the Movement of Fresh Produce to Urban Centers," Seattle: University of Washington, Department of Geography, 1957 (unpublished Master's Thesis).

______________. "Geographical Analysis of Customer Movements and Retail Business Locations; (1) Theories, (2) Empirical Patterns in Cedar Rapids, Iowa, and (3) A Simulation Model of Movement," Seattle: University of Washington, Department of Geography, 1959 (unpublished Doctoral Dissertation).

______________. "A Simulation Model of Interurban Travel," in W. L. Garrison, ed., Quantitative Geography (forthcoming).

______________, and Dacey, Michael F. "A Graph Theory Interpretation of Nodal Regions Papers and Proceedings, Regional Science Association, Vol. 7 (1961), 29-42.

Odum, Eugene P. Fundamentals of Ecology. Philadelphia: Saunders, 1959.

Olderham, C. H. G. and Sutherland, D. B. "Orthogonal Polynomials: Their Use in Estimating the Regional Effect," Geophysics, Vol. 20, No. 2 (1955), 295-306.

Ollson, G. and Perrson, A. "The Spacing of Central Places in Sweden," Paper to be presented at the Third European Congress, Regional Science Association, Department of Geography, University of Lund, Lund, Sweden, August 26-29, 1963.

Olson, E. C. and Miller, R. L. Morphological Integration. Chicago: University of Chicago Press, 1958.

Onat, E. T. and Stepak, W. "A Note on the Economic Design of a Transportation Network," Providence, Rhode Island: Brown University, Division of Applied Mathematics, Progress Report No. 4, January 25, 1958.

Orden, Alex. "The Trans-shipment Problem," Management Science, Vo. 2 (1956) 227-285.

PASCHINGER, V. "Die Relativen Hohen von Karnten," Peterm. Mitt., 1934, 331-333 and 367-368.

Patton, Clyde P. and Pitts, Forest R. "Distance-Azimuth Program for IBM Card 1620," University of Oregon, Department of Geography, 1962 (forthcoming).

Pearson, K. "On Lines and Planes of Closest Fit to Systems of Points in Space," Philosophical Magazine, 6th Series, Vol. 2 (1901), 559-572.

Pedersen, Harald A. "Cultural Differences in the Acceptance of Recommended Practices, Rural Sociology, Vol. 16, No. 1 (March, 1951), 37-49.

Pedoe. The Gentle Art of Mathematics. New York: Macmillan Co., 1962.

Peguy, Ch. P. "Principes de morphometrie alpine," Rev. Geogr. Alpine, 1942, 453-482.

___________. "Introduction a l'emploi des methodes statistiques en geographie physique, these complementaire," Rev. Geogr. Alpine, 1948, 1-103.

___________. Elements de Statistique Appliquee aux Sciences Geographiques. 1957. (available for $2.50 from Centre de Documentation Universitaire, 5 Place de la Sorbonne, Paris).

Peltier, L. C. Area Sampling for Terrain Analysis. Paper read at American Assoc. for the Advancement of Science meetings, Section E, Chicago, Illinois, 1959.

Penck, A. "Morphologie der Erdoberflache," Bibl. Geogr. Hanbucher, Herausgegeb, V. F. Ratzel, 1894.

Perle, E. D. "Time Series Analysis of Transportation Development: A Pilot Study of the Demand for Transportation in the United States," A portion of a report to be submitted October 31, 1962, to U. S. Army Transportation Research Command, Fort Eustis, Virginia, by the Transportation Center at Northwestern University under contract: DA-44-177-TC-685.

Peters, W. S. "A Method of Deriving Geogrpphic Patterns of Associated Demographic Characteristics within Urban Areas," Social Forces, Vol. 35 (1956), 62-68.

Petrie, W. M. Flinders. Social Life in Ancient Egypt. London: Constable and Company, 1923.

Pitts, Forrest R. "The Land Use Model of Von Thünen," an address prepared for delivery at Ewha Women's University, Seoul, Korea, in December, 1960.

_______________, ed. Urban Systems and Economic Development: Papers and Proceedings. Eugene, Oregon: University of Oregon, School of Business Administration, 1962.

_______________. "A Graph Theoretic Approach to Historical Geography," A paper presented to the annual meeting, Association of Pacific Coast Geographers, Los Angeles, June 12-14, 1963 (Hectographed).

_______________. "Problems in Computer Simulation of Diffusion," Papers and Proceedings, The Regional Science Association (forthcoming).

Platt, John R. "Functional Geometry and the Determination of Pattern in Mosaic Receptors," General Systems, Vol. 7 (1962), 103-120.

Plaza, Jose A. "On the Optimum Program of Highway Improvements: A Mathematical Programming Approach," Evanston, Illinois: Northwestern University, Department of Civil Engineering, 1962 (Master's Thesis).

Pollack, Maurice and Weibenson, Walter. "Solutions of the Shortest Route Problem--A Review," Journal of the Operations Research Society of America, Vol. 7 (1960), 224-230.

Ponstein, J. "Matrix Description of Networks," Journal of the Society for Industrial and Applied Mathematics, Vol. 9 (1961), 233-268.

Porter, Herman. "Models of Flow: of People, Commodities and Messages," Evanston, Illinois: Northwestern University, Department of Geography (Doctoral Dissertation in preparation).

Porter, P. "Putting the Isopleth in Its Place," Proceedings, Minnesota Academy of Science, Vols. 25-26 (1957-58), 372-384.

Porter, R. "Approach to Migration Through Its Mechanism," Geografiska Annaler, Vol. 38, No. 4 (1956), 317-343.

Pöyhönen, Pentti. "An Econometric Investigation of City Land Market Prices," Helsinki The State Institute for Technical Research, 1955.

Prager, William. "Derivation of Results Stated in Progress Report No. 1," Providence, Rhode Island: Brown University, Division of Applied Mathematics, Progress Report No. 2, December 14, 1956.

_______________. "Economic Design for Manhattan Geometry," Providence, Rhode Island: Brown University, Division of Applied Mathematics, Progress Report No. 5, January 27, 1958.

_______________. "On the Design of Communication and Transportation Networks," presented at the Symposium on Theory of Traffic Flow organized by General Motors Research Laboratories, Detroit, December 7-8, 1959.

Prager, William and Onat, E. T. "Optimum Trunk Layout for Model Discussed in Progress Report No. 2," Providence, Rhode Island: Brown University, Division of Applied Mathematics, Progress Report No. 3, March 25, 1957.

Pred, Allan, R. "The Enternal Relations of Cities During 'Industrial Revolution'," Chicago: University of Chicago, Department of Geography, Research Paper No. 76.

Prihar, Zvi. "Topological Properties of Telecommunication Networks," Proceedings of the Institute of Radio Engineers, Vol. 44 (1956), 927-933.

Proudfoot, Malcolm J. "Sampling with Transverse Traverse Lines," Journal of the American Statistical Association, Vol. 37 (1942), 265-270.

QUANDT, Richard E. "Models of Transportation and Optimal Network Construction," Journal of Regional Science, Vol. 2, No. 1 (1960), 27-45.

RAISZ, Erwin. "Block-File System of Statistical Maps," Economic Geography, Vol. 15 (1939), 185-188.

___________. "Our Lopsided Earth," Journal of Geography, Vol. 43 (1944), 88-91.

___________ and Raisz, H. J. "An Average Slope Map of Southern New England," Geographical Review, Vol. 27 (1937), 467-472.

Rao, C. Radhakrishna. "The Utilization of Multiple Measurement in Problems in Biological Classification," Journal of the Royal Statistical Society, Series B 10 (1948), 159-203.

__________________. Advanced Statistical Methods in Biometric Research. New York: John Wiley, 1952.

Rannells, John. The Core of the City. New York: Columbia University Press, 1956.

Rashevsky, N. "Imitation Effects as a Function of Distance," Bulletin of Mathematical Biophysics, Vol. 12, No. 3 (September, 1953), 197-234.

Reed, I. and Stewart, R. "Note on the Existence of Perfect Maps," Transactions on Information Theory, Institute of Radio Engineers, January, 1962, 10-12.

Reichenbach, H. Space and Time. New York: Dover, 1958 (original publication, 1927).

Rein, G. "Die Quantitativ-Mineralogische Analyse des Malsburger Granit-Plutons und ihre Anwnedung auf Intrusionsform und Differentiation Sverlauf," Jhareshefte Geologischen Landesamt in Baden-Wurtemberg, Vol. 5 (1961), 53-115.

Reynolds, Robert B. "A Test of the Law of Retail Gravitation," Journal of Marketing, Vol. 42 (1953), 276 -

________________. "Statistical Methods in Geographic Research," The Geographical Review, Vol. 46 (1956), 129-132.

Reynolds, R. V. and Pierson, A. H. "Tracking the Saw-Mill Westward," American Forests Vol. 31 (1925), 647.

Richards, Paul I. "Shock Waves on the Highway," Journal of the Operations Research Society of America, Vol. 4, No. 1 (1956), 42-41.

Richardson, Lewis F. Arms and Insecurity: A Mathematical Study of the Cases and Origins of War. Nicolas Rashevsky and Ernesto Trucco, eds., Pittsburgh: The Boxwood Press; Chicago, Quadrangle Books, 1960.

__________________. Statistics of Deadly Quarrels. Quincy Wright and C. C. Lienau, eds., Pittsburgh: The Boxwood Press; Chicago: Quadrangle Books, 1960.

__________________. "The Problem of Contiguity," General Systems, Vol. 6 (1961), 139-188.

Richardson, Moses. Fundementals of Mathematics. New York: The Macmillan Co., 1958.

Richmond, Samuel B. "Interspatial Relationships Affecting Air Travel," Land Economics, Vol. 33, No. 1 (February, 1957), 67-73.

Riedel, J. "Neue Studien uber Isochronenkarten," Petermann s Geographische Mitteilungen, Vol. 52 (1911), 281-284.

Riemann, B. "Uber di Hypothesen welche der Geometrie augrunde liegen," 1854, (Doctoral Dissertation). Reprinted in H. Weyl, Das Kontinum und Andere Monographien. New York: Chelsea, 1958.

Ringwood, A. E. "A Model for the Upper Mantle, 2," Journal of Geophysical Research, Vol. 67 (1962), 4473-4477.

______________. "A Model for the Upper Mantle," Journal of Geophysical Research, Vol. 67 (1962), 857-867.

Robinson, Arthur H. "A Method for Producing Shaded Relief from Areal Slope Data," Annals, Association of American Geognaphers, Vol. 36 (1946), 248-252.

__________________. Elements of Cartography. New York: John Wiley & Sons, Inc., 1953.

__________________. "The Necessity of Weighing Value in Correlation Analysis of Areal Data," Annals, Association of American Geographers, Vol. 46 (1956), 233-23[illegible]

__________________. "Mapping the Correspondence of Isarithmic Maps," Annals, Association of American Geographers, Vol. 52,(1962), 414-425.

__________________ and Bryson, Reid A. "A Method for Describing Quantitatively the Correspondence of Geographical Distributions," Annals, Association of American Geographers, Vol. 47 (1957), 379-391.

__________________, Lindberg, J. H. and Brinkman, L. "A Correlation and Regression Analysis Applied to Rural Farm Population Densities in the Great Plains," Annals, Association of American Geographers, Vol. 51 (1961), 211-222.

__________________ and Thrower, Norman J. W. "A New Method of Terrain Representation, Geographical Review, Vol. 47 (1957), 507-520.

Roder, W. and Berry, Brian J. L. "Direct Factor Analysis of Urban Flood Plain Data," University of Chicago, Department of Geography, February, 1960.

Rogers, C. A. "The Packing of Equal Spheres," Proc. London Math. Soc., Vol. 8 (1958), 609-620.

Rogers, J. "The Use of Equal-Area and Other Projections in the Statistical Treatment of Joints," Bulletin of the Geological Society of America, Vol. 63 (1952), 427-430.

Rosander, A. C. Elementary Principles of Statistics. New York: Van Nostrand, 1951.

Rosenfeld, L. "The Determination of an Optimal Trucking Route Through a Given Traffic Congestion Pattern," Proceedings of the Western Joint Computer Conference, 1956, 80.

Ross, F. A. "Editor's Note on the Center of Population and Point of Minimum Land," Journal of American Statistical Association, Vol. 25 (1930),

Ruhe, Robert V. "Graphic Analysis of Drift Topographies," American Journal of Science, Vol. 248 (1950), 435-443.

Rund, Hanno. Differential Geometry of Finsler Spaces, Berlin: Springer, 1959.

SABBAGH, Michael E. and Bryson, Reid A. "Aspects of the Precipitation Climatology of Canada Investigated by the Method of Harmonic Analysis," Annals, Association of American Geographers, Vol. 52 (1962), 426-440.

Sakharov, N. D. "Scholastics Instead of Science," Soviet Geography: Review and Translation, Vol. 3, No. 7 (1962), 16-22.

Salisbury, Neil Elliot. "A Generic Classification of Landforms of Minnesota," Minneapolis, Minnesota: University of Minnesota, Department of Geography, 1957 (unpublished Doctoral Dissertation).

Samuelson, Paul A. "Spatial Price Equilibrium and Linear Programming," The American Economic Review, Vol. 42 (1952), 283-303.

San Diego Metropolitan Area Transportation Study. Estimating Future Auto and Truck Trips. San Diego, 1959.

Sauer, Carl O. "The Education of a Geographer," Annals, Association of American Geographers, Vol. 46 (1956), 288-299.

Saushkin, Yu. G. "On the Subject Matter of Doctoral Dissertations in Geography," Soviet Geography: Review and Translation, Vol. 4, No. 1 (January, 1963), 47-53.

________________. "V. A. Anuchin's Doctoral Dissertation Defense," Soviet Geography: Review and Translation, Vol. 4, No. 1 (January, 1963), 53-59.

Savage, Leonard J. The Foundations of Statistics. New York: John Wiley, 1954.

Savigear, R. A. G. "Some Observations on Slope Development in South Wales," Transactions and papers, Publ. No. 18, Institute of British Geographers, 1953, 31-51.

Scates, D. E. "Locating the Median of the Population of the United States," Metron., Vol. 11 (1933), 49-65.

____________ and Van Nortwich, L. M. "The Influence of Restrictive Routes upon the Center of Minimum Aggregate Travel," Metron, Vol. 13 (1937), 78-81.

Schaefer, Fred K. "Exceptionalism in Geography: A Methodological Examination," Annals, Association of American Geographers, Vol. 43 (1953), 226-249.

Scharlau, Von K. "Ackerlagen und Ackergrenzen flurgeographische Begriffsbestimmungen, in Geographisches Taschenbuch 1956-1957. Wiesbaden, Germany: Franz Steiner Verlag Gmbh., 1958, 449-452.

Scheidegger, A. E. Theoretical Geomorphology. Englewood Cliffs, New Jersey: Prentice-Hall, 1961.

Schmid, Calvin F. and MacCannel, E. H. "Basic Problems, Techniques, and Theory of Isopleth Mapping," Journal, American Statistical Association, Vol. 50 (1955), 220-239.

Schmitt, Robert C. "Estimating Daytime Populations," Journal, American Institute of Planners, Vol. 22, No. 2 (May, 1956), 83-85.

Schneider, Morton. "Gravity Models and Trip Distribution Theory," Papers and Proceedings, Regional Science Association, Vol. 5 (1959), 51-56.

Schnore, Leo F. "The Separation of Home and Work: A Problem for Human Ecology," Social Forces, Vol. 32 (May, 1954), 336-343.

Schumm, St. A. "Evolution of Drainage Systems in Badlands at Pearth Amboy, New Jersey," Columbia University, Department of Geology, 1954.

____________. "Evolution of Drainage Systems and Slopes in Badlands at Pearth Amboy, New Jersey," Off. of Naval Research, Project NR 389-402, Tech. Report No. 8, New York, 1954.

Scott, G. D. "The Packing of Spheres," Nature, Vol. 188 (1960), 908-911.

Sebestyen, J. "Some Thoughts on Spatial Models for Development Purposes," Institute for Farm Economics, Budapest, paper to be presented at the Third European Congress, Regional Science Association, Department of Geography, University of Lund, Lund, Sweden, August 26-29, 1963.

Segre, B. and Mahler, K. "On the Densiest Packing of Circles," American Math. Monthly, Vol. 51 (1944), 261-270.

Sendler, Gerhard. "Verkehrsgeographische Ubersicht der Erde," Pettermanns Geographische Mitteilungen, Vol. 103, No. 2 (1959), 106-111.

Seshu, Sundaram and Reed, Myrall B. Linear Graphs and Electrical Networks. Reading, Massachusetts, Addison-Wesley, 1961.

Sherratt, G. G. "A Model for General Urban Growth," Management Sciences: Models and Techniques, Vol. 2 (New York: Pergamon Press) 1960, Proceedings of the 6th International Meeting of the Institute of Management Sciences, 1959.

Shevky, E. and Williams, M. The Social Areas of Los Angeles, Berkeley: University of California Press, 1948.

Shimbel, Alfonso. "Structural Parameters of Communication Networks," Bulletin of Mathematical Biophysics, Vol. 15 (1953), 501-507.

Shubik, Martin. Readings in Game Theory and Political Behavior. New York: Doubleday, 1954.

_____________. "Bibliography on Simulation Gaming, Artificial Intelligence and Allied Topics," Journal, American Statistical Association, Vol. 55 (1960), 736-751.

Shycon, Harvey N. and Maffei, Richard B. "Simulation - A Tool for Better Distribution Harvard Business Review, Vol. 38 (1960), 65-75.

Simpson, S. M., Jr. "Least Squares Polynomial Fitting to Gravitation Data and Density Plotting by Digital Computers," Geophysics, Vol. 19 (1954), 250-257.

Skellam, J. G. "Studies in Statistical Ecology: I, Spatial Pattern," Biometrika, Vol. 39 (1952), 346-362.

Sloane, Charles S. "Center of Population and Median Lines, and Centers of Area, Agriculture, Manufacturés, and Cotton, etc.," 14th Census of the U.S. 1920, Washington, D. C.: 1923, Vol. I, 12-41.

Sloss, L. L. "Stratigraphic Models in Exploration," Bulletin, American Association of Petroleum Geologists, Vol. 46 (1962), 1050-1057.

Smirnov, N. "Table for Estimating the Goodness of Fit of Empirical Distributions," Annals of Mathematical Statistics, Vol. 19 (1948), 279-281.

Smith, G. C. "Lorenz Curae Analysis of Industrial Decentralization," Journal, American Statistical Association, Vol. 42 (1947), 591-596.

Smith, G. H. "The Relative Relief of Ohio," Geographical Review, Vol. 25 (1955), 272-284.

Smith, K. G. "Standards for Grading Texture of Erosional Topography," American Journal of Science, Vol. 248 (1950), 655-668.

__________. "Erosional Processes and Landforms in Badlands National Monument, S. Dakota," Bulletin, Geol. Soc. Am., Vol. 69 (1958), 975-1008.

Smock, Robert. "A Comparative Description of a Capa-city Restrained Traffic Assignment," Detroit Area Traffic Study (in progress).

Solomon, H. "Distribution of the Measure of a Random Two-dimensional Set," Annals, Math. Stat., Vol. 24 (1953), 650-656.

Sommer, Robert. "Leadership and Group Geography," Sociometry, Vol. 24, No. I (March, 1961), 99-110.

Sonklar, C. Allgemeine Orographie, die Lehre von den Reliefformen der Erdoberflache, 1873.

Stafford, Howard A., Jr. "Factors in the Location of the Paperboard Container Industry," Economic Geography, Vol. 36, No. 3 (July, 1960), 260-266.

Steiner, Rodney. An Investigation of Selected Phases of Sampling to Determine Quantities of Land and Land Use Types. Seattle: University of Washington, 1954 (unpublished Doctoral Dissertation).

_______________. "Some Problems in Designing Samples of Rural Land Use," Yearbook Association of Pacific Coast Geographers,Vol. 19 (1957), 25-28.

Steinhaus, Hugo. "Length, Shape and Area," Colloquium Mathematicum, Vol. 3 (1954), 1-13.

Stevens, Benjamin H. "A Review of the Literature on Linear Methods and Models for Spatial Analysis," Journal, American Institute of Planners, Vol. 26, No. 3 (August, 1960), 253-259.

_______________. "An Application of Game Theory to a Problem in Location Strategy," Papers, Regional Science Associaiton, Vol. 7 (1961), 143-157.

_______________ and Coughlin, R. "A Note on Inter-Areal Linear Programming for a Metropolitan Region," Journal of Regional Science, Vol. 2, No.1 (1959), 75-83.

Stewart, Charles T., Jr. "The Size and Spacing of Cities," The Geographical Review, Vol. 48 (1958), 222-245.

_______________. "Migration as a Function of Population and Distance," American Sociological Review, Vol. 25, No. 3 (1960), 347-356.

Stewart, John Q. "An Inverse Distance Variation for Certain Social Influences," Science, Vol. 93, No. 2404 (January 24, 1941), 89-90.

_______________. "The Use and Abuse of Map Projections," The Geographical Review, Vol. 33, No. 4 (1943), 589-604.

_______________. "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population," Geographical Review, Vol. 37 (1947), 461-485.

_______________. "Demographic Gravitation: Evidence and Applications," Sociometry, Vol. 11, No. 1-2 (Feb. - May, 1948).

_______________. "Concerning 'Social Physics'," Scientific American, Vol. 178, No. 5 (May, 1948), 20-23.

_______________. "Mathematical and Physical Methods in Social Science," Notes on Conference in Princeton, October 12 and 13, 1949.

_______________. "The Development of Social Physics," American Journal of Physics, Vol. 18, No. 5 (May, 1950), 239-253.

_______________. "Potential of Population and Its Relationship to Marketing," in Reavis Cox and Wroe Alderson, eds., Theory in Marketing. Chicago: Richard D. Irwin, Inc., 1950.

Stewart, John Q. "Futher Development of the Dimensions of Society," Paper read at American Association for Advancement of Science, Philadelphia, December 27, 1951.

______________. "A Basis for Social Physics," Impact of Science on Society, Vol. 3, (1952), 110-133.

______________. "The Kinds of Energy," Abstract of paper presented January, 1953, at Cambridge meeting of American Phys. Soc.

______________. "Urban Population Densities (review)," Geographical Review, Vol. 43, No. 4 (1953), 575-6.

______________ and Warntz, William. "Macrogeography and Social Science," Geographical Review, Vol. 48 (1958), 167-184.

______________ and ______________. "Physics of Population Distribution," Journal of Regional Science, Vol. 1, No. 1 (1958), 99-123.

______________ and ______________. "Some Parameters of the Geographical Distribution of Population," The Geographical Review, Vol. 46 (1959), 270-272.

Stone, Kirk H. "World Air Photo Coverage," The Professional Geographer, Vol. 11, No. 3 (1959), 2-6.

Stone, Richard. "A Comparison of the Economic Structure of Regions Based on the Concept of Distance," Journal of Regional Science, Vol. 2, No. 2 (Fall, 1960), 1-20.

Stouffer, Samuel. "Intervening Opportunities: A Theory Relating Mobility to Distances," American Sociological Review, Vol. 15 (1950), 845-867.

______________. "Intervening Opportunities and Competing Migrants," Journal of Regional Science, Vol. 2, No. 1 (1960), 1-26.

Strahler, Arthur. "Equilibrium Theory of Erosional Slopes Approached by Frequency Distribution Analysis," American Journal of Science, Vol. 248 (1950), 673-696, 800-814.

______________. "Dynamic Basis of Beomorphology," Bulletin Geological Society of America, Vol. 63, (1952), 923-938.

______________. "Hypsometric (area-altitude) Analysis of Erosional Topography," Bulletin, Geological Society of America, Vol. 63, (1952), 1117-1142.

______________. "Quantitative geomorphology of Erosional Landscapes," Congr. Geol. Int. Alger. Sec. 8, fasc. 15 (1954), 341-354.

______________. "Statistical Analysis in Geomorphic Research," Journal of Geology, Vol. 62 (1954).

______________. "Basic Principles of Quantitative Geomorphology," (abstract) Annals, Association of American Geographers, Vol. 46 (1956), 275.

Strahler, Arthur. "Quantitative Slope Analysis," Bulletin, Geological Society America, Vol. 67 (1956), 571-596.

_______________. "Quantitative Analysis of Watershed Geomorphology," Trans. Am. Geophys. Union, Vol. 38, No. 6 (1957), 913-920.

_______________. "Dimensional Analysis Applied to Fluvially Eroded Landforms," Bulletin Geological Society of America, Vol. 69, (1958), 279-300.

_______________ and Koons, D. "Objective and Quantitative Field Methods of Terrain Analysis," Columbia University, Department of Geology, Report Project Nr. 387-021 ONR (1960).

Strong, C. L. "The Amateur Scientist: A Sundial that shows how any Spot on Earth is Lit by the Sun at Any Time," Scientific American,(August, 1959), 37-144.

Struik, Dirk J. Lectures on Classical Differential Geometry,(2nd ed.). Reading, Massachusetts: Addeson-Wesley Publishing Co., 1961.

Studies in Rural-Urban Interaction. Lund, Sweden: Royal University of Lund, Studies in Geography, Series B, Human Geography, No. 3, 1951.

Sviatlovsky, E. E. and Eells, W. C. "The Centrographical Method and Regional Analysis," Geographic Review, Vol. 27 (1937), 240-254.

Taaffe, Edward J. "A Map Analysis of United States Airline Competition Part II - Competition and Growth," Journal of Air Law and Commerce, Vol. 25, No. 4 (Autumn, 1958).

_______________. "Trends in Airline Passenger Traffic: A Geographic Case Study," Annals, Association of American Geographers, Vol. 49, No. 4 (December, 1959), 393-408.

_______________ and Garner, B. J. "A Geographic Consideration of the Journey-to-Work to Peripheral Employment Centers," in a study prepared by the Transportation Center at Northwestern University under a grant from the Automobile Manufactures Association, 1960.

Tate, R. F. "Correlation Between a Discrete and a Continuous Variable: Point-Biserial Correlation," Annals of Mathematical Statistics, Vol. 25 (1954), 603-607.

__________. "The Theory of Correlation Between Two Continuous Variables When One is Dichotomized," Biometrika, Vol. 42 (1955), 205-222.

Taylor, George Rogers. The Transportation Revolution, 1815-1860. New York: Rinehart, 1951.

Taylor, Griffith. "Environment, Village and City: A Genetic Approach to Urban Geography; With Some Reference to Possibilism," Annals, Association of American Geographers, Vol. 32 (1942), 1-67.

Thauer, W. "Neue Methoden der Berechnung and Darstellung der Relief-energie," Peterm. Mitt., 1955, 8-13.

Thiessen, A. H. "Precipitation Averages for Large Areas," Monthly Weather Review, July, 1911, 1086.

Thomas, Edwin N. "Areal Association Between Population Growth and Selected Factors in the Chicago Urbanized Area," Economic Geography, Vol. 36, No. 2 (1960), 158-170.

______________. "Maps of Residuals from Regression: Their Characteristics and Uses in Geographic Research," Iowa City, State University of Iowa, Department of Geography, No. 2, 1960.

______________. "Some Comments on the Functional Bases for Small Iowa Towns," Iowa Business Digest, Winter, 1960, 10-16.

______________. "Toward an Expanded Central Place Model," Geographical Review, Vol. 51 (1961), 400-411.

______________. "The Stability of Distance-Population-Size Relationships for Iowa Towns from 1900-1950." In Knut Norborg, ed., Proceedings of the IGU, Symposium in Urban Geography, Lund, 1960, Lund, Sweden: Royal University of Lund, Studies in Geography, Series B, Human Geography, No. 24, 1962, 13-31.

______________. "Spatial Behavior of a Dispersed Non-Farm Population," Papers and Proceedings, Regional Science Association (forthcoming).

Thomas, Frank, H. The Denver and Rio Grande Western Railroad: A Geographic Analysis," Evanston, Illinois: Northwestern University, Studies in Geography, No. 4, 1960.

Thompson, D'Arcy. On Growth and Form (abridged edition, J. Bonner, ed.). Cambridge: University Press, 1961.

Thompson, H. R. "Distribution of Distance to Nth Neighbor in a Population of Randomly Distributed Individuals," Ecology, Vol. 37 (1956), 391-394.

Thompson, W. R. "The Geometric Properties of Microscopic Configurations I, II," Biometrika, Vol. 24 (1932), 21-38.

Thrower, Norman J. W. "Animated Cartography," The Professional Geographer, Vol. 11, No. 6, (1959), 9-12.

Tinbergen, Jan. "The Appraisal of Road Construction: Two Calculation Schemes," Review Economics and Statistics, Vol. 39 (1957), 241-249.

______________ and Bos, H. C. Mathematical Models of Economic Growth. New York: McGraw-Hill, 1962.

Tissot, M. A. Memoire sur la Representation des Surfaces et les Projections des Cartes Geographiques. Paris: Gautier-Villars, 1881.

Tobinson, R. M. "Arrangement of 24 Points on a Sphere," Mathematical Annals, Vol. 144 (1961), 17-48.

Tobler, Waldo. Map Transformations of Geographic Space. Seattle, Washington: University of Washington, Department of Geography, 1961 (Unpublished Doctoral Dissertation, Available at University Microfilms, 313 First St., Ann Arbor, Michigan). Part originally appeared as "An Analysis of Map Projections," University of Washington, Department of Geography, March, 1960 (unpublished).

____________. "A Classification of Map Projections," Annals, Association of American Geographers, Vol. 52, No. 2 (1962), 167-175.

____________. "Studies in the Geometry of Transportation," a portion of a report to be submitted October 31, 1962, to U. S. Army Transportation Research Command, Fort Eustis, Virginia, by the Transportation Center at Northwestern University under Contract: DA-44-177-TC-685.

____________. "Curvilinear Coordinates," University of Michigan, Department of Geography, 1962 (unpublished).

____________. "Formulation of a Non-Euclidean Central Place Pattern," January, 1963 (unpublished).

____________. "Geographical Area and Map Projections," The Geographical Review, Vol. 53, No. 1 (1963), 59-78.

____________. "Map Projection Research by Digital Computer," Paper to be presencted A.A.G. meeting, Denver, September, 1963.

____________. "A Polynomial Representation of Michigan Population," Papers and Proceedings, Michigan Academy of Science, Arts, and Letters (forthcoming).

Toth, L. Fejes. Lagerungen in der Ebene Auf der Kugel, und im Raum. Berlin: Springer, 1953.

____________. "Filling of a Domain by Isoperimetric Circles," Publicationes Mathematicae, December, 1957, 1-2.

____________. "An Arrangement of Two-Dimensional Cells," Ann. Univ. Sci. Budapest, Eötvös. Sect. Math., Vol. 2 (1959), 61-64.

Toulmin, Stephen. The Philosophy of Science. London, England: Hutchinson House, 1953.

Trakhtenbrot, B. A. Algorithms and Automatic Computing Machines. In series: Topics in Mathematics, Boston: D. C. Heath and Company (forthcoming).

Tricart, J. and Muslin, I. "Idees des recherches: L'etude Statistique des Versants," Revue de Geomorphologie Dynamique, No. 4b (1951).

Tryon, R. C. Identification of Social Areas by Cluster Analysis. Berkeley: University of California Press, 1955.

Tschierske, Hilmar. "Raumfunktionelle Prinzipien in Einer Allgemeine Theoretischen Geographie, Axiometische und Empirische Bestandtiile in Ihr," (The Place of Functional Principles of Distribution in a General System of Geographical Theory), Erdkunde, Band 15, 92-109.

Twery, R. J. and Rossman, M. J. "A Model of Heterogeneous Transportation Costs," (unpublished).

Tyron, F. G. "The Changing Distribution of Resources," in Carter Goodrich, Phil, ed., Migration and Economic Opportunity, The Report of the Study of Population Redistribution, 1936, Ch. 4, 253.

ULLMAN, Edward L. "The Railroad Pattern of the United States," The Geographical Review, Vol. 39 (1949), 242-256.

_______________. "Rivers as Regional Bonds: The Columbia-Snake Example," The Geographical Review, Vol. 41 (1951), 210-225.

_______________. "Human Geography and Area Research," Annals, Association of American Geographers, Vol. 43 (1953), 60.

_______________. "The Role of Transportation and the Bases for Inter-action," in William L. Thomas, Jr., ed., Man's Role in Changing the Face of the Earth. Chicago: University of Chicago Press, 1956.

_______________. American Commodity Flow. Seattle: University of Washington Press, 1957.

_______________. "The Expansion of Urban Areas," mimeographed notes in connection with the Meramec Basin Research Project, Washington University, St. Louis, 1960.

U. S. Army Engineer. Technique for Preparing Desert Terrain. (Analogues, Technical Report No. 3-506). Vicksburg, Mississippi: U. S. Army Engineer, Waterways Experiement Station, 1959.

U. S. Bureau of the Census. "A Chapter in Population Sampling," Washington, 1947.

Unknown. "Derivation of Formula for Optimum Expressway Spacing," appendix, Chicago Area Transportation Study, Final Report, Vol. 3, Transportation Plan, 1962, Chicago, 121-123.

Unknown. "Discriminatory Analysis as a Tool for Geographic Research," (unpublished).

Uyemura, F. (Japan, Kagawa University). "Regional Input-output Model of the Shikoku Area and the Economic Effects of the Proposed Seto Great Bridge," paper read at the Hague meetings of the Regional Science Association, September, 1961.

VACQUIER, J., et. al. "Interpretation of Aeromagnetic Maps," Memoir 47, Geological Society of America, 1951.

Vadnal, A. "La Localization du Reseau de Routes Collectrices Dans Le Cercle," paper presented at the Congress European D'Econometic, Bilbao, 1958 (unpublished

Van Arsdol, M. Emperical Evaluation of Social Area Analysis. Seattle: University of Washington, 1957 (Doctoral Dissertation, unpublished).

Vance, James E., Jr. "Labor-Shed, Employment Field, and Dynamic Analysis in Urban Geography," Economic Geography, Vol. 36 (July, 1960), 189-220.

Van Paasen, C. The Classical Tradition of Geography. Groningen, the Netherlands: J. B. Wolters, 1957.

Verblunsky, S. "On the Shortest Path Through a Number of Points," Proc. Amer. Math. Soc., Vol. 2 (1951), 904=913.

Verburg, M. C. "Some Aspects of Location Analysis," Economisch Technologisch Instituut voor Zeeland, Netherlands, paper to be presented at the Third European Congress, Regional Science, Department of Geography, University of Lund, August 26-29, 1963.

Vidale, Marcello L. "A Graphic Solution of the Transportation Problem," Journal of the Operations Research Society of America, Vol. 4 (1956), 193-203.

Vining, Rutledge. "A Description of Certain Spatial Aspects of an Economic System," Economic Development and Cultural Change, Vol. 3 (1955), 147-195.

_______________ and Koopman, T. C. "Methodological Issues in Quantitative Economics Review of Economic Statistics, Vol. 31 (1949), 77-94.

Vinski, I. "Regional Growth of Fixed Assets in Yugoslavia, 1946-1960," Ekonomski Institut, Zagreb, paper to be presented at the Third European Congress, Regional Science Association, Department of Geography, University of Lund, August 26-29, 1963.

Von Thünen, Johann Heinrich. Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. Hamburg, 1826.

Voorhees, Alan M. "A General Theory of Traffic Movement," Papers and Proceedings, Institute of Traffic Engineers, Vol. 26 (1955), 46-56.

_______________. "Use of Mathematical Models in Estimating Travel," Journal of the Highway Division, American Society of Civil Engineers, Vol. 85, No. HW4 (December, 1959), 129-141.

WARNTZ, William. "Geography of Prices and Spatial Interaction," Papers and Proceedings, Regional Science Association, Vol. 3 (1957), 118-129.

_______________. "Transportation, Social Physics, and the Law of Refraction," The Professional Geographer, Vol. 9, No. 4 (1957), 2-7.

_______________. "Macrogeography and the Census," The Professional Geographer, Vol. 10 (1958), 6-10.

_______________. Toward a Geography of Price: A Study in Geoeconometrics. Philadelphia: University of Pennsylvania Press, 1959.

_______________. "Transatlantic Flights and Pressure Patterns," The Geographical Review, Vol. 51, No. 2 (1961), 187-212.

_______________ (unsigned). The Earth: Shape and Magnetism. Geography Learning Laboratory No. 172. Princeton, New Jersey: The Learning Center, Inc., 1962.

_______________ and Neft, D. "Contributions to a Statistical Methodology for Areal Distributions," Journal of Regional Science, Vol. 2, No. 1 (Spring, 1960), 45-66.

Watson, J. W. "Geography -- A Discipline in Distance," The Scottish Geographical Magazine, Vol. 71, No. 1 (1955), 1-13.

Waugh, Frederick J. Graphic Analysis in Economic Research. Washington, D. C.: U. S. Department of Agriculture, Agricultural Marketing Service, Agricultural Handbook No. 84, June, 1955.

_______________. Graphic Analysis in Agricultural Economics. Washington, D. C.: U. S. Department of Agriculture, Agricultural Marketing Service, Agricultural Handbook, No. 128, July, 1957.

Weaver, John C. "Crop-combination Regions in the Middle West," Geographical Review Vol. 44 (1954), 175-200.

_______________. "Isotope and Compound: A Framework for Agricultural Geography," Annals, Association of American Geographers, Vol. 44 (1954), 286-288 (abstract).

Webb, J. W. "Basic Concepts in the Analysis of Small Urban Centers in Minnesota," Annals, Association of American Geographers, Vol. 49 (1959), 55-72.

Webb, Wells Alan. "Analysis of the Martian Canal Network," Publications of the Astronomical Society of the Pacific, Vol. 67 (1955), 283-292.

Weiss, H. K. "The Distribution of Urban Population and an Application to a Servicing Problem," presented at the 19th Annual Meetings of the Operations Research Society of America, 1961.

Wendel, Bertil. A Migration Schema. Lund, Sweden: Royal University of Lund, Studies in Geography, Series B, Human Geography, No. 9, 1953.

Wentworth, C. K. "A Simplified Method of Determining the Average Slope of Land Surfaces," Am. Jour. of Sci., Vol. 20 (1930), 184-194.

Weyl, Hermann. Symmetry. Princeton, New Jersey: Princeton University Press, 1952.

_______________. "Matematische Analyse des Raumproblems," in Barcelona Lectures. Berlin: J. Springer, 1923. (Reprinted in H. Weyl, Das Kontinum und Andere Monographien. New York: Chelsea, 1958.)

Whitten, E. H. T. "Composition Trends in a Granite: Modal Variation and Ghost-Stratigraphy in Part of the Donegal Granite, Eire," Journal of Geophysical Research, Vol. 64 (1959), 835-848.

_______________. "Quantitative Evidence of Palimpsestic Ghost-Stratigraphy from Modal Analysis of a Granitic Complex," Report of the International Geological Congress (Norden), Part 14 (1960), 182-193.

_______________. "A New Method for Determination of the Average Composition of a Granite Massif," Geochim. Cosmochim. Acta, Vol. 26 (1962), 545-560.

_______________. "Application of Quantitative Methods in the Geochemical Study of Granite Massifs," Royal Society of Canada Symposium, 1962 (in press, 1963).

Whitten, E. H. T. A Surface-Fitting Program Suitable for Testing Geological Models which Involve Areally-Distributed Data. Technical Report No. 2, O.N.R. Task No. 389-135, contract No. NR 1228(26), Office of Naval Research, Geography Branch, Prepared at Northwestern University, Evanston, Illinois, 1963.

______________. "A Reply to Chayes and Suzuki," Journal of Petrology, Vol. 4 (1963) (in press).

Whittlesey, Derwent. "The Regional Concept and the Regional Method," in Preston E. James and Clarence F. Jones, eds., American Geography: Inventory and Prospect. Syracuse, New York: Syracuse University Press, 1954.

Wildbaur, H. "Die Reliefenergie in der morphographischen Karte," Peterm. Mitt., 1952.

Wilkinson, R. I. "Abridged Bibliography of Articles on Toll Alternate Routing," Bell System Technical Journal, Vol. 35 (1956), 597.

Willmore, T. J. An Introduction to Differential Geometry. Oxford, England: Clarendon Press, 1959.

Wilson, John Cook. On the Traversing of Geometrical Figures. Oxford, England: Clarendon Press, 1905.

Winchell, H. "A New Method of Interpretation of Petrofabric Diagrams," American Mineralogist, Vol. 22 (1937), 15-36.

Wingo, Lowdon, Jr., "An Economic Model of the Utilization of Urban Land," Proceedings, Regional Science Association, Vol. 7 (1961), 191-206.

______________. Transportation and Urban Land. Washington, D. C.: Resources for the Future, 1961.

Winterbotham, H. "Dots and Distributions," Geography, Vol. 19 (1934), 211-213.

Witheford, David K. "Comparison of Trip Generation by Opportunity Model and Gravity Model," Highway Research Board (Washington, D. C.), 1961.

Wolfe, Roy I. and Hickok, Beverly. An Annotated Bibliography of the Geography of Transportation. Berkeley, California: University of California, Institute of Transportation and Traffic Engineering, Information Circular No. 29, October, 1961.

Wolman, M. G. "The Natural Channel of Brandywine Creek, Pennsylvania," U.S.G.S. Prof. Paper 271, 1955, 1-56.

____________ and Leopold, L. B. "River Floodplains: Some Observations on their Formation," U.S.G.S. Prof. Paper 282-C, 1957.

Wonnacott, Ronald J. "Manufacturing Costs and the Comparative Advantage of United States Regions," Study Paper No. 9, Upper Midwest Economic Study, 1963.

Wood, Walter F. "Use of Stratified Random Samples in Land Use Study," Annals, Association of American Geographers, Vol. 45, No. 4 (1955), 350-367.

Wood, W. F. and Snell, J. B. "The Dispersion of Geomorphic Data around Measures of Central Tendency and Its Application," Research Study Report EA-8, Quartermaster Research & Development Center, Natick, Massachusetts, 1957.

__________ and ___________. "Predictive Methods in Topographic Analysis I: Relief, Slope and Dissection on Inch-to-the-mile maps in the United States," Technical Report EP-112, Quartermaster Research & Engineering Center, Natick, Massachusett 1959.

__________ and ___________. "Predictive Methods in Topographic Analysis II. Estimating Relief from World Aeronautical Charts," Technical Report EP-114, Quartermaster Research & Engineering Center, Natick, Massachusetts, 1959.

__________ and ___________. "Preliminary Investigations of a Method to Predict Line-of-Sight Capabilities," Research Study Report EA-10, Quartermaster Research & Engineering Center, Natick, Massachusetts, 1959.

Woytinsky, W. S. and Woytinsky, E. S. World Population and Production: Trends and Outlook. New York: Twentieth Century Fund, 1953.

Wright, John K. "A Method of Mapping Densities of Population with Cape Cad as an Example," Geographical Review, Vol. 26 (1936), 103-110.

______________. "Some Measures of Distributions," Annals, Association of American Geograp;ers, Vol. 27 (1937), 177-211.

Wyllie, P. J. "The Petrogenetic Model, An Extension of Bowen's Petrogenetic Grid," Geological Magazine, Vol. 99 (1962), 558-569.

YAGLOM, A. M. and Yaglom, I. M. Nonelementary Problems in Elementary Exposition (two volumes). San Francisco: Holden-Day, Inc. forthcoming.

Yaglom, I. M. Geometric Transformations (translated by Allen Shields). New York: Random House and L. W. Singer and Company for the Monograph Project of the School Mathematics Study Group, 1962.

Yates, Frank. Sampling Method for Censuses and Surveys. London: Charles Griffin and Co., Ltd., 1949.

Yeates, Maurice. "The 'Transportation Problem' in Geographical Research," Northwestern University, Department of Geography, Discussion Paper No. 2.

Yoshimura, S. "Methods of Chorometry and some Examples: Areal Relations Between Cultural Landscape and Surface Features," Geographical Review of Japan, Vol. 6, No. 11 (1930), 1-30, and No. 12 (1930), 12-47 (in Japanese).

Yuksel, H. "On Networks of Minimum Construction Cost," Providence, Rhode Island: Brown University, Division of Applied Mathematics, Progress Report, No. 6 (August 4, 1958).

Yule, George Udny and Kendall, M. G. An Introduction to the Theory of Statistics (14th edition). London: C. Griffin, 1950.

ZABORSKI, Bogdan. "How a Geographer Visualizes a Street Pattern in a Planned City," Annals, Association of American Geographers, Vol. 39 (1949), 79-80.

Zaustinsky, Eugene M. "Spaces with Non-Symmetric Distance," No. 34, American Mathematical Society, Memoirs, 1959.

Zettel, R. M. and Carll, R. R. Summary Review of Major Metropolitan Area Transportation Studies. Berkeley, California: University of California, Institute of Transportation and Traffic Engineering, 1962.

Zipf, George K. "The P_1P_2/D Hypothesis on the Intercity Movement of Persons," American Sociological Review, Vol. 11 (December, 1946), 677-686.

______________. Human Behavior and the Principle of Least Effort. Cambridge, Massachusetts, Addison-Wesley Press, Inc., 1949.

Zobler, Leonard. "Man-land Relations in Salem County, New Jersey: A Study in Quantitative Geographic Regionalization," New York: Columbia University, 1953 (Doctoral Dissertation).

______________. "Statistical Testing of Regional Boundaries," Annals, Association of American Geographers, Vol. 47 (1957), 83-95.

______________. "Decision Making in Regional Construction," Annals, Association of American Geographers, Vol. 48 (1958), 140-148.

______________. "The Distinction Between Relative and Absolute Frequencies in Using Chi Square for Regional Analysis," Annals, Association of American Geographers, Vol. 48 (1958), 456-457.

Zubrzycki, S. "Remarks on Random, Stratified and Systematic Sampling in a Plane," Colloquim Mathematicuum, Vol. 6, 251-264.

"Patterns are Morphological Laws"

Fred K. Schaefer

PATTERNS OF LOCATION

William Bunge

Schaefer's "morphological laws" are receiving increased attention by theoretical geographers. We explore the subject.

I. Introduction

This paper is intended to be a second supplement to the book, Theoretical Geography.[1] In the basic book, geography was related to science after Christaller's primary substantive example and Schaefer's primary methodology. The central problem of theoretical geography was tentatively identified as "the nearness problem," that is, "to located interacting objects as near to each other as possible." The first supplement to the basic book is an article entitled "Spatial Relations: The Subject of Theoretical Geography."[2] There the focus was sharpened by a hard examination of the methodological comments of select contemporary theoretical geographers and by a systematic review of some of the mathematics of space as applied to geography.

II. Three Types of Location

Predictive or theoretical geography has important old roots in geography; certainly Davis and Köppen were predictive geographers and a close historical examination might reveal much earlier antecedents. As in prospecting for oil, so in history, we usually find traces of what we seek if we look hard enough. But the main historical upthrust of the American-Swedish school of predictive geography has been through Christaller. Schaefer crystallized the methodological implications of Christaller's and related substantive work. Schaefer described the essence of predictive geography as the discovery of predictive patterns. Schaefer's insistence on spatial patterns (or the earth's geometry, or spatial structure, or whatever term you prefer) seems to run head on into predictive geographers' interest in movements (or circulations or spatial process or whatever term you prefer). The history of this important methodological contribution is reported elsewhere.[3] Tentative conclusions are that most locations on the earth's surface are explained, which is logically equivalent to predicted, on the basis of optimal movements (or geodesics, or least effort paths or whatever term you favor). Earlier the author has identified this recurring theme as the nearness problem - the problem of locating interacting objects as near to each other as possible. The catholic claim made for this approach is that it cuts across all branches of traditional systematic geography. Also, it is admitted outright that the probabilistic aspects of predictive geography do <u>not</u> conform to optimal arrangements.[4] With this important admission behind us the remainder of this paper will contain no further comment on location theory based on probabilities.

The "Schaefer versus almost all others" argument in favor of patterns rather than movements as the subject of theoretical geography dissolves to nothing when it is realized that particular optimal movements result in particular patterns, that the geometry and movements are intertwined in spatial harmony. In this paper, we will approach our subject from the point of view of Schaefer's patterns rather than interactions or movements of Ullman and others.[5] In some ways Schaefer's pattern approach to this dialectic is the more appealing of the two since the patterns are so easy to "map." Of course, the "maps" are not true geographic maps since they do not apply to any particular locations on the earth's surface; rather, they are idealized "maps" such as Köppen's Hypothetical Continent.

Very recent substantive work, especially by Boyce and Clark, Dacey, Nordbeck, and Tobler, is sharply in the spirit of a Universal Systematic Geography.[6] Dacey especially has been engaged in work on patterns. But the basic inspiration for this paper was received from Christaller and one of Tobler's favorite nongeographic teachers, Thompson.[7]

Christaller makes two isotropic assumptions; (1) that the density of his base object is uniform (even distribution of rural population) and (2) that his space is not twisted (uniform transportation in all directions). Thompson's work, dealing with the spatial but nongeographic aspects of life forms, contains most suggestive material. The very title of Thompson's book, On Growth and Form, can be paraphrased On Movement and Geometry without too much violence to his intended meaning.

Startling resemblances between geographic problems and those in biology might tempt readers to conclude, using our own argument of conservation of academic effort, that a really efficient science of location would include the spatial aspects of biology, physics, and so forth, as a study in General Systems. I think not. The very range of spatial scale between various sciences introduces worlds so weird that spatial experts in one have little to say to spatial experts in another. Subatomic physics is at a scale so different from the earth's surface as the home of man that the space is qualitatively almost totally foreign. Subatomic geometries have been superficially apposed to astronomical geometry by comparisons between the orbits of planets and electrons. In spite of early and prolonged attempts at such synthesis, efforts have been unsuccessful and the persistent experience of failure has been codified as "reductionism." However, while grandiose comparisons have been general failures some cross inspiration can be expected. For instance, Eigen values obtained from the characteristics of matrices have been used to explain ring jumps of electrons. In geography we might explore the characteristic of linear transformations as revealed in their matrices to see if Christaller's settlements occur in rings.[8] This exploration in the geometry of symmetry might shed light on Christaller's crucial "fixed k" assumption.

Electrical engineers, biologists, economists and others will learn much from us, and we from them, but this merely expresses the ultimate universality and interconnection of all knowledge. The logical conclusion to be drawn from successful borrowing of spatial notions in biology and elsewhere is not necessarily the horror of feeling compelled to claim all spatial problems are geography simply because all geography seems to be spatial. Throughout this paper references are made to nongeographic subjects where geographers can expect both to learn and to teach, but no wild territorial claims are made.

Returning to the main point of Thompson's insights, he noticed that one could accurately measure the length of a given specie of fish by measuring its weight. He assumed a constant density and shape to the fish. The outline of an object can be considered to be a property of density. Where the object ends its density suddenly drops to zero, a discontinuity to its density. Therefore, internal density and external density (shape) are in fact essentially the single element of density. To a modern geographer the resembalances and differences to Christaller are immediately apparent. Christaller avoided the problems that the irregular density of shape introduce, that is, he avoided the boundary problem, by the often used device of imagining an infinite plain. Both Thompson and Christaller assumed uniform internal density. Christaller also assumed a uniform transportation surface while Thompson simply ignored the internal circulation of his fish. Furthermore, one of the ways in which we have powerfully extended Christaller's original work is in the direction of understanding the basic "dimensional tension" involved in the problem of trying to spread points over an area.[9] So that while Christaller himself did not explore the dimensional problems implicit in his work, his intellectual descendents were compelled too. Thus, when Thompson emphasized the dimensional aspects of his fish, another correspondence to Christaller's work is established.

To translate biology into geography, consider the problem of predicting the volume of water at a river's mouth (weight of fish). If the rainfall is constant over the valley basins (fish are of uniform density) and the shapes of the basins (fish) the same, then the lengths of the valleys (fish) are proportional to the capacities (weights) of the rivers (fish). In addition, the pattern of the river system (veins and arteries) depends on the slope (internal transportability) of the terrain (fishes body). We now identify the spatial elements involved as dimension, morphology and density. The three elements all appear to be mathematical groups and therefore highly independent of one another.[16] At this speculative point in our understanding it also appears that there is an order to their fundamentality. Dimensions seem to be the most basic followed by morphology and lastly density. Christaller, Thünen and others exhibit a tendency to want to reduce the problems to elemental dimensional form. Geographers are fairly skilled in shifting dimensions. A dot map of elevation or a dot map of mean annual rainfall has every bit as much claim to legitimacy as hidebound isoline representations, but these are transformations within the set of dimensions. The transformations between dimensions and morphology and/or density are much more violent and more rare.

The truly elemental nature of our discussion strikes home when we notice that the three elements are actually three fundamentally different ways of predicting location. Therefore, I believe, we are at the heart of theoretical geography.

A. Patterns of Dimensional Location

In this section of our discussion only dimensions will be allowed to vary. Morphology and density are assumed to be uniform.

To obtain more accuracy and cover certain areas in classical climatology and oceanography and new problems in political geography, we could deal with three dimensions (volumes) as well as zero (points), one (lines) and two (areas) but it would merely clutter the argument so we will not include three or higher space.

1. "Locate points on lines so that they are as near to the line as possible." The dual and equivalent statement is, "Arrange points on lines so that they are as far from each other as possible."

The pattern is a uniform distribution of points along a line. (Figure 1.)

Figure 1. Points space uniformly along a line.

The scale of the spacing in this and subsequent examples depends on the ratios of the objects to each other and is trivial in terms of our discussion. Notice that if only one point is involved the median center is the solution; therefore, the problem can be thought of as one of "multiple medians." The dual statement suggests an analogue computer based on magnets.[11] If small bar magnets are skewered through corks and all the magnets turned the same way so as to mutually repell each other and then the corked magnets are allowed to float in a long narrow trough of water, they will form a uniform pattern along a line.

One method of obtaining a grasp of the power of the pattern is to stare at the unlabelled pattern and ask yourself "Of what is this a map?" Some possible answers include filling stations along a highway, major volcanic peaks along the Cascades and the distribution of ice cream vendors along a beach. Notice that these suggested applications to the earth's surface are more than shallow spatial coincidences. For instance, the total travel cost along a beach for the consumer of ice cream is minimized by such a pattern.[12] The volcanic pattern minimizes the movement of magma in the fissure, or put in another way, the uniform distribution marks points of the greatest internal pressure.

Since we are approaching our subject this time from Schaefer's point of view of patterns, in the remainder of the article we will not point out the rather obvious examples of minimized movements which <u>each</u> pattern represents.

2. "Locate straight lines in an area so that the lines are as near to the area as possible." The dual; "Arrange straight lines in an area so that the lines are as far from each other as possible." The pattern is one of straight parallel lines evenly spaced. (Figure 2.)

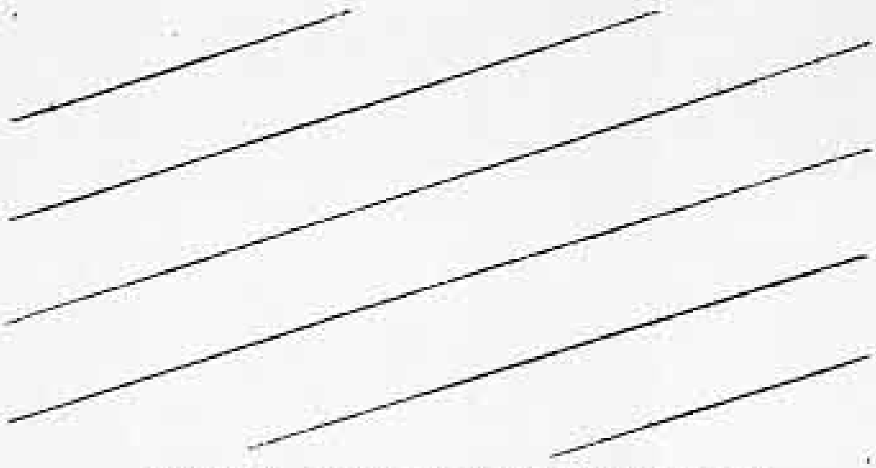


Figure 2. Parallel lines evenly spaced.

Straight lines are implied since we are allowing no morphology in our dimensional locations. Again the problem might be thought of in terms of self-repulsing floating magnets. This time the magnets are all tied together by strings to form self-repulsing lines. Notice that a single line "trying to get away from itself" is a straight line. The dual statement of placing objects as far from each other as possible is becoming rather tedious and will be dropped in most subsequent examples. It is well known that any extremum problem can be stated as a maximum or a minimum but still the reader might gain insight by conducting the exercise by himself.

Viewing the pattern as a map, such obvious translations as ridge and valley topography come to mind. I believe that the parallel line pattern is the most basic line-in-area pattern and that it is the <u>fundamental</u> river system, railroad pattern, etcetera.

3. "Locate points in an area as near to the area as possible." The pattern is the equilateral triangular distribution made so famous by Christaller. (Fig. 3a)

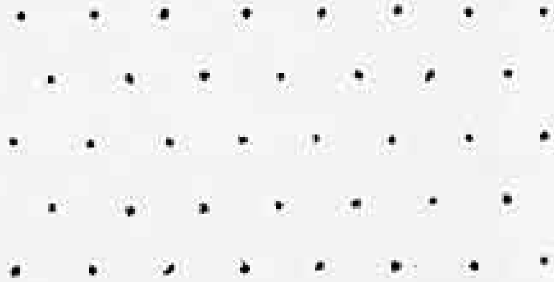


Figure 3a. Equilateral triangular distribution of points.

Possible maps include a range of objects such as classical settlement patterns, distributions of wild animals and efficient arrangements of oil wells.

So far we have drawn attention to the dialectic between the geometry and the movements but the dialectic can be even further expanded. Notice that the equilateral triangular distribution formally implies a hexagonal net. (Figure 3b.)

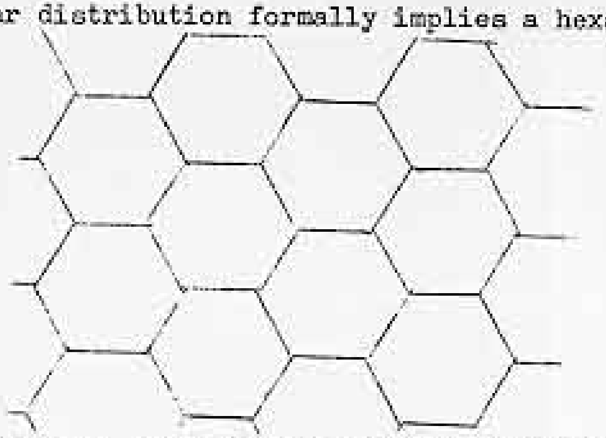


Figure 3b. Net of regular hexagons without points.

That is, we can establish the center of each cell in the net and thus generate the equilateral triangular distribution, or we can ask for the locus of all points halfway between the equilateral triangular distribution and immediately generate the net of hexagons. Stated in central place terms, the location of the central places determines the trade area and vice versa. It is really redundant to give both. These geometric duals appear often and, in combination with the implied movements, give a rich understanding from little. For instance, in Christaller's theory, with the single exogenous variable of k, say equals seven, consider the rich map we can draw. We can place settlements of various sizes, know their range of goods, their market boundaries, the movements to and from the centers to their hinterlands and even a great deal about the structure of their prices. The richness of Christaller's theoretical concepts have their foundations in the various formal mathematical dualities.

4. "Locate (straight) lines as near to points uniformly (equilateral triangularly) distributed in an area as possible." The pattern is, not too surprisingly, a set of parallel evenly spaced lines. (Figure 4.) The points must be

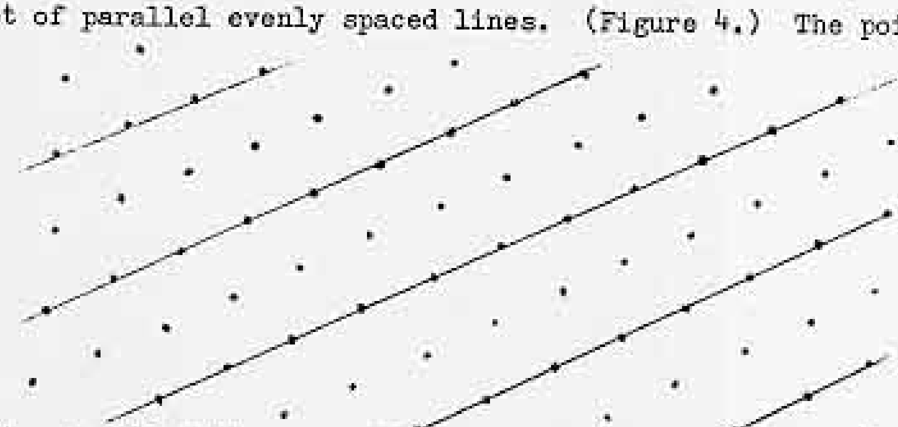


Figure 4. Parallel lines evenly spaced among an equilateral triangular distribution of points.

arranged in a uniform pattern or they will take on a variable density. Since one can project a methematical surface into a density surface and conversely, obviously, only a uniform distribution of points will be "flat."

If we consider just one line notice how closely the problem resembles that of regression. In the first example and in the immediately previous example, the problems can be thought of as one of finding multiple medians. Here the problem is one of multiple regression in the sense of simultaneously fitting many lines to the "data."

This last example has introduced objects of three different dimensions at once--lines and points located in an area. There is no reason to stop here. We could have four different kinds of points, say representing farmsteads, villages, towns and cities all simultaneously located as near to each other as possible. What is evidently needed is some notation system for expressing the possible dimensional combinations. Often good notation systems are suggestive of mathematical relationships that further simplify and provide deeper insight, but a search for notation would place us ahead of ourselves. Instead, let us continue an informal exploration of the half seen theoretical landscape swirling out of the fog.

B. Patterns of Morphological Location

Points can not be shaped but both lines and areas can. Some mist still hangs over what we mean by shaped space. An area of disuniform transport cost can be thought of as stretched and puckered so that circles of equal cost would not draw on the area as circles of metric distance. These circles can be thought of as not just Tissot's circles which preserve circles in the small in conformal maps, but circles in the large as well so that both small and large angles, and thus all shapes, are preserved. A true "conformality" in the broadest application of the word, We refer to these space puckerings and stretchings as "internally shaped." Notice, again that lines as well as areas have the possibility of being internally shaped.

As an aside, we should be dissatisfied with our over reference to cost-miles and time-miles and even with Tobler's sophisticated utile-miles.[13] Perhaps our almost exclusive concern with such space-warpers is due to the disproportionate influence of economic geography in current theoretical work. We need a grisly "death-miles" distance to explain human migration of a gross planetary sort. In climatology there exists the space twister of Coreolis acceleration. Coreolis-miles are just as legitimate as cost-miles.

All the previously mentioned patterns are seriously affected if we introduce the twisting of space. Obviously, if two points are at half the real miles as compared to the earth miles, they are located twice as close as an areal photograph would indicate. The reader can readily imagine many examples but some are explicated here because they seem to shed genuine insight of a rather startling sort.

1. "Locate finite areas of different real-miles as near to a single point as possible." The pattern is one of concentric circles whose radii increase in unequal increments. (Figure 5) Treated as a map, Thünen rings of agriculture

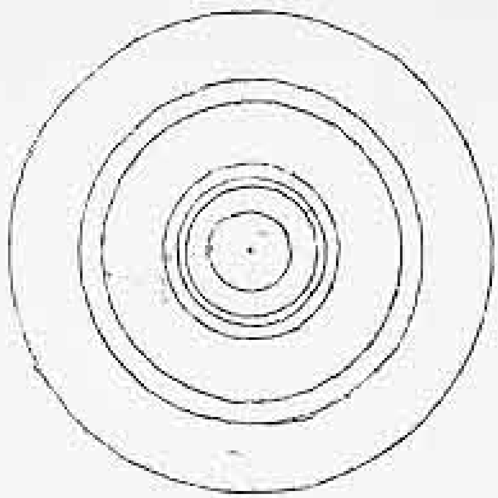


Figure 5. Concentric circles whose radii increase in unequal increments to produce irregular rings.

immediately come to mind. The rings suggest ecological circles of animals, from frogs to camels around a desert water hole. Rings of volcanic debris are caused by varying transportability of expelled material.

2. "Locate finite areas of different real miles as near to a straight line as possible." (Figure 6.) Thünen strips, with the central point replaced by a

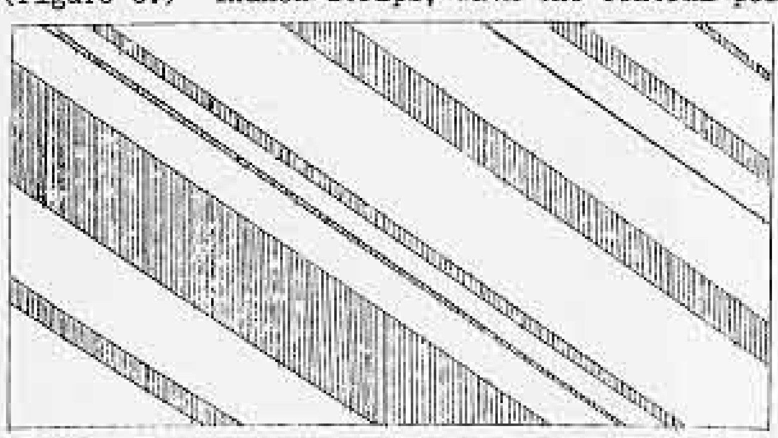


Figure 6. Strips of unequal width parallel to a base line.

central line, the most obvious extension imaginable but to my knowledge completely neglected, is apparent. Besides the strips of agricultural land use along wilderness roads, perhaps littoral-using ocean animals are similarly arranged with a band of short flight birds toward the inside and salmon and seals in the farthest band.

We have accepted that much of the movement on the earth's surface can be explained by least-effort paths. Therefore, the shape of the space, the morphological pattern to the space, can yield infinite patterns of objects. If the space is symmetrically shaped, the objects often form beautiful patterns. Let us examine some of the patterns of flow due to fairly symmetric space twisting. The brilliant Beckmann has produced several "maps" showing the flow of economic goods with various twistings.[14] (Figure 7.) Prager, in the Beckmann tradition

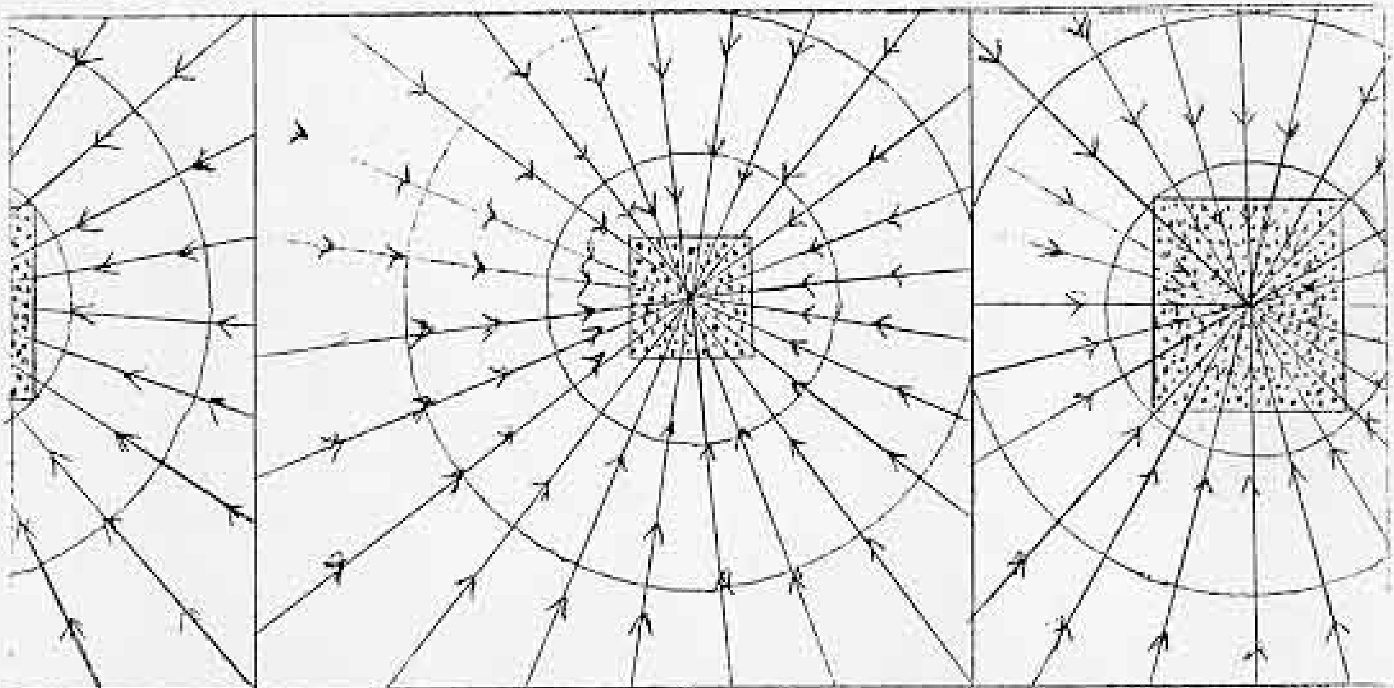


Figure 7. Beckmann patterns of flow. Stippled region indicates area of consumption.

has produced "maps" of telephone or any-other-type-of-network patterns.[15] (Fig. 8.)

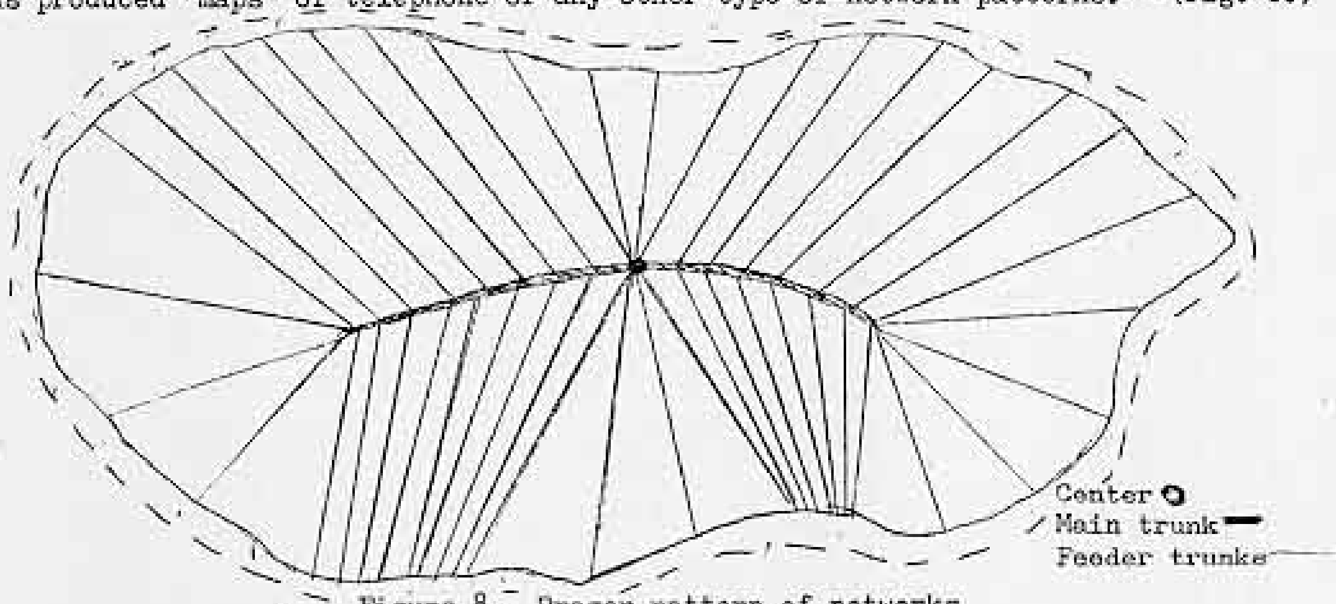


Figure 8. Prager pattern of networks.

There are several examples that are available to geography from our traditional area of interest in physical geography. Figure 9 shows the pattern of wind flow

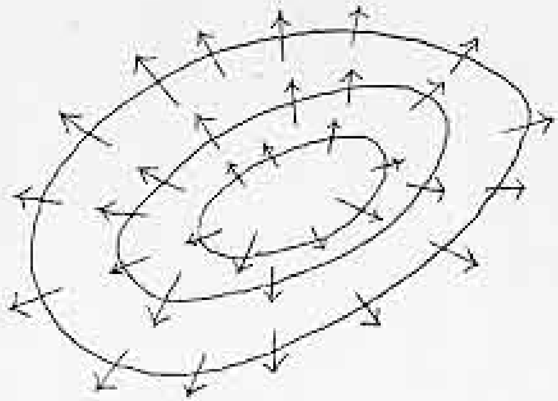
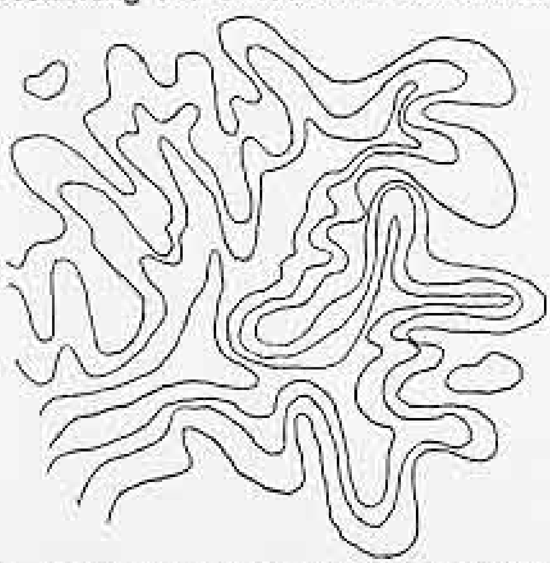
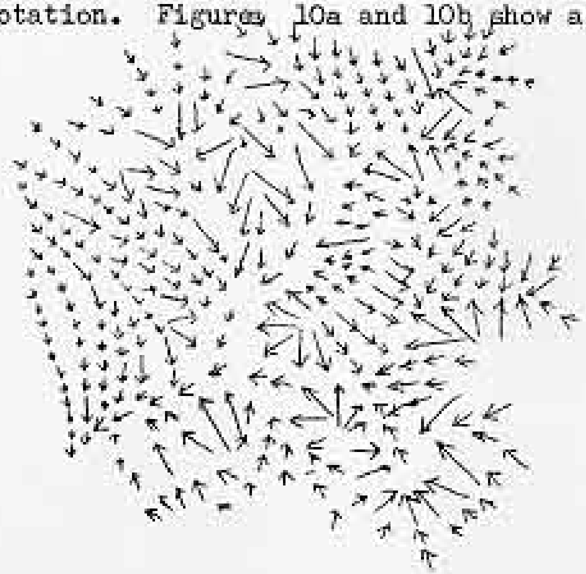


Figure 9. Wind pattern in a high, while discounting earth's rotation.

discounting the effects of the earth's rotation. Figures 10a and 10b show a



10a Contours of Stream Valley Pattern



10b Velocity Vectors - Length of arrows proportional to slope.

typical stream pattern literally flowing down the potential surface. Maybe it occurred to traditional physical geographers that their slopes were merely space-twisters varying the transportability of water. Maybe it occurred to them that the flow of water over the terrain is a vector flow over the potential map. But after Lehman in 1874 presented the geographic world with the hachure map, why is it not generally known among geographers almost a hundred years later that his hachures were merely a crude vector representation of the potential (contour) dual? Flow diagrams such as streams, are density vectors as opposed to velocity vectors. Actually the density vectors can be computed from the velocity vectors but not the other way around. That is, Figure 10a is reproducible from Figure 10b but not from Figure 11.

4. REPRINT

Imperfections in the Uniform Plane

Michael F. Dacey

with Forewords by John D. Nystuen, The University of Michigan

In this section, *Solstice* Board member, John D. Nystuen, selects a paper from the collected papers of the Michigan Inter-University Community of Mathematical Geographers (MICMOG) (of which he is Editor) to reprint here, some 30 years after its initial presentation. In addition to the reprint of work of Michael Dacey, Nystuen's original Foreword, and introduction of Dacey and his work to the assembled MICMOG group, is also reprinted. In addition, a new Foreword by Nystuen takes a look at the Dacey paper in retrospect. The paper is reprinted with permission of Nystuen, on behalf of the Michigan Inter-University Community of Mathematical Geographers.

Foreword, December, 1994

John D. Nystuen

Thirty years ago Michael Dacey contributed to the development of spatial statistics in highly original ways. Many of the ideas he used and introduced to the literature in the 1960s are now part of generally accepted spatial theory. For example, he was one of the first to use the idea of a dimensional transformation to permit evaluations of the spatial association of point and area phenomena. The transformational approach proved useful as a general concept as Keith Clarke has demonstrated in his interesting book (Clarke, 1990). Arthur Getis, a colleague of Dacey's, and Barry Boots used many of Dacey's ideas in their book (Getis and Boots, 1978) about modelling spatial process.

Today, vigorous effort is being expended on incorporating spatial analysis functions into Geographic Information Systems (GIS) software. We are re-issuing one of Dacey's seminal works to bring to the attention of contemporary scholars an important source of many of the concepts now becoming accessible to general uses of GIS technology. Dacey's work now speaks to another generation.

References

Clarke, Keith C. 1990. *Analytical and Computer Cartography*, Prentice-Hall, Englewood Cliffs, NJ.

Getis, A. and Boots, B. 1978. *Models of Spatial Processes*, Cambridge University Press, Cambridge.

Foreword, May, 1964

John D. Nystuen

We are pleased to present to our readers a paper by Professor Michael F. Dacey. Many of us are aware, if only vaguely, of his provocative and voluminous writings. Professor Dacey has penetrated deeply into realms where few, if any, have gone before. He travels alone and has left but a thin trail of mimeographed papers as scent. The track is now long and difficult to follow and he does not rest. He has allowed one of his works to become discussion paper #4 of our series. We hope this will expose his activities to a wider audience. Some may be inspired to join him in the new work that he is doing. I hope so. Certainly we must keep in contact with him. Regrettably many of his results depend upon his previous statements now difficult to obtain. I will attempt in this foreword a short review of the pertinent ideas by way of a summary of this paper. I have also added, with his permission, a glossary of symbols at the end of the paper.

Michael Dacey has for several years explored abstract spatial patterns using probabilistic methods. This paper is one of a series of such studies. Most of the work provides empirical examples of the concepts. The contrast in methodologies displayed between discussion paper #3 (W. Bunge, "Patterns of Location") and this one is marked. Professor Bunge turns away from probabilistic formulations (see page 3 of "Patterns of Location") and Professor Dacey rejects deterministic models (see page 1 below). I believe the relative worth of these two broad approaches to abstract geography will receive increasing attention in the literature. There is much precedent for concern over this question in other disciplines. Clearly Dacey accepts the value of a probabilistic approach.

It may aid the reader if the paper is viewed as consisting of six parts.

1. Professor Dacey first describes an abstract model of imperfections in a uniform plane. The characteristics of this model are specified in a general way. I believe that Professor Dacey is the first to suggest models where non-random patterns are disturbed by random variables (see Dacey and Tung, 1962).
2. The point pattern which results from the above mentioned model is to be summarized quantitatively in such a fashion that it can be compared with some actual geographic point pattern. Professor Dacey calls upon his previous extensive investigations of nearest neighbor statistics to do this job [1]. He specifies how measures of the distances to the 1st nearest, 2nd nearest, ... kth nearest neighbors of a sample of points in the point pattern may be used to describe the point pattern by probability distributions of these lengths. The strategy is to then compare the probability distributions of the model with a geographic pattern using a simple χ^2 statistic. Professor Dacey is aware that nearest neighbor methods may be used to compare point-to-area relations as well as point-to-point relations. A point pattern is not simply a set of points. The points occupy a space for which a metric is defined. The metric makes possible distance measures between the points. The fact that there is a space creates the boundary problems mentioned in the text. The original purpose of these statistics was to test if points were more clustered or more even than random. Imagine a study area which is mostly empty but has in one small region an even distribution of points. Measuring distances between points and using the nearest neighbor test would indicate a point pattern more even than random. In one sense, however, they are clustered for they occupy only a small section of the study area. There is a strategy for this situation. Use another point set to represent the area. This may be done by using an even distribution of points in the area or by assigning points to the area at random. The second set of points now represents the study area. The area has been abstracted into a point pattern and the nearest neighbor method may be used. Measures between the two point sets now reveals the original point pattern to be clustered. The decision concerning which method to employ depends upon whether the phenomenon studied has a postulated interaction of point-to-point or point-to-area. The text indicates the procedure for using either method.
3. Theoretical order distances are specified by equations (16) and (17). The probability functions are made more explicit and operational by assuming each lattice point is disturbed by the same two dimensional normal variate. Professor Dacey has ample evidence

that these particular probability distributions are useful for this purpose.[2]

4. Solutions of the equations in the previous section would yield an analytic solution regarding expected order distances for various disturbance models. However, these equations prove very difficult to evaluate. Recourse to a simulated solution is sought. An *almost periodic disturbance model* is postulated. Its parameters are estimated from data on an actual pattern of urban places in Iowa. Using these parameters, a set of points conforming to the structure of the theoretical model is generated with random digits and tables of normal deviates. This artificial pattern is one of many possible representations of the theoretical pattern. It is presumed to display the type of pattern expected from an analytic solution if one could be found.
5. The author now has two patterns: one, a simulated theoretical pattern which conforms to the structure of the model; and the other, an actual urban place pattern in Iowa. He also is able to make the appropriate nearest neighbor measures which characterize each pattern. The frequency distributions are then compared using the χ^2 statistic.
6. In an addendum, the author presents further testing of his model by taking advantage of a computer program which generates the distance measures required. The paper ends.

It must be clear to the reader from the contents of this paper that Michael Dacey has indeed traveled over much ground. He has previously developed many of the results needed in this study. Many of his solutions and applications are ingenious. He exhibits an understanding of the theoretical implications of his work. He has a wide knowledge of the literature on probability and is able to adopt simulation methods and computer technology to his purpose. All he lacks is someone to talk to.

Endnotes

1. Examples of his statements on nearest neighbor measures include: "Analysis of Central Place Patterns by Nearest Neighbor Method," Seattle, May 1959, mimeographed; "Analysis of Central Place and Point Patterns by a Nearest Neighbor Method," *Proc. of IGU Symposium in Urban Geography*, Lund, 1960, pp. 55-75; "Identification of Randomness in Point Patterns," (with Tze-hsiung Tung), Philadelphia, June 1962; mimeographed. (Dacey and Tung is now forthcoming in the *Journal of Regional Science*, v. 4.

2. See references at the end of the paper and also: "Order Neighbor Statistics for a Class of Random Patterns in Multidimensional Space," *Annals*, Association of American Geographers, v. 53 (Dec. 1963): 505-515, "Certain Properties of Edges on a Polygon in a Two Dimensional Aggregate of Polygons Having Randomly Distributed Nuclei," Philadelphia, June 1963, mimeographed.

Imperfections in the Uniform Plane

Michael F. Dacey

Wharton School of Finance and Commerce
University of Pennsylvania
See end of article for additional information

A statistical formulation of the spatial properties of central place system is proposed. Currently, the theoretical locations of central places are specified by geometric or algebraic quantities. This type of statement leads to certain rejection of central place models, for it is inconceivable that any observed pattern of central places corresponds exactly to the specified geometry. A probabilistic formulation is preferred for empirical analysis because deviations from the precise locations are contained within the statement of the model.

In the classical theory of Christaller (1933) and Lösch (1939) central places form a honeycomb pattern or hexagonal lattice on the undifferentiated, unbounded plane. A probabilistic statement of this location pattern incorporates deviations from the precise lattice locations, and the deviations are subject to stochastic processes. This initial formulation of a probabilistic central place distribution uses the concept of imperfections in the uniform plane to define these deviations. Imperfections may be combined with the central place geometry in many ways. Here one basic formulation and two closely related models are proposed. The models possess some properties of the Christaller-Lösch system and evidently are not inconsistent with the spirit of central place theory.

This report has two purposes. First, a general model of imperfections in the uniform plane is constructed. Second, the application of a particular model to a map pattern is evaluated.

The map pattern of urban places in Iowa has been selected for an initial examination of the imperfection concept. The empirical test involves interpretation of parameters of the model in terms of phenomena commonly studied by geographers and estimation of these parameters from the Iowa map pattern. Because the formal statement of the model contains equations that are difficult to evaluate analytically, this initial study has used a simulation technique to obtain summary measures on theoretical patterns. Properties of a fabricated pattern are compared with the Iowa map pattern, and the level of agreement is found acceptable to the first approximation.

The Christaller Spatial Model

The theoretical distribution of central places may be expressed in terms of a plane lattice. Let P represent a plane symmetry lattice. Choosing any arbitrary point of this lattice as an origin point O, the location of any other given lattice point can be defined with respect to this origin by a vector T

$$T = ut_1 + vt_2 \tag{1}$$

where u and v are integers. The vector notation implies that the plane is constructed as a linear lattice having a translation period t_1 which is repeated periodically at an interval t_2. The translation periods t_1 and t_2 may be regarded as vectors separated by the angle g. Using K to denote a collection, the lattice points of P are defined by

$$P = KT = K(ut_1 + vt_2). \tag{2}$$

Central place theory conventionally uses a hexagonal lattice for which the translations t_1 and t_2 are of the same unit length and the angle of periodic rotation is $g = \pi/3$.

A more general discussion is obtained by not restricting attention to the hexagonal lattice. In this report P represents any plane lattice which may have a three-, four-, or six-fold axis. In applying the lattice to a particular problem, the translation periods t_1 and t_2 and the angle of rotation g need specification.

Types of Imperfections in the Uniform Plane

Three types of imperfection in the uniform plane are studied in this report. These imperfections are closely related to certain kinds of imperfections found in nearly perfect crystals. An introduction to crystal imperfections is found in Van Bueren (1961, especially Chapters 2-4) and an excellent synthesis of the concept of imperfection in the solid state is given by Seitz (1952). The basic principles of our formulation draw heavily upon concepts used in the study of crystals and the solid state; the mathematical formulation is, however, quite different.

The imperfections under consideration are identified as (i) dislocations or disturbances, (ii) vacant lattice sites and (iii) interstitial points. These three types of imperfections are most easily defined by considering two maps containing point symbols. For the present purposes assume the maps have identical area and number of points. One map represents a finite domain of the lattice P. The other map, called S, may show fabricated locations or the positions of actual objects. Figure 1 is "good" map S overlaid on a square P.

i. The term dislocation is more descriptive of the first imperfection, but it has a definite meaning in crystallography and solid state physics; so we shall call this imperfection a disturbance. A disturbance occurs when the location of a point is not exactly at a theoretical lattice site but is 'sufficiently' close so that with high degree of certainty a disturbed point is correctly associated with its theoretical location.

ii. A vacant lattice site occurs where no point is 'close' to a theoretical lattice site. Where two or more points occur in the vicinity of a lattice site, it is not called a vacant lattice site even though the one point correctly associated with that theoretical location may not be identifiable.

iii. An interstitial imperfection occurs in the uniform plane where a point is not identified with any lattice site. Interstitial locations occur where a point is too distant from a theoretical location to be associated with high degree of certainty with a particular lattice site, or where two or more points are located 'close' to a lattice site and the one point correctly assigned to that theoretical location is not identifiable.

These imperfections are not given precise definitions. In constructing the imperfection model more precise definitions are given.

A Model of the Imperfect Plane

One basic formulation and two modifications are described. All imperfections under consideration are the result of stochastic processes, in the space rather than the more common time dimension. The principal feature of an imperfection model is the imperfection in pattern related to disturbances or shocks from geometrically exact locations (Figure 1). While this single type of imperfection is adequate for many physical systems, it is probably too restrictive to encompass patterns formed by economic, social or cultural systems. To handle complex map patterns two additional types of two dimensional stochastic processes were studied. One type of imperfection generates interstitial points and is defined by a two dimensional, uniform, random variable. The other type of imperfection generates clusters of points and is defined by spatially contiguous probability distributions. Because the pattern of urban places in Iowa is relatively homogeneous and contains no examples of large metropolitan centers, it was not necessary to incorporate a contagious process in a model for the Iowa map pattern. For this reason, only the first two types of imperfections are discussed in this report.

The Disturbance Effect

Each lattice point of P is associated with a stochastic variable ξ. The ξ is the disturbance variable and defines the realized location of a point with respect to its theoretical lattice site. It is convenient to separate ξ into its two polar components: a distance ρ and a rotation angle θ. So, $\xi \equiv (\rho, \theta)$.

The displacement of the point s_{ab} from its equilibrium position $(at_1 + bt_2)$ is given by the random variable ξ_{ab}. So, the disturbed position of this point is

$$s_{ab} = at_1 + bt_2 + \xi_{ab}. \tag{3}$$

It is assumed that the same stochastic variable is associated with each lattice site. Then, if a point is disturbed from each lattice site the collection of randomly disturbed points is

$$S_1 = K(ut_1 + vt_2 + \xi_{ab}), \tag{4}$$

u and v integers. This notation indicates that ξ has translation period t_1 which is repeated periodically at an interval t_2. In this sense the stochastic variable is carried through space and is associated in turn with each lattice site. Accordingly, in point set S_1 each lattice site $(at_1 + bt_2)$ has exactly one corresponding disturbed point s_{ab}.

Vacant Lattice Sites

It is not necessary to apply a disturbance to each lattice site. Instead a lattice site and the variable ξ_{ab} may be taken in conjunction with a binary or on-off operator which nullifies the vectors defining some disturbed points so that the corresponding lattice sites are vacant. As a consequence, there is a sparser network of disturbed points than lattice sites. Because a disturbed point is not associated with each lattice site, the disturbance term is said to be repeated almost periodically. A more precise definition of the almost periodic disturbance is given.

A binary operator to produce vacant lattice sites is defined for $(at_1 + bt_2)$, denoted in symbols by β_{ab}, such that for $0 \leq \lambda \leq 1$,

$$\beta_{ab} = 1, \quad \text{with probability } \lambda$$

$$\beta_{ab} = 0, \quad \text{with probability } 1 - \lambda. \tag{5}$$

The vectors defining location of the disturbed point s_{ab} are multiplied by β_{ab} so that the disturbed point is realized with probability λ and is not defined with probability $(1 - \lambda)$. In more precise form, the location of the disturbed point having equilibrium position $(at_1 + bt_2)$ is

$$s_{ab} = \beta_{ab}(at_1 + bt_2 + \xi_{ab}) \tag{6}$$

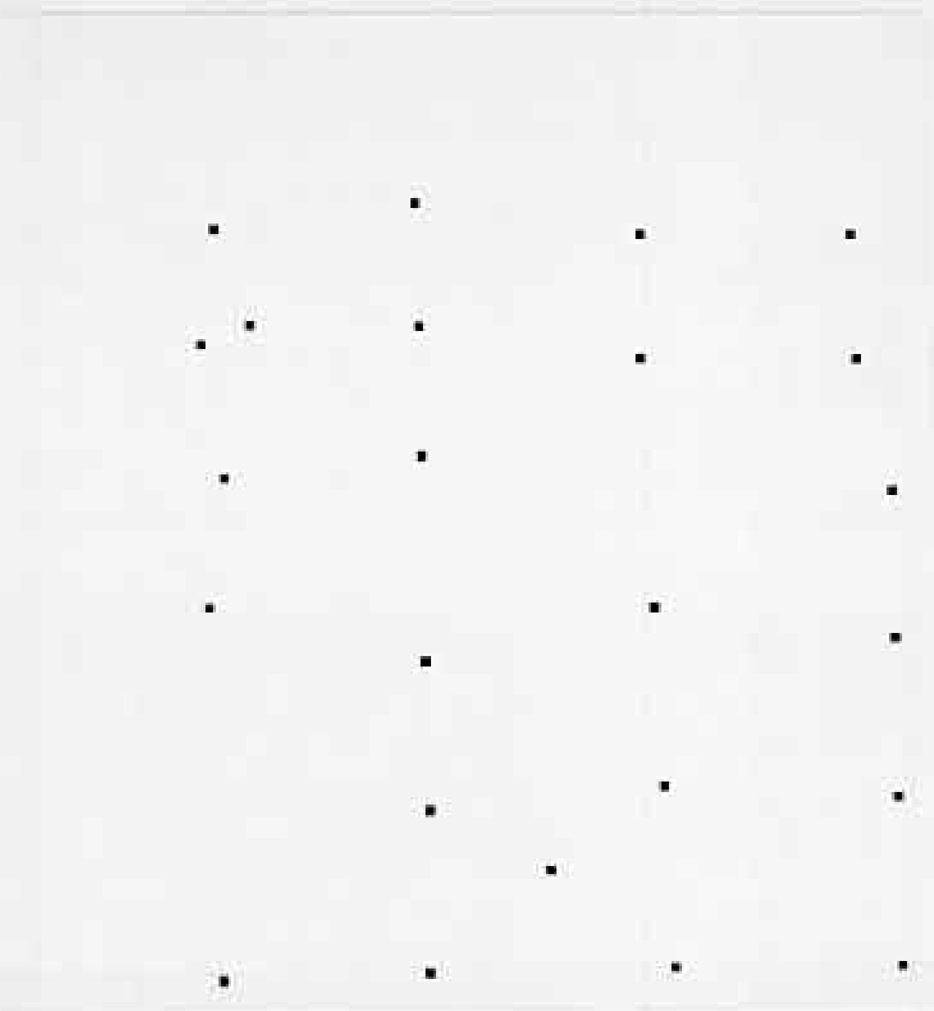


Figure 1. Map of imperfection model. Most symbols show disturbance effect on a square lattice. There are two vacant lattice sites, and two examples of interstitial points. Most map patterns are, of course, not this regular. This figure shows a six by four square lattice which has been altered as suggested.

with the usual convention that $s_{ab} = 0$ does not define a point at the lattice site 0. So, for $\beta_{ab} = 0$ the disturbed point s_{ab} does not exist, while for $\beta_{ab} = 1$ location is found precisely in the manner for the period disturbance.

Each lattice site is associated with the same stochastic variable and with the same binary operator. Accordingly, the relation (6) is carried through space with translation period t_1 repeated periodically at interval t_2. The collection of points generated by the almost periodic disturbance is

$$S_2 = K(\beta_{uv}(ut_1 + vt_2 + \xi_{uv}) \tag{7}$$

u and v integers. The S_2 is completely identified by the underlying lattice P, the probability λ, and the parameters specifying the components ρ and θ of the stochastic variable ξ. It is summarized by the parameter set $S(t_1, t_2; \lambda, \xi)$.

Uniform Random Disturbance

This collection of points, denoted by R, is a random point set. To make the definition explicit, an arbitrary origin is selected and the lattice point O of P is convenient. The R is specified by the theoretical frequency of points within distance r of the origin. Where the parameter γ is the expectation that a unit area contains a point belonging to R, put

$$p = \pi\gamma r^2 \tag{8}$$

where $\gamma > 0$. The frequency p describes any arbitrary disk of radius r, so that the distribution ξ is independent of the specified origin. It is a property of R, Feller (1957) that the distribution conforms to a Poisson process. The probability of finding exactly j points of R within any disk of radius r is $p^j e^{-p}/j!$.

Definition of the Basic Model

The model to be considered in this report is defined by the combination of an S and the R point sets; call this model M and

$$M = S \cup R. \tag{9}$$

This model is summarized by the parameter set $M(t_1, t_2; \lambda, \xi; \mu)$, where $\mu = (\lambda + \gamma)$. For a model containing S and R points only, μ is the mean density of points per unit area.

Several interesting formulations of M are defined by special values of the parameters λ and γ.

The *periodic disturbance model* M_1 is given by $\lambda = 1$, for one disturbed point is associated with each lattice site. A *complete periodic disturbance model* also has $\gamma = 0$, for each point is disturbed from a lattice site and there are no random points from R.

The *almost periodic disturbance model*, called M_2, is given by $0 < \lambda < 1$. The magnitude of γ determines if M_2 has a one-to-one correspondence of points to lattice sites or if M_2 has more or less points than lattice sites. If $\gamma = 1 - \lambda$ the theoretical density of points belonging to S_2 and R equals the density of lattice sites. If $\gamma > 1 - \lambda$ the expected number of points exceeds the number of lattice sites, while the expected number of points is less for $\gamma < 1 - \lambda$.

The point set given for $\lambda = 0$ is a random point pattern. It is of course recognized that R is only one of many point sets that could be combined with S_1 or S_2 disturbed points.

Description of Pattern

The disturbance models are described by the underlying lattice P, the density measures λ and γ and the disturbance process ξ. The combination of these parameters produce disturbed and interstitial points and vacant lattice sites in the uniform plane. In a formal sense a model is completely specified by the lattice parameters and the several probability functions. This specification of a model does not, however, describe or summarize in any useful fashion the point pattern generated by a particular model. But, numerical summary of point pattern M is prerequisite to test of the hypothesis that an observed map pattern is similar to an imperfection pattern.

To measure the level of correspondence between observed and theoretical patterns there is need for (i) measurements on one or more properties of the observed pattern and (ii) theoretical values for the same properties on the pattern defined by the model. In addition, if parameter values for the model are estimated from the observed pattern, the properties for test of similarity between observed and theoretical patterns should be independent of the properties initially used to estimate parameters.

In this report pattern is summarized by two classes of order distance statistics. The methods are described briefly and then their utility as descriptive measures of pattern are indicated.

Point to Point Order Distances

Let i represent any arbitrary point in a point pattern Q. The measured map distance from i to the j nearest point is represented by R_{ij}. J measurements are taken from i and are ordered to satisfy the inequalities

$$R_{i1} < R_{i2} < \cdots < R_{ij} < \cdots < R_{iJ} \tag{10}$$

and the R_{ij} is called the j order distance. For description of a bounded map pattern the j order distance is recorded only if R_{ij} is less than the distance from i to the nearest map boundary. The chance of bias due to the influence of boundaries is reduced by this constraint, but there is loss of information to the pattern description because all distance relations are not utilized.

The R_{ij} measurements reflect the arbitrary map metric. The dimensional constant which eliminates effect of scale is $d^{1/2}$, where d is the density of points in Q. Measurements in Q are reduced to standardized distance by the transformation

$$r_{ij} = d^{1/2} R_{ij}. \tag{11}$$

Standard distances are used in this report to describe all patterns.

Let I denote a collection of points in Q, and $i \in I$. One description of Q uses standard distances from each origin point $i \in I$ to the J nearest points.

Locus to Point Order Distances

A second description of pattern uses distance measurements from coordinate locations to points. Let L define a set of locations in Q and in general a locus $\ell \in L$ is not a point symbol of Q. The measured distance in Q from locus ℓ to the h nearest point is denoted by $R_{\ell h}$. The measurements from ℓ are ordered by distance and put in standard form; in symbols

$$r_{\ell 1} < r_{\ell 2} < \cdots < r_{\ell h} < \cdots < r_{\ell H} \tag{12}$$

$$r_{\ell h} = d^{1/2} R_{\ell h}. \tag{13}$$

The second description of Q uses standard distances from each locus $\ell \in L$ to the H nearest points. The boundary constraint pertains to these distances also.

Sampling Methods

The elements of I may consist of all or a sample of points in Q. For this study a census was taken, largely because of small pattern size.

The loci in L necessarily constitute a sample, and these locations may be designated by random, stratified or uniform sampling methods. The most efficient mesh for plane sampling has been studied by a number of writers, as Zubrzycki (1961) and Dalenius, Hajek, and Zubrzycki (1961), but there are no general conclusions. This study used random sampling, largely because the patterns of interest contain high degree of uniformity in spacing and random sampling is probably less sensitive to this type of spatial bias. However, this topic requires study.

Summary Description of Pattern

A point pattern may be summarized by (i) the lower moments of the j and h order distances or (ii) the frequency distributions of these order distances. The j order point to point distances provide a quantitative summary of the arrangement of points with respect to other points of the pattern, but these distances do not explicitly reflect the arrangement of points with respect to the map space. The complementary h order locus to point distances provide a quantitative summary of the arrangement of points with respect to the loci in L. To the degree the sample mesh of L is a measure of the map space, h order distances also summarize the arrangement of points with respect to the map space. Because these two classes of distances reflect two different aspects of pattern, this type of summary statement captures many of the subtle characteristics composing a point pattern.

Comparison of Map Patterns

The descriptive measures provide a basis for evaluating the degree of similarity between two or more patterns. patterns are called similar if the order distances summarizing each of the patterns have"the same statistical parameters. The standardized distances allow direct comparison of any two point patterns, for the distances represented by the variable r (either r_{ij} or $r_{\ell h}$ are normalized to account for differences in scale, unit measurement and density of points. Using either means or frequency distributions of order distances, the hypothesis that two or more sets of measurements belong to the same statistical population may be tested by standard procedures.

Theoretical Order Distances

This paragraph considers the basic derivation of order distances for imperfection models. The derivations are simplified by studying (i) lattices for which $t_1 = t_2$, (ii) nearest neighbor situations only, and (iii) the stochastic variable ξ defined by the normal law.

Two nearest neighbor lattice sites are separated by the distance t $(= t_1 = t_2)$. Let the random variable X denote the distance between two disturbed points associated with any two nearest neighbor lattice sites. It requires only elementary geometry to show that the distance between points (ρ_1, θ_1) and (ρ_2, θ_2) is

$$x = ((\rho_1 \cos\theta_1 - \rho_2 \cos\theta_2 + t)^2 + ((\rho_1 \sin\theta_1 - \rho_2 \sin\theta_2)^2)^{1/2}. \tag{14}$$

The simplest derivation of order distances is for the complete periodic disturbance model $(\lambda = 1$ and $\gamma = 0)$ on the hexagonal lattice. Let m $(= 6)$ denote the number of nearest neighbors to each lattice site. We consider the distances from an arbitrary point i at $(at_1 + bt_2 + \xi_{ab})$. It is assumed that the m nearest points to i are disturbed from nearest neighbor lattice sites only. The x_k is the distance from point i to the k $(= 1, 2, \ldots, m)$ nearest point.

If the disturbance term is identical and independent for each lattice site, the m distances from i may be interpreted as m independent observations in a sample of size m from the population defined by the random variable X. Because the observations are ordered from shortest to longest, x_k is the kth order statistic. It is well known that the distribution function of the kth order statistic is given by

$$\Psi(x_k) = \frac{m!}{(k-1)!(m-k)!} F^{k-1}(\omega) F^{k-1}(\omega)(1-F(\omega))^{m-k} f(\omega) \qquad (15)$$

where $f(\omega) = dF(\omega)$ and the variable X, after making the probability transformation for a specified $f(\rho)$ and $f(\theta)$, is substituted for ω. The z crude moment of the k order statistic for the complete periodic disturbance model is

$$\mu_z{}'(x_k) = \frac{m!}{(k-1)!(m-k)!} F^{k-1}(\omega) \int_0^\infty \omega^z F^{k-1}(\omega)(1-F(\omega))^{m-k} f(\omega) d\omega. \qquad (16)$$

The derivation is far more complex if the lattice is not hexagonal and undoubtedly requires more advanced concepts than provided by elementary probability methods. Moreover, even in this simplified case, numerical evaluation of (16) is not necessarily possible by elementary procedures.

In the statement of disturbance models the normal law was interpreted in polar coordinates by the folded half-normal distribution; that is, the distribution function for location about a lattice site is

$$F(\xi) = F(\rho, \theta) = \int_0^\rho \int_0^\theta f(\rho) f(\theta)\, d\rho\, d\theta \qquad (17)$$

where

$$f(\rho) = \sqrt{2}\exp(-\rho^2/2\sigma^2)/(\sigma\sqrt{\pi}) \qquad \rho > 0$$
$$f(\theta) = (2\pi)^{-1} \qquad 0 < \theta < 2\pi.$$

It seems appropriate to accept that $f(\xi)$ is identical for each lattice site so that the parameter σ is constant throughout the lattice space. Using (17) to define (14) and substituting the resulting probability transformation into (16) gives an expression for order statistics that, for me, is totally intractable.

Some simplification is gained by interpreting the normal law by the bivariate or circular normal distribution. In this case the distance variable X has a well known form. It may be shown that the distribution function is

$$F(x) = 1/2\exp(-t^2/2\eta^2) \int_0^{(x/\eta)^2} e^{-x/2} I_0(tx^{1/2}/\eta)\, dx \quad x > 0 \qquad (18)$$

where $\eta = 2\sigma^2$ and $I_0(\bullet)$ is the modified Bessel function of the first kind of zero order. This expression is recognized as the integral of the non-central χ^2 with two degrees of freedom. In a slightly different form it occurs as a basic distribution function in bombing or coverage problems, Germond (1950). By substituting (18) for $F(\omega)$, (16) gives the z crude moment of order statistics from a non-central χ^2 distribution; however, tables of values have not been published.

It is apparent that even the simplest imperfection model yields equations that are difficult to evaluate. Where $\lambda \neq 1$ and/or $\gamma \neq 0$ the equation systems are immensely more complex and numerical evaluation may be considered, for any practical purpose at this time, impossible. In order to circumvent these mathematical problems the imperfection model has been evaluated by simulation of an equation system for a given set of parameter values.

Analysis of the Pattern of Urban Places in Iowa

The imperfection models were designed to produce types of patterns and distributions studied in the social sciences. Moreover, the particular class of patterns motivating the present formulation are formed by map representations of urban places. As a partial evaluation of the adequacy of the imperfection model to replicate town and city patterns, the distribution of urban places in Iowa, 1950, is studied.

Many parameters of the Iowa distribution are already available in Dacey (1963a). These data provide empirical estimates of parameters for application of the imperfection model to the Iowa pattern. Using estimated parameters, the degree of correspondence of M_2 with the observed pattern of urban places is analyzed. Simulation is used to evaluate the theoretical imperfection model.

Almost Periodic Disturbance Model

The almost periodic disturbance model M_2 is specified by three sets of parameters:

t_1, t_2 and g identify the underlying lattice P,

ξ specifies the disturbance term generating the point set S_2 and

λ and γ are the scale densities for the point sets S_2 and R, respectively.

These three sets of parameters are given numerical values by relating the imperfection concept to structural features of the Iowa map pattern. In this construction, each parameter is described in terms of the corresponding property of the Iowa pattern. Since the theoretical pattern is synthetically fabricated, the definitions and interpretations of parameters are biased toward operational statements.

Lattice Parameters

The M_2 is fabricated as a rectangular map space containing the domain of a square lattice. The domain is of dimensions 12 by 18 and contains 96 points. Thus, the parameters are $t_1 = t_2 = 1$, $g = \pi/2$.

The primitive cells of the square lattice have an abstract correspondence to counties, and in this context lattice points represent the geographic center of counties. This lattice has some resemblance to the Iowa map. In gross form Iowa is roughly a rectangle and most counties in Iowa are approximately square. However, the counties do not form a square grid, largely because of surveying adjustments for the earth's curvature. An alternative, and possibly a closer, approximation to the Iowa structure is the diamond lattice.

The lattice has 96 squares while Iowa has 99 counties. There is no formal advantage to using a lattice of approximately the same dimensions as the study area.

For specification of other parameters the following relations are established between M_2 and the Iowa map:

i. square lattice cells of M_2 are equated with Iowa counties,

ii. lattice points of M_2 are equated with geographic centers of counties,

iii. S_2 and R points are equated with urban places.

Using this dictionary (α) the distribution function for distance from lattice site to S_2 point is estimated from the observed distances from geographic center of counties to nearest urban place and (β) the frequency distribution of points in primitive lattice cells is estimated from the observed frequency distribution of urban places in counties. These two properties are evidently independent of the order distances used to summarize observed and theoretical patterns.

Disturbance Variables

In my earlier study of Iowa it was shown that for interior counties containing an urban place the distance from the geographic center to nearest urban place was closely approximated by the folded half-normal distribution, as defined for $f(\rho)$ in (17), with scale parameter $\sigma = 0.2286$. Observed and calculated frequency distributions are compared in Table 1.

The angular component θ of the disturbance term is taken as a uniform random variable, as defined in (17). No evidence is presented for this assumption, so the uniform variable is entered into the model on the theoretical consideration that a completely chance factor occurs in the disturbance process. However, in examining the location of places with respect to geographic centers I found no evidence of directional bias.

On the basis of these estimates, the vector component ρ and the angular component θ of the disturbance variable ξ are defined for M_2 by the folded, uniform bivariate distribution (17).

Scale Variables

The remaining two parameters of M_2 are the density measures λ and γ. Because M_2 contains only S_2 and R points, the density of all points is $\mu = \lambda + \gamma$. For the Iowa map pattern there are 93 places and 99 counties, so the estimated density of total points in M_2 is $(93/99) = \mu$.

The individual densities λ and γ were estimated from the frequency distribution of urban places among Iowa counties, Table 2. A two parameter probability density function that gives a good fit to the observed frequencies has been stated by Dacey (1963b). By assuming that each disturbed point in S_2 is always located in the primitive cell of its theoretical lattice site and that each random point in R has an equal probability of occurring in each primitive cell, the probability that a cell contains x points is

$$f(x;\lambda,\mu) = (\gamma^{x+1}e^{-\gamma}/x!) + (x\lambda\gamma^{x-1}e^{-\gamma}/x!) \tag{19}$$

where $\gamma = \mu - \lambda$ and $x = 0, 1, \ldots$. The parameter λ was estimated by the method of moments from the distribution of urban places among Iowa counties. Table 2 compares observed and expected frequencies for the parameters $\lambda = 0.74$, $\gamma = 0.20$ and $\mu = 0.94 \cong 93/99$.

Comparison of M_2 and Iowa

A synthetic pattern was constructed from the pattern M_2 for the parameters

$$t_1 = t_2 = 1 \quad g = \pi/2$$

$$\sigma = 0.2286 \quad \lambda = 0.7396 \quad \gamma = 0.1979.$$

These parameters were applied to a space containing 96 lattice sites, so that M_2 contained 71 S_2 points and 19 R points. Tables of random digits and standard normal deviates were used to generate a synthetic M_2. Because of the small pattern size, random digits and normal deviates were tested for randomness.

The M_2 and Iowa patterns were described by (i) distances from origin points to the 10 nearest neighbors and (ii) distances from loci to the 10 nearest points. The boundary constraint was applied so that the number of recorded measurements tends to decrease as the order of neighbor increases.

Order mean distances are listed in Table 3 for point to point measurements and in Table 4 for locus to point measurements. The tabulated data on M_2 give mean distances for the 10 lower order neighbors and the number of recorded measurements for each order. Distances obtained from the Iowa map were standardized by multiplying each observed mean order distance by the square root of the density of urban places. The tabulated data on Iowa give the standardized mean distances and approximate miles for the 10 lower order neighbors. Also tabulated are the absolute and percentage differences between the observed and calculated mean order distances. Many other properties of M_2 and Iowa were collected but are not included in this report.

There are many reasons for not conducting an elaborate analysis for goodness-of-fit of the M_2 data to the Iowa data. Important reasons include the small size of the fabricated M_2 and difficulty in transforming frequency distributions into the normal form. These and similar problems could, largely, be handled in a more careful experimental design. More control was not exercised because I wanted a fast, crude evaluation of an imperfection model to determine whether it possessed any empirical reference; and, hence, merited detailed consideration. A fair test of the imperfection approach to urban systems requires a substantially more sophisticated model than M_2

Though recognizing the 'imperfections' in M_2, it seems sufficiently provocative to justify release of this highly preliminary report. While statistical methods were used to evaluate hypotheses of no difference between M_2 and Iowa (which were not rejected by the available data), reports on levels of significance and other statistical findings do not seem particularly critical at this stage of development.

Evaluation

The synthetic pattern M_2 reproduces with considerable fidelity the Iowa map pattern of urban places. The correspondence between M_2 and Iowa is a statistical rather than a cartographic similarity. This criterion of similarity determines the type of conclusions that can be drawn from the present study.

Both patterns were summarized by sets of distance measurements. These distances represent, however, quite different conceptualizations. The Iowa pattern refers to an observed distribution that exists in the real world, and at a point in time a study area has a single pattern of urban places. In contrast, the synthetic pattern represents a probabilistic model that is an abstract construction. This model does not describe one map pattern. Instead, the model defines a set of theoretical values. It is possible to interpret the model and synthetically construct a pattern that is representative of the model; yet, the model generates only one of an infinity of different patterns that correspond precisely to the statement of the model.

In more formal terms, the reduction of the distribution of urban places to order distances in a one-to-one mapping but the reduction of the model to a pattern is a one-to-many mapping. So, for the Iowa distribution only one pattern is formally possible (all representations must be conformal) while the mapping of the model is multi-valued. Consequently, while a single map describes the Iowa pattern, there is no cartographic summary of the pattern contained within the theoretical model.

While we reduce a map to a set of numbers we do not return a corresponding set of numbers to the map form. The cost of reducing the Iowa map pattern to a system of equations describing an imperfection model is the loss of the map description of that pattern. Whether this loss is compensated by the substantially greater analytical utility of a mathematical construction is a question that each student must resolve for himself.

In evaluating these questions the role of simulation should be correctly interpreted. Simulation was used only after all parameters of the model were estimated. This is not general in social science investigations of large, complex systems by means of simulation. Often, the model is simulated many times, each run using a different set of parameter values. The model being simulated is then adjudged successful if some set of parameters provides a good fit to the data at hand. This iterative approach is based upon an a priori acceptance of the model. In this application the simulation is used primarily to study properties of a complex model, but it does not provide any independent means of verifying the model itself. Simulation was not used for this purpose; for the imperfection concept simulation serves as the poor man's (mathematically poor, that is) numerical integration of a completely specified probabilistic model which can not be evaluated by analytic methods.

Table 1

Frequency Distributions of Observed and Calculated Standardized Distances, c_1, from Geographic Center of Interior Counties Containing an Urban Place to Nearest Urban Place

Distance	Freq.	Dist.	Error	$\frac{(f_0-f_c)^2}{f_0}$
c_1/σ	f_0	f_c	$f_0 - f_c$	
0- .243	11	11.72	- .72	0.471
- .486	11	11.04	- .04	0.000
- .729	11	9.82	1.18	0.127
- .972	8	8.23	- .23	0.005
-1.215	6	6.51	- .51	0.237
-1.458	3	4.85	*-1.85*	*0.052*
-1.701	5	3.39	*1.61*	
-1.944	2	2.28	**- .28**	**0.265**
-2.187	2	1.41	**.59**	
-2.430	2	.83	**1.17**	
Over 2.430	0	.92	**- .92**	
Total	61	61		1.157 ($\equiv \chi^2$)
				df=4

$.90 > Pr(\chi^2 = 1.157) > .75$

Iowa data, f_0 from Dacey (1963a). The standard deviation is $\sigma = 0.2286$. The calculated frequency, f_c, is from the unit half-normal distribution.

Table 2

Comparison of Observed Distribution of Urban Places per County in Iowa, 1950, with Expected Distribution of Points per Primitive cell of M_2

Number of Places x	Frequency Distributions Observed $g(x)$	Expected $E(x)$
0	21	21.1
1	64	64.2
2	13	12.4
3	1	1.2
≥ 4	0	.1

Observed values are from Dacey (1963a). Expected values are computed from (20) with $\lambda = .74$ and $\gamma = .2$.

Table 3

Comparison of j Order Distances for M_2 and Iowa Maps

Order	M_2		Iowa		Error	
j	n_j	$\bar{r}_j$	$d_0^{1/2}\bar{R}_j$	Mi.	$\bar{r}_j - d_0^{1/2}\bar{R}_j$	As % of Iowa
1	65	0.63	0.66	16	-.03	4.7
2	58	0.84	0.84	21	.00	
3	56	0.98	0.99	25	-.01	1.4
4	55	1.12	1.12	28	.00	
5	53	1.24	1.24	31	.00	
6	46	1.35	1.36	34	-.01	1.0
7	44	1.46	1.49	37	-.03	2.1
8	41	1.54	1.60	40	-.06	4.0
9	37	1.65	1.68	42	-.03	2.0
10	36	1.74	1.78	44	-.04	2.0

Iowa data are from Dacey (1963a).

Table 4

Comparison of h Order Distances for M_2 and Iowa Maps

Order	M_2		Iowa		Error	
h	n_j	$\bar{r}_h$	$d_0{}^{1/2}\bar{R}_h$	Mi.	$\bar{r}_h - d_0{}^{1/2}\bar{R}_h$	As % of Iowa
1	40	0.42	0.41	10	.01	4.7
2	36	0.72	0.72	18	.00	
3	32	0.97	0.93	23	.04	4.2
4	31	1.07	1.13	28	-.06	4.8
5	29	1.21	1.26	31	-.05	4.0
6	28	1.32	1.39	35	-.07	4.8
7	28	1.43	1.45	36	-.02	1.8
8	27	1.55	1.56	39	-.01	0.8
9	22	1.62	1.65	41	-.03	1.6
10	20	1.71	1.74	43	-.03	1.9

Iowa data are from Dacey (1963a).

October 14, 1963 Philadelphia, Pennsylvania

This original paper by Dacey, when printed in the *Papers* of the Michigan Inter - University Community of Mathematical Geographers, was supplemented with an 'Addendum' reflecting computer programs current at the time by Professor Duane F. Marble and Mr. Marvin Tener, and a second examination of the Iowa data by Dacey (December 13, 1963). A Glossary by Nystuen offered expanded explanations of complicated material for readers uncomfortable with notation. The added materials are not reprinted here.

* The support of the Regional Science Research Institute and of the National Science Foundation is gratefully acknowledged.

** Current address: Department of Geography Northwestern University Evanston, IL

References

Christaller, W. 1933. *Die zentralen Orte in Süddeutschland*. Jena: Fischer.

Dacey, M. F. 1963a. *Iowa: The Classic Plane or Croupier's Table.* Mimeographed.

Dacey, M. F. 1963b. *A Poisson-Type Distribution for Dispersed Population.* Mimeo.

Dacey, M. F. 1963c. "The Status of Pattern Analysis: Identification of Problems in the Statistical Analysis of Spatial Arrangement," paper presented at the Regional Science Association meetings, Chicago, 1963.

Dalenius, T., J. Hajeck, and S. Zubrzycky. 1961. On Plane Sampling and Related Geometrical Problems. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1. Berkeley: University of California, 125-150.

Feller, W. 1957. *An Intorduction to Probability Theory and its Applications, Vol. I.* New York: Wiley. Second edition.

Germond, H. H. 1950. *The Circular Coverage Function.* Santa Monica: RAND, Memorandum 330.

Lösch, A. 1939. *Die Raumliche Ordnung der Wirtschaft.* Jena: Fischer. (Translated by W. H. Woglom and W. F. Stolper as the *Economics of Location*, New Haven: Yale University Press, 1954.)

Seitz, F. 1952. "Imperfections in Nearly Perfect Crystals: A Synthesis," *Imperfections in Nearly Perfect Crystals* (W. Schockley, J. H. Hollomon, R. Maurer, F. Seitz, eds.) New York: Wiley.

Van Bueren, H. G. 1961. *Imperfections in Crystals.* New York: Interscience.

Zubrzycki, S. 1961. "Concerning Plane Sampling." *Second Hungarian Mathematical Cong.* Budapest: Akademiai Kiado.

Readers of *Solstice* might also be interested to note the following additional references to Dacey's work, not noted in the MICMOG publication.

Dacey, M. F. 1960. A note on the derivation of nearest neighbor distances, *Journal of Regional Science*, 2, 81-87.

Dacey, M. F. 1960. The spacing of river towns, *Annals* Association of American Geographers, 50, 59-61.

Dacey, M. F. 1962. Analysis of central place and point pattern by a nearest neighbor method, *Lund Studies in Geography* 24, 55-75.

Dacey, M. F. 1963. Order neighbor statistics for a class of random patterns in multidimensional space, *Annals* Association of American Geographers, 53, 505-515.

Dacey, M. F. 1963. Certain properties of edges on a polygon in a two dimensional aggregate of polygons having randomly distributed nuclei. Mimeo.

Dacey, M. F. 1964. Two-dimensional random point patterns: A review and an interpretation, *Papers*, Regional Science Association, 13, 41-55.

Dacey, M. F. 1964. Modified Poisson probability law for point pattern more regular than random, *Annals* Association of American Geographers, 54, 559-565.

Dacey, M. F. 1965. Order distance in an unhomogeneour random point pattern, *The Canadian Geographer*, 9, 144-153.

Dacey, M. F. 1966. A compound probability law for a pattern more dispersed than random and with areal inhomogeneity, *Economic Geography*, 42, 172-179.

Dacey, M. F. 1966. A county seat model for the areal pattern of an urban system, *Geographical Review*, 56, 527-542.

Dacey, M. F. 1966. A probability model for central place location, *Annals*, Association of American Geographers, 56, 550-568.

Dacey, M. F. 1967. Description of line patterns, *Northwestern Studies in Geography*, 13, 277-287.

Dacey, M. F. 1968. An empirical study of the areal distribution of houses in Puerto Rico, *Transactions*, Institute of British Geographers, 45, 15-30.

Dacey, M. F. 1969. Proportion of reflexive n-th order neighbors in spatial distributions, *Geographical Analysis*, 1, 385-388.

Dacey, M. F. 1969. A hypergeometric family of discrete probability distributions: Properties and applications to location models, *Geographical Analysis*, 1, 283-317.

Dacey, M. F. 1969. Some properties of a cluster point process, *Canadian Geographer*, 13, 128-140.

Dacey, M. F. Similarities in the areal distributions of houses in Japan and Puerto Rico, *Area*, 3, 35-37.

Dacey, M. F. 1973. A central focus cluster process for urban dispersion, *Journal of Regional Science*, 13, 77-90.

MICHIGAN INTER-UNIVERSITY

COMMUNITY OF MATHEMATICAL GEOGRAPHERS

DISCUSSION PAPER

Number 5

"A SIMULATION STUDY OF BARRIER EFFECTS IN SPATIAL DIFFUSION PROBLEMS"

by

Robert S. Yuill

University of Michigan

April 1965

FOREWORD

Every other Wednesday night the Michigan Community of Mathematical Geographers holds a joint seminar in Brighton, Michigan. Brighton is a small community closest to the point of minimum aggregate travel from the three universities involved.

Academic credit is given students who attend the seminar. Mr. Yuill is a student of the seminar and developed his interest in theoretical geographic barrier problems and in simulation techniques as an outgrowth of discussion of that topic presented to the seminar by Professor Nystuen in the Spring of 1964. The paper Mr. Yuill produced and which is herein presented was accepted as his Master's Thesis at the University of Michigan. It also received limited distribution as Technical Report No. 1, Spatial Diffusion Study, (O.N.R. Task No. 389-140, Contract No. 1228(33)).

We are grateful to Professors Duane F. Marble (Northwestern University) and Forrest R. Pitts (University of Pittsburgh) directors of that project for underwriting the cost of typing and preparation of the plates for reproduction.

The Editor, April 1965

CONTENTS

A SIMULATION STUDY OF BARRIER EFFECTS IN SPATIAL DIFFUSION PROBLEMS

Robert S. Yuill
Department of Geography
University of Michigan
Ann Arbor, Michigan

INTRODUCTION

Barriers play an integral role in restricting and shaping human activity, and in determining the distributional patterns of many phenomena. However in geography as well as in many other social sciences, the term "barrier" is frequently used and accepted at a very superficial level. An extremely general definition may be used such as that found in dictionaries: e.g., "Any obstruction; anything that hinders approach; any limit or boundary ..."[1] Although barriers in geography are usually considered with respect to questions of spatiality, this restriction still does not produce a more precise definition. There has been a strong tendency in geography to regard a barrier as an impediment to all activities regardless of their nature. Some of the devious paths into which this question may lead are apparent in the writings of the environmental determinists. For example, Semple in her American History and Its Geographic Conditions states: "At the end of the first century of permanent settlement they (the English settlers) found themselves in possession of a narrow strip of coast, shut off from the interior of the country by an almost unbroken mountain wall ..."[2] "The Appalachian system which,

1. ____, Webster's New Collegiate Dictionary, G. & C. Merriam Co., Springfield, Massachusetts, 1953.

2. Semple, Ellen Churchill, American History and Its Geographic Conditions, Houghton Mifflin Co., New York, 1903, p. 37.

together with the ice-worn highlands of New England, presented such an insuperable barrier to the early colonists, extends from the Green Mountains of Vermont to the pine-covered hills of Alabama."[3] This view however, is called into question by Brown and others: "Early maps of North America showed high mountain ranges occupying much of the interior. These fancied mountains gradually disappeared from the maps and were replaced by the various "chains" of the Appalachians. At first, the Appalachians were thought to be so formidable as to prohibit easy communication across them, but this belief too was abandoned in the light of exploratory accounts."[4] Brown's statement raises a serious question about regarding a barrier as something immutable. If the mountains were no longer regarded as complete barriers to the interaction of the people, did they then cease to be barriers at all? This dilemma again appears in Wolfe's recent work: "Precise definition of the terms 'barrier' and 'corridor' as they are used in the present context would be desirable, but it is not possible to provide them: barriers do not always serve as barriers nor corridors as corridors."[5]

One solution to the question of the nature of a barrier is indicated by Mackay: "The inhabitants separated by a boundary do not, except in very unusual circumstances, live in complete isolation from each other. On the contrary, a constant stream of human interaction flows back and

---

3. Ibid, p. 38

4. Brown, Ralph H., Historical Geography of the United States, Harcourt, Brace & World Inc., New York, 1948, p. 96

5. Wolfe, Roy I., Transportation and Politics, D. Van Nostrand & Co., New Jersey, 1963, p. 15.

forth across a boundary . . . . If we can estimate, with reasonable precision, the effect of physical and cultural boundaries (e.g. a river or political boundary) upon each type of interaction, we will possess a powerful tool for regional analysis and boundary studies."[6]

A logical consequence of this is that if the effect of a barrier is to be understood, then this effect must be viewed within the context of the activity to which the barrier is supposedly an impediment. The barrier then may be actually defined in terms of the activity; a functional relationship. This concept of dynamic process appears to be crucial to the study of barriers in a spatial context. If barriers are studied only in relation to static distributions, then the whole character of the barrier may be easily misinterpreted or distorted. It is within the context of this paragraph that the study reported upon here is undertaken.

In geography, distributions of observed items in the real world are generally not manipulated at the level at which they are originally obtained. Normally, these distributions are represented in an abstract symbolic fashion on maps. The mapped (abstracted) representation is then manipulated rather than the raw data, and the extraneous facets of the real world are ignored. What has been done is that the raw data has been abstracted to form a simple model in which only the essential characteristic of the observations, their spatial distribution, has been considered.

It is, of course, possible to represent a barrier in this abstract

---

6. Mackay, J. Ross, "The Interactance Hypothesis and Boundaries in Canada", The Canadian Geographer, No. 11, 1958, pp. 1-8.

sense. Further, it may then be possible to deal with the effects of barriers within an abstract two-dimensional space where the barriers also are represented by symbols. (See Figures I and II). It is easily seen that by abstracting only the functional aspect of the barrier from the observed data, two seeming dissimilar barriers may be represented symbolically in much the same fashion. This, of course, raises an important question for the researcher--are barriers, which are observed to be physically dissimilar, actually functionally similar? If so, then there may exist a class of common spatial effects of barriers despite severe differences in their original nature.

Origin of Barriers

Although barrier situations are encountered in many types of geographic research, there appear to be only two principal forms when their spatial representation is considered. The first, and most obvious, is what will be called an external barrier. The origin of these barriers is completely independent of the activity to which the barrier is functionally related. An example might be some natural physical feature, e.g., a swamp, and its impact upon interpersonal communications. (See Figure III).

Internal barriers on the other hand, are a direct result of the spatial arrangement of the activity under consideration. This is most common when the activity represents an interaction among a particular set of points which are in turn a subset of a much larger point set. An example of this might be interaction among members of some ethnic group which is unevenly mixed throughout a larger population. (See Figure IV).

It is important to recognize that both internal and external

FIGURE I

MOUNTAIN RANGE AS A PHYSICAL BARRIER

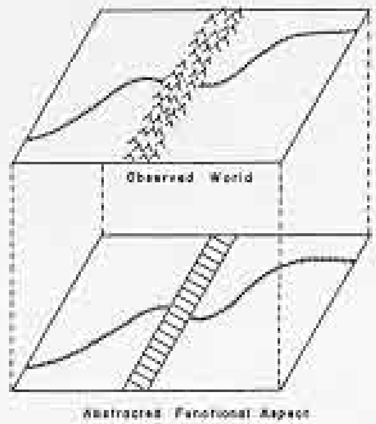


FIGURE II

POLITICAL BOUNDARY AS A CULTURAL BARRIER

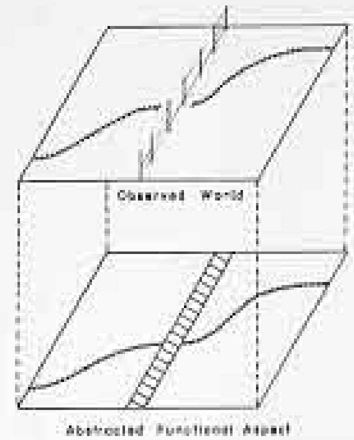


FIGURE III

EXTERNAL BARRIER

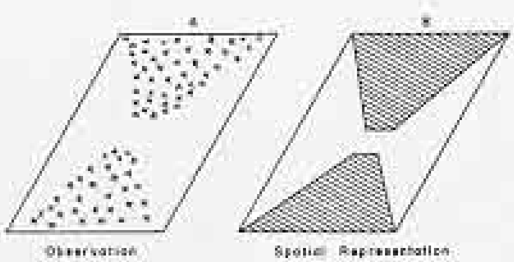
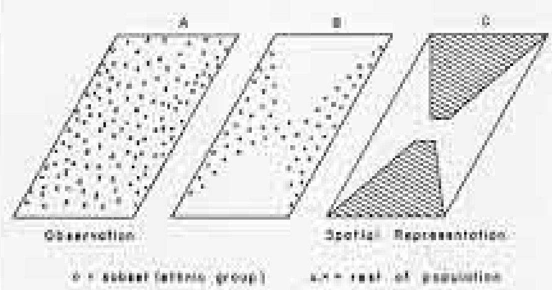


FIGURE IV

INTERNAL BARRIER



barriers may be spatially abstracted in an identical fashion (Figures III-b, IV-c). But, though these diverse barriers may be represented abstractly in the same symbolic fashion, the question of the similarity of their spatial effect upon their respective activities must be considered. While a certain similarity may seem evident at first glance, further consideration indicates that the case is somewhat more complex--that the effect of a barrier is really composed of two elements: functional effects and shape effects.

Functional Character of Barriers

One may readily imagine a multifold proliferation of the possible functions of a barrier. However if consideration is limited to the spatial aspects, it becomes possible to examine only a few essential characteristics.

Since a barrier is the creation of a dynamic spatial process, its functional effects must be also defined in terms of the process. These effects fall within three general categories: _absorbing_, _reflecting_, and _permeable_ barriers.

Absorbing Barriers

This is defined as a barrier that absorbs energy from the activity which comes into contact with it. The action is turned neither away nor back, but is completely stopped in its original direction of movement. As a consequence, the total amount of energy in the interaction system is diminished by the amount absorbed by the barrier.

Reflecting Barriers

This barrier, unlike the absorbing, does not affect the total sum of energy of the activity. Instead, that portion of the energy which comes into contact with the barrier undergoes an abrupt change of direction without change in momentum.

Permeable Barriers

Permeability affects the spatial process or activity only in that some of the activity which contacts the barrier is able to pass through it. However Nystuen has pointed out an important dichotomy in the action of permeable barriers. They may serve as filters, allowing the penetration of only certain portions of an activity, e.g., selective tariffs. On the other hand the barrier may act as a screen in which case those portions of the activity which possess sufficient energy are able to penetrate the barrier.[7]

Combinations

It is of course highly unlikely that a barrier can be adequately represented in terms of only a single function. Almost any example which one might specifically choose to illustrate one of the particular functions mentioned above, would be found to contain at least one of the other functional effects. Further, although these three types may be discussed independently, they are all compatible within the same barrier. Figure V-a represents individuals (or unit elements) of some process encountering first single function barriers, then a multifunction barrier. If the total energy of the spatial activity is then considered, Figure V-a may be further symbolized as in Figure V-b.

---

7. Personal communication from Dr. John D. Nystuen.

FIGURE V-a

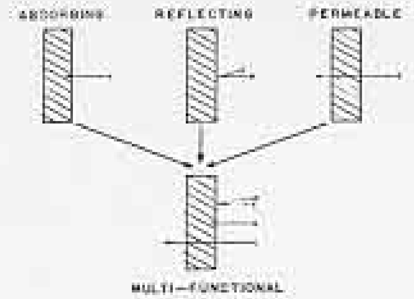


FIGURE V-b

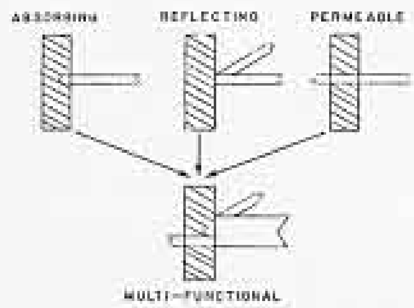


FIGURE VI

GRAPH OF BARRIER COMBINATIONS

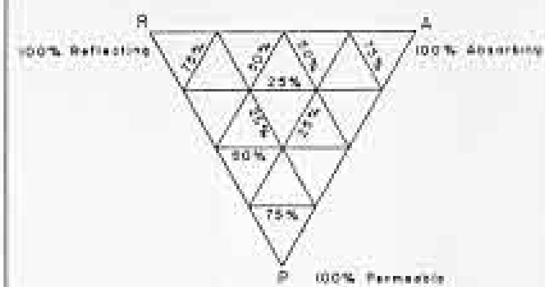


FIGURE VII

ABSORBING BARRIERS — TYPES

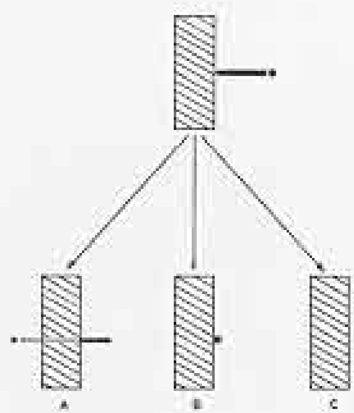
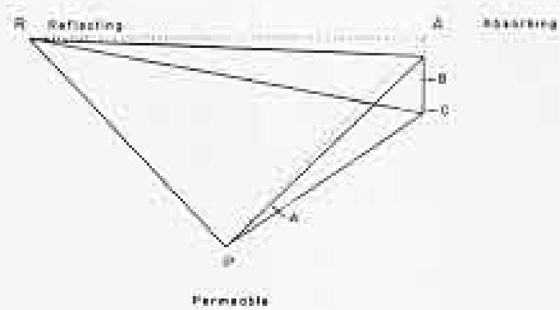


FIGURE VIII

COMBINATION UNIVERSE OF TOTAL ENERGY



If it is assumed that all portions of the activity which contact a barrier must be accounted for within the three functional categories, then the following relationship holds for the distribution of the available energy:

$$E_t = E_a + E_r + E_p$$

Where: E_t = the energy contained within the portion of the spatial activity which contacts the barrier.

E_a = the fraction of E_t which is absorbed by the barrier.

E_r = " " " " reflected by the barrier.

E_p = " " " " which penetrates the barrier.

This relationship is expressed graphically in Figure VI.

Variations within E_r and E_p have no effect upon E_t for by definition they only change the spatial distribution of the activity. With E_a however, the internal variations are significant for they directly control the defined decrease of energy. Figure VII illustrates three major variations of an absorbing barrier starting in each case with an individual or unit element encountering a barrier. In (a) the individual is able to continue through the barrier but his velocity is reduced. E_t is consequently diminished by that energy loss. In (b) the motion of the individual is completely halted resulting in a greater diminution of E_t than was true in (a). In (c) the individual itself is destroyed and E_t is reduced to zero. This range of possibilities may be expressed by adding another dimension to Figure VI to represent the change in E_t. This added dimension (see Figure VIII) is perpendicular to the plane ARP representing zero energy change in E_t. Since any combination which includes an absorbing barrier causes some decrease in E_t the resulting solid figure is entirely below the ARP plane except at the line RP. This solid figure represents the universe from thich the functional barriers of this study are drawn.

Shape of a Barrier

The second principal property of a barrier is connected solely with its spatial representation - its shape. This is in many cases the most important property of a barrier. However although shape is an integral part of almost every geographic distribution, little has been done to provide a systmeatic classification of shape. Many times a shape is compared to its nearest Euclidian counterpart or again, a classification may be more or less arbitrary as Bunge indicates: "Another common method of treating shapes is to devise a classification and subjectively assign shapes to it. Oxbow, circular, shoestring, and rectangular are typical classes. Often these classifications are so vague that it is difficult to have much confidence in the assignment of shapes of the various categories. Another difficulty is that the classes are so arbitrary."[8] However a barrier can, in many cases, be represented by one or a combination of simple Euclidian shapes without much loss of accuracy. It is beyond the scope of this investigation to attempt a solution to the problem of classification of shape. A few barriers with simple shapes will be considered, with the hope that the results will provide a basis upon which more complex analogies may subsequently be constructed.

As mentioned earlier a barrier may be abstractly represented in terms

8. Bunge, William, Theoretical Geography, Lund Studies in Geography, Series C, General and Mathematical Geography, No. 1, Gleerup, Lund, Sweden, 1962

of its functional and spatial (shape) characteristics. This study is then to be conducted entirely upon this basis, considering only the elements of function and shape which are common to all barriers. Hopefully, knowledge gained from investigations in this area will be applicable to all spatial barrier situations in geography, whether they be physical, political, or cultural.

It is impossible to deal with all the ramifications of barriers in a paper of this limited scope. Both the range of material investigated and the investigative techniques themselves must be severely limited. Hence the types of barriers to be considered here will be very simple ones.

A variety of methods exist for dealing with barrier situations. Certain approaches may utilize analogies drawn from the physical sciences such as those from hydraulics or electrical networks. Many of these, although very useful in specific cases, suffer from limited flexibility.

The technique chosen here is a Monte Carlo simulation of a simple diffusion situation. Monte Carlo methods comprise that branch of experimental mathematics which is concerned with experiments on random numbers....In the case of a probabilistic problem, the simplest Monte Carlo approach is to observe random numbers, chosen in such a way that they directly simulate the physical random processes of the original problem, and to infer the desired solution from the behavior of these random numbers.[9]

---

9. Hammersley, J.M., and D.C. Handscomb, Monte Carlo Methods, London Methuen and Co., 1964, p. 2.

Since the diffusion model used in this paper is derived principally from the work of the Swedish geographer Hägerstrand, a brief resume of that important work is in order. In the following condensation all quotations are taken from one of Hägerstrand's papers.[10]

Hägerstrand's Spatial Diffusion Models

A highly significant process which may explain many nebula distributions is the diffusion of techniques and ideas through the network of social contacts. This process of diffusion was studied in a rural area in southern Sweden with two innovations, a government subsidy to improve pasture and control of bovine tuberculosis.

Since the vehicle for the diffusion process was the network of social contacts, these were closely investigated. "A good many observations suggest that this network has a definite spatial structure which probably is rather stable, that is the links connect different places with probabilities which presumably change only slowly and thus to some extent are predictable."

Interpersonal communications were judged to be most important for information transmission at the local level. The interaction level was measured via such surrogates as telephone traffic and local migration, which were felt to be measures independent of the diffusion data. It was concluded that "the communication links of the average individual on the local plane must very rapidly decrease in number with increasing distance or in the sample region roughly with the square of the distance."

---

10. Hägerstrand, Torsten, "On Monte Carlo Simulation of Diffusion", in <u>Quantitative Geography</u> (William Garrison, ed.) forthcoming.

Using this information the diffusion of an innovation was simulated by Monte Carlo methods. For the first model an even population distribution and an ideal transportation surface were assumed. The rules governing the operation of the system were the following:

1. Only one person carries the item at the start.
2. The item is adopted at once when heard of.
3. Information is spread only by telling at pairwise meetings.
4. The telling takes place only at certain times with constant intervals (generation intervals) when <u>every</u> carrier tells one other person, carrier or non-carrier.
5. The probability of being paired with a carrier depends upon the geographic distance between teller and receiver in a way determined by empirical estimate.

Based upon the network of social contacts a mean information field (contact field) was constructed for the area, and it was assumed that this field represented the probability of contact, as a function of spatial separation, for any two persons. For simplicity the MIF was modified to the form of a rectangular grid of 25 cells (5 cells on a side) with the innovation carrier located in the center cell. (This grid coincides exactly in orientation and dimension with the grid division of the study area). The probability of each cell being contacted by the innovation carrier (center cell) was approximately $1/\text{distance}^2$.

Following the rules of the model the diffusion was simulated over a number of generations then compared to the actual diffusion in the study area. Although a comparison was made only in very general terms, there were striking visual similarities in the spatial patterns.

In later variations of this basic model, Hägerstrand changed some of the parameters and operating procedures in order to obtain a more realistic representation of the diffusion process. Two important problems with which he had to deal were those of study area boundaries and barriers. The boundary problem revolved around the treatment of cells near the edge of the study area so that they would have the same opportunity of being contacted by an innovation carrier as cells in the interior. Hägerstrand's solution was to add a border two cells in width to the study area (Figure IX). Individuals in the border could be contacted and could contact their counterparts in the study area. This permitted the innovation to move outside the main study area and then come back in again. In this manner the probability of contact was the same for each cell within the study area.

The barriers were topographic restrictions (long lakes and forests) which distorted the previously assumed even transportation surface. Hägerstrand recognized two categories of barrier here: weak in which communication was reduced; and absolute in which communication was completely cut off. Figure X illustrates the representation of this situation in the Hägerstrand simulation. The barriers were placed between the cells, a one-dimensional representation. There is no possible interaction over an absolute boundary. Across a weak boundary the communication was cancelled about half the time by a random process.

Although Hägerstrand was not primarily concerned with barriers in his study, he recognized their importance and devised a method for investigating their effects. It is his work which has given impetus to the present study of barrier effects.

STRUCTURE OF THE SIMULATION MODEL

The diffusion model used in this study is based upon and is somewhat similar to the one proposed by Hägerstrand. However in order to meet the requirements of this particular study certain structural changes were necessary. These arose from the limited objectives of the present study, as well as from technical limitations of the specific computer program which was employed.

Since the primary emphasis in the present instance is upon the barrier rather than on the diffusion itself, the model may be simplified by omitting certain details relevant only to specific diffusion situations. Several important changes are made in terminology. Many terms were borrowed from Hägerstrand for they are equally applicable to an abstract diffusion model. However it must be noted that there is no generic carry-over of terms. They apply as used here, only to the spatial functions.

Cell: This is the areal unit into which the study area is subdivided. The cell either contains a number of items or members, or it is empty, or it contains some suppresion device. In the latter case, the cell is designated as a barrier and reacts in a specified manner upon the spatial course of the diffusion. Therefore a barrier is considered as an area-based function rather than as a line-based function as in Hägerstrand's work.

Transmitter: A member of the population (in a cell) which in interacting according to the rules of Monte Carlo simulation, attempts to contact members in the same or another cell.

Acceptor: A member of the population of a cell which has been contacted by a transmitter and which becomes a transmitter itself in all succeeding generations.

One further change involves the ommission of Hägerstrand's operational rule number one since it will not be relevant in the present context.

Since each cell represents a unit of interaction or a barrier, the floating grid (MIF) which determines the areal probability of contact was modified. The five-square grid or probability field of Hägerstrand was changed to a smaller three-square grid of nine cells, and the probabilities were adjusted so that their sum in the nine cells totaled one. This contraction of the grid made possible the use of barriers and borders which were only one cell wide. The attendant ease in programming is obvious.

Further, the potential adopters within a cell were not considered individually but statistically. That is, a certain proportion of the population is a cell is considered to be acceptors, rather than singling out specific individuals.

Operational Procedure

The resultant computer simulation model works in the following manner:

For a transmitter in any cell, there is a given probability of contact with the surrounding cells in addition to contact within its own cell. Figure XI shows the symmetric mean information field which surrounds the transmitter. This field (or grid as it is represented here) is centered upon each cell which has transmitters in any given generation. The Monte Carlo process is applied as many times to a cell as there are transmitters, then the system moves on to the next cell containing transmitters.

When a cell has been contacted by a transmitter, a further probability function is called into play. The probability of contacting a member of that cell's population not previously contacted is equal to $(1 - N_a/N_p)$, where N_a is the number of acceptors in that cell and N_p is the total population of the cell.

Once a valid contact is made, a member becomes an acceptor and transmitter beginning with the next generation. If (statistically) an already contacted member of the population of a cell is again contacted, that transmission is regarded as lost.

Functional Barriers

Since it is impossible to incorporate infinite variability into the program, some limitations of the barriers tested had to be imposed. Operationally, this consisted of limiting the functions of the barriers to five, which are described as they function within the computer program. Given the transmission situation in Figure XII, the effect of each functional barrier is easily demonstrated.

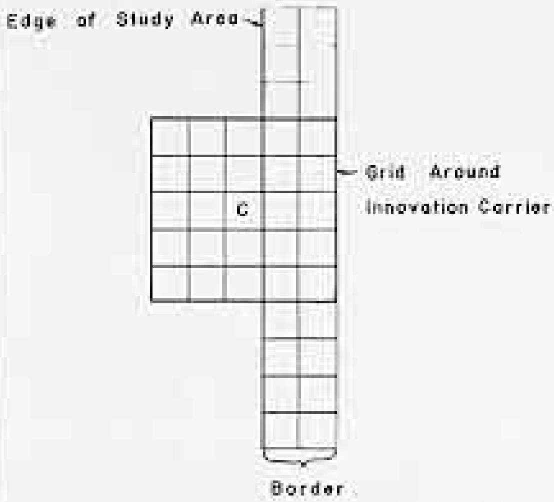
<u>Super Absorbing Barriers</u> (Figure XIII)

With this class of barrier (represented approximately by (c) in the universe of Figure VIII), the transmitters are temporarily destroyed but may be replaced by subsequent contact from other transmitters.

<u>Absorbing Barriers</u> (Figure XIV)

When an absorbing barrier is contacted the transmission is lost for that generation, but the transmitter itself is unchanged by the operation. Its position in the functional universe (Figure VIII) is

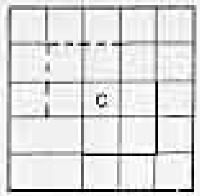
FIGURE IX



C = Innovation Carrier

FIGURE X

HAGERSTRAND'S BARRIER REPRESENTATION



Solid Line — Absolute Barrier

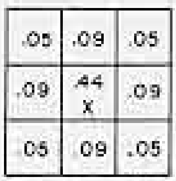
Broken Line — Weak Barrier

C= Innovation Carrier

FIGURE XI

PROBABILITY OF CONTACT

BY TRANSMITTER



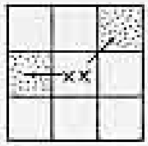
X = Transmitter

FIGURE XII

Initial Conditions:

Two Transmitters in the Center Cell

Arrows show the Cells the Transmitters Attempt to Contact



x = transmitter

o = acceptor

▒ — Barrier Cells

FIGURE XIII

SUPER ABSORBING BARRIERS

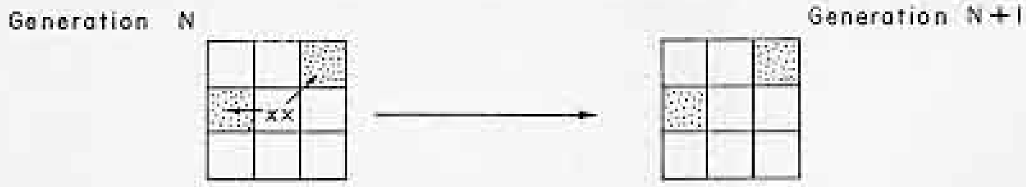


FIGURE XIV

ABSORBING BARRIERS

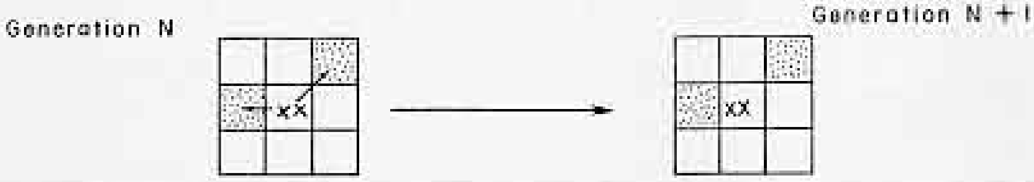


FIGURE XV

REFLECTING BARRIERS

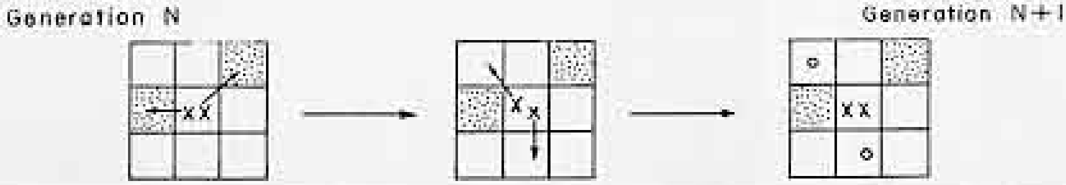


FIGURE XVI

DIRECT REFLECTING BARRIERS

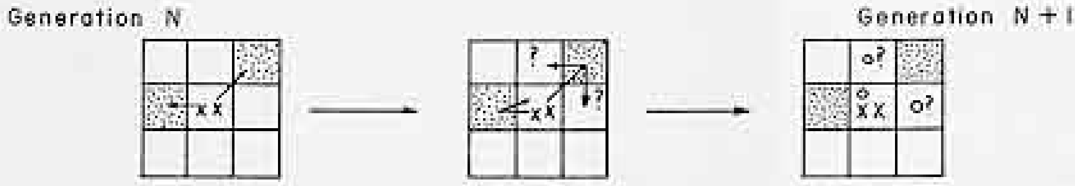


FIGURE XVII

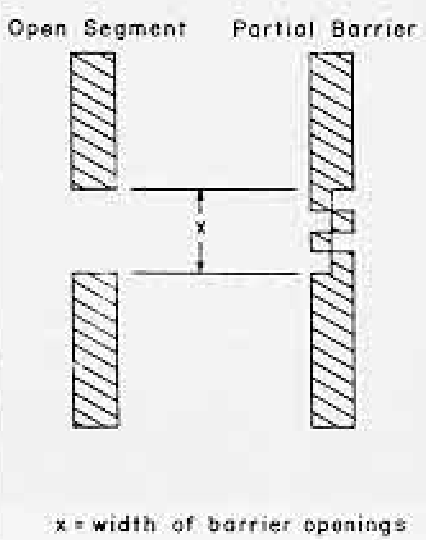
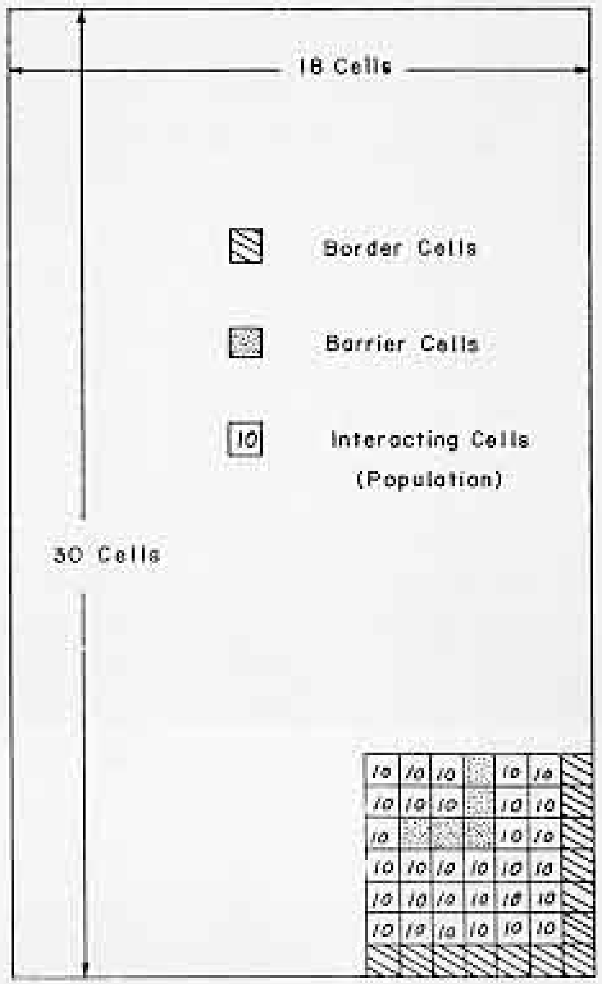


FIGURE XVIII

SIMULATION AREA



about that of (b).

Reflecting Barriers (Figure XV)

In this barrier (R in Figure VIII), the transmitter is given an opportunity to transmit again within the same generation.

Direct Reflecting Barriers (Figure XVI)

In this barrier (R in Figure VIII), the contact is deflected to the nearest cell toward the transmitter. (Distance is measured from cell center to cell center). In the case of a corner cell, which has two equally near neighbors, the choice between them is a random one. Since the mean information field is symmetrical, the deflection by all the corner cells will be alike and those of all side cells alike.

Permeable Barriers

Permeability, in comparison with other barrier functions, is a negative quality since the more permeable a barrier, the more it loses its function as a barrier. In the theoretical universe depicted by Figure VIII the permeability of a barrier can vary from zero to 100 percent. However in the computer program developed for this study, a permeable barrier is somewhat difficult to represent since each cell is assumed to be either a barrier of a specified nature, or a cell with interacting members. This obviously precludes intermediate positions in the permeable spectrum. To partially obviate this difficulty an attempt was made to approximate an intermediate permeable barrier through a structural spatial distribution of barrier cells. Figure XVII shows such a distribution, a checkerboard arrangement, together with a completely open segment of a barrier for comparison.

Size of the Study Area

Since the cell is the basic spatial unit in this computer simulation, the dimensions of the simulation area are likewise expressed in these units. For this investigation the process was confined to an area of 30 by 18 cells or a total of 540 cells. Of this however, the 92 cells which form the outside edge are excluded since their function is solely to confine the simulation to the specified area and to provide "border bounce" capability. This size of area was considered the best compromise between an area of sufficient extent and efficient use of computer time. Figure XVIII shows the simulation area. The computer program utilized in this study is detailed in Appendix A.

RESULTS OF THE SIMULATION STUDY

The simulation model was applied to a variety of barriers in order to test the effects of each characteristic in insolation - or more precisely, to hold the others constant. The characteristics examined here were the three functional effect series (absorbing, reflecting and permeable) plus some simple elements of shape.

In attempting to analyze the results of the simulations it was found that identical criteria could not be used in all situations. The disparity of shape among the barriers investigated made it impossible to find one type of measurement which would satisfy all requirements. Therefore the problem of comparing different measures arose. This is not as grave as would appear since each barrier situation was intended to supplement the information obtained from the others. Further, all the measures were based directly upon the spatial response of the diffusion simulation so they are genetically closely related.

The immediate results and deductions from each group of tests are first given separately (following), then in combination.

Separate Group Results

Group I: Parallel Barriers Perpendicular to a Linear Diffusion Wave Front

The purpose of this group was to test the effect on the diffusion wave of 1) varying the width of the channel between two parallel barriers, and 2) differences arising out of absorbing and reflecting functional effects; permeability was held constant at zero. (See Figure XIX).

The measure used here of the influence or effect of these barriers, was the rate of advance of the diffusion wave front between the barriers. The distance between the two barriers (x) took on the successive values of 1, 2, 3, 4, 6, & 9 cells, and the functional nature of the barriers varied from absorbing to reflecting.

Approximately 20 generations of the simulation were required for the diffusion front to pass the length of the barriers. For each generation the rate of advance was calculated and the means of the various combinations were found. Figure XX shows these means plotted against the width between the barriers for each of the four functional forms tested.

From a general knowledge of spatial and functional restrictions certain results can be reasonably predicted. The most obvious of these are:

a) As the distance (x) between the barriers increases, there is a decreasing effect of the barriers upon the rate of advance of the diffusion wave front.

b) In narrower channels, the absorbing and super absorbing barriers retard the diffusion wave the most.

Although these predictions are generally substantiated, there appear unexpected variations:

a) The reflecting and direct reflecting barriers apparently have little effect upon the rate of advance of the diffusion wave. Figure XX shows little variation with distance between barriers, and in fact, the variations may be entirely due to the nature of the Monte Carlo process.

b) In contrast, the absorbing and super absorbing barriers do show a response to the width of channel, but the response is non-linear. In Figure XX, there appears to be a definite break in the trend at a channel width of four cells, with the rate of advance remaining approximately constant for greater widths. This aspect will be discussed further in a subsequent section.

Group II: Passage Around A Bar Barrier Parallel to a Linear Diffusion Wave.

The spatial configuration encountered here is probably one of the most common representations of a barrier. The purpose of tests on this group is to examine the spatial effect of the barrier in terms of its length and function; as in Group I the permeability is held constant at zero. Since the effect of the barrier is particularly spatial, rate of advance as a measurement is not applicable because it would tend to obscure some of the more important aspects of the situation. For this reason the measurement criteria used here are the differences in time (Δt) between the generation in which the linear diffusion wave

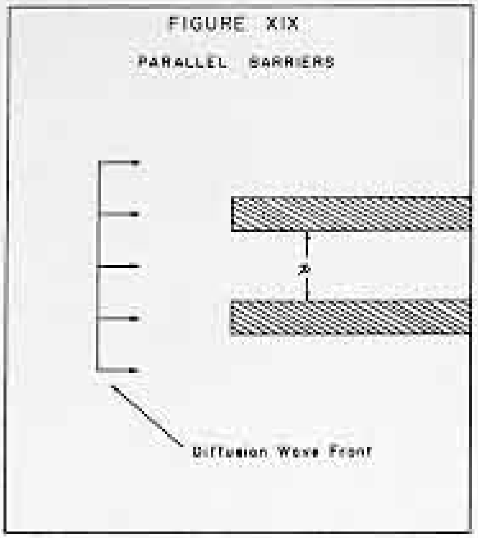


FIGURE XIX
PARALLEL BARRIERS

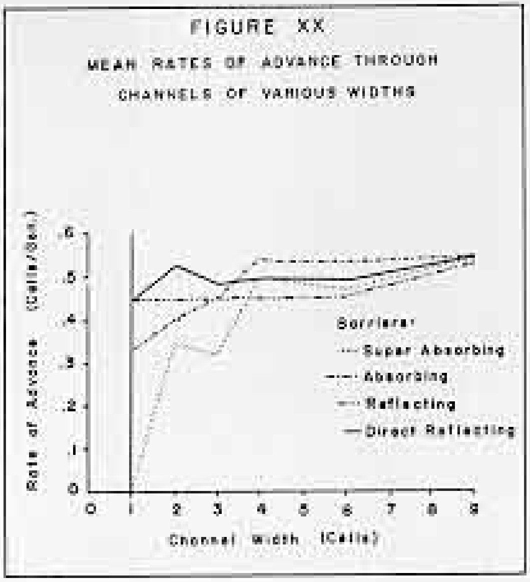


FIGURE XX
MEAN RATES OF ADVANCE THROUGH CHANNELS OF VARIOUS WIDTHS

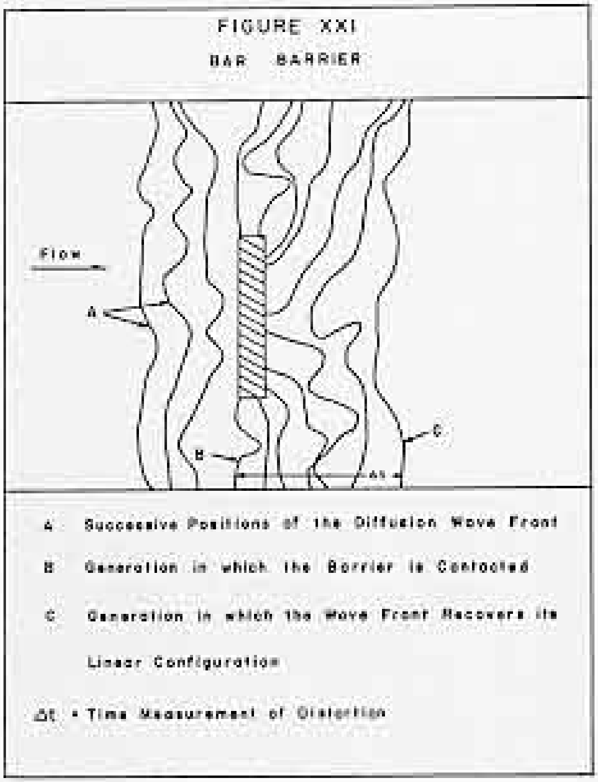


FIGURE XXI
BAR BARRIER

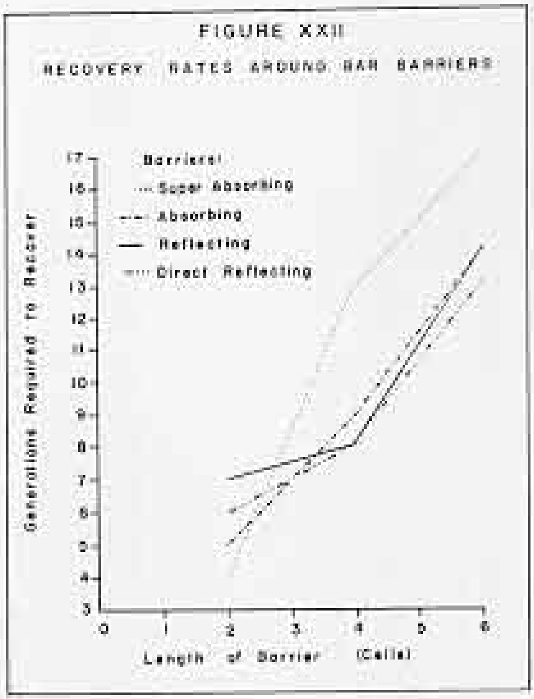


FIGURE XXII
RECOVERY RATES AROUND BAR BARRIERS

first encounters the barrier (Figure XXI - b), and the generation in which the wave recovers its linear form after passing the barrier (Figure XXI - c). Figure XXI shows an actual pattern of diffusion around a barrier.

The spatial distortions caused by all members of this group are alike - varying only in a matter of degree-- and as a consequence, the various functions and lengths of barriers may be readily compared. Since the recovery time (Δt) of the plane diffusion wave is also dependent upon the amount of its distortion, Δt may be used as a measure of this distortion. (The units of Δt are computer generations--equal are arbitrary.)

The effects of four absorbing and reflecting barriers were tested on barriers of 2, 4, and 6 cells in length. Again as in Group I, there are certain expected effects. Figure XXII indicates that in general the longer the barrier the greater the recovery time (Δt) required by the diffusion wave (due to greater distortion). Also absorbing barriers introduce more distortion than the reflecting barriers.

For each length of barrier the functional types were ranked in order of decreasing recovery time (Δt) as shown in Table I. These ranks were summed for each type of barrier giving the results shown in Table II.

This result conforms to expectations as the absorbing barriers have a higher rank (lower totals) and a longer recovery time. However an examination of the data generated for barriers two cells in length shows the rank order completely reversed, producing a singular exception to the general pattern. This fact, although by itself of no great significance, will be discussed further in connection with other unexpected deviations.

TABLE I

Barrier Type	Length: 2 cells R.T.	R.O.	4 cells R.T.	R.O.	6 cells R.T.	R.O.
Super Absorbing	4	4	13	1	17	1
Absorbing	5	3	9	2	14	2.5
Reflecting	7	1	8	3.5	14	2.5
Direct Reflecting	6	2	8	3.5	13	4

R.T. -- Recovery Time in Generations R.O. -- Rank Order

TABLE II

Barrier Type	Sum of Ranks*
Super Absorbing	6.0
Absorbing	7.5
Reflecting	7.0
Direct Reflecting	10.5

* From Table I

Group III: Passage of a Linear Diffusion Wave Through an Opening in a Barrier Parallel to the Wave Front

This group represents a barrier situation of common occurrence, particularly in the physical sciences, and its common occurrence has led to the accumulation of a great deal of empirical information on this particular spatial restriction. Hence the general results here can be predicted with moderate precision. This is fortunate, since no good measure of the diffusion wave shape was found.

For this study the barriers were two cells in width, impermeable and the size of opening (x) varied from two to six cells. Figure XXIII illustrates the effect of a five cell wide barrier opening upon a linear diffusion wave.

From Figure XXIII it is seen that the resulting curvilinear form in the stages after passage is about as expected. It was further found that as the opening in the barrier became narrower, the shape of the wave front approached that created by a point source. Definite observed differences in the spatial pattern of the diffusion did appear when absorbing and reflecting barriers were contrasted, as shown in Figure XXIV.

Since the rate of advance in the original direction of the plane wave (as shown by the arrow) remained approximately the same, the differences in shape are due to the portions of the wave near the barrier (a and a'). There are no other factors to consider; the functional effects of the barriers must be responsible for the difference. To test this hypothesis, the distances traveled by the diffusion wave for reflecting and super absorbing barriers were compared in the cells immediately adjacent to the barriers. Figure XXV is a plot of distance against time (expressed in generations) for the two cases.

Root mean square calculations of regression lines (rate of advance of the diffusion wave along the barrier in the region of a and a') gave slopes of 0.56 for the reflecting barrier and 0.39 for the super absorb-

FIGURE XXIII

PASSAGE THROUGH BARRIER OPENING

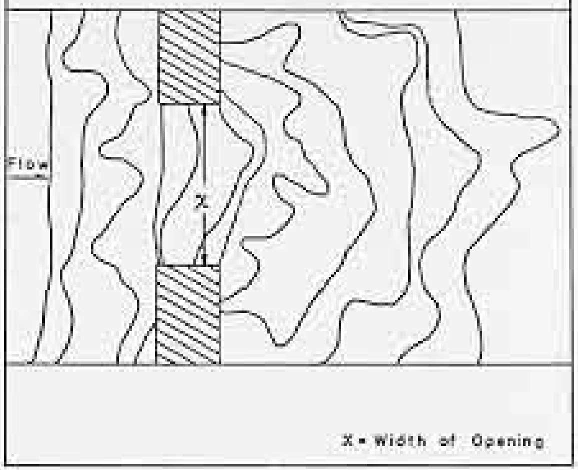


FIGURE XXIV

DIFFERENCES IN SPATIAL PATTERNS

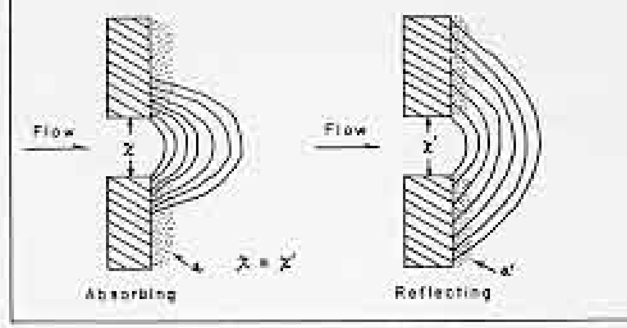


FIGURE XXV

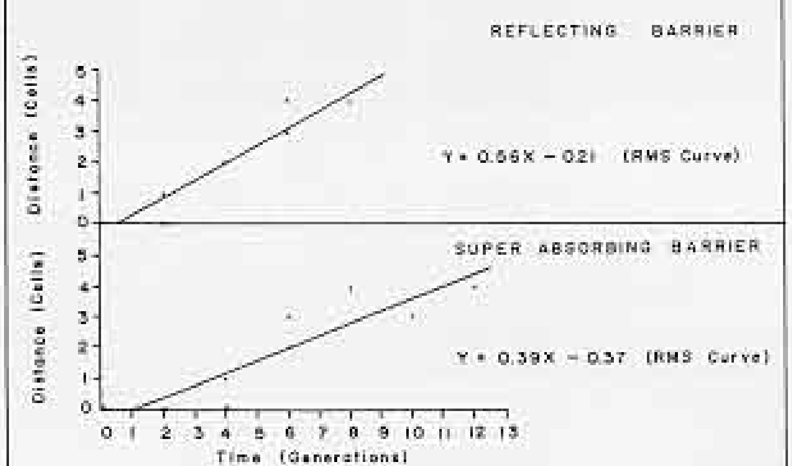


FIGURE XXVI

TEST OF PERMEABLE BARRIERS

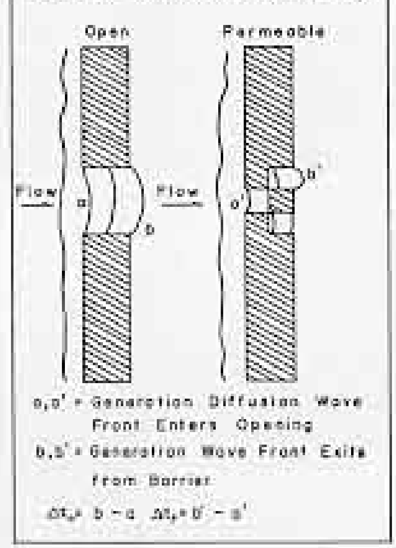
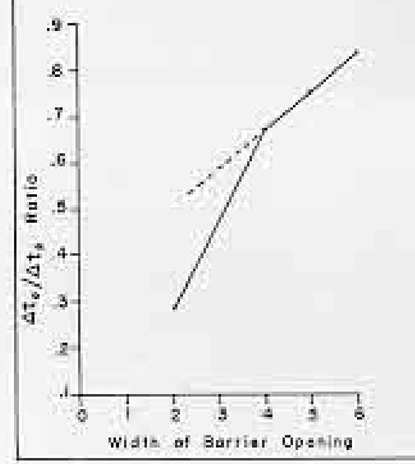


FIGURE XXVII

FLOW THROUGH PERMEABLE BARRIER



ing barrier. When these rates of advance are compared with Figure XX it is seen that the rate of advance for the reflecting barrier is approximately that of an unobstructed diffusion wave, while that for the super absorbing barrier is significantly smaller. This substantiates by measurement a conclusion which is indicated by the functions of the barriers - that there is friction between the absorbing barrier and the passing wave of diffusion which acts as a drag upon the wave and distorts its shape.

Group IV: Permeable Barriers

This group was measured in conjunction with the barrier configurations of Group III. Openings of 2, 4, 5, and 6 cells were tested using a series of super absorbing barriers. The permeable barriers were approximated by the checkerboard configuration of Figure XVII.

There was some difficulty in finding a suitable measure for this group. Although not entirely satisfactory, the best measure was to compare the time required for the diffusion wave front to pass from one side of the barrier to the other in permeable and open barriers. The time difference was expressed in generations (Figure XXVI).

The time difference for each size of barrier opening was expressed as a ratio $\Delta t_o/\Delta t_p$ (See Figure XXVI), which was then plotted against the width of barrier opening in Figure XXVII.

If the points are connected with straight line segments there appears to be a definite regularity among the upper three. This would seem to indicate that the longer the permeable barrier the greater the rate of advance of the diffusion wave through it. In actuality, Figure XXVII shows only that the longer the barrier the greater is the probability that the diffusion wave will pass through it at some point with the same rate of advance as if there were no barrier there. But this conclusion is even more unsatisfactory in terms of the objective, for little has been shown about the real effect of a permeable barrier. Furthermore the shape of the diffusion wave was too often determined by only the point at which the wave first passed through the barrier. Therefore this particular spatial approximation of the barrier function of permeability does not seem to be a particularly fruitful line of inquiry.

Figure XXVII does indicate something of importance, however. The ratio $\Delta t_o / \Delta t_p$ of the barrier opening of two cells does not follow the general trend of the other points. If there is a relationship between the ratio and the width of barrier opening, then it may be a non-linear function. There even may be more than one relationship involved.

Results and Implications

Results

The most obvious feature of the combined results is the dominant effect of shape in the diffusion pattern. This response to shape is not surprising, and it is one that is readily observable in the physical sciences - particularly in the flow of fluids. However there are significant variations in this general pattern which are generated by the functional effect of the different barriers. These are as follows:

Reaction to Function. On the basis of functional effects, the absorbing and reflecting barriers are naturally expected to display certain differences. Figure XXV illustrates such a difference. However, referring to Figure XX it is seen that the difference is due to only the absorbing barriers performing as expected. The absence of any noticeable reflecting barrier effect upon the wave front advance is at first surprising, but is actually due to the operation of the computer program. Since only probabilities within the floating grid are changed and not the boundaries, there is little spatial effect by these reflecting barrier functions.

Persistent Irregularity in Data. In considering three of the investigated groups there is found a curious pattern of irregularity. (See Figures XX, XXII, and XXVII). Between the values of two and four cells, there appears either a definite change in slope or in the case of a multi-curve graph (Figure XXII), a reversed order. This is especially significant in that this irregularity is found among widely different barrier configurations measured by different methods. It was concluded that the anomaly was most likely related to the floating grid since the dimensions of the grid (three cells square) fall exactly in the center of the critical area of change (two - four cells).

If indeed, the anomaly can be related only to the grid, then there must be a causal relationship. Such a relationship might be structured as follows:

Let the floating grid be represented by a circle with a diameter of three arbitrary units. When the circle encounters parallel barriers of Group I as shown in Figure XXVIII the spatial relationship is more apparent. Part 'a' shows that though the circle may touch a barrier it is not necessarily affected by it since its movement is really not restricted. When two sides of the circle come into contact with the barriers ('b') restriction begins, for movement is then curtailed in the direction perpendicular to the flow vector. Part 'c' represents an even more drastic curtailment of movement. The absorbing and super absorbing barriers of Figure XX show exactly this effect. (The reflecting barriers do not show the effect for the reasons mentioned earlier).

In a related manner, the expected functional order of Group II (See Figure XXII) breaks down completely at barrier lengths smaller than the floating grid diameter. Figure XXIX illustrates how the ratio of grid diameter to barrier length is a critical factor in determining the effect of that barrier.

Figure XXVII of Group IV further indicates a much greater spatial interference for barrier openings of less width than the grid diameter.

By implication and logical extension of the above results, the following conclusions were reached:

1. A circle (representing the probability field of an individual) is really only affected by the barriers with which it comes into contact. Further, the degree of influence or effect is proportional to the amount of contact between the circle and the barrier.

2. When contact is made with a barrier there is a spatial effect on the circle depending upon the function of the barrier. Figure XXX shows these different spatial effects of the absorbing and reflecting barriers.

If the area of contact is 'x' in part 'a' then for an absorbing barrier ('b') the area of the circle is reduced by that amount. The reduction in area however, is generalized as a reduction in the radius of the circle.

For the reflecting barrier ('c') the effect is not a reduction

FIGURE XXVIII

UNIT CIRCLE ENCOUNTERING PARALLEL BARRIERS

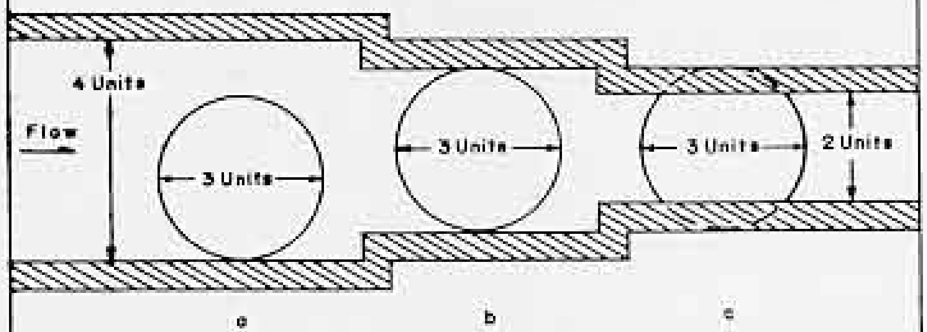


FIGURE XXIX

UNIT CIRCLE ENCOUNTERING BAR BARRIER

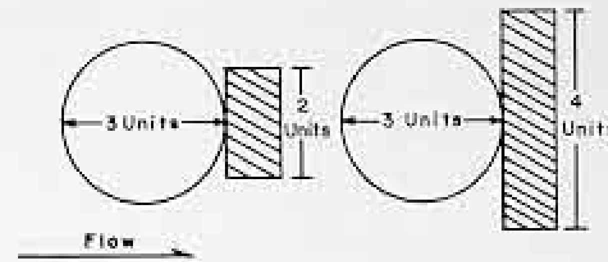
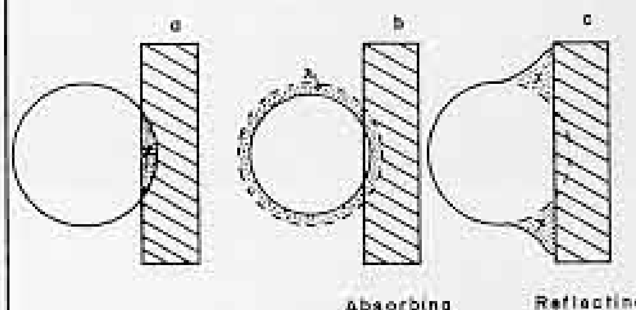


FIGURE XXX

SPATIAL EFFECT OF BARRIERS ON UNIT CIRCLE



but a distortion of area with the center of the circle remaining stationary.

3. As far as permeability is concerned, the results here offer few definitive guidelines. However since retardation of the diffusion wave front was noted, there are several logical spatial representations:

a/ The floating grid diameter is reduced while passing through the barrier.

b/ The grid size is untouched but the barrier is enlarged to the proper effective distance via a map transformation.

APPLICATION OF RESULTS

There is almost an infinite variety of barrier situations to which these results may be applied. A few will be given here for the purpose of illustration. First however, the abstract processes must be put into a more concrete form.

The floating grid, it must be remembered, is based upon the mean information field of an individual. It is further related to real situations by means of the concept of uncertainty of location. Rarely in the dynamics of human events is the individual stationary. It is not reasonable, then, to consider him as a point location in a dynamic simulation. A more accurate representation is a closed curve which delineates the area of the individual's most probable occurrence. Within this areal domain however, there are important distinctions in the probability distribution. For a traveler or nomad there is almost an equal probability of his being at any particular point within the area. Bur for an individual with a fixed place of residence there is a much greater probability of his being near that residence. For convenience the circle is adopted as the closed curve representing the individual.

The first illustration is the use of the results to simulate travel rates by land (in the United States in the early 19th. century) by using only the gross physical features. Figure XXXI shows various time distances from New York City using very few and very simple assumptions on rates of travel and permeability. For the model these assumptions consist of dividing the area under consideration into three regions, each of which is considered homogeneous with respect to rate of travel. The regions are: 1) An area of unrestricted movement, corresponding roughly to the Atlantic coastal plain, the piedmont, and associated valleys. 2) An absorbing but permeable barrier which reduces the energy (rate of travel) to one-third that of the unrestricted region. This corresponds to the Appalachian System. 3) An impermeable, reflecting barrier corresponding to the Adirondack Mountains.

Figure XXXII is a map of actual recorded rates of travel in 1800. The shape of the theoretical model agrees well with the actual although there are some distinct differences in the time rates to reach the same point. The importance of this illustration however, is that this very rudimentary simulation of a functional barrier can produce reasonable spatial results.

Barriers are also very important in the spatial configuration of cities. Probably the most efficient area for a city is a circle, but functional barriers often transform this shape during the process of growth. A common case is a river of sufficient width so as to hinder easy passage across it. In the settlement and growth of a city, a river of this magnitude may be regarded as a reflecting barrier (See Figure XXX-c).

and for interchange across it, an absorbing but permeable barrier. Neglecting changes in modes of transportation the generalized resulting pattern of a city is shown by Figure XXXIII.

For comparison some diagrams of actual cities on rivers are shown in Figure XXXIV. The prevalence of cities similar to this all over the world (Bay City, Michigan; Düsseldorf, Germany; Sault Ste. Marie, Canada- to mention just a few) affords an excellent starting place for a solid study of city shapes and the response of the growth mechanisms to barriers.

CONCLUSIONS

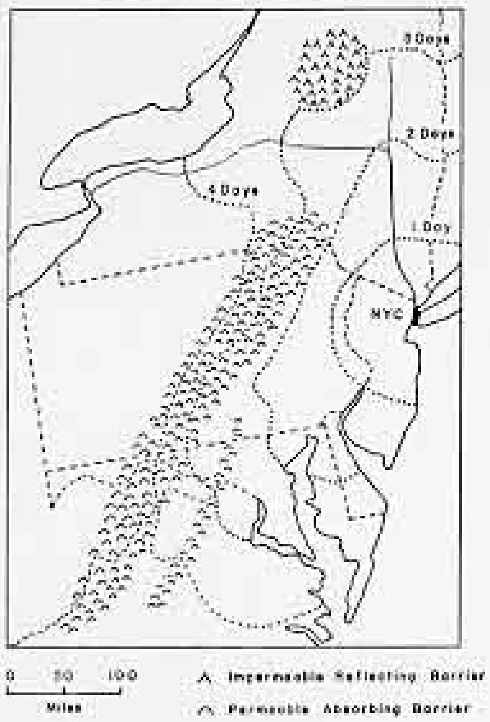
The avowed aim of this study has not been completely fulfilled. There are perhaps more unanswered questions about the spatial nature of barriers than before. However several significant points have emerged. The first is the outgrowth of considering a barrier as a function of a process or activity, and revolves about the idea that the range of a barrier may be limited by the range of an individual interactor or unit of the activity involved. That is, the barrier is only effective when it is in spatial conflict with the closed curve describing the dynamic location of an individual unit. Although this may scarcely be regarded as established or proven, the results in this study give strong evidence in support of it.

The second point is the significance of shape in the measurement of barrier effects. No common measure could be found which was satisfactory for all the barrier shapes tested. Instead each shape necessitated a different form of measurement ranging from rate of travel - a linear measure, to distortion of a plane wave - an areal

FIGURE XXXI

THEORETICAL TRAVEL TIME FROM NEW YORK CITY ABOUT 1800

Mode of Travel — Stagecoach

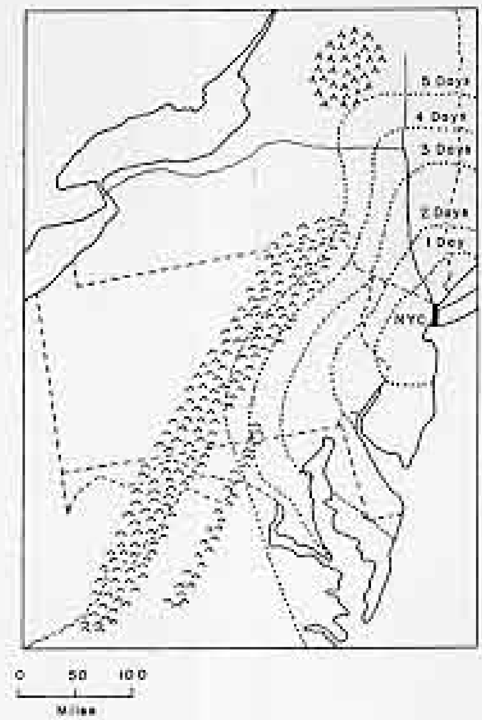


Barrier Absorbs 2/3 the Velocity of the Vehicles

FIGURE XXXII

RATES OF TRAVEL — 1800

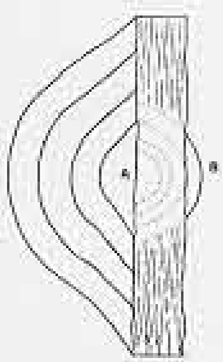
(Rate from New York City by Stage)



after Brown, Op Cit., p101

FIGURE XXXIII

GENERALIZED CITY GROWTH PATTERN ON RIVER



A — Start of City

B — Related Settlement

FIGURE XXXIV

DIAGRAMS OF RIVER CITIES

or two dimensional function. This indicates that for direct comparison of barriers, shape is a most important property. A further implication is of a hierarchy of barrier properties with shape being primary and functional effects being secondary.

The third point is that though different measures were required by different barrier configurations, the quality actually measured in each case was the spatial distortion of the advance of the diffusion wave front. In each case time was the principal element (rate of advance, recovery time, etc.) required to measure the diffusion. The implication is that for geographic problems of movement in space, time is an essential element. This is especially true for diffusion models for which time is both a regulating and a limiting factor. Time is therefore inextricably connected with space in all diffusion processes. The question now arises about the place of time in a steady state flow around barriers such as quotas in international trade. Since time is not an essential part of this flow, are the results gained from this diffusion model valid for such flow? It is suggested that this would be an interesting line of inquiry for further investigation.

A few primitive rules or ideas have been advanced in the study for dealing with spatial effects of barriers. For more exact and better definition of these concepts, work is needed on many facets: the effect of size changes in the floating grid, spatial distortion of non-linear diffusion wave fronts, specific combinations of barriers, interference caused by multiple diffusion, and many others. With refinements introduced by further investigation along the lines mentioned here, these notions will become useful tools in geographic research.

REFERENCES CITED

Brown, Ralph H., Historical Geography of the United States, Harcourt, Brace & World Inc., New York, 1948

Bunge, William, Theoretical Geography, Lund Studies in Geography - Series C., General and Mathematical Geography, No. 1, Royal University of Lund, Sweden, 1962

Hägerstrand, Torsten, "On Monte Carlo Simulation of Diffusion", in Quantitative Geography (William Garrison, ed.), forthcoming

Hammersley, J. M. and Handscomb, D. C., Monte Carlo Methods, Methuen & Co., London, 1964

Mackay, J. Ross, "The Interactance Hypothesis and Boundaries in Canada", The Canadian Geographer, No. 11, Ottawa, 1958 pp. 1-8

Semple, Ellen Churchill, American History and Its Geographic Condition, Houghton Mifflin Co., New York, 1903

Wolfe, Roy I., Transportation and Politics, D. Van Nostrand & Co., New Jersey, 1963

APPENDIX A: Program Listing

The Monte Carlo model used in this study was programmed for the IBM 7090 computer at the University of Michigan. The program is written in MAD (Michigan Algorithmic Decoder); the reader who is unfamiliar with this language is referred to the MAD manual which is available from the University of Michigan Computing Center. The principal parts of the program are as follows:

Data Input

The population distribution, barrier structure, values of the floating grid (MIF), and the point of termination are read in and stored. In this process certain sequences must be preserved: The population must be read in before the barriers, and the storage locations for acceptors must be cleared to zero before the initial acceptors are read into the memory. Note that use is made of the simplified input-output option of MAD.

Simulation Initiation

This is the starting point for the two basic loops which make up the main portion of the program. The major loop (JUMP1 - LOOP1) centers the MIF sequentially on each cell in the simulation field. The inside loop (JUMP2 - LOOP2) allows each cell to transmit once per generation for each acceptor in that cell. The destination of each transmission is also determined here.

Reaction of the Individual Cell to a Transmission

Depending upon the population of the cell, the transmission encounters either a barrier or a cell with interacting members. (Cell Populations $\leqq$ zero indicate barriers).

Termination

At the end of each generation, a current status report is printed which shows the spatial distribution of acceptors by cells. The simulation terminates when the number of acceptors in a specified cell passes a predetermined level.

The subroutine RANDOM provides a means of generating random numbers which are uniformly distributed over the interval $0 \leqq x \leqq 1$. RANDOM employs the power residue method of random number generation, and the periodicity is $2^{35} - 1$.

```
          R SYSTEMATIC BARRIER STUDY, MONTE-CARLO SIMULATION.
          R TOTAL FIELD 18 BY 30 CELLS.  THE MEAN INFORMATION FIELD IS
          R A SQUARE GRID OF 9 CELLS.  THE GRID CENTER IS ON CELL(N) AS
          R (N) VARIES FROM 1 TO 540.  VALUES OF THE GRID ARE BASED
          R UPON HAGERSTRAND MODELS.

           PRINT COMMENT $1$
           INTEGER I,N,X,QTS,QTA,QTR,QTD,Z,IAX,SAB,ABSORB,REF,DREF
           DIMENSION POP(540),ACCEPT(540),RECEPT(540),Z(1),AT(1),QTS(1),
          1QTA(1),QTR(1),QTD(1),IAX(1),SAB(100),ABSORB(200),REF(100),
          2DREF(100),FOLK(1),FIELD(10),CELL(50)

          R DATA INPUT

          R FIELD SPECIFIES VALUES OF FLOATING GRID (OR MEAN INFORMATION
          R FIELD).

           THROUGH ALPHA, FOR I = 2, 1, I .G. 9
ALPHA      READ DATA FIELD(I)

          FOLK SPECIFIES POPULATION OF INTERACTING CELLS.

           READ DATA FOLK
           THROUGH SET1, FOR N = 1, 1, N .G. 540
           POP(N) = FOLK
SET1       ACCEPT(N) = 0.

          R THE FOLLOWING READS IN SUBSCRIPTS OF THE CELLS WHICH ARE TO
          R BE BARRIERS.  CARD FORMAT IS 16 INTEGER FIELDS OF 5 COLUMNS
          R EACH.

          R BARRIER              PROGRAM DESIGNATION     CELL POP VALUE
          RSUPER ABSORBING             SAB                       -3.
          RABSORBING                   ABSORB                    -2.
          RREFLECTING                  REF                       -1.
          RDIRECT REFLECTING           DREF                       0.

           READ DATA QTS
           READ FORMAT CARD, SAB(1)...SAB(QTS)
           VECTOR VALUES CARD = $16I5*$
           THROUGH BETA, FOR I = 1, 1, I .G. QTS
BETA       POP(SAB(I)) = -3.
           READ DATA QTA
           READ FORMAT CARD, ABSORB(1)...ABSORB(QTA)
           THROUGH GAMMA, FOR I = 1, 1, I .G.QTA
GAMMA      POP(ABSORB(I)) = -2.
           READ DATA QTR
           READ FORMAT CARD, REF(1)...REF(QTR)
           THROUGH DELTA, FOR I = 1, 1, I .G. QTR
DELTA      POP(REF(I)) = -1.
           READ DATA QTD
           READ FORMAT CARD, DREF(1)...DREF(QTD)
           THROUGH KAPPA, FOR I = 1, 1, I .G. QTD
KAPPA      POP(DREF(I)) = 0.

          R THE FOLLOWING READS IN THE INITIAL ACCEPTORS .

           READ DATA IAX
           THROUGH SET2, FOR I = 1, 1, I .G. IAX
```

```
SET2      READ DATA CELL(1), ACCEPT(N)

          R Z = SUBSCRIPT OF CELL TERMINATING THE PROGRAM.
          R AT = VALUE OF ACCEPTORS IN CELL(Z) AT WHICH PROGRAM ENDS.

           READ DATA Z, AT

          R INPUT DATA VERIFICATION.  PRINTS OUT 18 BY 30 CELL ARRAYS
          R SHOWING INITIAL ACCEPTORS AND BARRIERS.

           PRINT COMMENT $1  INITIAL ACCEPTORS$
           PRINT FORMAT OUTPUT, ACCEPT(1)...ACCEPT(540)
           V'S OUTPUT =$1H4,1F3.0,29F4.0/(1H0,1F3.0,29F4.0)*$
           PRINT COMMENT $1 POPULATION AND BARRIER DISTRIBUTION $
           PRINT COMMENT $0 POP OF -3 IS SUPER ABSORBING BARRIER $
           PRINT COMMENT $  POP OF -2 IS ABSORBING BARRIER $
           PRINT COMMENT $  POP OF -1 IS REFLECTING BARRIER $
           PRINT COMMENT $  POP OF 0 IS DIRECT REFLECTING BARRIER $
           PRINT FORMAT OUTPUT, POP(1)...POP(540)

          R MAIN PROGRAM

          R GEN = NUMBER OF GENERATIONS.
          R RECEPT = ACCEPTORS OF EACH CELL IN ANY ONE GENERATION.
          R ACCEPT = CUMULATIVE NO. OF ACCEPTORS IN EACH CELL.
          R Q = TRANSMITTERS IN CELL(N)

           RNO = 0.
           GEN = 0.
JUMP1      THROUGH NULL, FOR N = 1, 1, N .G. 540
NULL       RECEPT(N) = 0.
           THROUGH LOOP1, FOR N = 1, 1, N .G. 540
           WHENEVER POP(N).E.0..OR.POP(N).E.-3..OR.POP(N).E.-2..OR.
          1POP(N).E.-1..OR.ACCEPT(N).LE.0.
           TRANSFER TO LOOP1
           END OF CONDITIONAL
           Q = 0.

          R THE FOLLOWING SPECIFIES THE CELL OF THE MEAN INFORMATION
          R FIELD WHICH A TRANSMITTER ATTEMPTS TO CONTACT.

JUMP2      Y = RANDOM.(RNO)
           WHENEVER Y .GE. FIELD(9)
           TRANSFER TO GRID9
           OR WHENEVER Y .GE. FIELD(8)
           TRANSFER TO GRID8
           OR WHENEVER Y .GE. FIELD(7)
           TRANSFER TO GRID7
           OR WHENEVER Y .GE. FIELD(6)
           TRANSFER TO GRID6
           OR WHENEVER Y .GE. FIELD(5)
           TRANSFER TO GRID5
           OR WHENEVER Y .GE. FIELD(4)
           TRANSFER TO GRID4
           OR WHENEVER Y .GE. FIELD(3)
           TRANSFER TO GRID3
           OR WHENEVER Y .GE. FIELD(2)
           TRANSFER TO GRID2
           OTHERWISE
```

```
         TRANSFER TO GRID1
         END OF CONDITIONAL

        R GRIDS(1-9) SPECIFY REACTION OF THE CELL CONTACTED BY THE
        R TRANSMITTER.  WITH A POP OF ZERO OR LESS. THE CELL ACTS AS
        R A BARRIER.  A POPULATION GREATER THAN ZERO INDICATES A CELL
        R WITH INTERACTORS WHICH MAY BE CONTACTED TO BECOME ACCEPTORS.
        R THE PROBABILITY OF TRANSMITTING TO AN UNCONTACTED INTER-
        R ACTOR OF A CELL DECREASES AS THE RATIO ACCEPTORS/CELL POP
        R OF THAT CELL INCREASES.

GRID1    WHENEVER POP(N-31) .E. -3.
         ACCEPT(N) = ACCEPT(N) - 1.
         TRANSFER TO LOOP2
         OR WHENEVER POP(N-31) .E. -2.
         TRANSFER TO LOOP2
         OR WHENEVER POP(N-31) .E. -1.
         TRANSFER TO JUMP2
         OR WHENEVER POP(N-31) .E. 0.
         Y = RANDOM.(RNO)
         WHENEVER Y .GE. 0.500 .AND. POP(N-30) .G. 0.
         TRANSFER TO GRID2
         OTHERWISE
         TRANSFER TO GRID4
         END OF CONDITIONAL
         OTHERWISE
         CONTINUE
         END OF CONDITIONAL
         QUOT = ACCEPT(N-31) / POP(N-31)
         Y = RANDOM.(RNO)
         WHENEVER Y.GE.QUOT.AND.RECEPT(N-31)+ACCEPT(N-31).L.FOLK
         RECEPT(N-31) = RECEPT(N-31) + 1.
         END OF CONDITIONAL
         TRANSFER TO LOOP2
GRID2    WHENEVER POP(N-30) .E. -3.
         ACCEPT(N) = ACCEPT(N) - 1.
         TRANSFER TO LOOP2
         OR WHENEVER POP(N-30) .E. -2.
         TRANSFER TO LOOP2
         OR WHENEVER POP(N-30) .E. -1.
         TRANSFER TO JUMP2
         OR WHENEVER POP(N-30) .E. 0.
         TRANSFER TO GRID5
         OTHERWISE
         CONTINUE
         END OF CONDITIONAL
         QUOT = ACCEPT(N-30) / POP(N-30)
         Y = RANDOM.(RNO)
         WHENEVER Y.GE.QUOT.AND.RECEPT(N-30)+ACCEPT(N-30).L.FOLK
         RECEPT(N-30) = RECEPT(N-30) + 1.
         END OF CONDITIONAL
         TRANSFER TO LOOP2
GRID3    WHENEVER POP(N-29) .E. -3.
         ACCEPT(N) = ACCEPT(N) - 1.
         TRANSFER TO LOOP2
         OR WHENEVER POP(N-29) .E. -2.
         TRANSFER TO LOOP2
         OR WHENEVER POP(N-29) .E. -1.
         TRANSFER TO JUMP2
```

```
           OR WHENEVER POP(N-29) .E. 0.
           Y = RANDOM.(RNO)
           WHENEVER Y.GE. 0.500 .AND.POP(N-30) .G. 0.
           TRANSFER TO GRID2
           OTHERWISE
           TRANSFER TO GRID6
           END OF CONDITIONAL
           OTHERWISE
           CONTINUE
           END OF CONDITIONAL
           Y = RANDOM.(RNO)
           QUOT = ACCEPT(N-29) / POP(N-29)
           WHENEVER Y.GE.QUOT.AND.RECEPT(N-29)+ACCEPT(N-29).L.FOLK
           RECEPT(N-29) = RECEPT(N-29) + 1.
           END OF CONDITIONAL
           TRANSFER TO LOOP2
GRID4      WHENEVER POP(N-1) .E. -3.
           ACCEPT(N) = ACCEPT(N) - 1.
           TRANSFER TO LOOP2
           OR WHENEVER POP(N-1) .E. -2.
           TRANSFER TO LOOP2
           OR WHENEVER POP(N-1) .E. -1.
           TRANSFER TO JUMP2
           OR WHENEVER POP(N-1) .E. 0.
           TRANSFER TO GRID5
           OTHERWISE
           CONTINUE
           END OF CONDITIONAL
           Y = RANDOM.(RNO)
           QUOT = ACCEPT(N-1) / POP(N-1)
           WHENEVER Y.GE.0.500.AND.RECEPT(N-1)+ACCEPT(N-1).L.FOLK
           RECEPT(N-1) = RECEPT(N-1) + 1.
           END OF CONDITIONAL
           TRANSFER TO LOOP2
GRID5      Y = RANDOM.(RNO)
           QUOT = ACCEPT(N) / POP(N)
           WHENEVER Y.GE.QUOT.AND.RECEPT(N)+ACCEPT(N).L.FOLK
           RECEPT(N) = RECEPT(N) + 1.
           END OF CONDITIONAL
           TRANSFER TO LOOP2
GRID6      WHENEVER POP(N+1).E. -3.
           ACCEPT(N) = ACCEPT(N) - 1.
           TRANSFER TO LOOP2
           OR WHENEVER POP(N+1) .E. -2.
           TRANSFER TO LOOP2
           OR WHENEVER POP(N+1) .E. -1.
           TRANSFER TO JUMP2
           OR WHENEVER POP(N+1) .E. 0.
           TRANSFER TO GRID5
           OTHERWISE
           CONTINUE
           END OF CONDITIONAL
           Y = RANDOM.(RNO)
           QUOT = ACCEPT(N+1) / POP(N+1)
           WHENEVER Y.GE.QUOT.AND.RECEPT(N+1)+ACCEPT(N+1).L.FOLK
           RECEPT(N+1) = RECEPT(N+1) + 1.
           END OF CONDITIONAL
           TRANSFER TO LOOP2
GRID7      WHENEVER POP(N+29) .E. -3.
```

```
          ACCEPT(N) = ACCEPT(N) - 1.
          TRANSFER TO LOOP2
          OR WHENEVER POP(N+29) .E. -2.
          TRANSFER TO LOOP2
          OR WHENEVER POP(N+29) .E. -1.
          TRANSFER TO JUMP2
          OR WHENEVER POP(N+29) .E. 0.
          Y = RANDOM.(RNO)
          WHENEVER Y.GE.0.500 .AND.POP(N-1) .G. 0.
          TRANSFER TO GRID4
          OTHERWISE
          TRANSFER TO GRID8
          END OF CONDITIONAL
          OTHERWISE
          CONTINUE
          END OF CONDITIONAL
          Y = RANDOM.(RNO)
          QUOT = ACCEPT(N+29) / POP(N+29)
          WHENEVER Y.GE.QUOT.AND.RECEPT(N+29)+ACCEPT(N+29).L.FOLK
          RECEPT(N+29) = RECEPT(N+29) +1.
          END OF CONDITIONAL
          TRANSFER TO LOOP2
GRID8     WHENEVER POP(N+30) .E. -3.
          ACCEPT(N) = ACCEPT(N) - 1.
          TRANSFER TO LOOP2
          OR WHENEVER    POP(N+30) .E. -2.
          TRANSFER TO LOOP2
          OR WHENEVER POP(N+30) .E. -1.
          TRANSFER TO JUMP2
          OR WHENEVER POP(N+30) .E. 0.
          TRANSFER TO GRID5
          OTHERWISE
          CONTINUE
          END OF CONDITIONAL
          Y = RANDOM.(RNO)
          QUOT = ACCEPT(N+30) / POP(N+30)
          WHENEVER Y.GE.QUOT.AND.RECEPT(N+30)+ACCEPT(N+30).L.FOLK
          RECEPT(N+30) = RECEPT(N+30) + 1.
          END OF CONDITIONAL
          TRANSFER TO LOOP2
GRID9     WHENEVER POP(N+31) .E. -3.
          ACCEPT(N) = ACCEPT(N) - 1.
          TRANSFER TO LOOP2
          OR WHENEVER POP(N+31) .E. -2.
          TRANSFER TO LOOP2
          OR WHENEVER POP(N+31) .E. -1.
          TRANSFER TO JUMP2
          OR WHENEVER POP(N+31) .E. 0.
          Y = RANDOM.(RNO)
          WHENEVER Y.GE. 0.500 .AND.POP(N+30) .G. 0.
          TRANSFER TO GRID8
          OTHERWISE
          TRANSFER TO GRID6
          END OF CONDITIONAL
          OTHERWISE
          CONTINUE
          END OF CONDITIONAL
          Y = RANDOM.(RNO)
          QUOT = ACCEPT(N+31) / POP(N+31)
```

```
          WHENEVER Y.GE.QUOT.AND.RECEPT(N+31)+ACCEPT(N+31).L.FOLK
          RECEPT(N+31) = RECEPT(N+31) + 1.
          END OF CONDITIONAL

         R TERMINATION AND OUTPUT

         R THE FOLLOWING ALLOWS EACH CELL TO TRANSMIT ONLY ONCE FOR
         R EACH ACCEPTOR IN EACH GENERATION.

LOOP2     Q = Q + 1.
          WHENEVER Q .L. ACCEPT(N)
          TRANSFER TO JUMP2
          END OF CONDITIONAL
LOOP1     CONTINUE
          GEN = GEN + 1.
          THROUGH ADD, FOR N = 1, 1, N .G. 540
ADD       ACCEPT(N) = ACCEPT(N) + RECEPT(N)

         R THE FOLLOWING SPECIFIES FORM OF THE OUTPUT.  THE ACCEPTORS
         R ( OF EACH OF THE 540 CELLS) ARE PRINTED OUT IN AN 18 BY 30
         R CELL ARRAY WITH THE GENERATION NUMBER PRINTED ABOVE THE
         R ARRAY.

          PRINT FORMAT COUNT, GEN
          V'S COUNT = $1H1, H*  GENERATION NUMBER =  *, F3.0 *$
          PRINT FORMAT OUTPUT, ACCEPT(1)...ACCEPT(540)

         R THE FOLLOWING CARDS TEST FOR THE CONDITION WHICH TERMINATES
         R THE PROGRAM.

          WHENEVER ACCEPT(Z) .L. AT
          TRANSFER TO JUMP1
          END OF CONDITIONAL
          END OF PROGRAM
```

FOREWORD

Last winter, Dr. William Warntz, research associate with the American Geographical Society, presented a guest lecture to the members of MICMG. In that lecture, Dr. Warntz spelled out quite clearly, among other things, the necessity for a closer association by geographers with its sister geo-science, geometry, in order to more fully understand the complexities of spatial distributions.

Since Dr. Warntz's comments appeared particularly germane to the interests of MICMG discussion paper readers, he agreed to expand this topic into a formal discussion paper. Thus, MICMG is pleased to present, as our second guest author, Dr. Warntz's views on an essential and timely topic in the expanding area of mathematical geography.

The Editor, June, 1965

A NOTE ON SURFACES AND PATHS AND APPLICATIONS TO GEOGRAPHICAL PROBLEMS

By William Warntz

Recent gains in the development of general geography as a science at the theoretical - predictive level have been accomplished largely, perhaps, as a result of the simultaneous increase in both the sophistication and the naiveté with which geographers view the phenomena of the real world. This seeming paradox is explainable, for example, with reference to mathematics. More than ever before, geographers are using the tools of calculus, probability, topology, symbolic logic, the various algebras, geometries, for example, are being taken more literally than ever before.

We are aware of the topological nature of problems in non-spatial sciences like chemistry involving connectives or bondings of molecules, and of those concerning hierarchial chains of command, responsibility, and interactions as recognized in certain of the non-spatial social sciences. In addition, in general classificatory science and in the mathematical set theory underlying it, use is made of Venn diagrams, which utilize topological properties of an idealized space to portray graphically such relations as are implied in subsets, intersects, and unions, etc. But, we can take Venn diagrams in a far more literal sense than they were originally intended and by substituting real space and attendant phenomena for ideal space and by insisting upon utilization of all of the geometric properties involved as well as just the topological ones, geographers can reinterpret, add to, and refine the conventional concepts in the methodology of uniform regional geography and provide it with a basis in logic. Further discussion of this important topic lies outside the present paper and is reserved for presentation elsewhere. Of course, geographers also take the spatial considerations in topology literally in analysis of highway networks, river systems, and so on.

With regard to geometries, one also can cite numerous examples of geometrical solutions to problems which are not inherently spatial or in which the problems have been abstracted from space, and geometry is employed <u>only</u> by analogy. Included would be such things as one approach in economics to consumer tastes, prices, and preferences. Indifference curves are used to portray a surface of satisfaction. Various paths on this surface have meaning with regard to the income effect and the substitution effect. Other examples abound in economics.

In chemistry the use of surfaces and paths to show relationships in non-spatial thermodynamics is due to the nineteenth century American scientist, Josiah Willard Gibbs. His methods of geometrical representation of thermodynamic properties of substances by means of surfaces showed, for example, how to diagram water as it undergoes changes from solid to liquid to gas. So impressive was this work that the brilliant British scientist, Clerk Maxwell, saluted Gibbs by building for him a plaster model entitled, "a statue of water." In general the mathematics of response surfaces, supported by appropriate statistical measures now appear as well in many non-spatial sciences, e.g., psychology, learning theory, and so on.

Today geographers are taking the <u>geo</u> in geometry literally and the study of earth related surfaces and paths has now been expanded far beyond its original application to such things as land forms, contour mapping, drainage patterns, temperatures, pressures, precipitation, and the like in physical geography alone.

The modern geographer conceives of surfaces based also on social, economic, and cultural phenomena portraying not only conventional densities but other things such as field quantity potentials, probabilities, costs, times, and so on. Always, however, these conceptual surfaces may be regarded as overlying the surface of the real earth and the geometric and topological characteristics of these surfaces, as transformed, thus describe aspects of the geography of the real world.

Elaboration of additional important ideas may be found in William Bunge's <u>Theoretical Geography</u>, Lund, Sweden, 1962. Bunge has pointed out the necessity and the efficiency of recognizing the inseparability of <u>geometry</u> as the mathematics, i.e., the language, of spatial relations and <u>geography</u> as the science of spatial relations.

Illustrations of the application of the ideas, the geometry of surfaces and paths to a number of geographical problems can readily be given. The examples following have been selected because of their non-spatial diversity.

Our first example concerns any conformal map projection with the one exhibited here in figure 1, the well known Mercator projection - equatorial case. An exact mathematical isomorphism exists between the paths of light rays in an isotropic medium with an index of refraction varying from point to point but constant in all directions around any given point and the least distance paths or great circle arcs on the earth's spherical surface as represented on a conformal map. Let f be the scale of a conformal map at any point - expressed as the fractional ratio of distance on the map to actual distance on the earth, and let ds be the infinitesimal line element measured on the map. Then, the value $\int \frac{1}{f} ds$ is a minimum for the great circle track between any two given points as compared with the values obtained by integrating along any nearby alternative paths. This can be stated as a calculus of variations problem.

The designation of great circle tracks can also be accomplished by graphical portrayal of a surface and a gradient path. Shown in figure 1 is a Mercator map with a distance surface based on London depicted by isodistance lines in statute miles with a constant interval. On the real earth such lines would be concentric circles at first with increasing circumferences and then decreasing to zero at the antipodal point. When portrayed on the Mercator projection these circles are transformed, for although the Mercator is conformal in the small, it cannot show

DISTANCE SURFACE AND GEODESICS BASED ON LONDON

MERCATOR PROJECTION

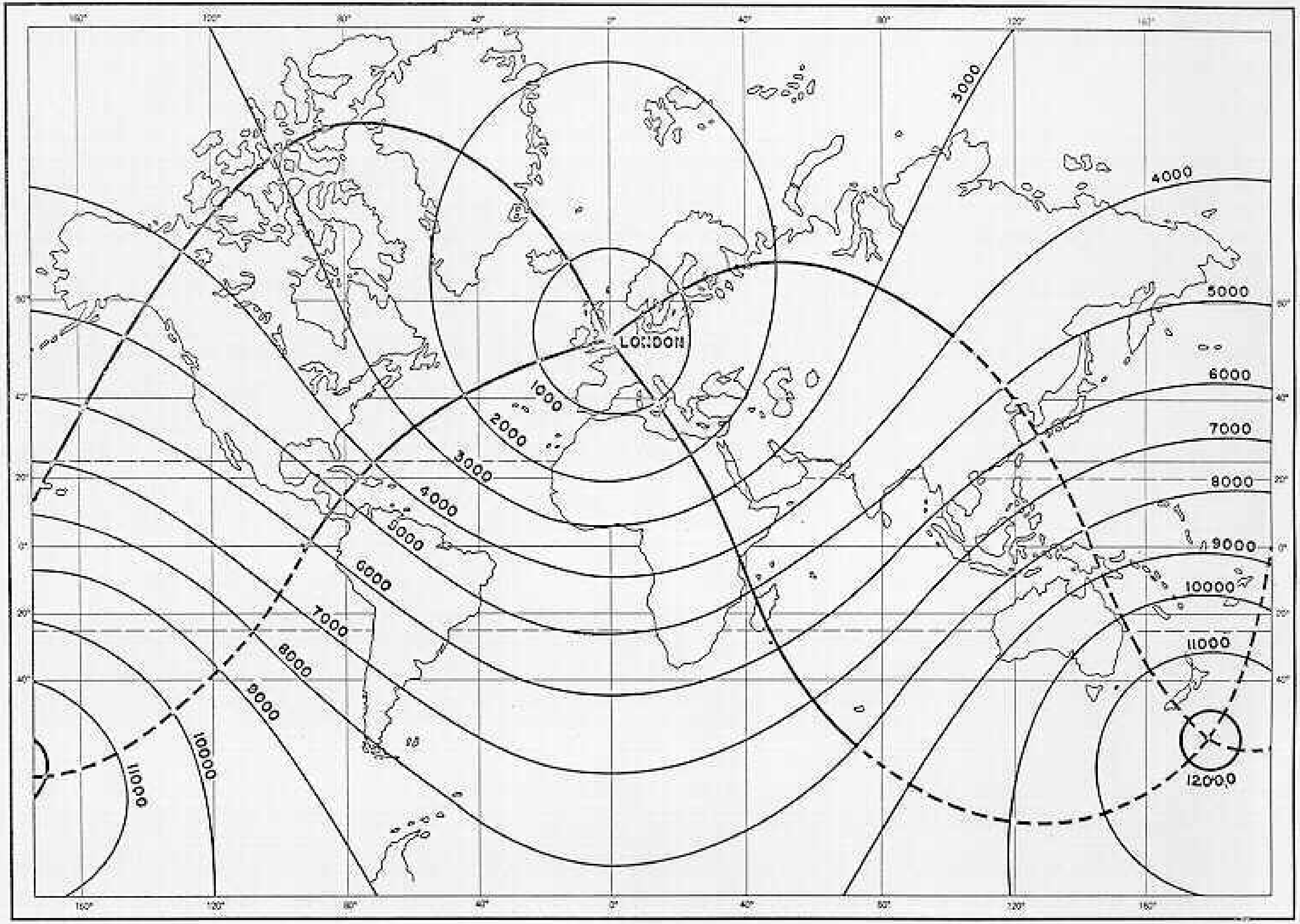


FIGURE I

correct shapes in the large because of scale change. On a spherical earth great corcles from all points to a given point would be orthogonal trajectories to the iso-distance circles centered on the given point. On this conformal map that property is retained but the great circles must bend to achieve it. Notice then, the presentation in figure 1 of the isolines of the distance surface and the orthogonal trajectories or gradient paths providing one solution to the problem of great circle determination. (Note that an actual geodesic on the map plane, the straight line, traces out a rhumb or constant heading line on the earth.)

Figure 1A shows another conformal map - a stereographic with the center of projection, in this case, at the South Pole. Shown on this map are certain iso-distance circles in statute miles and great circle paths based on Salisbury, Southern Rhodesia. Again the gradient path for great circles on the distance surface is found. The rule of orthogonal trajectory applies to this as to all other conformal maps.

The stereographic has a number of interesting properties. All great or small circles, or arcs of circles, on the earth's surface map as circles or arcs of circles on the stereographic. But concentric circles on the earth, while mapping as circles, do not, in general map as concentric circles. The exception, of course, is the case for the center of the projection. In figure 1A (assuming a perfectly spherical earth) small circles of latitude (these may be regarded as iso-distance lines on the earth) do map as concentric circles about the South Pole. However, let us look at the iso-distance circles about Salisbury. These are concentric circles on a spherical earth but on the plane stereographic projection they map as non-concentric circles - or non-concentric arcs of circles since the entire earth is not shown on one stereographic projection.

The centers of these mapped iso-distance circles about Salisbury march steadily away from Salisbury along the straight line on the plane of the projection from the

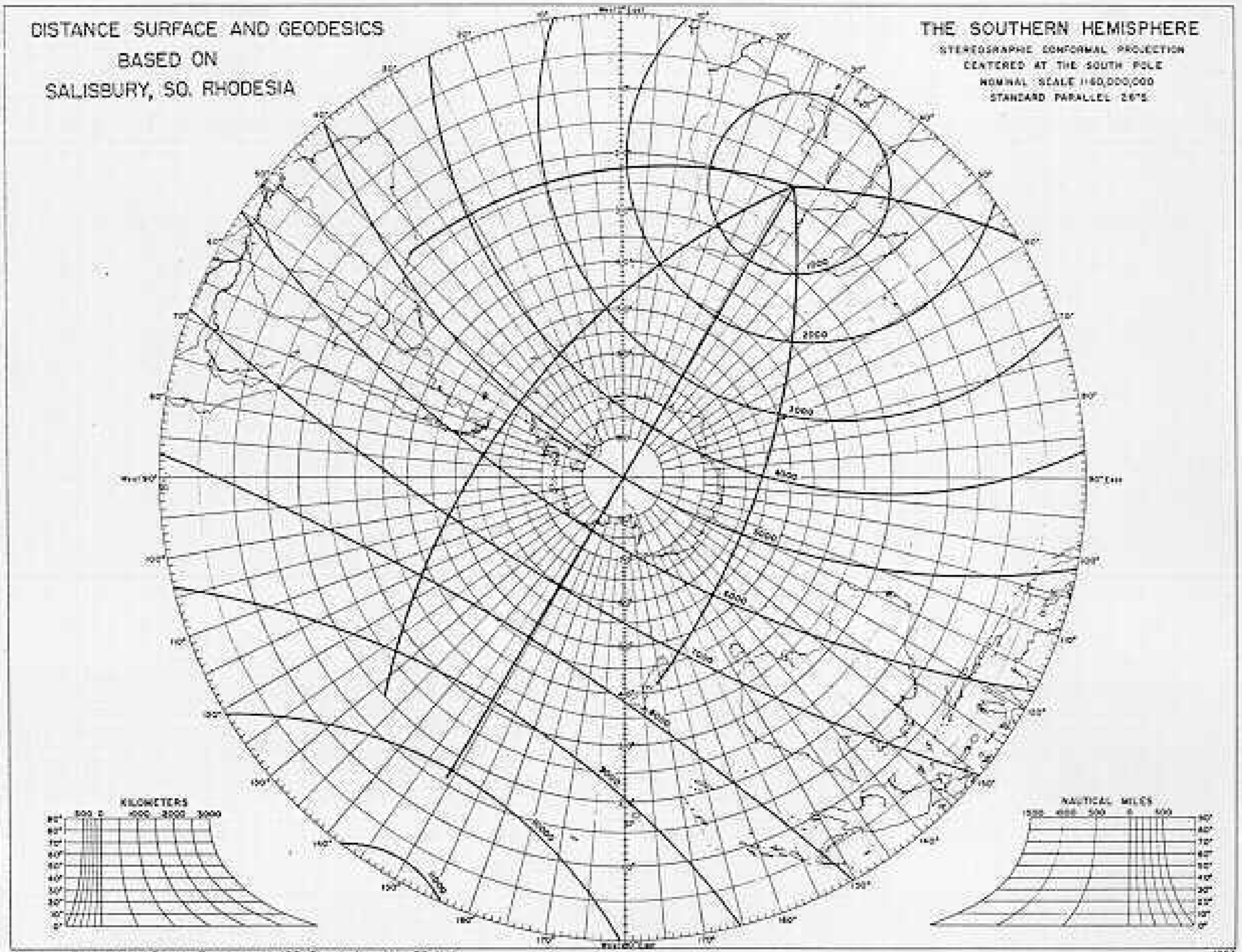

DISTANCE SURFACE AND GEODESICS
BASED ON
SALISBURY, SO. RHODESIA
THE SOUTHERN HEMISPHERE
STEREOGRAPHIC CONFORMAL PROJECTION
CENTERED AT THE SOUTH POLE
NOMINAL SCALE 1:60,000,000
STANDARD PARALLEL 26°S
KILOMETERS
NAUTICAL MILES
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
American Geographical Society, Broadway at 156 Street, New York 32, N.Y.

FIGURE IA

center of the projection (the South Pole) extended through Salisbury. The direction of movement, of course, is away from the center of the projection. The iso-distance circle that passes through the antipodal point of the center of the projection will map as a straight line. For this map the antipodal point of the center of the projection is the North Pole. (This is also the center of perspective for this projection.) So, the iso-distance circle about Salisbury that passes through the North Pole maps as a straight line and its center on the plane of projection lies at infinity, in fact at both plus and minus infinity if we adopt that convention. The next iso-distance circle out from Salisbury will map with negative curvature and its center on the plane of the projection will lie between the point at negative infinity and the antipodal point of Salisbury. The centers of the succeeding iso-distance circles based on Salisbury will march ever closer to Salisbury's antipodal point. This path, of course, is the straight line on the plane from the point of minus infinity to the antipodal point of Salisbury and which when extended would also pass through the South Pole (and beyond through Salisbury and to the point of plus infinity noted above).

Once the positions of the selected iso-distance circles have been determined the drawing of the least earth distance path or geodesic from any point to Salisbury is a simple matter. It is the orthogonal trajectory or gradient path on this conformal map. All places are relative sources for Salisbury as the single sink. The refraction analogy holds again and the reciprocal of the map scale at any point may be regarded as an index of refraction for the geodesic there.

It is interesting to note that the family of geodesics passing through Salisbury which, of course, are great circles on the earth's surface also map as circles on the stereographic.

On the surface of the spherical earth, the latitude and longitude circles may be regarded, respectively, as distance circles about, and geodesics passing through, the geographic poles. Thus, our Salisbury case resembles the polar case on any non-polar stereographic.

The next example considers the problem of determining the path for minimum time enroute for an aircraft between two airports separated by a broad expanse with winds of varying direction and velocity. Almost never will the least distance path, i.e., great circle route, afford the minimum time enroute. Generally, it is possible to deviate from this great circle route to enhance the speed over the earth's surface by utilizing more favorable winds. This will be attempted, ordinarily, so long as the addition to speed is proportionally greater than the addition to distance. An elegant mathematical solution to this calculus of variations problem exists which, however, is not practical. Essentially the problem is one of the path which minimizes the value, $\int \frac{1}{v} ds$ when v is the speed of the aircraft over the earth's surface and ds is as above. Difficulty arises from the fact that velocity and direction of the aircraft over the surface depends upon the vector of heading and true air speed of the aircraft combined with that of wind direction and velocity. The resulting tail wind or head wind component at any given point in the wind field is not independent of the aircraft's actual heading. But, a graphical solution can be used to develop a family of isotime lines and one also of course lines based on a given departure point and showing the times and routes for minimum time flights to all destinations. The method involved is that modified by including wind vector considerations from the seventeenth century Dutch scientist, Huygens, relating to sound and light refraction based on wavelets, envelope curves, and time fronts. Figure 2 shows some of the family of minimum time routes on the time surface integrated about New York City for DC-8 jet aircraft flying to Europe at a constant pressure altitude of 300 millibars on October 17, 1960, and Figure 2A shows how to create such a surface and attendant paths.

When portrayed in full, the situation depicted in figure 2 reveals cusps or

TIME SURFACES AND ROUTES FROM NEW YORK CITY, DC-8, OCTOBER 17, 1960

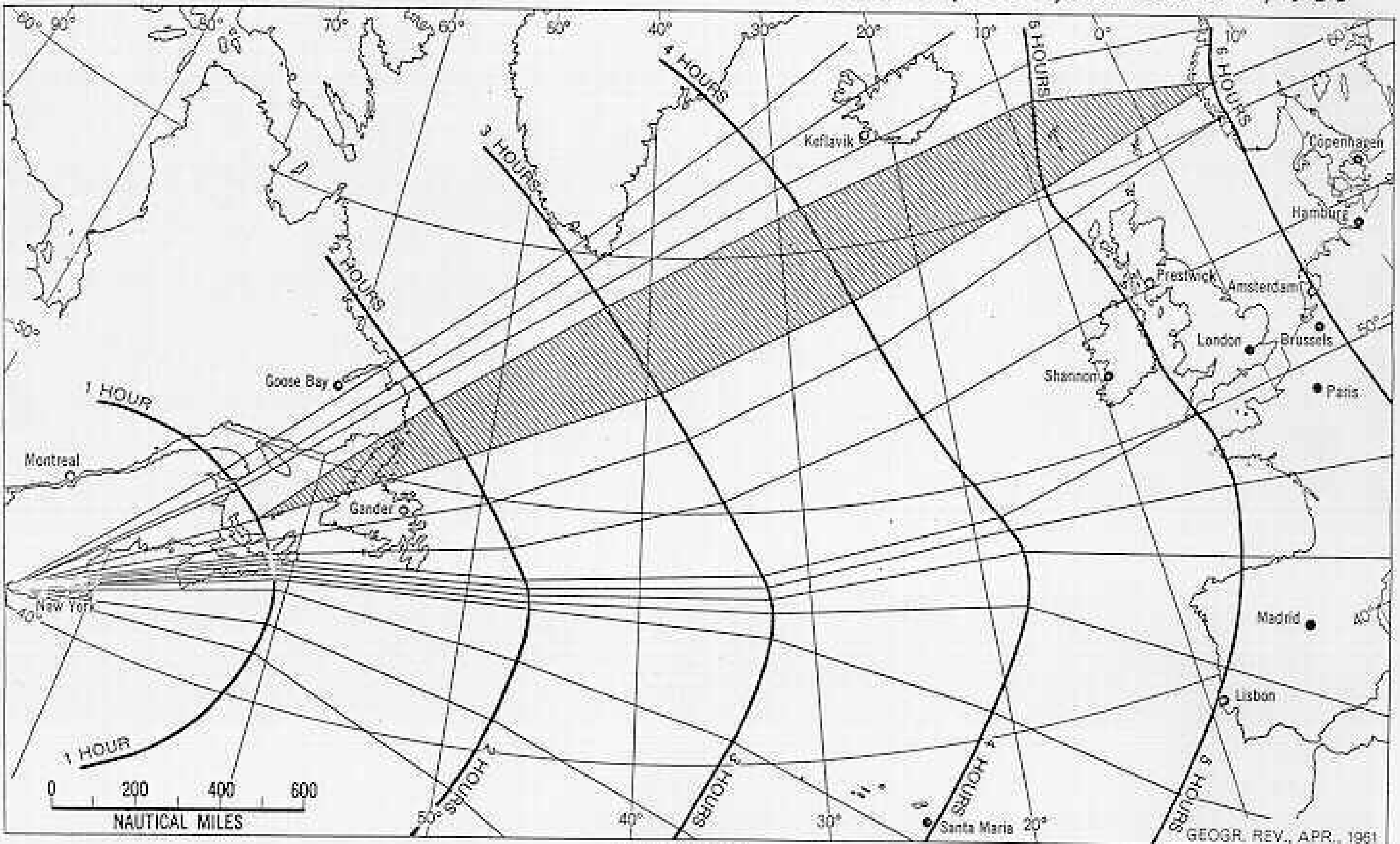
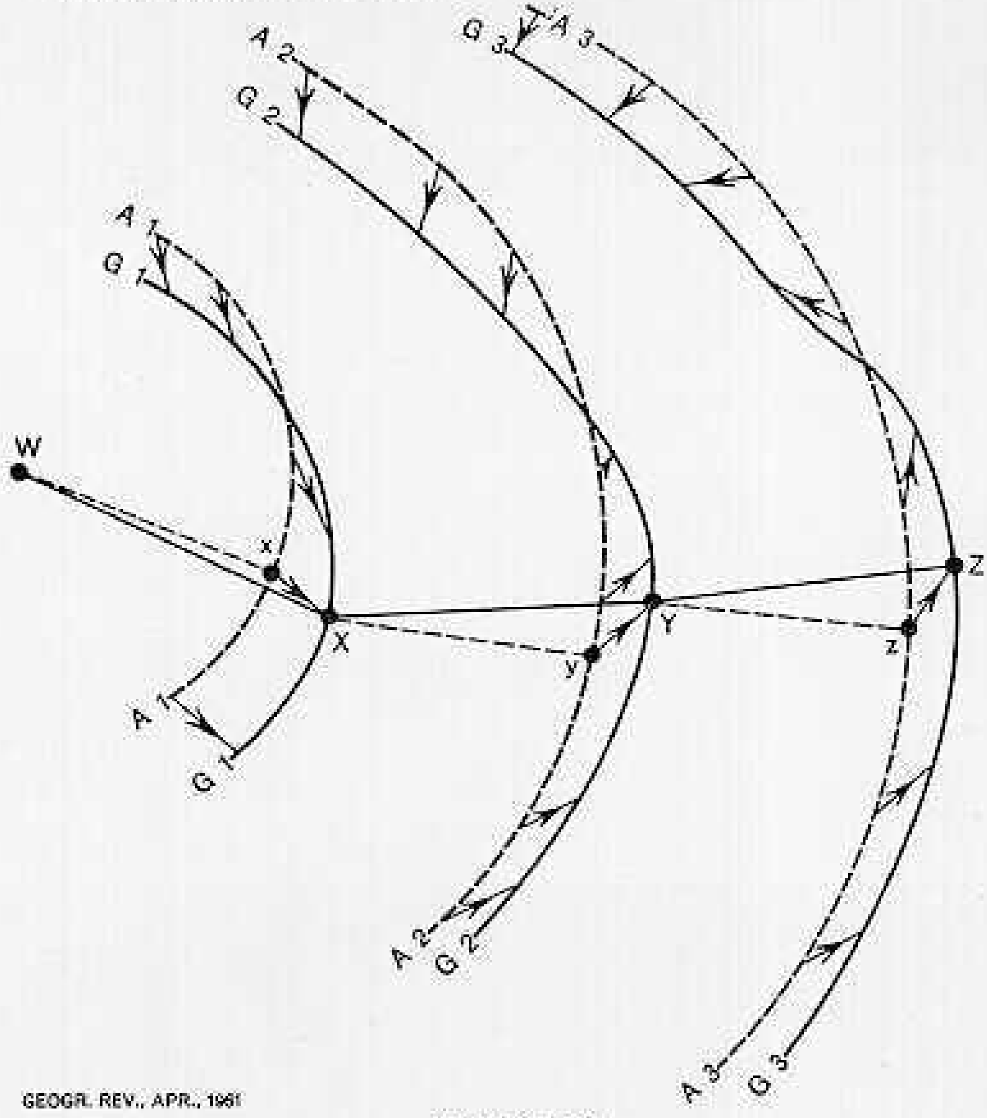


FIGURE 2

GRAPHICAL DETERMINATION OF LEAST TIME PATH



GEOGR. REV., APR., 1961

FIGURE 2A

focal points in the least time routes and thus provides alternate equal minimum-time paths to certain destinations. Such a situation refers to the induction of caustics and occurs at the rear of a closed vortex associated with a strong high or low or in a region of strong wind shear as in a jet stream or a frontal zone.

Under certain pressure distributions complicated systems of time fronts and routes exist, and, as with their isomorphic counterparts in optics, reflection, diffraction, refraction, and other patterns can be ascertained to exist in addition to, and in agreement with, the caustics.

With departure point W in figure 2A as the center, draw a circle with a radius representing the distance the aircraft can fly in one hour at its cruising true air speed. Only an arc of this circle is shown here, covering the general direction of the intended flight. This arc of the circle, labeled A1, indicated the maximum distances the aircraft could fly in the complete absence of wind. The effect of wind is estimated by drawing the appropriate wind vectors representing velocity and direction from a number of points on the air position line A1. The heads of these wind vectors can be connected by the smooth curve G1, which designates the position of the time front after one hour and gives the farthest ground positions the aircraft could achieve in that time in the existing wind field.

From as many points as desired on curve G1, arcs may be swung off each with a radius again equal to the true air speed. The curve A2 is drawn as the envelope of these arcs. Displacement of A2 by the appropriate wind vectors yields G2, the position line of the time front at the end of the second hour after departure.

This procedure can be continued until the time front closest to required destination Z is obtained. In the case portrayed here destination lies exactly on a time front. If, as is usually the case, destination does not lie on a time front, interpolation will yield an adequate estimate of the least time in which the flights can be accomplished.

To obtain directly the specific route form W to Z that affords this minimum flight time, plot backward from destination to departure point. From Z (in this case on G3) the wind vector is plotted in reverse, giving the point z on A3. From this point the perpendicular is dropped to G2. This new point, Y, is on the required optimum flight path.

Repetition of this procedure yields point X on G1 and finally W as the departure point. The optimum flight path, here the minimum-time path, is then the curve WXYZ. This course represents the connected series of ground-speed-true-course vectors each perpendicular to its attendant air position curve. Likewise, the true-air-speed-true-heading vectors are perpendicular to their ground position curves, the time fronts.

The methods used here result in approximations of the desired result. Such graphical procedures involve arbitrary discontinuities and averages. Specifically, the accuracy achieved apart from the adequacy of the meteorological forecast tends to vary inversely with the length of the time interval chosen to portray wind velocities, air speeds, ground speeds, and the successive positions of the time front.

The final examples utilize a potential of population map based on 1960 census counts for the conterminous United States and computed and compiled jointly by the American Geographical Society and Princeton University using the IBM 7090 computer. Figure 3 shows this map, the most detailed such map yet produced. Potential of population is shown as a macrogeographic spatially continuous surface by means of isolines indicating geographical variation in the values on this surface. Potential of population measures the aggregate accessibility of the entire population of the country to all points when the value at any given point represents the summed contributions from all members of the population when each contribution is proportional to the reciprocal of the distance of the person away from the given point. More specifically, the value of potential of population at a given point is $\int \frac{1}{r} D dA$ when D is the population density over any infinitesimal element of area, dA, and r is

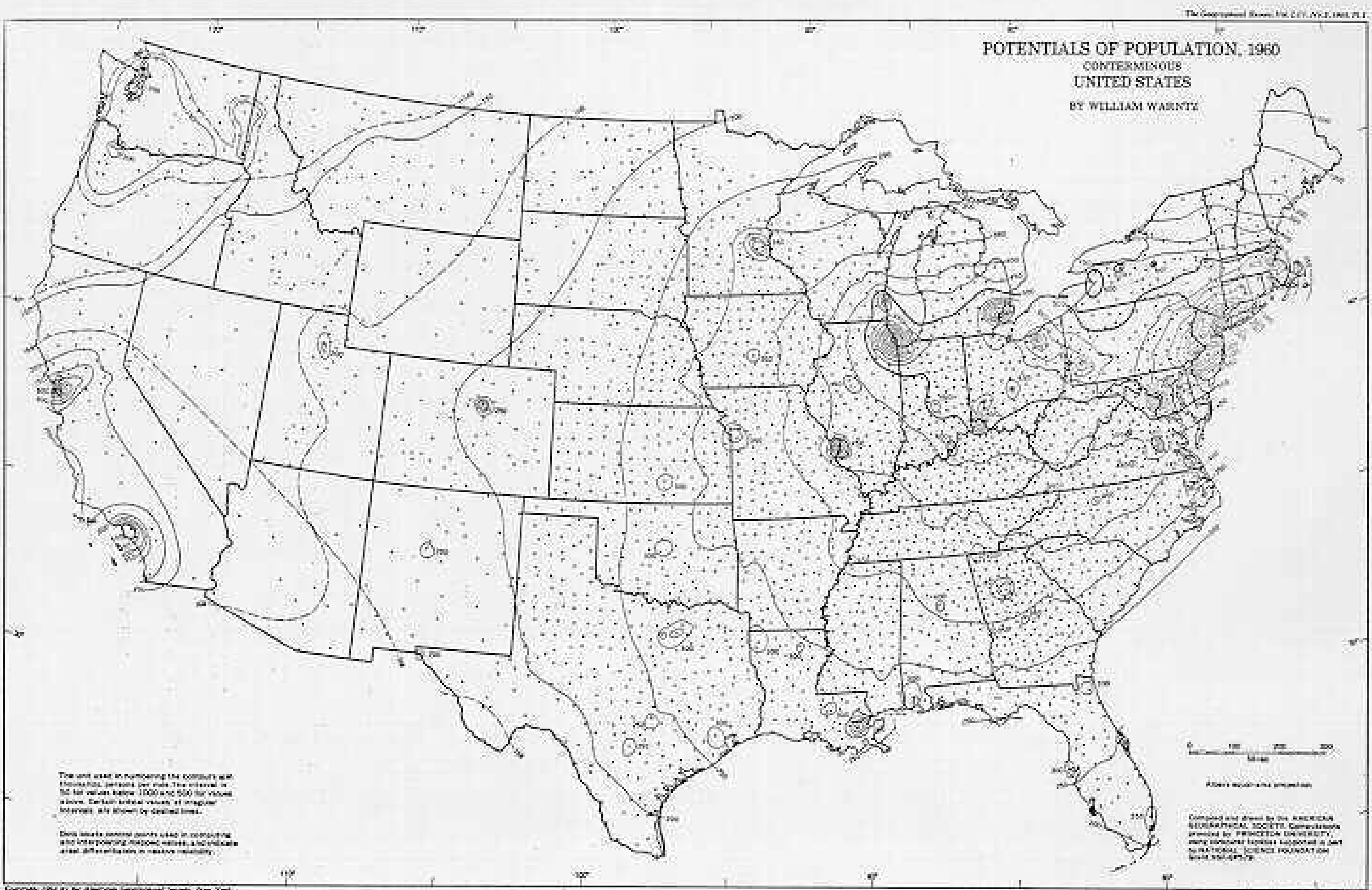

POTENTIALS OF POPULATION, 1960
CONTERMINOUS
UNITED STATES
BY WILLIAM WARNTZ

FIGURE 3

the distance of each such element from the given point. The integration is extended to all elements when D is not zero. The value of potential of population can be determined for as many points as required to facilitate mapping the surface by the contouring technique. Units of potential of population are in persons times distance to the minus one power. On figure 3 values are in thousands, persons per mile. To convert to thousands, persons per kilometer, the mapped values may be multiplied by the factor 0.622.

Maps of this sort indicate spatial structuring of a wide range of economic and social phenomena and are based upon formalization arising from empirical formulas describing certain spatial processes, i.e. flows that are of economic and social importance.

With regard to spatial structure it is important to note that values of non-urban land in the United States vary directly with and in close agreement with the potential of population surface. Specifically for 1960, land value in dollars per acre = 6.104^{-6} times potential of population raised to the 1.6 power when potential was in units of thousands, persons per mile. The coefficient of correlation was 0.863 based on state averages and thus was quite high. For the number of degrees of freedom obtaining here any value of the coefficient of correlation exceeding 0.288 is to be deemed significant at the five per cent fiducial level.

Urban land values increase with potential to an even higher power than for non-urban land so that one simple linear logarithmic function is inadequate. Even so, however, one may closely approximate a continuous land value surface for the United States by transforming the potential surface everywhere by the factor of proportionality and exponent of potential given above due to the peaking of potential locally in urban areas. Such a procedure will permit us to provide the illustration intended below.

Let us imagine that it is the task to establish from any given place the routes to all other places in the country and in each case the route is the one for which the total land acquisition cost would be at a minimum. A map of such a family of routes and the attendant isocost surface on which these routes represent gradient paths can be produced by means of the simple Huygens' graphical method. Figure 4 represents an isocost surface integrated about Lewistown, Montana. Lewistown, in this case, is a "sink" and all other places are relative "sources." The minimum land acquisition cost route to any other given point in the United States may be found by plotting the orthogonal trajectory to the isocost lines from the given point back to Lewistown. Hence, the routes are gradient paths. Several such paths are shown on the map. Note the divide between the northerly and southerly routes emanating from this town in Montana. To further emphasize the nature of such surfaces and the paths on them, an additional graphical presentation is supplied in figure 5 based on Murfreesboro, Tennessee.

Note particularly by comparing figures 4 and 5 that the path from Lewistown to Murfreesboro is precisely the same as that from Murfreesboro to Lewistown. Thus, although the sets of iso-cost contours differ greatly the tangents to both sets of contours coincide along this path. Because the original values on the land value surface are independent of the directions of lines through points on that surface, the conditions obtains that between any two points the least cost paths coincide. Other than this, however, a separate map is required for each point to obtain the minimum cost paths from all other points to the given point. The possibility of producing one general map to accomplish all such paths is being investigated.

For the minimum time paths for aircraft as shown in figure 2 the same path does not exist in opposite directions between any two points. As noted earlier, the effect at a point of the wind on the speed and the heading of an aircraft depends upon the heading and the speed with which the aircraft approaches the point.

COST SURFACE AND MINIMUM LAND ACQUISITION COST ROUTES BASED ON LEWISTOWN, MONTANA

(IN HUNDREDS OF THOUSANDS OF DOLLARS)

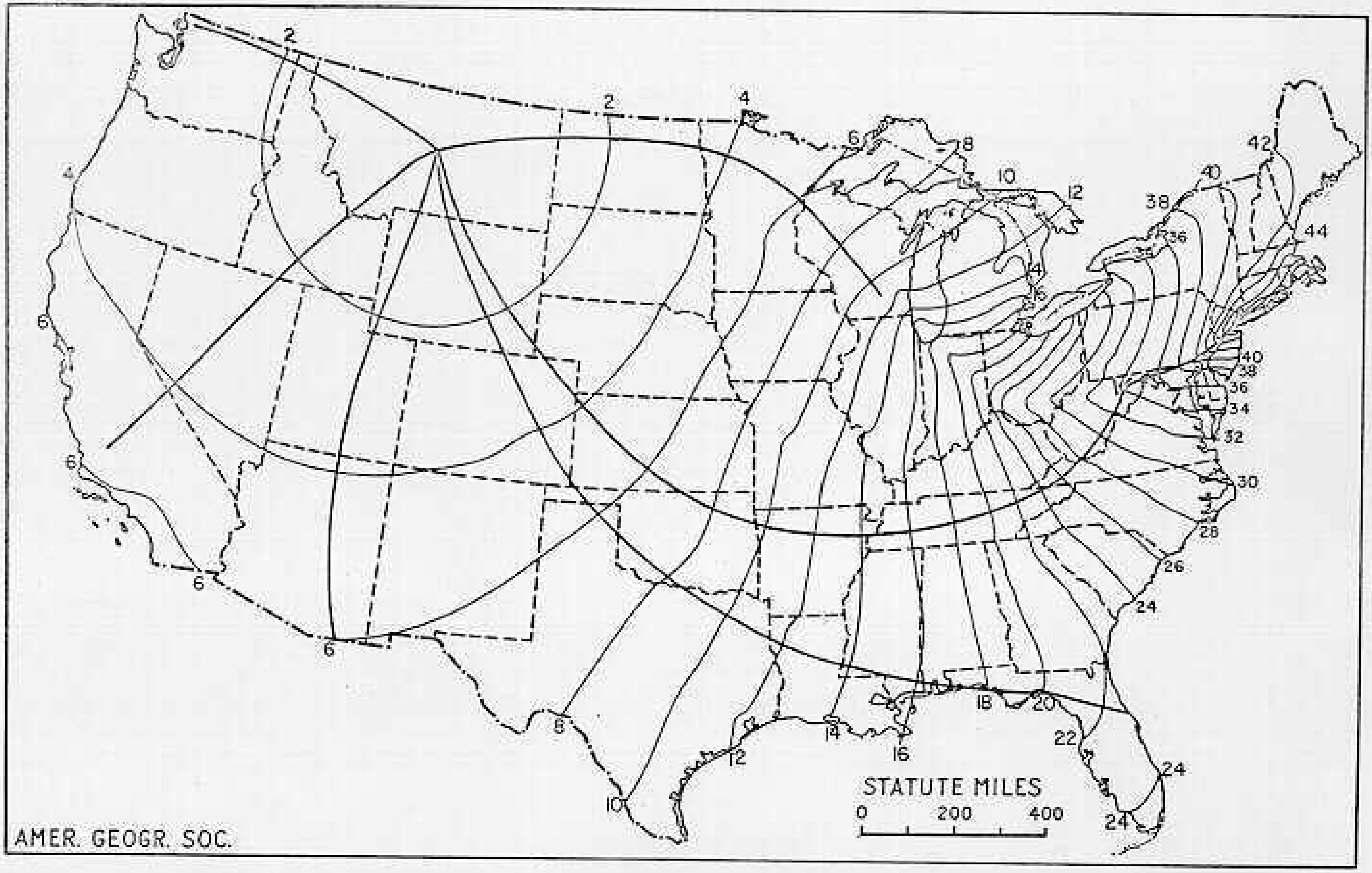


FIGURE 4

COST SURFACE AND MINIMUM LAND ACQUISITION COST ROUTES BASED ON MURFREESBORO, TENNESSEE

(IN HUNDREDS OF THOUSANDS OF DOLLARS)

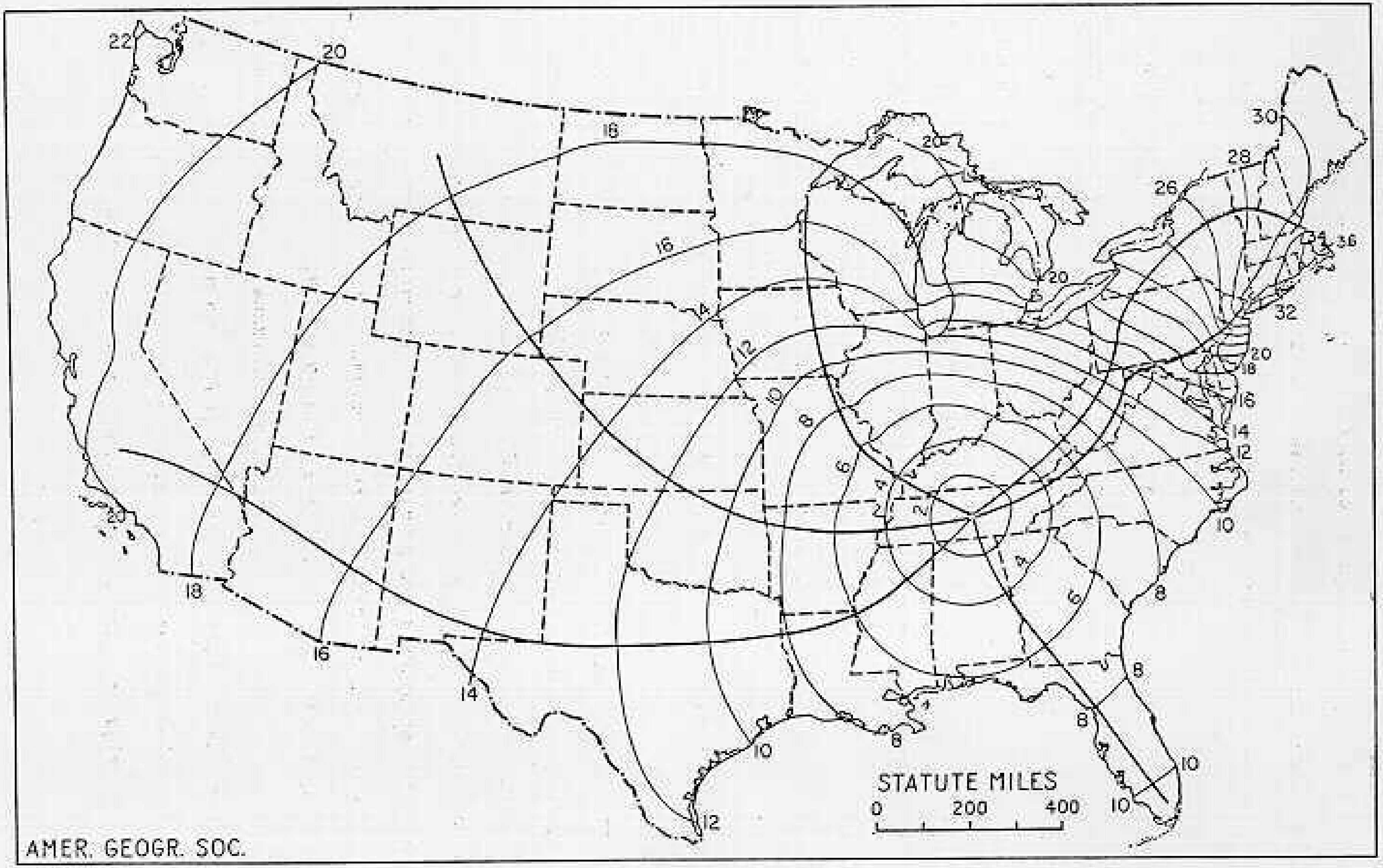


FIGURE 5

The procedure used to achieve the surfaces represented in figures 4 and 5 involved a decision as to the finite width assumed for the minimum cost routes to be shown. This is indicated in the legend of figure 6 showing the graphical method employed to determine these surfaces and paths.

Many other examples might be given in addition to the three kinds given above, but the distance surfaces, time surfaces, and cost surfaces that these represent are indicative of the wide application of the basic ideas of surfaces and paths in geography generally. This, of course, includes a part of the general theory of geodesics, and intensive study and the literal application of it to the real earth and especially the various conceptual surfaces covering it seems called for at this stage in the development of geography and its growth toward a truly unified discipline.

We have noted that mathematical and graphical solutions exist for our kinds of problems. In some cases, the mathematics is intractable, however, and graphical solutions, or rather approximations, do not exist or have not yet been invented.

A third kind of solution exists as an extension of the graphical method and that is to build three dimensional models of the various surfaces. For example, one could make a model of the equatorial mercator projection of the iso-distance surface based on London. A convenient scale could be established between values on the distance surface and height above the plane on which London was regarded to exist. Thus London would become the pit with all others points elevated above it and with heights increasing in a monotonic and continuous fashion outwards until the one peak on the surface, the antipodal point for London was reached.

One could make this into an operational model, indeed a geographical analogue computer, for finding the great circle path on the surface from any point through London by allowing a ball to roll freely on the surface from that point. It will roll "down hill" to the pit of London along the gradient path. A difficulty exists

GRAPHICAL METHOD FOR DETERMINING COST SURFACE ABOUT A POINT

S AND S' REPRESENT SUCCESSIVE ISOCOST LINES; $r_1 r_2 r_3$ = DISTANCE RADII ACHIEVABLE FOR OUTLAY OF $100,000 ASSUMING THAT 15 ACRES OF LAND ARE REQUIRED FOR EACH MILE OF HIGHWAY LENGTH

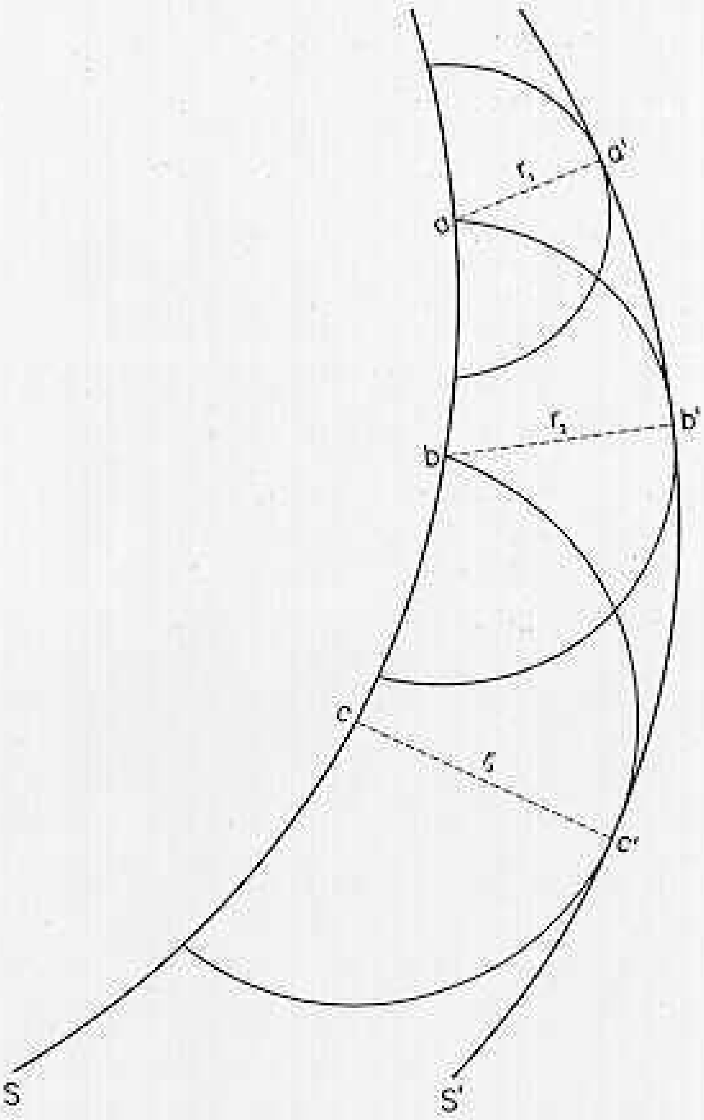


FIGURE 6

because the ball has mass and gained momentum may cause it to depart from the locally steepest slope at some places. However, if the model's vertical scale is selected so that no slopes are too steep the problem is minimized. Introducing additional friction to the surface could help also.

Boundary problems exist also. The interruption of the map along a meridian is easily overcome by making the model periodic in an east-west direction. The fact that the geographical poles lie at infinity poses another less easily solved problem.

Models for the other phenomena presented in this paper are possible too. Thus, elevation on the surface could be made proportional to time from New York, or cost from Lewistown, etc. Our geodesic-finding ball would approximate the required paths.

Geography, geometry, and graphics which had their first grand synthesis in cartography at the time of Ptolemy stand to benefit mutually in the recently established cooperation which greatly extends and intensifies those early benefits. This time spatial patterns in general are being considered with regard to social, economic, and cultural phenomena as well as physical phenomena with recognition that the <u>geo</u> in geometry and the <u>geo</u> in geography have more in common than we would have dared to dream a decade ago.

REFERENCES AND ACKNOWLEDGEMENTS

Part of this paper was presented to the International Geographical Congress, London, July 1964. The topics in it were also presented to the faculty and students of the Michigan Inter-University Community of Mathematical Geographers in November 1964. The author is deeply indebted to that group for their incisive comments and recommendations.

All figures have been drawn by the cartographic staff of the American Geographical Society. Figures 1, 1A, 4, and 5 appear here for the first time. The others have appeared previously in the Geographical Review and are reproduced here with permission.

Thanks are owed to the Editor of the Geographical Review for allowing the author to quote passages from his article, "Transatlantic Flights and Pressure Patterns," from Vol. 51, No. 2, 1961, pp. 187-212 of that journal.

John Z. Stewart in "The Use and Abuse of Map Projections," Geographical Review, Vol. 33, No. 4, 1943, pp. 589-604, pointed out the analogy between the inverse scale on any conformal map and the index of refraction in geometrical optics.

Various staff members at the American Geographical Society have advised the author. In particular, O. M. Miller, the assistant director of the Society, offered valuable advice.

At Princeton University Professor Steve Sláby and Mr. C. Ernesto Lindgren of the School of Engineering's Department of Graphics and Engineering Drawing have provided many enlightening comments. They are much interested in the modern revival of the cooperative assistance of geometry and geography and especially of the role that graphics plays in this.

Also at Princeton University, the sculptor in residence, Joe Brown,

advised the author and demonstrated for him certain of the techniques of three-dimensional model building. It was in Mr. Brown's studio and with his advice that the author completed the three-dimensional model of the 1960 Potential of Population Surface for the United States. This model is currently included in the American Geographical Society's exhibit at the New York World's Fair.

The gentlemen cited above must in no way be held accountable for the defects of this paper, however. Those came from the mind of the author alone. But, the author is deeply indebted to these many gentlemen and would like to make known his gratitude to them.

Forward

Stig Nordbeck was Visiting Professor of Geography at Wayne State University during the spring of 1964. His series of lectures at the Michigan Inter-University Community of Mathematical Geographers Seminar in Brighton, Michigan was most stimulating and we encouraged him to expand this portion of his remarks so that the world community of mathematical geographers would also have their benefit.

Those who had the good fortune to enjoy the Nordbeck lectures first hand had the opportunity to be struck anew with the benefits of international exchange. Here was a man from the Swedish school who thought on parallel yet separate tracks to the Americans at Brighton. The process was mutually stimulating.

The flavor of the Brighton Seminars is that the work itself dominates. The production of this series of discussion papers has no schedule whatsoever, as those who receive it are amply aware. We produce a paper when we feel we have something to say of interest to you. Our seminars are most informal and student and faculty distinctions are kept at a minimum. Intellectual criticism can come strongly from any quarter.

This paper by Stig Nordbeck is in this intellectual tradition. Nothing is sacred to Nordbeck's mind. If he finds his geographic law fits rivers and cities and volcanoes, he says so; indeed, he proves so, even in the face of the fact that every university catalogue "proves" one cannot mix the physical and social sciences. What are the pedagogical and philosophical implications of Nordbeck's work? We do not fully know; but they will be disruptive and, therefore, criticized. If Galileo cried to his critics, "But it moves!" Nordbeck might cry to the critics of the spatial patterns he has discovered on diverse maps, "But they fit!"

William Bunge
Wayne State University
June, 1965

The Law of Allometric Growth

by

Stig Nordbeck

The law of allometric growth was originally discovered by biologists. It states that the rate of relative growth of an organ is a constant fraction of the rate of relative growth of the total organism. Assume that y is the size of the organ and x that of the organism. Then the allometric growth law can be written as (1):

$$y = ax^b \qquad (1)$$

where a and b are constants (a is always positive). The formula (1) is derived by means of calculus as follows: It is known that the rate of growth of an organ at time t is $\frac{dy}{dt}$, where y is the size of the organ. The rate of relative growth thus is $\frac{dy}{dt}/y$ for the organ and $\frac{dx}{dt}/x$ for the organism. The principle of allometric growth now gives formula (2):

$$\frac{dy}{dt}/y = b\frac{dx}{dt}/x \qquad (2)$$

That formula (2) is equal to (3) is easily seen by substituting $\frac{dx}{dt}.\frac{dt}{dy}$ in (d) for $\frac{dx}{dy}$,

$$\frac{1}{y} = b\frac{dx}{dy}/x \qquad (3)$$

Integrating (3) gives equation (4), where log a is a constant integration.

$$\log y = \log a + b\cdot\log x \qquad (4)$$

That this equation (4) is identical to formula (1) may be proven by taking logarithms of (1).

The law of allometric growth is valid for phenomena other than organs and organisms. Beckman (1958), for instance, assumed that the total urban population in a country was an "organism" and the cities its "organs." These assumptions allowed him to verify the "rank size rule" for cities. This rule states that the size of city number n is approximately one <u>nth</u> of the size of the largest city. In other words: The rank-number n of the city times its size (population) is constant. This empirical law was first observed by Auerbach (1913) and Zipf (1941).

Horton (1945) and Strahler (1954) introduced another such system in their analysis of drainage basins. The smallest "finger-tip" tributaries are designated as order 1. Where two channels of order 1 join, a stream segment of order 2 is formed, where two of order 2 join, a segment of order 3 is formed, and so on (See figure 1). Horton's law of stream numbers in a drainage basin states that the number of stream segments of each order forms a geometric series. There is 1 segment of the highest order, s of the next highest order, s^2 of the next order and so on. This bifurcation ratio rule for streams is similar to the rank size rule for cities. According to Beckman's model, built upon (among others) Lösch's theory of location, cities (when divided into different size classes) show the following pattern: There is one largest city, there are s cities of the next largest size, s^2 cities in next class, and so on. The bifurcation ratio rule can also be derived by the law of allometric growth as the rank size rule for cities.

In order to make it more general and easier to use, the law of allometric growth will here be formulated in two other ways. The first new formulation is the following. If the growth of an individual follows the law of allometric growth according to formula (1), it is possible to

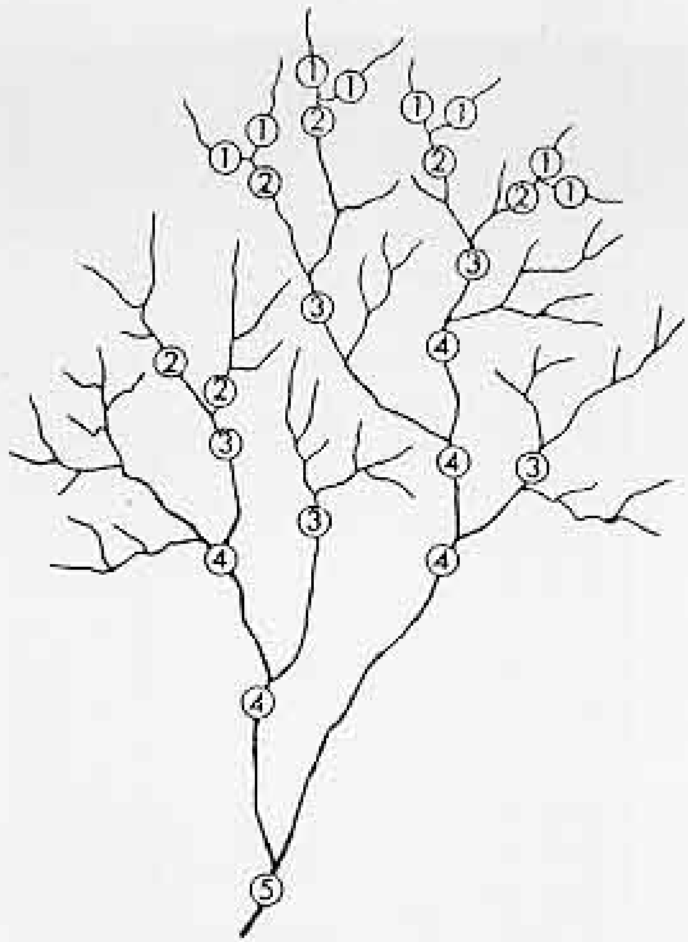


Figure 1
An example of stream branch orders as defined by Strahler and Horton.

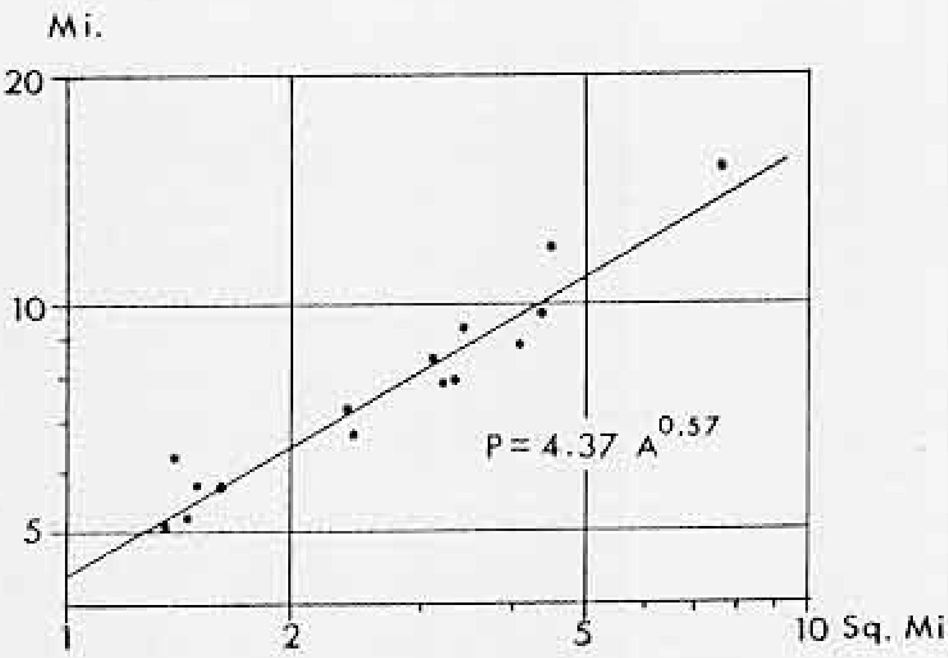


Figure 2
The relationship between perimeter and river length of drainage areas in fourth order basins of the Allegheny River.

measure the value of one variable at different times and to estimate the values of another variable by means of the measured values. This rule is used by parents when they measure their two-year old child's length and say that the length of the grown-up child shall be twice this length. The old Swedish farmers use this rule too, when they measure the body perimeters of their pigs or cattle. Knowing the length of the perimeter, a farmer can estimate the slaughter-weight of a pig with extreme exactitude. As a matter of fact such an estimation, based on measuring the perimeter of an animal, is much more accurate than an estimation based on the weight of the pig (cow) alive. A pig (cow) has the same (breast) perimeter when it is hungry and thirsty as when it is satisfied. Since the price of pork is highest for a special weight class, farmers try to slaughter their pigs when they can get the highest profits possible. It is therefore very important for them to calculate the slaughter-weight correctly. This estimation must be done 2-3 weeks before the slaughtering. The highest paid weight class varies in such a way that it is greater when there is a shortage of pork in Sweden than when there is a surplus. But the farmers, knowing the law of allometric growth, are able to allow for this changing weight standard.

The second new formulation of the law of allometric growth is the following: Instead of assuming the measurement of a growing individual at different times, it is assumed that a series of individuals all have the same shape but are of different size. In this case the law of allometric growth states that it is possible to estimate the values of a variable by means of the measured values of another variable. This rule makes it possible to measure a variable which is very difficult to measure. It is used by foresters when they first classify the trees they are going to cut down as very high trees, high trees, mean high trees etc. and then measure the diameter of each tree and use these values to estimate the total volume of the trees. It is also used by the aforementioned farmers. All animals do not have exactly the same shape. A pig can be very short compared with the mean pig. In this case the farmer takes this into consideration and allows the perimeters to become a little larger than in the mean case.

The law of allometric growth, and especially the version of it which is based on the same shape of a series of individuals, will here be applied to some geographical objects. Some simple examples and illustrations of this principle are given in table 1. In this table S means the shape of the individuals, M is the measured variable, DM is the dimension of M, L is a length variable, A is an area variable and V is a volume variable.

Rivers and the law of allometric growth.

All drainage areas ought to have the same "shape." The same forces are always at work when water runs from the land to the oceans. These forces are independent of the ground over which the water passes, but the drainage area will become larger if the ground is soft and smaller if it is hard. The "shape" of the drainage basins is, however, the same. Hence the law of allometric growth is valid for drainage basins. Strahler (1957) writes, that studies of actual drainage basins in homogeneous rock masses show that geometrical similarity is closely approximated when mean values are considered. He gives the term "geometrical similarity" the same meaning which here has been given to the expression "same shape." If the ground is geologically not homogeneous, the shape of drainage areas probably will be deformed. But this deformation follows allometric growth

S	M	DM	L	A	V
square	side = x	1	diagonal = d $d = \sqrt{2} \cdot x$	area = A $A = x^2$	-
cube	area = x	2	side = s $s = \sqrt{\frac{x}{6}}$	$A = x$	volume = V $V = \frac{x\sqrt{x}}{6\sqrt{6}}$
circle	radius = x	1	perimeter = p $p = 2\Pi x$	$A = \Pi \cdot x^2$	-
sphere	radius = x	1	$p = 2\Pi x$	$A = 4\Pi \cdot x^2$	$V = \frac{4}{3}\Pi x^3$
sphere	volume = x	3	radius = r $r = \sqrt[3]{\frac{3x}{4\Pi}}$	$A = \sqrt[3]{36\Pi x^{\frac{2}{3}}}$	$V = x$
hexagon	side = x	1	$p = 6x$	$A = \frac{3\sqrt{3}x^2}{2}$	-
cube	volume = x	3	$s = \sqrt[3]{x} = x^{0.33}$	$A = \sqrt[3]{x^2} = x^{0.67}$	$V = x$

Table 1. If the dimension of the measured value x is 1, the length variable $L(x) = a \cdot x$, the area variable $A(x) = a \cdot x^2$ and the volume variable $V(x) = a \cdot x^3$ where a is different constants. If the dimension of x is 2, $L(x) = a \cdot x^{0.5}$, $A(x) = a \cdot x$ and $V(x) = a \cdot x^{1.5}$. If the dimension of x is 3, $L(x) = a \cdot x^{0.333...}$, $A(x) = a \cdot x^{0.667}$ and $V(x) = a \cdot x$.

rules too, and the law of allometric growth is still valid. This is illustrated by the following examples.

The perimeter P of a drainage area is an one-dimensional variable, and its area A is a two-dimensional one. Thus the equation $P = aA^b$ is valid. This equation is equal to the linear equation $\log P = \log a + b \log A$, where b ought to be very close to 0.5. Morisawa (1959) gives values of area and perimeter of 16 fourth-order basins in the Allegheny River watershed. These data are plotted on a log-log diagram (figure 2), and the line $\log P = \log a + b \log A$ is calculated by means of the least squares smoothing technique. The resulting line $P = 4.37A^{0.57}$ is drawn in the diagram. The b-value was expected to be very close to 0.5, and the received b-value 0.57 is therefore too high. This can be explained in the following way: The short perimeters are too short, probably due to the method of measurement. They were all measured on maps of scale 1:24 000 and these maps showed, of course, the same degree of generalization for all basins, independent of their size, which caused the perimeters of the small basins to appear too short. More precise measurements would probably give a somewhat lower b-value (closer to 0.5). If the two largest basins and the second smallest one are excluded the b-value also is quite close to the theoretical value 0.5. These three basins give also a b-value equal to 0.5 if their data are smoothed by the least squares method. Hence, the data in figure 2 can belong to two different sets of basins, but the very low number of observations makes it impossible to find out if there are one or two sets of basin types. The correlation between log P and log A is very high. The correlation coefficient r is equal to 0.9642.

The length L of a drainage area can be defined as the length of the main river (the longest river) from its smallest source-stream to its mouth. It is easier to measure this length L correctly than to measure the perimeter P of a drainage basin. It follows then that the correlation between log L and log A will be higher than the correlation between log P and log A. The calculated b-value will not differ so much from the theoretical b-value 0.5 as it did in the example in figure 2.

In figure 3 the lengths of 91 rivers were plotted against the areas of their drainage basins. These rivers are of quite different size from the Amazon river which has the largest drainage area (7 050 000 sq.km) to a small Swedish stream (470 sq.km) flowing from the South Swedish highland (a tributary of the Lagan river). All Swedish rivers which were recognized on a map of scale 1: 2 000 000 were used in the diagram if it was possible to find data for them, a total of 49 rivers. The remaining 42 rivers are all from other continents. They were chosen by a method similar to that used for the Swedish rivers. No classification of the rivers was made before they were plotted in the diagram, and no river was excluded because it did not fit the curve.

The least squares method applied to data in figure 3 gives the equation $A = 0.1039L^{2.0009}$. The area A is given in sq.km. and the length L in km. The computed b-value (2.0009) is as close to the theoretical b-value (2) as could be expected. Consequently the correlation between log A and log L is very high. The correlation coefficient r is 0.9905.

By means of the residual method, the rivers of the diagram in figure 3 can be divided into three groups. A plain river has a large drainage area compared with its length. The Amazon river and the Swedish rivers Lidan and Fyrisån belong to this group. A valley river has a small drainage area compared with its length. The Mekong river and the Dnjestr river belong to this group. Most of the rivers in the example, figure 3, belong to the

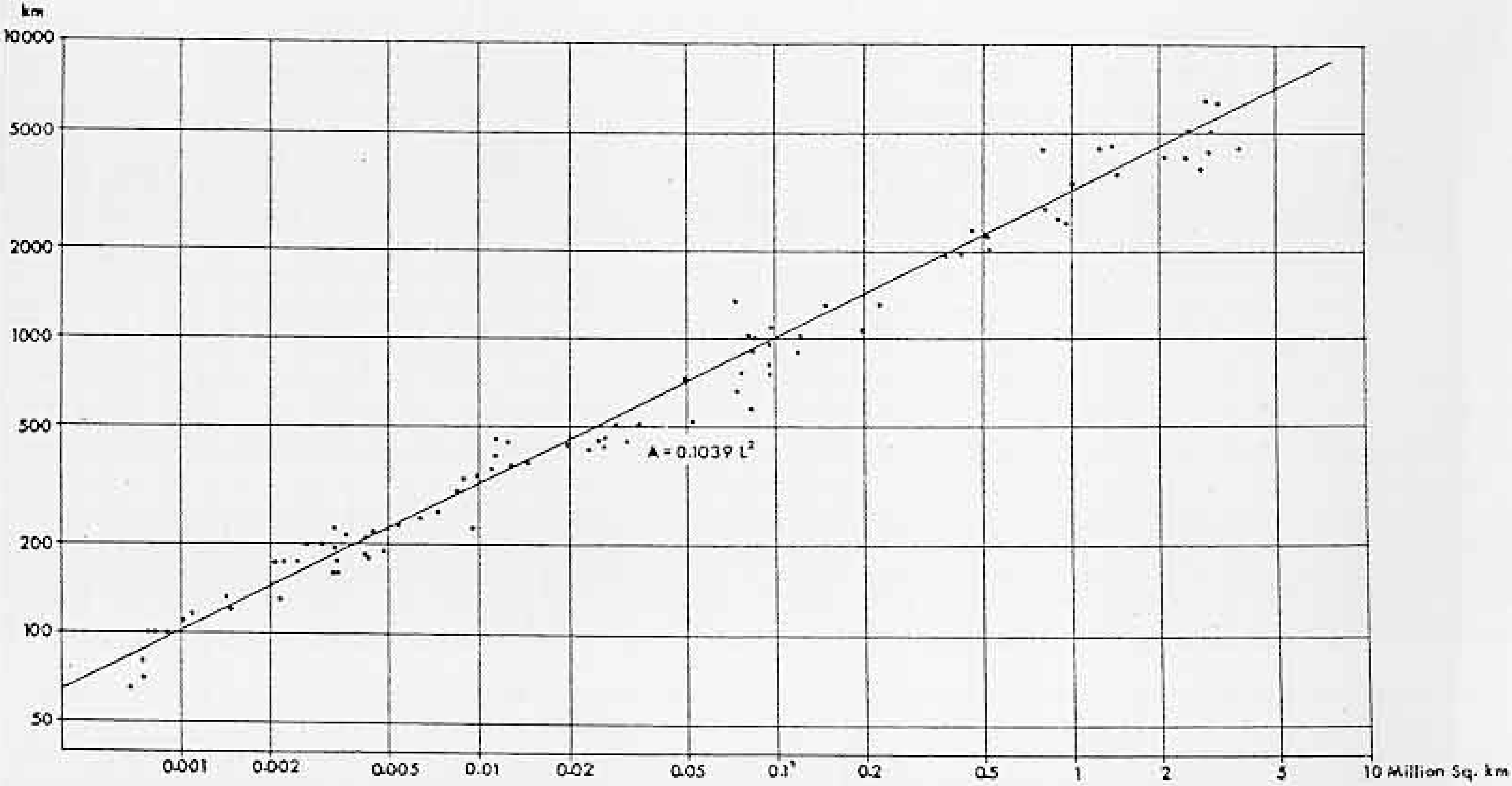


Figure 3
The relationship between river length, L, and drainage area (km^2).

third group, the mean rivers.

In the example above, the area of the drainage basins has been written as a function of the basin length. It is also possible to write the length as a function of the area. In this case the least squares method and the data in figure 3 give the formula $L = 3.44A^{0.49}$.

Allometric growth and meanders.

Leopold (1964) writes that nearly all natural channels are sinous to some extent. Not only do they exhibit a more or less regular aspect of sinousity, but the size of the curves assumed by a channel bears a constant relationship to the channel itself. Small channels wind in small curves and large channels in large curves. The meander length or wave length L is generally proportional to channel width W and so is the radius (r_m) of the curves which the channel exhibits.

All flowing water meanders. The melt-water running on glacier ice, and the water in the Gulf Stream both show the same meandering pattern as the water in rivers, in spite of the absence of sediment load and river banks. The nature of meanders and meandering forces make it very likely that all meanders have the same shape. Consequently, the law of allometric growth is valid for meanders. Figure 4 is a revised version of a figure in Leopold (1964). In this figure the meander length L in feet is plotted on a log-log diagram against the channel width W, in feet. Both L and W are one-dimensional variables. It follows then that the theoretical b-value in formula $L = aW^b$ is equal to 1. Leopold (1964) found the b-value to be 1.01 and the a-value 10.9, as is seen in figure 4. If the mean radius of curvature r_m is plotted against the meander length L, the b-value is 0.98(figure 5). In both these cases the calculated b-value is so close to the theoretical b-value as to be considered equal. The line

$L = 10.9W^{1.01}$ is drawn in figure 4 and the line $L = 4.7r_m^{0.98}$ in figure 5.

Volcanoes and allometric growth.

The shape of a volcano depends on the materials of which it consists. A viscous lava forms very high cone volcanoes because the slope angle of such a lava is very high. The angle of slope of a recently formed cinder cone ranges between 26° and 30°. Strato-volcanoes are more extensive than cinder cones and may have slope angles as high as 35°, the natural angle of slope of volcanic ash, but their slopes generally range between 20° and 30°. The lava of shield volcanoes is highly fluid and travels far down the low slopes, which do not usually exceed 4° or 5° (Strahler 1960). Hence, the shape of volcanoes can be considered a transportation problem. This can be compared with an economic application: The rings constructed by von Thunen. The shape of a volcano can be derived in the same way, using the same technique that he did when he got his rings.

All volcanoes belonging to the same class can be assumed to have the same shape. Consequently the law of allometric growth is valid for such volcanoes. In figure 6 the height H in meters of six volcanoes is plotted against their volume V in cubic meters. The least squares method gives the formula $H = 0.102V^{0.3946}$. The height H is an one-dimensional variable and the volume V a three-dimensional one. It follows then that the expected b-value is 0.3333... There are at least three different explanations for the fact that the calculated b-value for volcanoes differs so much from the theoretical b-value.

1. It is here assumed that the length of a length variable $L(x)$ can be calculated by means of the formula $L(x) = ax$ and an area variable by aid of $A(x) = ax^3$. This is true only in such cases where the height $H(x)$ is a length variable and thus the formula $H(x) = ax$ is valid. If the

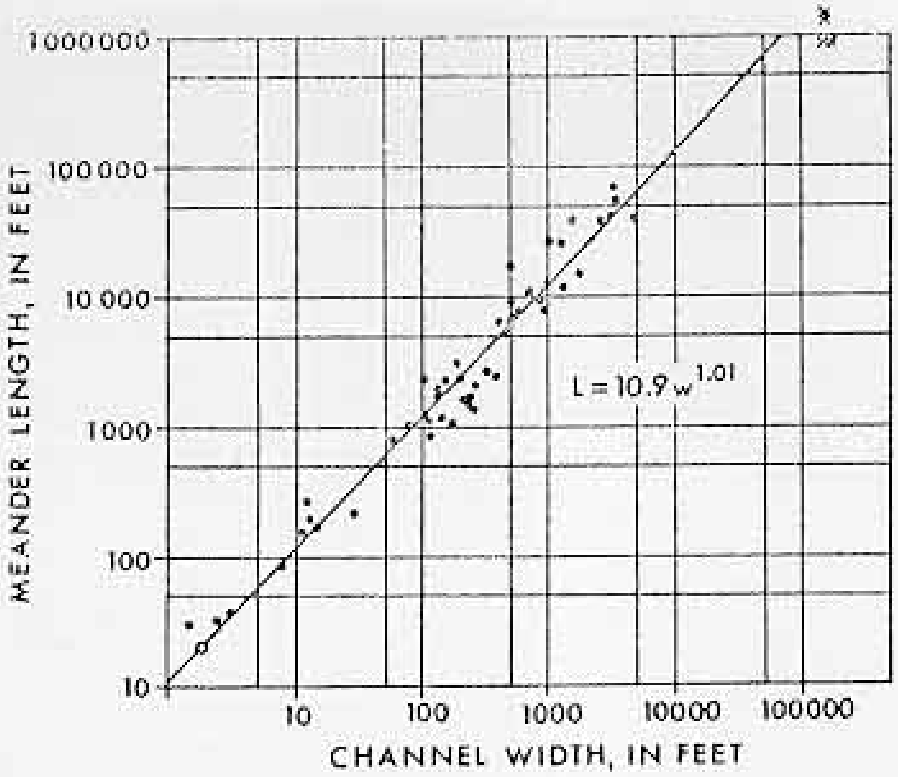


Figure 4

The relationship between meander length, L, and meander width, w.

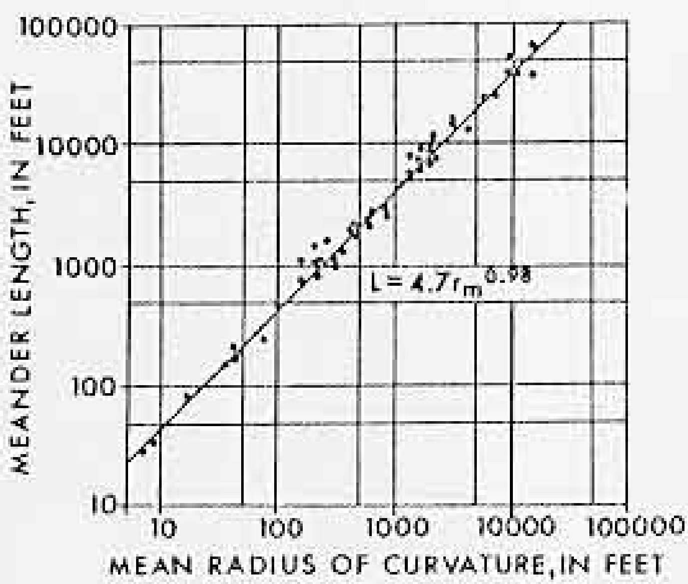


Figure 5

The relationship between meander length and the radius of curvature in channels.

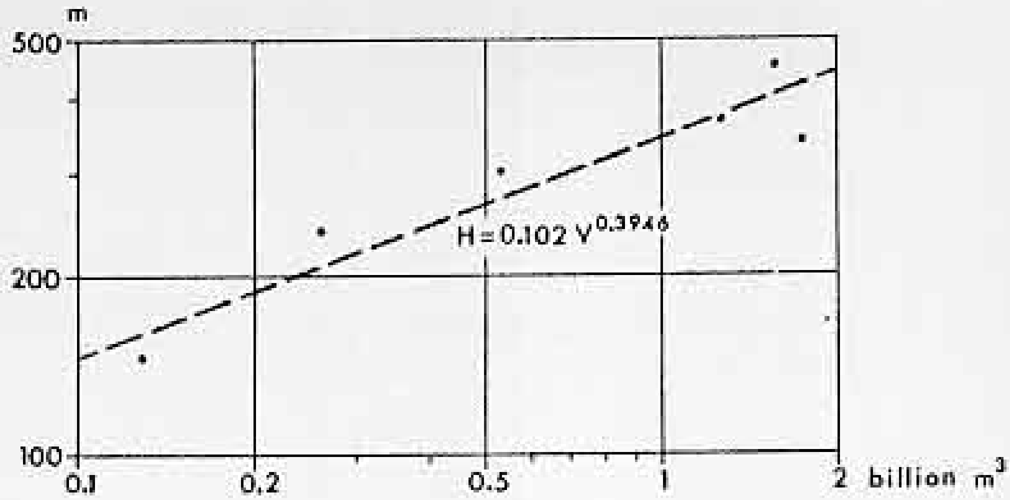


Figure 6

The relationship between vulcano height (meters) and vulcano volume (meters^3).

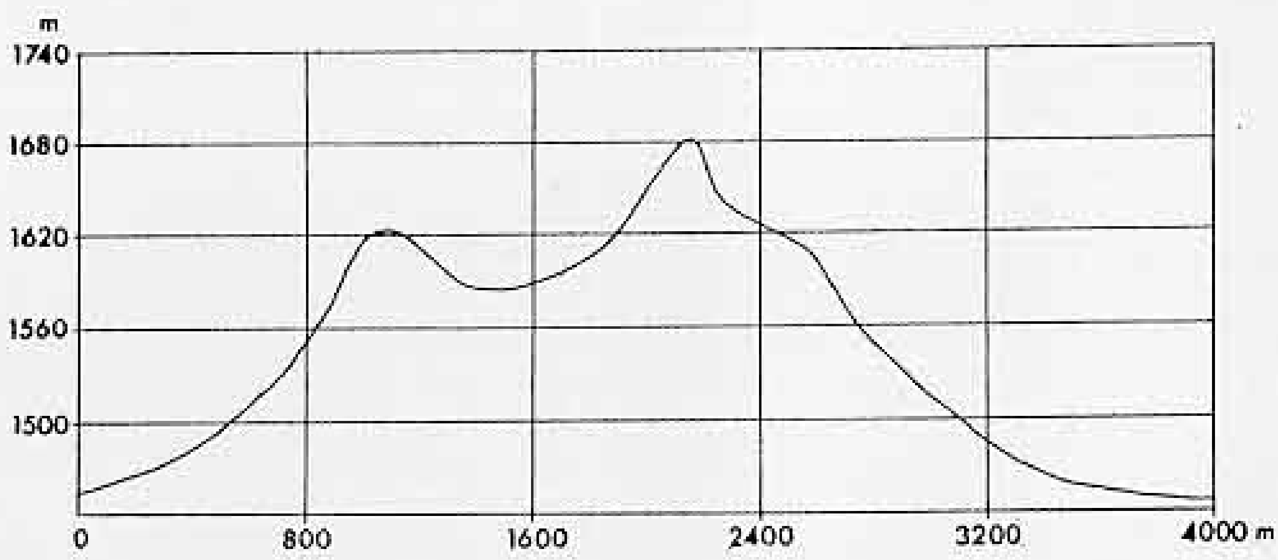


Figure 7

A section through Butte I, a vulcano in Idaho.

equation $H(x) = ax^{0.5}$ is valid the volume formula is changed to $V(x) = ax^{2.5}$. Hence, $L(x) = aV^{0.4}$, where V is the volume of the volcano. This b-value is equal to the calculated b-value in the example (figure 6), but there are no reasons why the height of a volcano should not be a length variable. Thus it is not true that the height formula $H(x) = ax^{0.5}$ is valid. There must be other reasons by which the over-estimated b-value can be explained.

2. Errors in the measurement of the height, area, or volume of the volcanoes. It is impossible to define a volcano correctly. Where does the volcano start? These errors are of the same kind, and independent of the size of the measured volcano. All measured values are probably too low but the errors of the lowest values are greater than those of the higher ones, if these errors are measured in per cent of the actual value.

3. The volcanoes examined were old volcanoes. Figure 7 shows a profile of one of them: Butte I of Menon Buttes, Idaho, U.S.A. Figure 8 is a profile of Butte II. These two old volcanoes are considered as having the same shape. This was true when they were young. But erosion has changed their shape. This erosional wane is not allometric but arithmetical. That means that the correct value can be calculated by the formula $y = e + cx$, where both y and x are one-dimensional variables and e is the erosion constant. The height of one volcano is H_1 and of the other H_2. The two volcanoes are of the same age and the erosion has been the same in both cases. Thus the correct heights are $H_1 + e$ and $H_2 + e$, where e is the erosion constant. Assume that H_1 is x and H_2 is cx, where c is greater than 1. Both H_1 and H_2 are too low, but the relative error of H_1 is greater than the relative error of H_2. Erosion has also made the volumes of the volcanoes too low. However, the relative error of the volume is not as

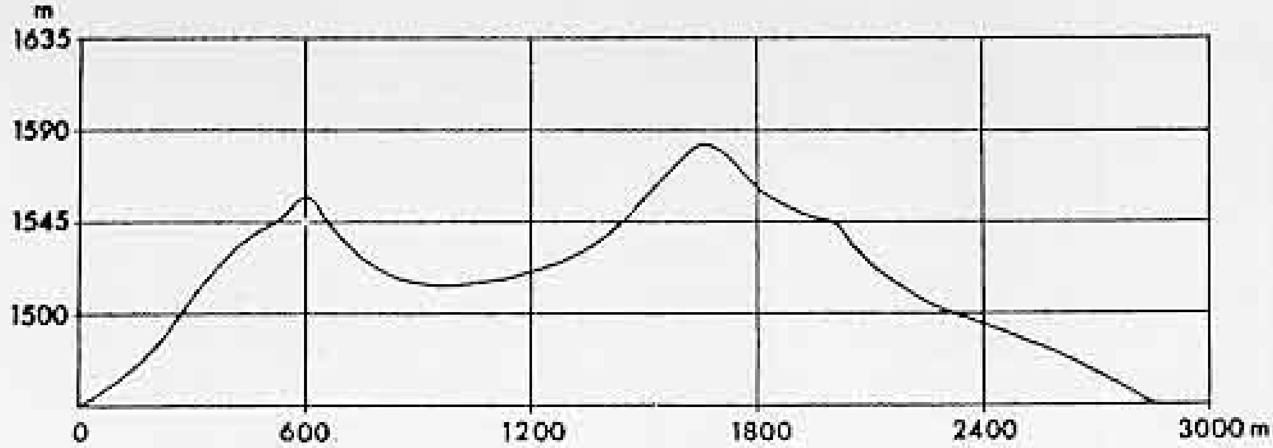


Figure 8
A section through Butte II, a vulcano in Idaho.

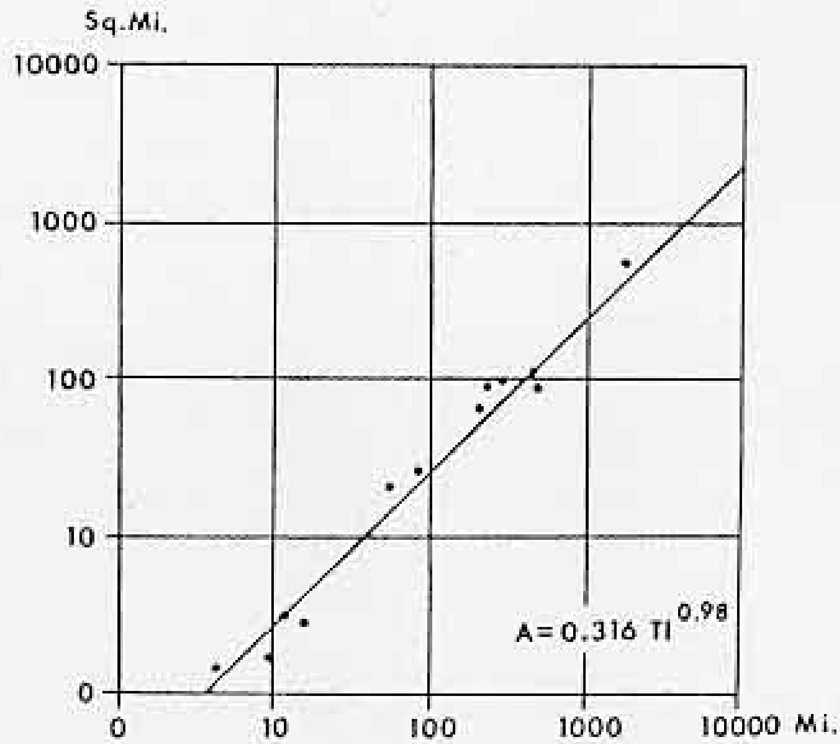


Figure 9
The total stream length, T_1, is a two-dimensional variable. Hence, the equation $A=aT_1$, where A is the area of a drainage basin, is valid. The least square method gives the equation $A=0.316T_1^{0.98}$ for this example.

great as the relative error of the height. It is only the transportation of materials away from the volcano that changes its volume. The transportation of materials from the top of the volcano to its base does not change its volume, but changes its height. Compared with the relative denudation of heights, the relative denudation of the volumes is low and can be considered equal to zero. Hence, the effects of erosion are that both the height and the volume of an old volcano are too low and that the relative error of the height of a small volcano is much larger than that of a bigger one. Consequently the b-value of the formula $H = ay^b$ will be higher than 0.333... which is the theoretical b-value if H and V are correctly measured and if the erosion has not changed these variables H and V. The correlation between log H and log V in the example, figure 6, is quite high. Thus the correlation coefficient r is equal to 0.9717. The line $H = 0.102V^{0.3946}$ is drawn in figure 6.

Recognition of dimensionality.

It is sometimes difficult to recognize dimensionality of the measured or of the estimated variable. The perimeter of a drainage basin was an one-dimensional variable (figure 2) and so was the length of the main river (figure 3). The total stream length of drainage areas is often erroneously recognized as an one-dimensional variable, too. An area has the dimensionality 2, a line 1 and a point 0. The total length is a sum of lines. However, it is not a length variable but an area variable as can be seen by the following examples. The size of an area A can be measured by aid of points arranged systematically or at random and by aid of a reference area B. The size of B is known, and so is the number of points which would belong to B if these points were arranged in exact same way as those belonging to A are. This number is here called N_B. The number of points belonging to A is N_A. It is possible to estimate the size of area A by means of the formula $A/B = N_A/N_B$ where A is the unknown size of A and B the known size of B. The error of this estimation depends upon N_A in such a way that increasing N_A will give a decreasing error. Consequently, the error of the estimation of the size of an area A by the aid of the formula $A = BN_A/N_B$ is very small if N_A is large enough. Hence, N_A is a two-dimensional variable. The number of street intersections in a city as a measurement of the area of that city is one application of this statement. Another one is the number of points in a drainage basin at which two streams join.

The size of an area A can also be measured by means of randomly or systematically arranged lines. These lines cover a reference area B, the size of which is known (area A usually belongs to area B). The total length of the lines which belong to A is called T_A and the total length of the lines belonging to B is named T_B. If T_A is large enough, the area of A can be calculated by aid of the formula $A = BT_A/T_B$. Consequently, the total length T_A is a two-dimensional variable. The total length of the streets in a city can be used as measurements of its area, since they are organized in a regular system. The rivers and streams of a drainage area also form a regular system. Hence, the total stream length is a two-dimensional variable.

In figure 9 the areas of some drainage areas are plotted against the total length of corresponding streams. The sizes of the areas are called A and the total lengths of the streams are called T_1. The least squares smoothing technique gives the equation $A = 0.316T_1^{0.98}$; the line is drawn in figure 9. This figure is a revised version of a figure in Morisawa (1959). Since A and T_1 both are two-dimensional variables the expected b-value in the formula $A = aT_1^b$ is equal to 1. The calculated b-value is equal to 0.98 and this is so close to 1 that the two b-values can be considered equal.

Densely populated areas and allometric growth.

It is very easy to get data about the population of built-up areas in Sweden. These data are given in the official statistics of Sweden if the population of the built-up area is larger or equal to 200 persons. But it is very difficult or impossible to get data about the area of the built-up areas. It would therefore be very excellent if the unknown size of a built-up area could be estimated by aid of the known population.

The population of a built-up area depends on three variables: the size of the area which is a two-dimensional variable (length and width) and the population density which is defined in the following way: A reference area of specified shape and size is moved over the built-up area and the population density of a point (x,y) is the number of persons belonging to the reference area that has this point as its central point (Nordbeck, 1964). The distribution of population is a two-dimensional distribution and the density function is the frequency function of this distribution. It is very easy to show that all density functions thusly defined are continuous by integration (double integrals) (Nordbeck, 1965). This verification of the continuity of the density functions is easiest to do if the reference area has a regular shape and if its size is constant. In this sense a square with its sides parallel to the axes of the coordinate system is the best reference area.

A three-dimensional diagram such as an isarithmic map of a built-up area is quite small compared with the built-up area. Such a three-dimensional diagram corresponds closely to a contour map of a volcano.

There is always a downtown with quite low population densities. This is a result of the very high rents in the downtown area. Only offices, stores, etc., can afford these high rents. In very small built-up areas the "downtown" area consists of a town square, a church, and some very small stores, or of a railway-station and stores, etc. Churches, squares, and so on, belong to these downtown areas because they were there first and not because they can pay high rents. The "uptown" has the highest population densities. These are the slum or apartment areas as in Sweden. The rents in such areas are lower than in the downtown area. The buildings can therefore be used as dwellings. Increasing transportation costs cause the population densities to decrease with increasing distance from the uptown area until the suburban areas grade into rural land use. If a profile is drawn along a line through a built-up area using the population density as dependent variable this profile will be quite similar to those of the volcanoes in figures 5 and 6.

The population of an administrative unit consisting of an old city decreases as the population of the total built-up area grows (Kant 1962, Clark 1954, Berry 1958). This was observed for Amsterdam as early as the 16th century (Dickinson 1961). This decreasing population is explained by the fact that the growing downtown with low population density has pushed the uptown with high density to the outside of the boundaries of the administrative unit.

The law of allometric growth is valid for built-up areas if it can be assumed that they all have the same "shape." Consequently it would be possible to estimate the size of a built-up area of known population using the well-known technique of measuring one variable and estimating another variable.

The population of Swedish towns are shown plotted against their areas (figure 10). These data refer to administrative units. In most cases both the population and the size of such a unit are different from the population and the size of the corresponding built-up area. This deviation between the data used in figure 10 and the correct data for built-up areas is so large that it is impossible to use them in the least squares method or in any other smoothing technique to determine the a- and b-values of the formula $A = aP^b$, where A is the size of the built-up area and P its population. Most of the areas of the administrative units are too high compared with the size of the built-up areas. The biggest town in the world is Kiruna (13,181 sq.km.) but the built-up area of Kiruna is less than 20 sq.km. On the other hand, the area of the administrative unit can be smaller than the built-up area. This is valid for the Stockholm area, Gothenburg area, and for the surroundings of Malmoe. Consequently, the data of figure 10 can not be used when the line $A = aP^b$ is calculated by means of an objective smoothing method. However, this line $A = aP^b$ can be determined by the aid of a subjective method. Such a subjective method uses only the data for those administrative units which coincide with their built-up areas. This method gives the line $A = 0.03P^{2/3}$. It can be assumed that all built-up areas are circles or can be approximated by circles. It follows then that the radius of such a circle can be estimated by means of the known population, since the area $A = \pi r^2$ where r is the length of the radius. The line $r = 0.1P^{0.333...}$ is drawn in figure 10. Both of these equations $A = 0.03P^{0.666...}$ and $r = 0.1P^{0.333...}$ can be used as an approximation of the area A (in sq.km.) amd the radius r (in km.) of the built-up areas in Sweden. These estimations are not exact but they are much better than estimations based upon the areas of administrative units.

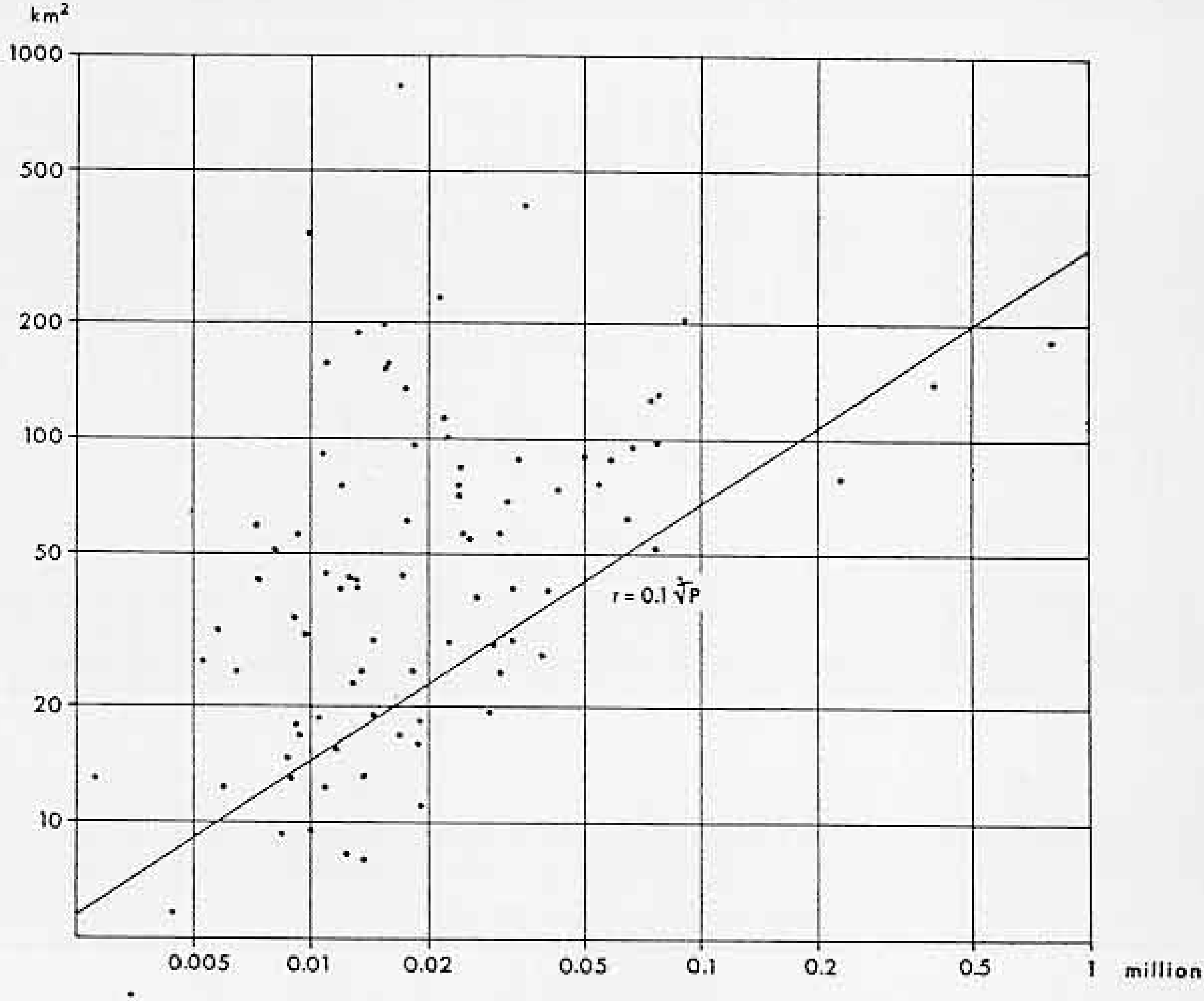


Figure 10
The relationship between area (km^2) and population of Swedish cities (administrative areas) in 1960.

The built-up area in Sweden (Swedish "tätort") is defined as an agglomeration of dwelling houses which are supposed to be connected by links. Each house of the "tätort" must have at least 1 such link the length of which is less than 200 meters. The population of all Swedish "tätorter" with more than 200 inhabitants is given in the Swedish census.

In the American statistics (Census, 1960) an urbanized area is defined as an area that contains at least one city of 50,000 or more inhabitants and the surrounding closely-settled incorporated places with more than 2 500 inhabitants, or with less than 2 500 inhabitants if such a place has a closely-settled area of 100 dwelling units or more. Enumeration districts in unincorporated territory with a population density of 1 000 inhabitants or more per square mile are also included in the urbanized area. Other enumeration districts with lower population density are also recognized as urbanized areas if including them eliminates enclaves or close indentations in the urbanized area (of 1 mile or less across the open end) or links outlying enumeration districts of qualifying density that were no more than 1.5 miles from the main body of the urbanized area. Counties which are classified as urban, towns in the New England States, and the townships in New Jersey and Pennsylvania are always included in the urbanized areas. This is a short summary of the definition of urbanized areas in the census in U.S.A. for 1960. In 1950 blocks were the smallest units used, instead of the larger enumeration districts. The 1950 criteria also permitted exclusion of portions of counties classified as urban, etc.

The definition of urbanized area in 1950 more closely approximated the actual area than did that of 1960 but there were significant differences between the urbanized area and the built-up area. The error resulting from the poor definition of urbanized area is a function of the perimeter of the area. Hence, it is an one-dimensional variable and the formula $e = ax$ is valid where e is the error and x a length variable. If the population P is a three dimensional variable the formula $P = cx^3$ is valid (c in the formula is a constant). In that case the relative error is equal to $(ax)/(cx^3) = dx^{-2}$; if x is very large the relative error is very small. The definition of urbanized area excludes some of the built-up area. The given data for the size of the urbanized areas is therefore a poor approximation of the size of the real built-up area. If the area is small the relative error will be too high and the area too low.

The population of built-up areas and the size of these areas for 1950 in the U.S.A. were plotted in a diagram with logarithmic scales (See figure 11). The least squares method gives the curve $A = 0.00126P^{0.86}$ where A is the area in square miles and P the population. If these areas are approximated with circles having the radius r in miles the following formula is received: $r = 0.020P^{0.43}$. The line $r = 0.020P^{43}$ is the dashed line in figure 11. The correlation between log A and log P is very high; the correlation coefficient r is equal to 0.9279.

Figure 12 shows the correlation between the area and the population of built-up areas in the U.S.A. for 1960. The correlation coefficient r in this case is equal to 0.9224. The least squares method gives the line $A = 0.00151P^{0.8757}$ or the line $r = 0.0219P^{0.44}$, which is the dashed line in figure 12 (r is the radius of a circle having the area A).

The solid line in figures 11 and 12 is the same line as was drawn in figure 10, that is the line $r = 0.1P^{0.333...}$ where r is in kilometers. The dashed line in figure 11 is $r = 0.030P^{0.43}$ km. and that of figure 12 is $r = 0.033P^{0.44}$ km. The dashed and the solid lines are not parallel. The b-values obtained are about 0.10 higher than the theoretical b-value 0.333... if the population density is a one-dimensional variable. This deviation can be explained in more than one way. The population density can be a variable with lower dimensionality than 1 and the poor definition of urbanized areas can give a higher b-value. It was pointed out when the volcanoes were discussed that a 0.5-dimensional height variable gave the theoretical b-value equal to 0.40. The population density ought to be a zero-variable. A man living in a big town needs the same area as a man living in a small town. However, figures 11 and 12 show that the dimensionality of the population density is closer to 1 than it is to 0.

The poor definition of urbanized areas has had the effect that the relative error of the smaller built-up areas was quite large, in such a way that the size of the area was too low. In 1950 the definition of built-up areas worked with blocks and in 1960 it used enumeration districts. This change in the definition of the urbanized areas had the effect of excluding some real built-up areas from the urbanized areas. In other words, the 1950 definition of an urbanized area more closely approximated the actual built-up areas than that of 1960. Consequently the b-value of 1950 is 0.01 lower than that of 1960. It is most likely that a perfect definition of urbanized areas would give a b-value which would be close to 0.333.... The population density of enumeration districts belonging to urbanized areas was 1000 inhabitants per square mile.

The b-value in the formula $r = aP^b$ will probably be higher than 0.44 if the definition of urbanized areas is changed in such a way that the lower limit of the population density of the enumeration districts is higher than 1000. Data for units designated "Densely-Inhabited Districts" are given in the 1960 census statistics of Japan. A densely-inhabited district is defined as an area consisting of contiguous enumeration districts with a population density of 4000 inhabitants or more per square kilometer. It is delineated within the boundary of city, town or village constituting an agglomeration of 5000 or more population as of October 1, 1959. The area and population of densely-inhabited districts in 1960 in Japan were plotted against each other in figure 13. The least squares method gives the equation $A = 0.000281P^{0.9146}$ sq.km., or $r = 0.00946P^{0.46}$ km. The correlation coefficient r is equal to 0.9665.

This definition of densely-inhabited areas means that large parts of the built-up areas are excluded. Consequently, the b-value will be higher than in the examples in figures 11 and 12. But the a-value is too low, too. However, as is seen of figures 11, 12, and 13, the cities in the U.S.A. are more spacious than the Japanese ones. This is explained by the existence of poor public transportation systems in the U.S.A. The American man has no choice. He must use a car in order to go from his home to downtown. The higher living standard in the U.S.A. can also explain this deviation. More people in the U.S.A. than in Japan can afford to live in suburbs and to have a car. It follows then that the population density of an American town is lower than that of a Japanese town.

The disadvantages of the approximation of built-up areas with urbanized areas and densely-inhabited districts can be eliminated in the following way. The size of most of these areas is too low. The areas are divided into different size classes. Let us now suppose that one area of each size class is equal to the corresponding built-up area. This area is the largest one of the size class compared with the population. In other words: The chosen area has the lowest population density of the areas

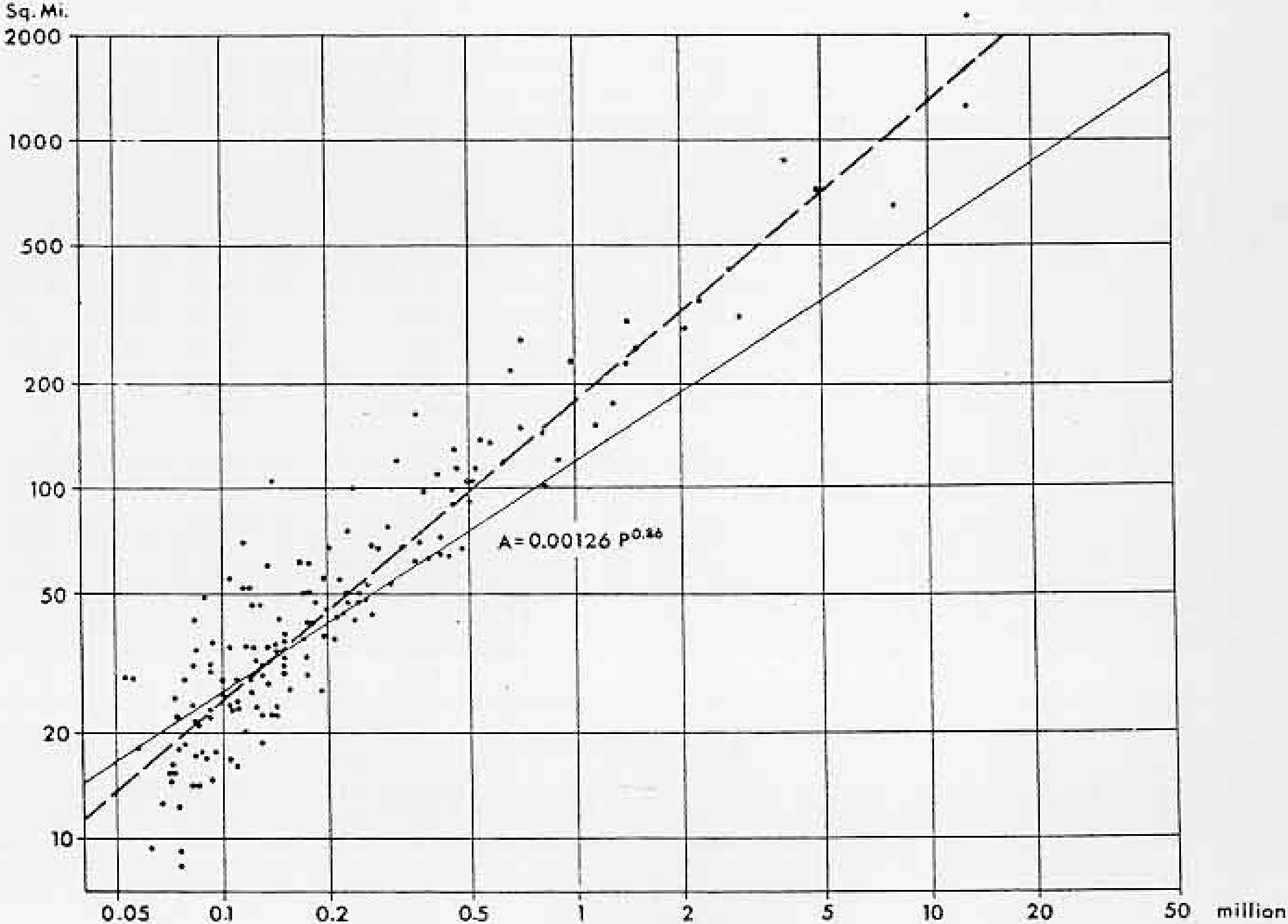


Figure 11

The relationship between area (miles^2) and population

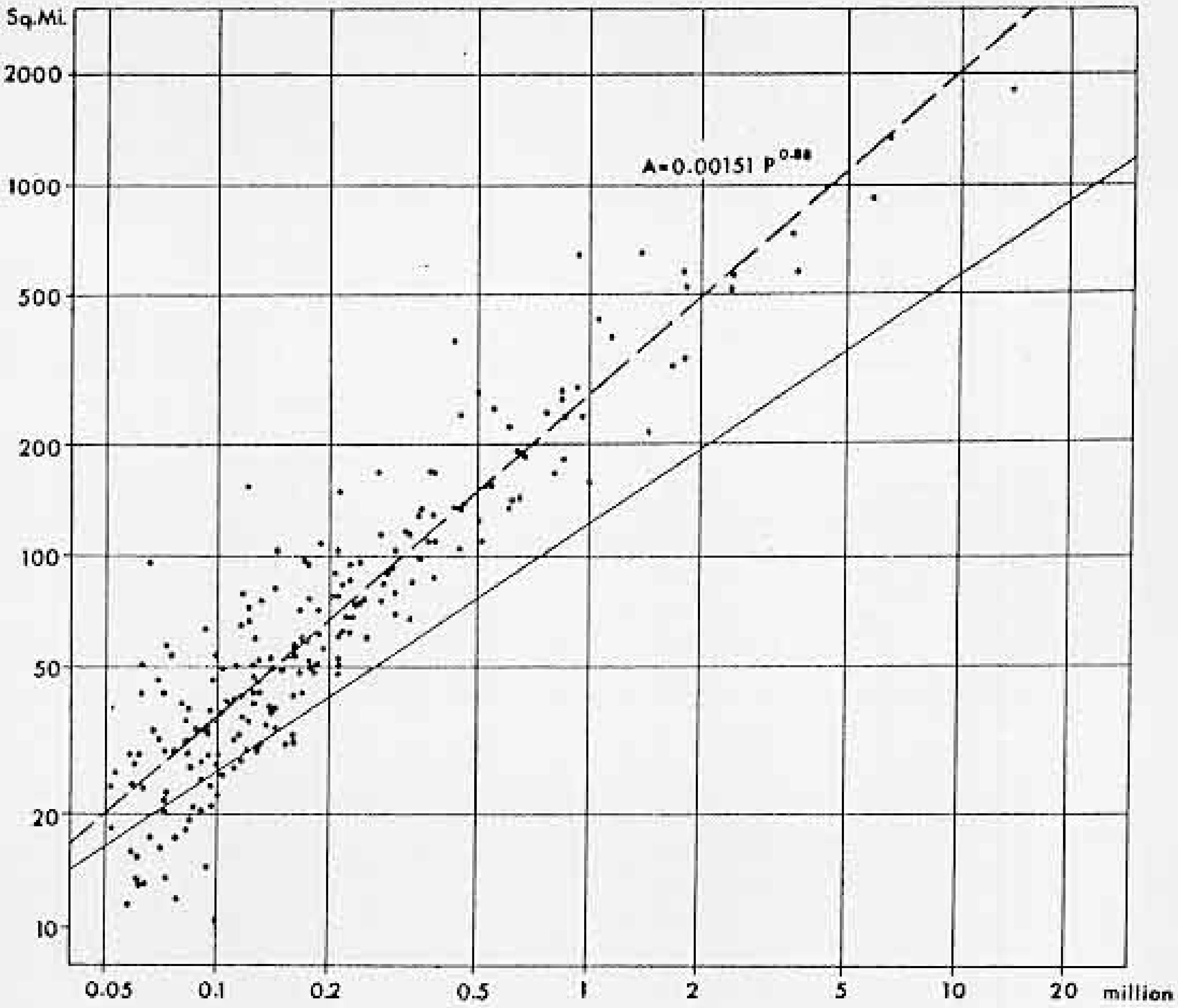


Figure 12.
The relationship between area (miles^2) and population of Urbanized Areas in U.S. in 1960.

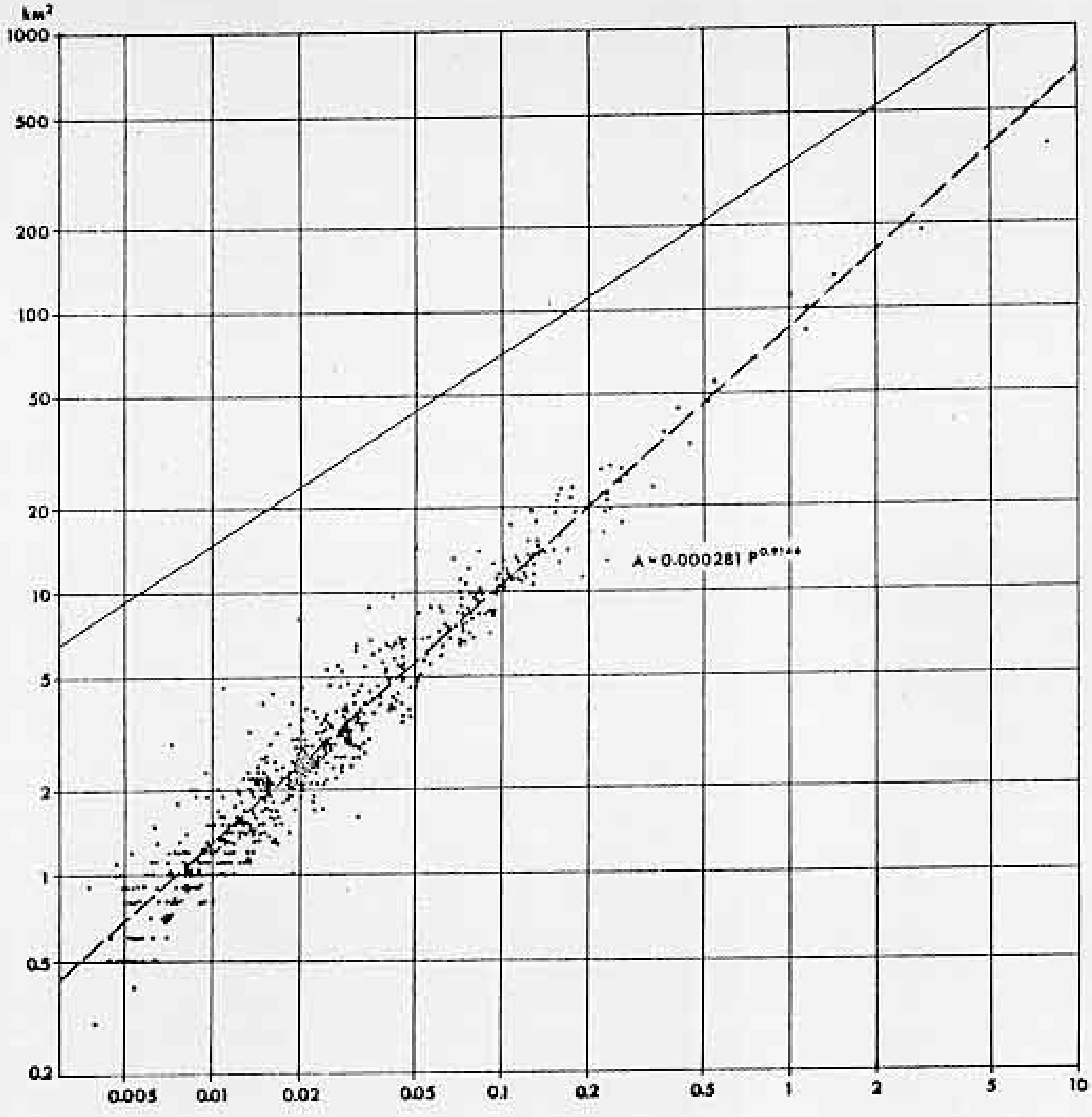


Figure 13

The relationship between area (km^2) and population of Densely Inhabitated Districts of Japan 1960.

belonging to the same size class. This can be done since there is so much data given (1960: U.S.A., 213; Japan, 518; 1950: U.S.A., 155 observations). It is seen in figures 11, 12 and 13 that these selected areas will give a b-value close to 0.33... The upper limitations of the dots in the figures are lines parallel to the line $r = 0.1P^{0.33...}$ which line is drawn in all three figures.

Built-up areas and allometric growth.

The data used in figures 10-13 are taken from the official statistics of Sweden, U.S.A. and Japan. The urbanized areas of U.S.A. and the densely-inhabited districts of Japan are poor approximations of the corresponding built-up areas. The b-values obtained were too high in spite of the high correlation between the log A (area) and log P (population). It was also pointed out that correct values of the built-up areas and of the population probably would give a lower b-value. This, however, is only a conjecture and is not very satisfactory. The areas of about 70 Swedish built-up areas (tätorter) are situated in South Sweden (Skåne and Blekinge). This part of the country was mapped in 1959-1961. Hence, the topographical map sheets show the situation of about 1960; the same time for which population data are available.

The area A and the population P of the measured built-up areas were plotted in figure 14. The correlation coefficient r is equal to 0.9762. Regression of log A on log P gives the line $A = 0.0085P^{0.664}$. If the area A is a circle with the radius r the length of r is given by the following formula: $r = 0.053P^{0.332}$. A is in square kilometers and r is in kilometers. The line $r = 0.053P^{0.332}$ can be compared with the line $r = 0.1P^{0.333...}$ drawn in figure 10 which gives too high an approximation of r. This approximation is quite better than such an one built on the administrative units.

The line $d = 0.1P^{0.333...}$ can be used as an approximation of the diameter d. The regression of population P on area A gives the line $A = 0.00687P^{0.696}$.

According to the diagram (figure 14), the built-up areas can be divided into three or more main groups:

1. Spacious built-up areas. Such areas have large areas compared with their populations. There are some towns belonging to this group which have many summer houses which have been erroneously classified as permanent dwelling houses. However, these areas can in most cases be recognized as the areas having the smallest population compared with size. There are only a few of the areas in figure 14 which belong to this group.

2. Mean built-up areas. This group includes most of the areas in figure 14. The area of a mean built-up area is less than $0.01P^{0.666...}$ and greater than $0.007P^{0.666...}$

3. Compact built-up areas. These have high populations compared with their areas. Old fishing villages belong to this group as do some railroad-station villages.

In the example in figure 14 the unknown size of the built-up area was estimated by means of the known size of the population. Of course this technique also permits estimation of the population based upon the known size of the built-up area. If the shape of a series of built-up areas is known, the formula for estimating the population also is known. The size of the areas can be measured from air photos. Using this method two men can estimate the population of 200-300 built-up areas in one day.

Population potentials.

The population potential V_i for place number i is defined as follows.

$$V_i = \sum_{j \neq i}^{n} (P_j / d_{ij}^{b}) + Oi \qquad (5)$$

where O_i is the contribution to the potential at i from the place i. P_j is

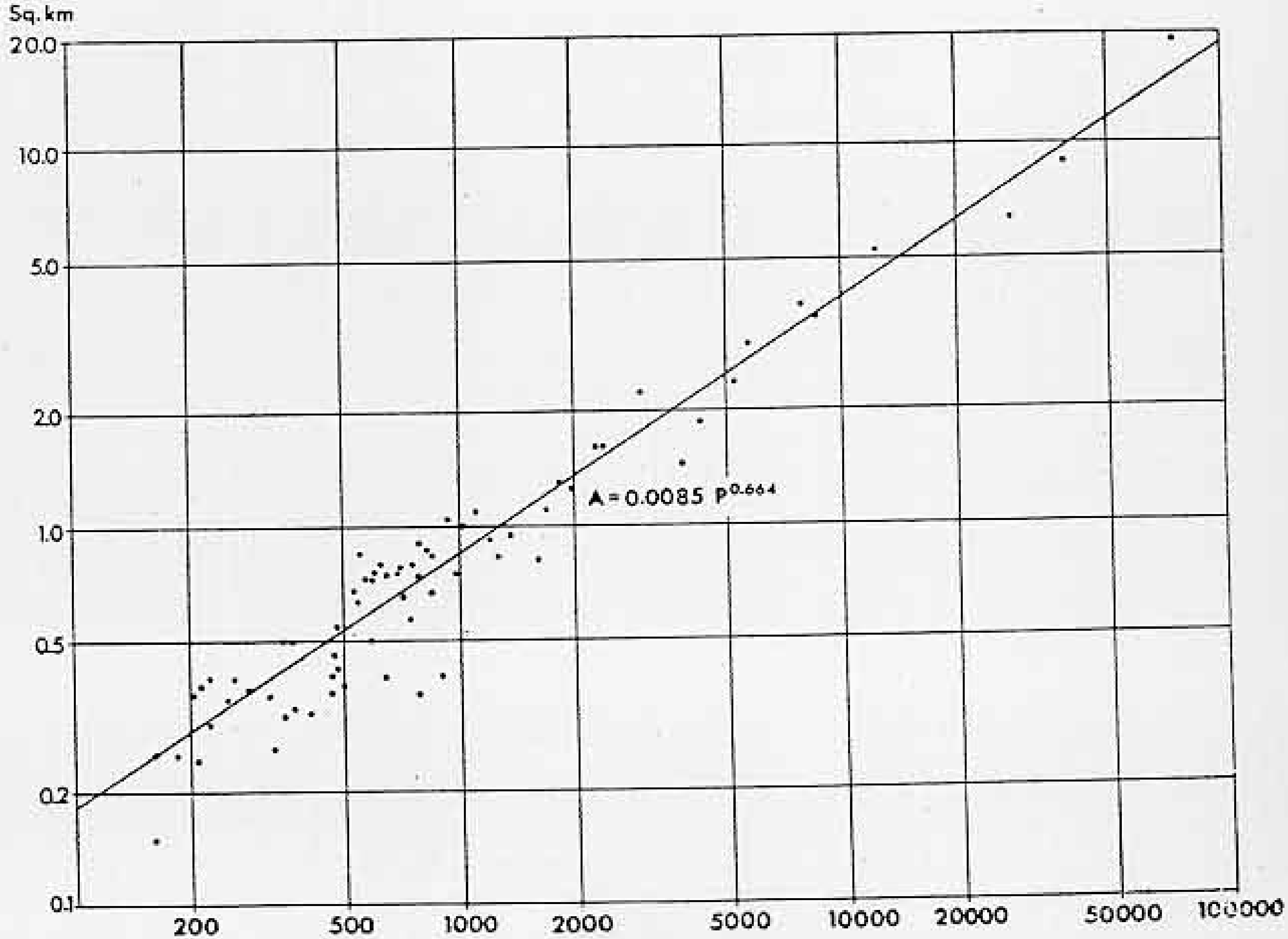


Figure 14

The relationship between area (km^2) and population of built up areas in Southern Sweden in 1960.

the population of place j, and d_{ij} is the distance between place i and place j, n is the total number of places, and b is a constant. If it can be assumed that place number i is close to a circle with the radius R and that the population is uniformly distributed over this circle, O_i is calculated by formula (6).

$$O_i = \frac{2P_i}{R^2} \int_0^R r^{1-b}\, dr \qquad (6)$$

If the constant b in formula (6) is equal to 1, O_i is equal to $2P_i/R$ which value usually is used as an approximation of O_i independent of the value of the constant b used in the summation of the first part of the potential formula (5). A b-value equal to zero makes the O_i equal to P_i. It must also be observed that b is not allowed to be equal to 2 or greater than 2, because the value of O_i in both these cases becomes infinite. The b-value in the sum of formula (5) has no such liminations.

The calculation of the population potentials involves a great number of numerical operations. Consequently, it is convenient to let a computer do these operations. Since the population of each place is known the computer can also determine the radius of each area by means of the allometric growth formula. Such an estimation of the radius is much better than one based upon administrative units. An administrative unit can be divided into two parts; A core equal to its central built-up area with the radius R_c, and one ring with the width R_d-R_c where R_d is the radius of the administrative unit. The population of the core is P_{ci} and that of the ring is P_{ri}. The contribution to the population potential coming from the place itself is determined by formula (7).

$$O_i = \frac{2P_{ci}}{R_c^2} \int_0^{R_c} r^{1-b}\, dr + \frac{2P_{ri}}{(R_d - R_c)^2} \int_{R_c}^{R_d} r^{1-b}\, dr \qquad (7)$$

It is here assumed that R_c is less than R_d.

This example shows how the radius of a built-up area can be determined by means of the allometric growth formula and how the size of this radius is used in population potential formulas. Other potentials, such as income potential are calculated in the same way as the population potential. In all these cases the radius of the place itself must be estimated as correct as possible since the greatest contribution to the potential generally comes from the place itself.

Deformation of shape.

There are no anthropoid giants over 4 meters tall. The length of the biggest mosquito is less than 2 inches. This depends on the fact that the constitution of a mosquito body does not allow it to grow too much. The wings and the legs of an insect cannot bear a too heavy a body. The tower of Babel could not have been as high as it is depicted. It is impossible to build a wooden house more than a few dozen meters high. Iron or concrete buildings can be many hundreds of meters high. The highest mountain of the world is less than 10 000 meters high. A mountain much more than 10 000 meters in height must change shape due to structural failure of its base.

The above examples indicate that the size of a growing individual having a special shape always has an upper limit. If the individual exceeds this limit its shape must be deformed. This deformation of the shape can easily be recognized in a diagram with logarithmic scales because there is a break in the point distribution corresponding to the change from one shape to another.

It is most likely that the cities and built-up areas cannot grow without any limitation. It is impossible to say the value of this limitation but as is seen in figures 11, 12 and 13 the large built-up areas such as New York, Los Angeles and Chicago have not yet passed this limitation. They all fit the curves perfectly. This is also true for Tokio and other large towns in Japan. Perhaps the shape of a built-up area is not deformed until its population is more than 100 million inhabitants.

Summary.

There are three main groups of growth:

1. Arithmetical growth which means that a growing individual increases with the same absolute value during each of a series of time periods all having the same length. Such a growth curve follows the formula $y = a + bx$. This is a straight line in a diagram with arithmetical scales.

2. Geometrical growth. A growing individual increases with the same relative value during the time periods. The formula $y = ab^x$ is valid for this kind of growth. In a diagram having the y-scale logarithmic and the x-scale arithmetic this equation is represented by a straight line.

3. Allometric growth. The relative growth of an organ y is equal to a constant fraction of the relative growth of the body x. This growth curve follows the formula $y = ax^b$ which is a straight line on a diagram with double logarithmic scales.

The streets of a town can be looked upon as organs and the town as the organism, etc. Thus the allometric growth formula has some very interesting geographical applications. It is formulated in such a way that a variable x can be estimated by means of an other variable y which is much easier to measure than x. This formulation of the law of allometric growth is valid for a series of individuals all having the same shape but of different sizes.

The rank-size rule for cities can be verified by means of this law (Beckman, 1958). The corresponding rule for incomes (Pareto, 1896-97) can also be derived by aid of a technique similar to Beckman's. The bifurcation ratio rule by Horton (1945) and Strahler (1952) is a rank-size rule valid for river channels and it can be verified in the same way as other rank-size rules.

The law of allometric growth is applied in this paper to rivers, meanders, volcanoes, drumlins, urbanized areas (U.S.A., 1950 and 1960), densely-inhabited areas (Japan, 1960) and built-up areas. In all these cases the correlation between log y and log x is extremely high. The correlation coefficients are all greater than 0.9. The constant b in the formula $y = ax^b$ shows the dimensionalities of the variables x and y. If y is an one-dimensional variable and x a two-dimensional one the theoretical b-value is equal to 0.5. The lengths of some rivers of widely varying sizes were compared with their corresponding drainage areas. The b-value obtained was very close to 0.5. This value can be compared with the value 0.6 which Leopold-Wolman-Miller (1964) refer to. The deviation between these two b-values can be explained in the following way: The b-value 0.6 was calculated by means of a few rivers and the difference in size between the largest and the smallest river was not great. A wider sampling of river and basin sizes would probably give a b-value closer to 0.5.

The a-value in the formula $y = ax^b$ can be used when classifying the given y-values. The rivers were divided into three groups: 1. Plain rivers with a short main river and a large drainage area. 2. Mean rivers. 3. Valley rivers with long main rivers compared with their drainage areas. The built-up areas of Southern Sweden were also divided into three types:

Compact built-up areas, Mean areas and Spacious built-up areas.

The formula $y = ax^b$ was used when the areas of the built-up areas in Sweden were estimated on the basis of calculated population and income potentials. It can also be used when the area of a built-up area is known and the population of this area to be estimated. In this case the area can be measured using large-scale air photos.

This paper has introduced some geographical applications of the allometric growth formula. This formula undoubtedly has many geographical applications which have not been mentioned here. The formula is always valid for an individual belonging to a series of individuals if all these have the same shape or almost the same shape. Consequently, if some individuals which are supposed to have the same shape do not fit the allometric growth formula they do not have the same shape. However, it must also be observed that there can be individuals which fit the formula in spite of having different shapes.

Acknowledgements

The volume of the volcanoes, figure 6, was measured by Mr. John Haake, Department of Geography, Wayne State University. The size of the built-up areas, figure 14, was measured by amanuens Thomas Hjortsjö, Department of Geography, University of Lund. The figures were drawn by Mr. Rezsö László. The calculation of the correlation coefficients and the a- and b-values in figures 11-13 were done by the Computer Department, Wayne State University, and by the Computing Center, University of Michigan. The English was corrected by Mr. Daniel Snyder, Department of Geography, Wayne State University. I thank all and especially professor William Bunge, Wayne State University, who gave me so much help.

Stig Nordbeck
University of Lund
December, 1964

Literature

Atlas of Sweden: Hydrology, Watercourses and Drainage-Basins, Territorial Waters. Stockholm 1957.

Auerbach, F., Das Gesetz der Bevölkerungskonzentration. Petermann's Geographische Mitteilungen 59. 1913.

Bagnold, R.S., Some aspects of river meanders. U.S. Geol. Survey Prof. Paper 282-E.

Beckman, M.J., City Hierarchies and the Distribution of City Size. Economic Development and Cultural Change. Vol. VI, No. 3, April 1958.

Bengtsson, B.-E. and Nordbeck, S., Construction of Isarithms and Isarithmic Maps by Computers. Nordisk Tidskrift for Informationsbehandling. (Bit). Hefte 2. 1964.

Berry, B. and Garrison, W.L., Alternate Explanations of Urban Rank-Size Relationships. Annals Association of American Geographers. Vol. 48. 1958.

Berry, B.J.L., Simmons, J.W. and Tennant, R.J., Urban Population Densities: Structure and Change. The Geographical Review. Vol. LIII, 1963.

von Bertalanffy and Pirozynsky, Ontogenetic and Evolutionary Allometry, Evolution, Vol. VI, No. 4.

Census of Population: 1960, Volume I, Characteristics of Population, Part 1, United States Summary.

Chorley, R.J., The Shape of Drumlins. Journal of Glaciology. 1958.

Clark, C., The location of industries and population. The Town Planning Review. Vol. XXXV. No. 3. 1964.

- Urban Population Densities, Journ. Royal Statist. Soc. Ser. A. Bd 114. 1951.

Dickinson, R.E., The West European City. London 1961.

Essays on Growth and Form Presented to D'Arcy Wentworth Thompson. Edited by W.E. le Gros Clark and P.B. Medawar. Oxford University Press. First edition 1945. Reprinted 1947.
Medawar, P.B., Size, Shape and Age.
Reeve, E.C.R. and Huxley, J.S., Some problems in the study of Allometric Growth.
Richards, O.W. and Kavanagh, A.J., The Analysis of Growing Form.

Förteckning över Sveriges vattenfall I-III. Statens meteorologisk-hydrologiska anstalt och Kungl. vattenfallsstyrelsen. Stockholm 1930, 1932, 1945.

Horton, R.E., Drainage-basin characteristics. Am. Geophys. Union Tranv. No. 13. 1932.

- Erosional development of streams and their drainage basins. Geol. Soc. Am. Built 56. 1945.

Huxley, J., Problems of Relative Growth. London 1932.

Kant, Edg., Zur Frage der inneren Gliederung der Stadt. Proceedings of the IGU Symposium in Urban Geography. Lund 1960. Lund Studies in Geography. Ser. B. No. 24.

Leopold, L.B., Rivers. American Scientist.

Leopold, L.B. and Langbein, W.B., The concept of entropy in landscape evolution. U.S. Geol. Survey. Prof. Paper 500-A.

Leopold, L.B. and Wolman, M.G., River Meanders. Geol. Soc. Am. Built. Vol. 71.

Leopold, L.B., Wolman, M.G. and Miller, J.P., Fluvial Processes in Geomorphology, San Francisco 1964.

Lösch, A., The Economics of Location. New Haven 1954.

Morisawa, M.E., Relation of Quantitative Geomorphology to Stream Flow in Representative Water-sheds of the Applachian Plateau Province. Technical Report No. 20. Project NR 389-042. Office of Naval Research, Department of Geology, Columbia University. New York 1959.

Nordbeck, S., Framställning av kartor med hjälp av siffermaskiner. Lund 1964.

Nordström-Johnsson-Norborg-Fourén, Geografiska tabeller med arbetsuppgifter. Stockholm 1964.

Norling, G., Folkräkningar och geografi. Ymer. Häfte 3. 1964.

1960 Population Census, Densely-Inhabited Districts. Bureau of Statistics, Office of the Prime Minister. Tokio 1961.

Pareto, V., Cours d'Economie Politique, Lausanne 1896-97.

Rapoport, A., Some theoretical consequences of the allometric growth equations. Bulletin of Mathematical Biophysics. Vol. 17. 1955.

Stahl, W.R., Similarity and Dimensional Methods in Biology. Science, Vol. 137, 1962.

Strahler, A.N., Dimensional analysis in geomorphology, Off. Nav. Res. Proj. NR 389-042, Tech. Rep. 7, Dept. Geol., Columbia Univ., N.Y. 1957.

- Quantitative Analysis of Watershed Geomorphology. Am. Geophys. Union Trans. 1957.

- Quantitative geomorphology of erosional landscapes. C.-R. 19th Intern. Geol. Cong., Algiers, 1952. 1954.

- Physical Geography. New York 1960.

Thompson, D.W., On growth and Form. Cambridge 1959.

von Thünen, J.H., Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. Hamburg 1826.

Zipf, G.K., National Unity and Disunity. Bloomington Ind., 1941.

Foreward

In these two short papers Professor Tobler demonstrates both the originality and analytical ability so typical of his work. Some of the ideas he expresses have been developed from conversations held by the members of the Michigan Community of Mathematical Geographers over the past year. The advantages of extending analytical techniques of one dimension to other dimensions is one example.

His analysis of map generalization by numerical means has produced surprises. I did not expect that some smoothing operators would prove to have inverses. A class of such operators is demonstrated here. Some of our readers may be surprised at the usefulness of computer produced maps. Professor Tobler has assembled programs which produce such maps with ease and speed and at costs under the costs of manually producing the same.

The two ideas I find most exciting in these papers are the emergence of an explicit operational definition of site or neighborhood conditions by use of local operators and the suggestion that by use of frequency filters and similar techniques periodicities in space may be revealed.

The Editor

January 1966

Numerical Map Generalization

W. R. Tobler
Department of Geography
University of Michigan
Ann Arbor

ABSTRACT

Mathematical procedures for the generalization of isarithmic maps are derived. It is shown that the level of generalization for different purposes may be specified in advance, and that under certain conditions the original map can be recreated from the generalized map.

ACKNOWLEDGEMENTS

The author is indebted to the computing center of the University of Michigan for support of this study. J. Harbaugh made available the two-dimensional Fourier analysis computer program. Colleagues and students in the Michigan Inter-University Community of Mathematical Geographers contributed numerous comments. The data employed for the examples came from a digitalization by R. Yuill of a contour map given on page 92 of Garnier (1963).

The process of map generalization bears a certain resemblance to the process of abstracting textual materials, such as books and articles. As a condensation one expects that only the most relevant and important items will be retained in the simplified version. It is of some theoretical interest to speculate on the extent to which this process can be formalized. As a form of "picture abstracting" the process of map generalization is probably a simple case of a more general problem of pattern analysis, which in turn may be considered to be the basis for the important inductive approach to any knowledge. In particular, it might be anticipated that a study of map generalization might also illuminate the process of scientific generalization, especially as practiced in fields which traditionally make frequent use of maps.

Mapped data may be considered to consist of point (or zero-dimensional) phenomena, and line, area, intensity, and flow symbols. Generalization consists of the application of a transformation which modifies the map data. Pillewizer and Topfer (1964) have considered generalization of point symbols, and Perkal (1958) has contributed to the problem of generalizing lines and areas. The present discussion examines the generalization of "three-dimensional" phenomena, which are assumed to be continuous. Information of this type is most often shown on maps by isarithms of various types; contours, isopleths, isobars, and so on.

As *a priori* criteria one would expect that generalization of isarithmic maps should somehow eliminate "small scale" features. If generalization is considered as a transformation it is natural to inquire whether the process can be reversed. If topographical maps are generalized the result should conform to appropriately modified accuracy standards. Finally, certain statistical parameters (mean, standard deviation, and so on) should be preserved by the transformation.

Trend analysis provides one form of map generalization. The geographical trend of population density (persons per square mile) in Michigan is (Tobler, 1964a)

$$D = 6314 - 68.7 * \text{latitude N} - 36.9 * \text{longitude W},$$

or, in words, the population declines towards the northwest. The "plane" defined by this equation is in some sense a geographical generalization of the distribution of population in the state of Michigan. The average population density is preserved and the generalization represents about 8% of the observations. A quadratic generalization of the same data,

$$D = 207529 - 4547 * \text{lat. N} - 2389 * \text{long. W} + 52 * \text{lat. N} * \text{long. W},$$

accounts for some 15% of the data. In both of these instances the generalization is extreme, though cases can be cited (Krumbein and Graybill, 1965) for which this technique of trend analysis is most useful.

Another possible method of contour generalization is to smooth the values by a moving average, as in the treatment of time series data. For expository convenience consider first only a profile. This can be represented by a single valued function $Z = f(x)$. From this profile select a finite sample of values Z_i at a set of discrete locations $x_1, x_2, x_3, \ldots, x_i, \ldots x_n$. For notational convenience let $\Delta x = x_i - x_{i-1}$ = a constant for all i. Smoothed values Z* can be obtained from the original values from the following linear equation

$$Z_i^* = \frac{\sum_{m=-k}^{m=k} W_m Z_{m+i}}{\sum_{m=-k}^{m=k} W_m}$$

except for obvious modifications at the ends of the interval. The W's are weights, so that Z* is a weighted moving average. This can be written, for k = 1 and with the sum of the weights equal to unity, as the matrix equation

$$Z^* = Z \cdot S$$

where Z is the (1 by n) vector of Z_i and S is an (n by n) smoothing matrix of the form

$$S = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 & & 0 & 0 & 0 \\ 0 & 1/4 & 1/2 & 1/4 & 0 & 0 & & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/2 & 1/4 & 0 & & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 & & 0 & 0 & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & & 1/2 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1/4 & 3/4 \end{bmatrix}$$

which represents a three point binomially weighted moving average. For $k > 1$ the matrix is simply less sparce. Note that a second smoothing can be obtained as

$$Z^{**} = Z^* \cdot S,$$

or, more generally

$$Z^M = Z \cdot S^M .$$

Repeated application of the smoothing operation of course results in increased generalization.

Under certain conditions the original values can be restored from the smoothed values by solving the above matrix equation to obtain

$$Z = Z^* \cdot S^{-1} .$$

In other words, a generalized profile can be ungeneralized ! The condition is the existence of the inverse matrix S^{-1}. The necessary theorems demonstrating the existence of this inverse are in the appendix. In general, the exact inverse requires an inverse weighting using all n of the smoothed values. An approximate, three weight, inverse for the matrix shown can be constructed from the weights -0.5, +2.0, -0.5, however (Holloway, 1958). This is less exact but is more practical for large n.

If the sum of the weights in the smoothing operation is equal to unity, then the mean of the series is invariant. The extremes are usually reduced, however. In some cartographical situations it seems important that the extrema be retained. Several simple operators can be devised which restore the extremes. For example, let

$$Z_i^r = A + BZ_i^*$$

where

$$B = \frac{Zmax - Zmin}{Z^*max - BZ^*min.}$$

This normalization restores the maximum and minimum of the series.

A numerical example is given in Table I and is illustrated in the accompaning figures.

TABLE I
PROFILE GENERALIZATION

i	Z	Z^*	Z^r	Z^{**}	Z'	Z"	$\hat{Z}$	$\hat{\hat{Z}}$
1	490	493	490	495	644	645	490	492
2	500	500	501	500	647	647	501	499
3	510	505	508	508	652	644	499	512
4	500	523	536	529	624	641	511	498
5	580	565	600	568	662	644	559	585
6	600	620	684	670	627	597	520	593
7	700	875	1073	881	472	553	863	709
8	1500	1155	1500	1051	992	751	1365	1489
9	920	1020	1294	986	547	681	1088	934
10	740	750	882	788	637	609	674	724
11	600	633	704	648	614	632	604	619
12	590	575	615	579	662	643	567	571
13	520	533	551	537	634	643	526	539
14	500	505	508	517	642	635	482	483
15	500	523	536	537	624	633	496	516
16	590	595	646	590	642	652	605	578
17	700	648	726	629	699	666	686	708
18	600	625	691	631	622	641	614	597

PROFILE GENERALIZATION

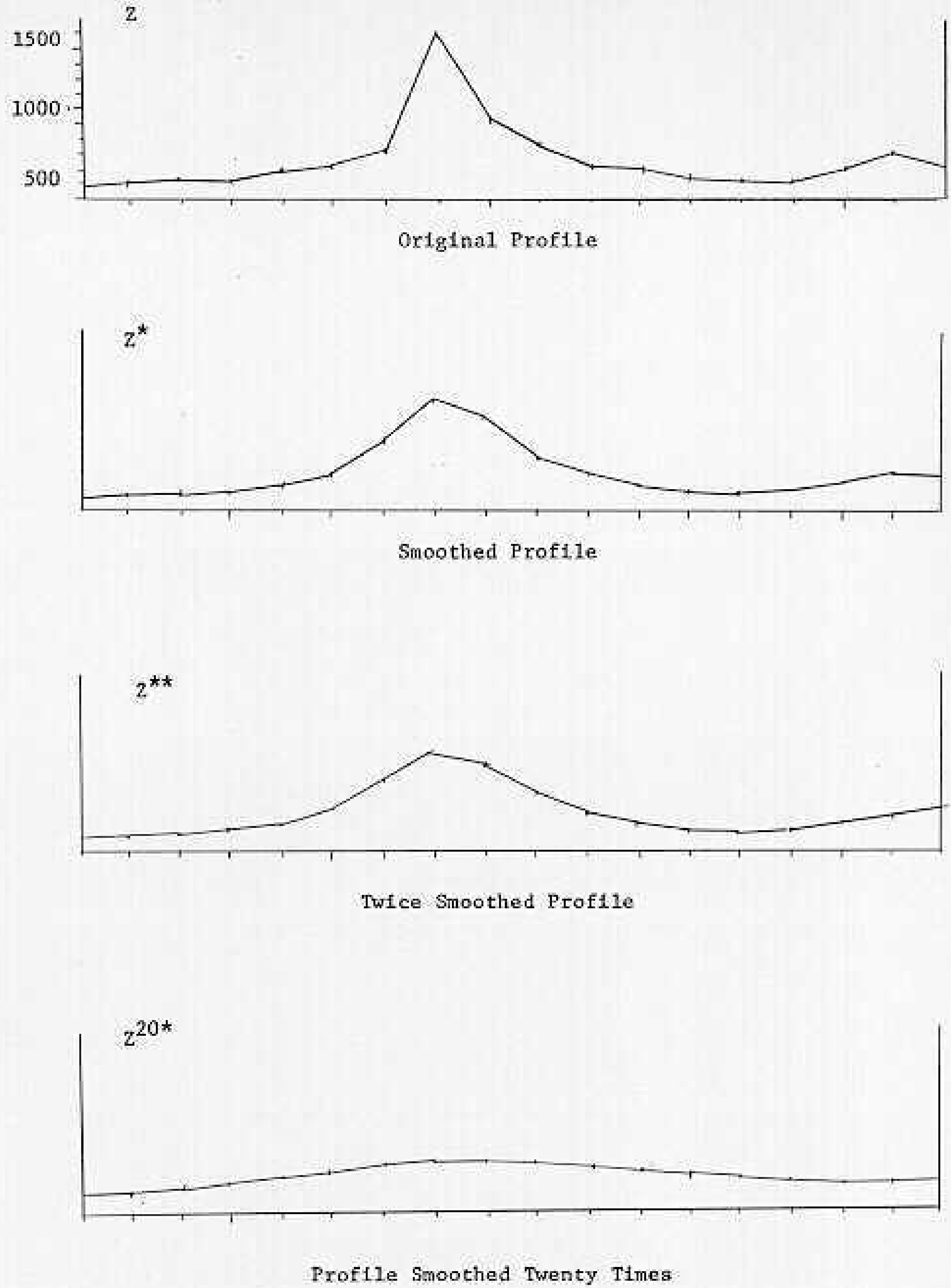
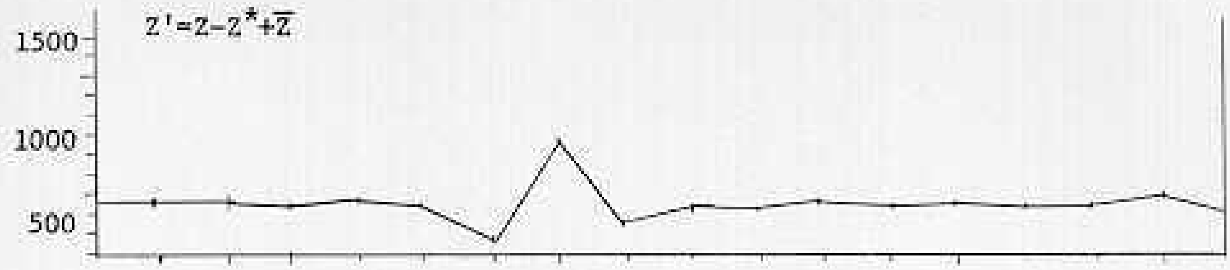
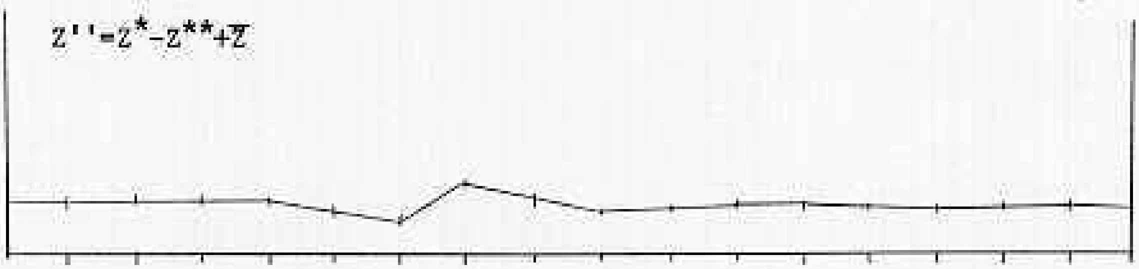


Figure 1a

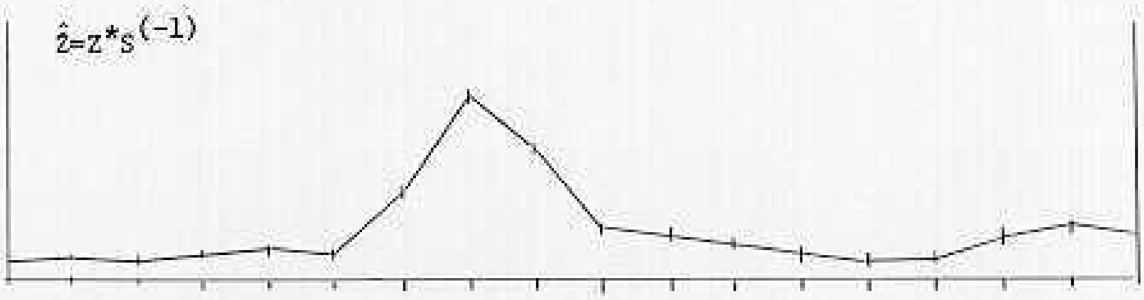
PROFILE GENERALIZATION



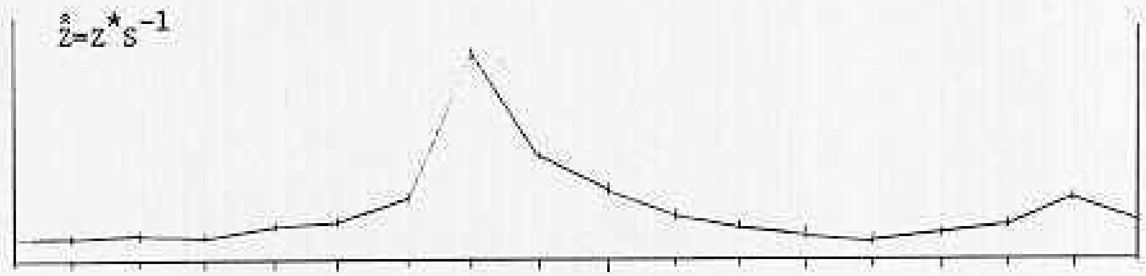
The Small Scale Features



The Medium Scale Features

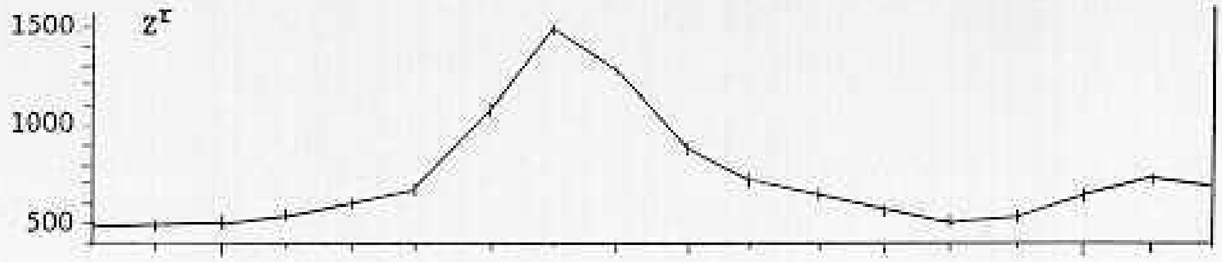


Restored Values obtained from Z^* by use of the 3 weight approximate inverse $S^{(-1)}$

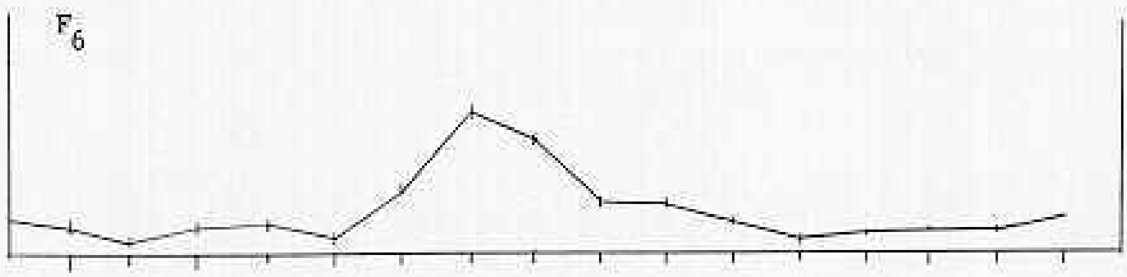


Restored values obtained from Z^* by use of the exact inverse S^{-1}

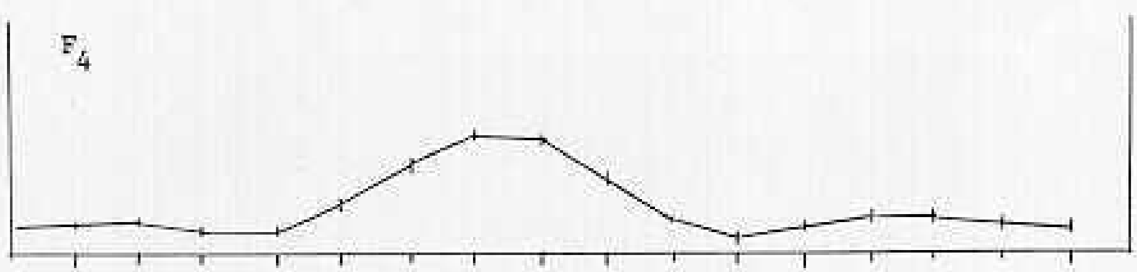
Figure 1b



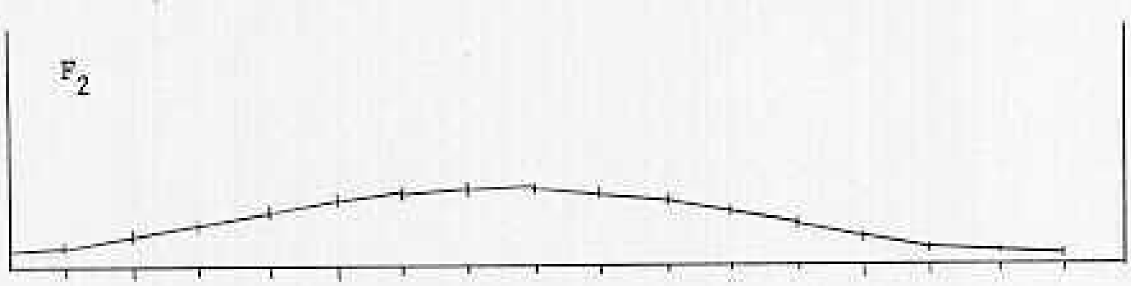
Smoothed, then Normalized



Fourier expansion of 12 terms



Fourier expansion of 8 terms



Fourier expansion of 4 terms

Figure 1c

The smoothing operation can be applied to generalize contour maps, not only profiles. Let the value of a surface at any point ij be defined by $Z_{ij} = f(x_{ij}, y_{ij})$. In most cartographical situations this is a single valued function. For convenience take $\Delta x = \Delta y =$ unity, and a k^{th} two-dimensional weighted moving average can then be defined by

$$Z^*_{ij} = \frac{\sum_{p=-k}^{p=k} \sum_{q=-k}^{q=k} W_{pq} Z_{i+p,\ j+q}}{\sum_{p=-k}^{p=k} \sum_{q=-k}^{q=k} W_{pq}}$$

This is easily implemented on a digital computer. In matrix form, the process can be written as

$$Z^* = S \cdot Z \cdot S,$$

where Z is now a matrix of "elevational" data, and S is a conformable smoothing matrix as in the previous case. Repeated smoothings can be obtained by repeated matrix multiplications. The inverse operation ("ungeneralizing, or roughening") is given by

$$Z = S^{-1} \cdot Z^* \cdot S^{-1}.$$

The data matrix Z, of course, need not be square, but the smoothing matrices are then of different orders.

A somewhat more practical procedure than the matrix formulation, and perhaps easier to visualize, is to consider the contour generalization to be the resultant of two operations taken in sequence, one applied to all the east-west profiles, then the other applied to all of the north-south profiles. The amount of smoothing in these two directions is usually the same, but could also differ. Construction of a nine-point two dimensional weighting function from the one dimensional case then follows by cross multiplication from a table such as given here:

	1/4	1/2	1/4
1/4	1/16	1/8	1/16
1/2	1/8	1/4	1/16
1/4	1/16	1/8	1/16

Construction of an approximate inverse weighting function is similar, e.g.:

	-1/2	2	-1/2
-1/2	1/4	-1	1/4
2	-1	4	-1
-1/2	1/4	-1	1/4

The common cartographic rational for map generalization is retention of legibility under reductions of scale. It is suggested here, however, that it may also be appropriate to generalize maps, not only because of scale changes, but also simply for different map purposes. In other words, the map information is to be filtered in accordance with the use of the map. A brief discussion of spatial frequency filters is therefore appropriate.

It is known that a single valued profile can adequately be represented as the Fourier sum of a large number of sine and cosine curves of varying frequency, amplitude, and phase. Similar results hold for the two dimensional case. The moving average operates as a transformation to modify the amplitudes and phases of the individual trigonometric terms (which are usually referred to as frequencies), and this modification is not the same for all terms. The amount (which can be calculated) by which the amplitudes and phases of individual frequencies are modified depends on the weights employed in the moving average. The binomial weighting employed in the previous examples has certain advantages (e.g., large scale features are retained and small scale features are filtered out, and phases are unchanged) but is only one of many possible weighting schemes. The simple (or equally weighted) moving average might, at first glance, appear more advantageous, but it has certain undesirable properties; for example, some peaks may become troughs. Since the frequency response of particular weighting schemes may be calculated, it might be expected that one could apply the procedure in reverse. To some extent filters (weights) can be devised to eliminate, or emphasize, features of any size (frequency), and approximate inverse filters can also be obtained. This suggests that it is possible to specify the level of generalization in an *a priori* fashion for different map scales. For example, at a map scale of 1:500,000 features having a wavelength of less than one half kilometer can probably not be shown. Holloway (1958) demonstrates explicitly the amount of generalization (that is, the weighting function) required (in his notation set $\sigma = 0.167$ kilometers). This example is elaborated in the appendix.

Also implicit in the construction of frequency filters is the notion that, if one knows the geographical scale to which some process is sensitive, one can mathematically generalize a map specifically for this purpose, and this map will differ from that for some other process. For example, topographic features of wavelengths greater than 25 kilometers and less than 0.10 meter are probably not important for the construction of airports. An appropriate filter would eliminate these large and small scale features but would leave medium sized features. Unfortunately we do not have much information concerning the wavelengths to which geographical phenomena are sensitive. It is also suggested that Perkal's (1958) ε- generalization procedure consists of a form of frequency cut-off (truncation of the Fourier series).

An example of contour map generalization by the procedures described is given in Table II and the accompaning illustrations.

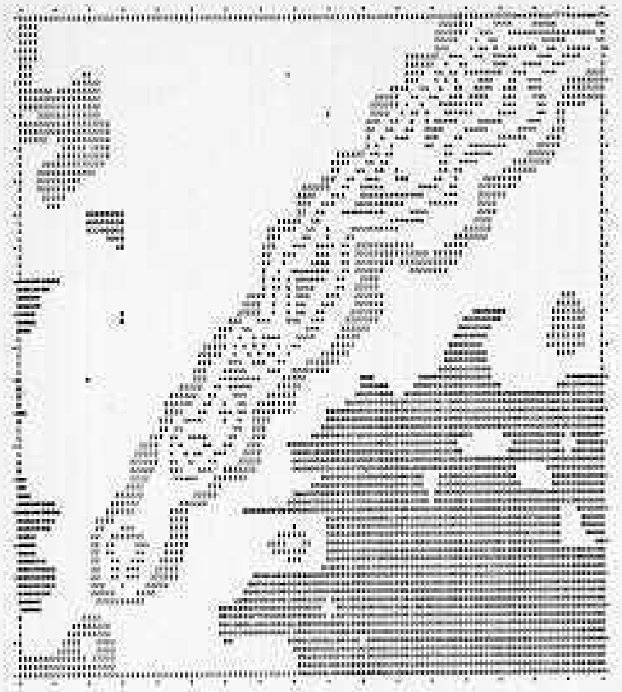
TABLE II

ELEVATIONS EMPLOYED IN CONTOUR GENERALIZATION[1]

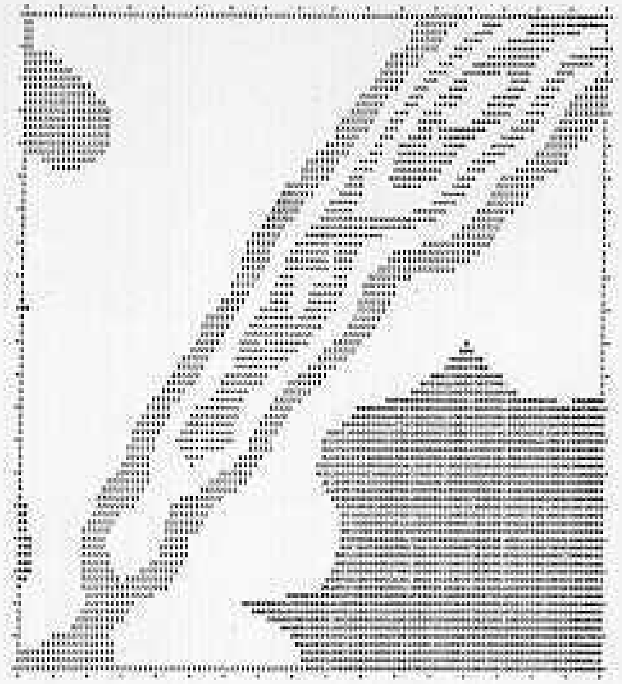
630	620	605	595	590	605	590	605	605	590	580	600	610	640	710	1050	1400	1340
640	605	580	560	610	595	598	590	600	600	575	610	620	860	1500	1380	980	800
640	600	650	570	570	600	580	605	570	580	590	585	850	1600	1300	910	720	660
800	690	640	600	570	600	585	605	570	630	580	800	1530	1400	1210	1030	595	598
580	590	750	570	570	560	550	560	570	600	595	1300	1370	995	900	750	580	590
540	730	590	575	560	550	560	570	595	630	1100	1380	1300	790	690	580	570	570
550	610	490	500	530	600	570	560	605	1050	1200	1050	850	850	590	570	580	610
540	520	510	500	530	570	560	595	900	1200	645	670	805	630	570	550	580	605
500	495	540	540	550	565	580	600	1250	1100	690	570	605	570	570	570	580	585
475	525	520	500	520	560	580	700	1250	850	700	580	550	505	490	550	700	550
520	505	510	505	550	570	700	1500	910	750	585	550	525	480	520	585	700	550
500	530	500	520	605	590	950	900	650	585	500	505	500	470	510	520	500	480
500	510	530	560	590	790	1200	650	620	495	475	500	500	475	475	480	490	500
505	510	550	595	650	1350	950	640	475	480	475	500	480	510	520	460	510	460
495	540	575	580	800	930	650	610	510	500	460	475	510	480	480	530	480	455
500	480	560	700	850	595	510	495	485	500	500	470	500	460	460	470	510	480
480	530	590	1100	800	600	520	575	800	470	460	450	500	480	460	470	510	500
485	485	605	1000	700	570	540	495	500	490	475	470	440	420	420	445	465	460
470	525	600	585	520	530	495	475	470	510	480	450	430	380	400	430	440	440
580	580	800	570	505	515	500	510	500	480	505	420	400	400	430	440	450	460
700	700	705	605	580	580	615	605	505	470	440	400	380	420	410	450	490	485

1) Elevations in feet, grid interval of 0.5 mile.

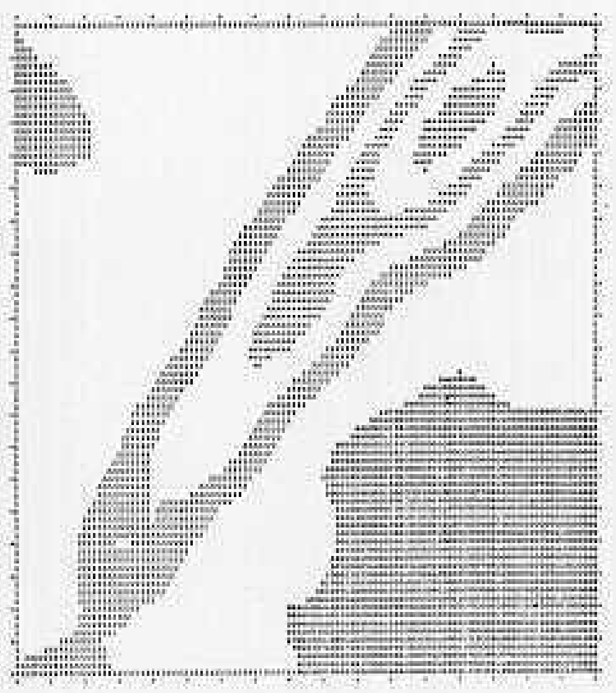
MAP GENERALIZATION



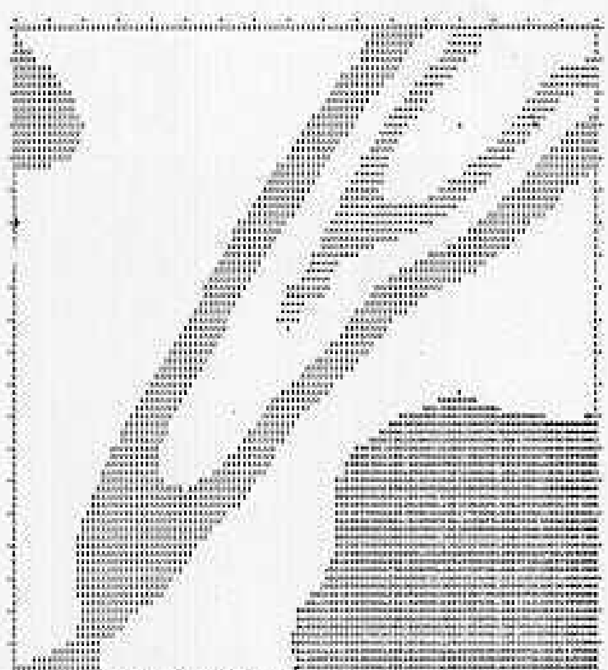
Original Map Z



Smoothed Map Z^{*}

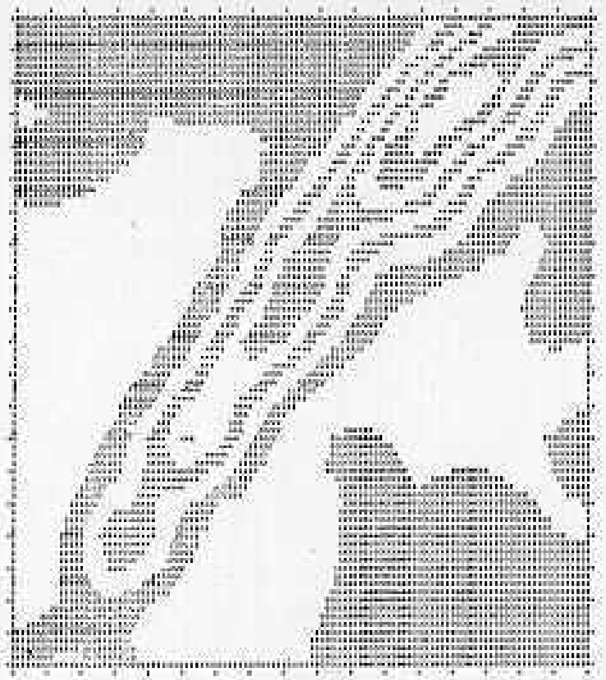


Twice Smoothed Map Z^{**}

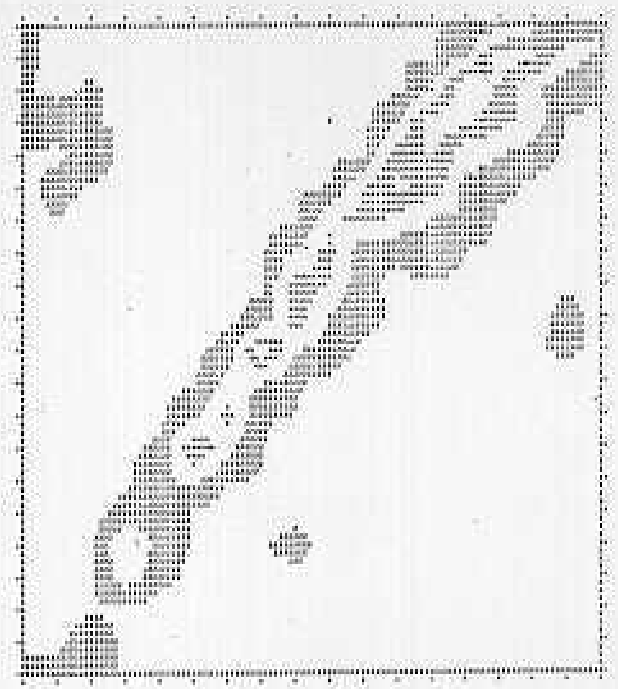


Thrice Smoothed Map Z^{***}

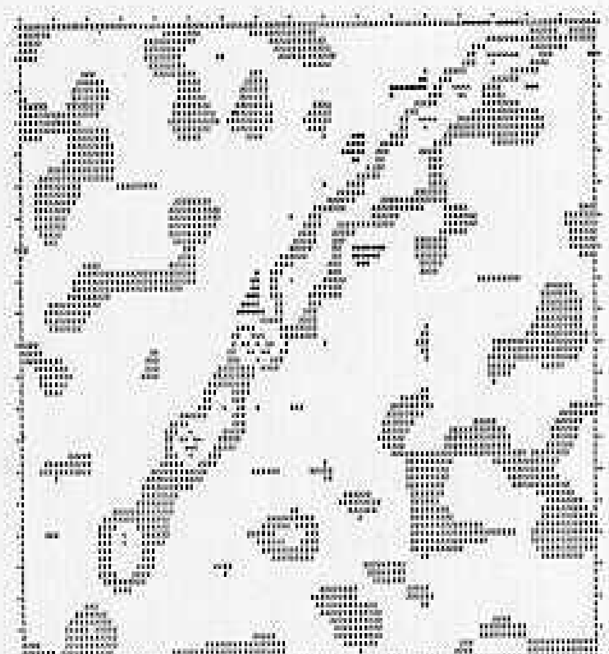
MAP GENERALIZATION



Smoothed - Then Normalized



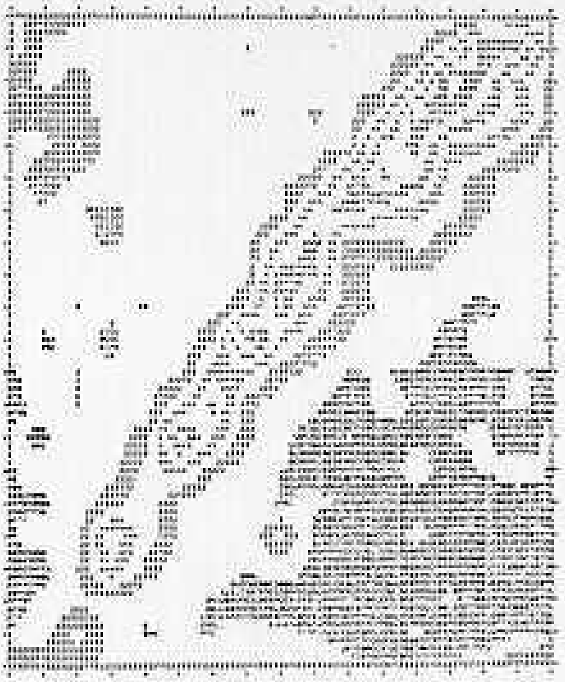
Original Map - Doubled Contour Interval



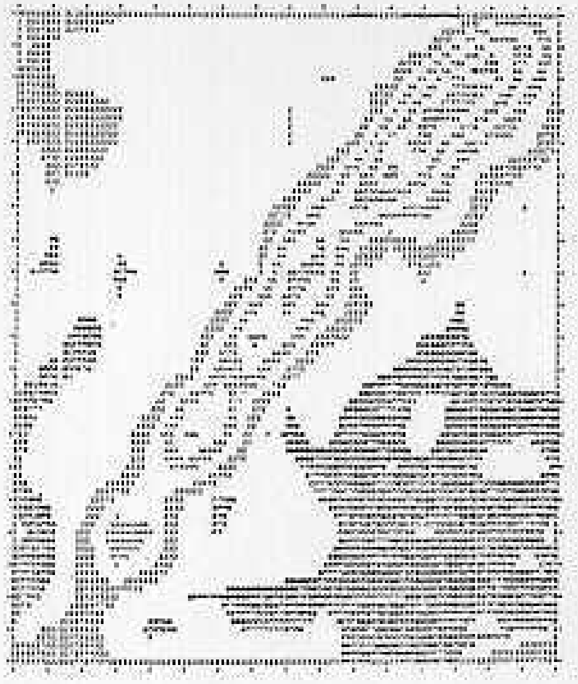
Small Scale Features
Removed During 1st Smoothing

MAP GENERALIZATION

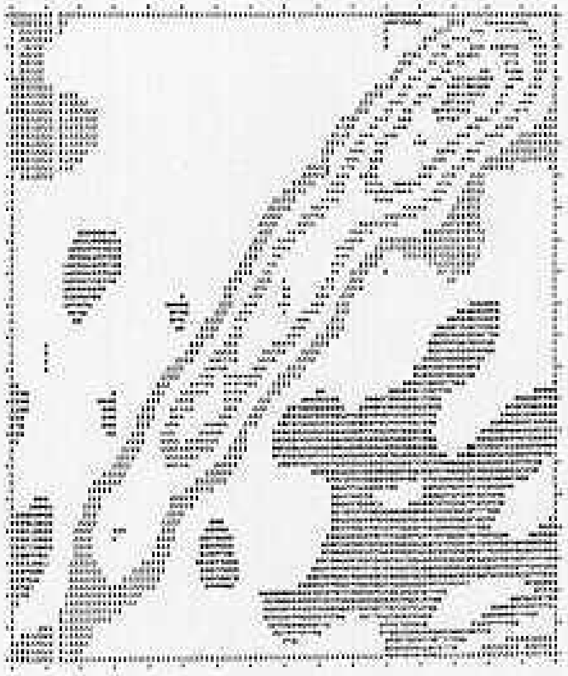
Using Trigonometric Polynomials



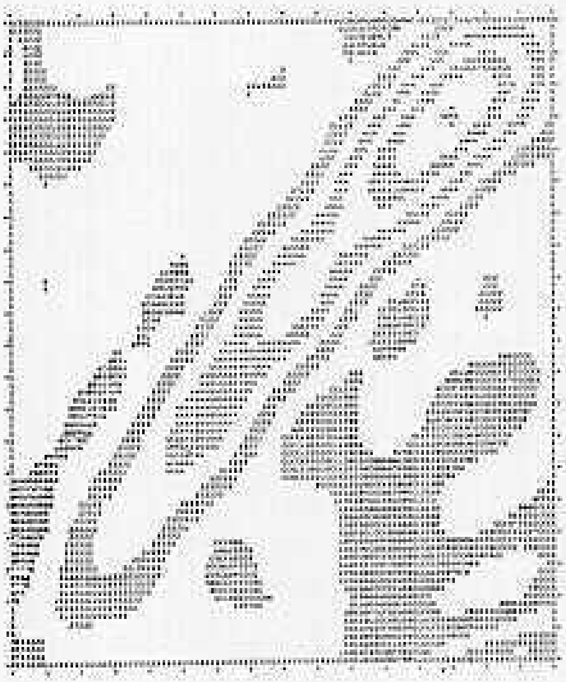
285 Terms



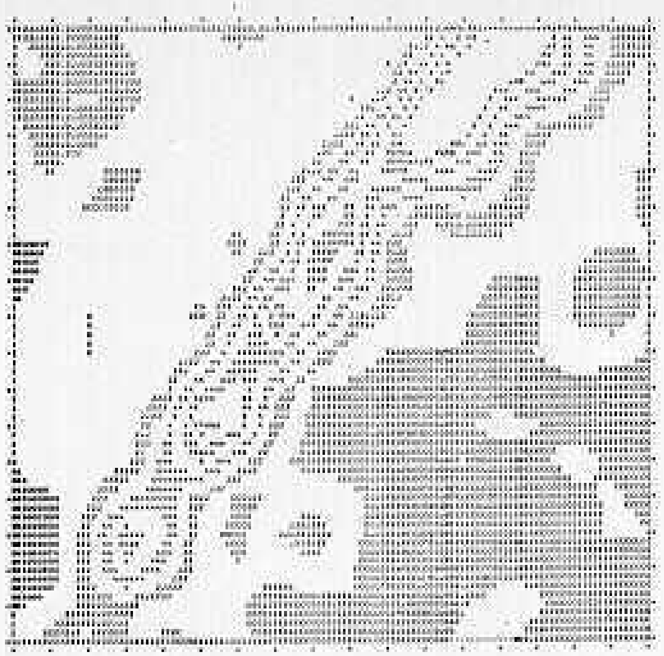
143 Terms



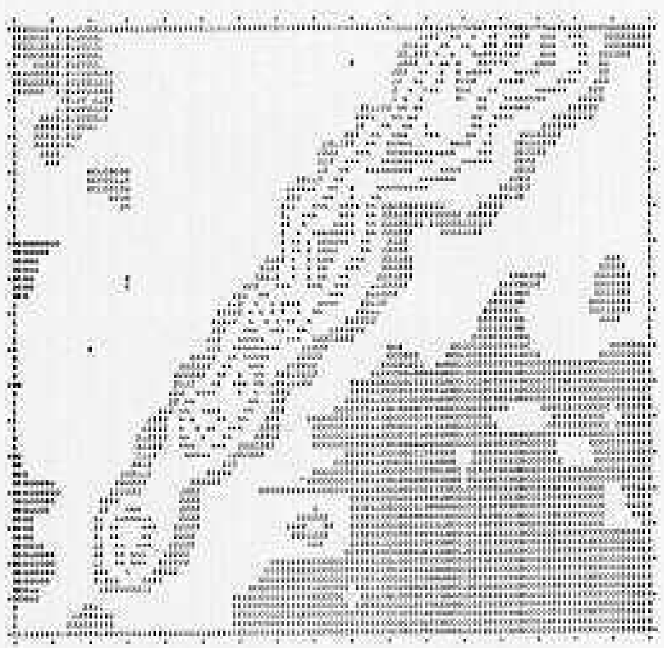
81 Terms



49 Terms



Map Restored Using Approximate Inverse



Map Restored Using Exact Inverse

It is also possible to consider any individual contour as a line, and attempt line-generalization, as of say, a coastline. As the map scale becomes smaller and smaller, less and less of the coastline detail can be shown. This situation is discussed at some length in the available literature. Alternately, as Perkal suggests, different generalizations of the coastline may be appropriate for different purposes. One computer technique for coastline generalization consists of thinning. Given a string of positional coordinates for sequential points along the coast, points j are eliminated if

$$D_{ij} < \epsilon$$

where $j = i + k$, $k = 1,2,\ldots,n$,

and in the cartesian case,

$$\left[D_{ij} = (x_i - x_j)^2 + (y_i - y_j)^2\right]^{1/2}$$

and ϵ is some minimal distance, for example, twice the minimal plotting size of the automatic drawing instrument. This approach has already been employed with considerable success and is more fully discussed elsewhere (Hershey, 1963; Tobler, 1964b). Thinned data cannot be restored, however.

If the coastline were a single valued function it could be generalized as a topographic profile. The coastline, however, is not single valued. It might be decomposed into single valued intervals but a parametric representation in terms of arc length s is more convenient. Let the coastline be represented by the complex function

$$Z = f(x + iy), \quad i^2 = -1,$$

where $s = s(s)$ and $y = y(s)$. For equal intervals Δs of s the j^{th} point of the coastline is

$$Z_j = x(s_j) + iy(s_j).$$

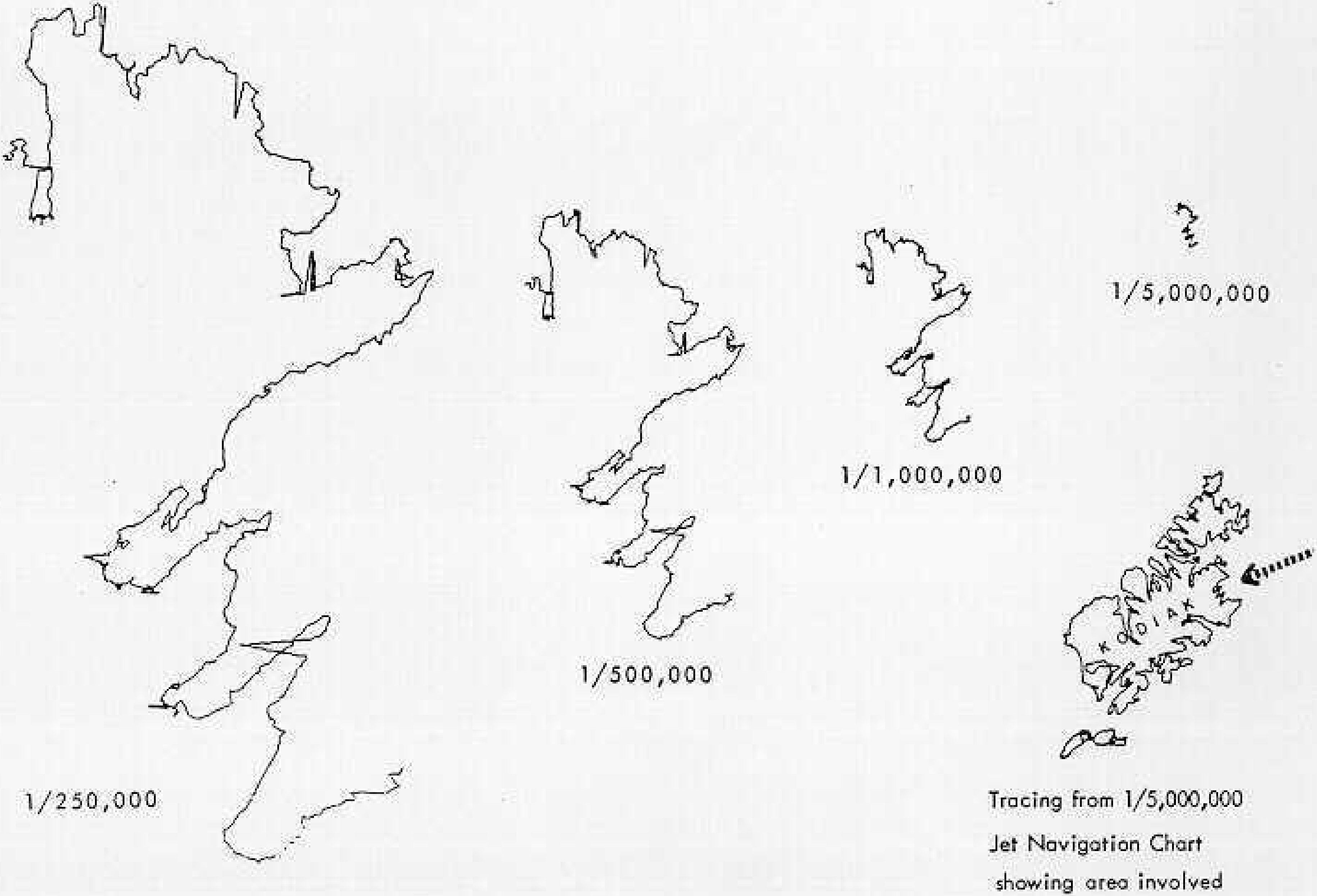
A smoothing can then be accomplished by taking

$$Z^* = Z \cdot S$$

where S is a smoothing matrix as before. Since the coastline closes, k values

Generalization of Coastline by Thinning. After Tobler (1964b).

Machine drawn from data compiled at 1/50,000



of Z_j could be added to the beginning and end of the complex vector Z, and S becomes an n + 2k by n + 2k matrix. Different smoothings could in fact be applied to the real and imaginary parts of Z, though it is difficult to imagine why this might be done.

Examination of a number of published contour maps at different scales but of the same area suggests that the practical cartographer generalizes by thinning after enlarging the contour interval in approximately the inverse ratio of the map scales. This process is easily formalized, e.g., on a digital computer, and appears suitable for the navigational usage of maps. The filtering procedure, on the other hand, seems to have advantages for more advanced theoretical investigations. For example, property values for individual parcels of real estate within a city (as might be obtained from the assessor's office) typically show considerable variation from one parcel to the next. The space smoothing technique can be employed to investigate the existence of geographical patterns of property values. The generalization facilitates the recognition of patterns because it appears to be true, as Holloway (1958, p.386) suggests, that we do "high-pass filtering in our 'mind's eye'."

Several additional smoothing techniques can be found in works on numerical methods, and these can generally be extended to the two-dimensional case fairly easily. Virtually all of these techniques make the assumption that "there is a strong antecedent probability that if the observations had been more accurate the curve would have been smooth" (Whittaker and Robinson, 1944), where smooth means that higher order differences vanish. In the topographic situation, at least for purposes of generalization, it can be assumed that the errors are negligible and the concept of a filter seems more appropriate. In terms of geographical theory one might argue that the pattern-generating processes are subject to random disturbances, which we wish to eliminate by the generalization procedure.

APPENDIX I

The Inverse of the Smoothing Matrix

Definitions

(1) A matrix $A = [a_{ij}]$ is diagonally dominant if

$$\left|a_{ii}\right| \geq \sum_{\substack{J=1 \\ J \neq i}}^{n} \left|a_{ij}\right| \quad , 1 \leq i \leq n$$

(2) A matrix A is reducible, $n \geq 2$, if there exists an n by n permutation matrix P such that

$$P \cdot A \cdot P^T = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

where A_{11} is an r by r submatrix, A_{22} is an (n-r) by (n-r) submatrix, $1 \leq r \leq n$. If such a P does not exist then A is irreducible.

(3) A directed graph is strongly connected if for any ordered pair of nodes N_j and N_k there exists a directed path $N_j N_{m1}, N_{m1} N_{m2}, \ldots, N_{mr} N_k$ connecting N_j and N_k.

(4) A matrix A is irreducibly diagonally dominant if A is irreducible and diagonally dominant with the strict inequality valid for at least one i.

Theorems

(I) A matrix is irreducible if, and only if, its directed graph is strongly conected.

(II) A matrix which is irreducibly diagonally dominant is non-singular.

Remarks

The 18 by 18 smoothing matrix S used in the generalization is irreducibly diagonally dominant. The inverse matrix S^{-1} can in fact be exibited explicitly, but is omitted here (interested parties can obtain copies from the author). An "inverse" $S^{(-1)}$ constructed from the approximate inverse weighting function (-.5, +2, -.5) has the property that

$$S^{(-1)} \cdot S \neq I$$

as was expected.

APPENDIX II

Deterimination of a Binomial Weighting Function

Procedure

(1) Given (in appropriate units, e.g. kilometers):

(a) The observation interval Δx

(b) The "cut-off" wavelength

(2) Set the cut-off wavelength = 3σ

(3) The interval between observations then corresponds to

$$\frac{\Delta x}{\sigma}$$

on the abscissa of the normal curve.

(4) Record (e.g., CRC Standard Math Tables, 12th ed., pp. 245-249) the ordinates of the normal curve from zero to at least $\pm 3\sigma$ in increments of $\frac{\Delta x}{\sigma}$.

(5) Adjust the values obtained to yield a weighting function which sums to unity by dividing by $\frac{\sigma}{\Delta x}$.

Example

(1) a) Suppose data are available at every 0.1 km. (e.g. $\Delta x = 0.1$).

b) On a 1/500,000 scale map features having a wavelength of less than one half km. cannot be shown. The cut-off frequency is 0.5 km.

(2) $3\sigma = 0.5 \rightarrow \quad \sigma = 0.167$ km.

(3) $\frac{\Delta x}{\sigma} = \frac{0.1}{0.167} \rightarrow \sim \frac{0.1}{0.2} = 0.5$

(4) From the table of the normal curve:

Abscissa	Ordinate	Normalized	Weights
0.0	0.3989	0.1994	Wo
0.5	0.3521	0.1760	$W_{\pm 1}$
1.0	0.2420	0.1210	$W_{\pm 2}$

Continued on next page

Abscissa	Ordinate	Normalized	Weights
1.5	0.1295	0.0648	$W_{\pm 3}$
2.0	0.0540	0.0270	$W_{\pm 4}$
2.5	0.0175	0.0088	$W_{\pm 5}$
3.0	0.0044	0.0022	$W_{\pm 6}$

The 13 weights obtained can be used to smooth the data. Note that by taking $\frac{\Delta x}{\sigma} = 0.5$, the map is over generalized, since $3\sigma = 0.6$ km not 0.5 km as specified in the initial conditions. This is hardly of consequence since perfect cut-off filters cannot be realized in practice (cf. Holloway).

References

B. J. Garnier, 1963, Practical Work in Geography, London, E. Arnold Co.

A. V. Hershey, 1963, The plotting of Maps on a CRT Printer, NWL Report No. 1844

J. Harbaugh and F. Preston, 1965, "Fourier Analysis in Geology", Computers and Computer Applications in Mining and Exploration, J. Dotson and W. Peters, eds., Tuscon, University of Arizona, Vol. I, pp. R1 - R46.

J. L. Holloway, 1958, "Smoothing and Filtering of Time Series and Space Fields", Advances in Geophysics, 4, New York, Academic Press, pp. 351-389.

E. Imhof, 1957, "Generalizierung der Hohenkurven", Kartographische Studien, Erganzungsheft Nr. 264, Pet. Geogr. Mitt., pp. 89-99.

W. Krumbein, and Graybill, 1965, An Introduction to Statistical Models in Geology, New York, McGraw-Hill.

D. H. Maling, 1963, "Some Quantitative Ideas about Cartographic Generalization" ICA Bulletin, No. 4.

W. D. Montgomery, and P. W. Brome, 1962, "Spatial Filtering" Journal of the Optical Society of America, 52, 11, pp. 1259-1276.

A. J. Pannehoek, 1962, "Generalization of Coastlines and Contours," International Yearbook of Cartography, II, pp. 55-74.

J. Perkal, 1958, "Proba Obiektywnej Generalizacji", Geodezja i Kartografia, VII, 2, pp. 131-142.

W. Pillewizer, and F. Topfer, 1964, "Das Auswahlgesetz: Ein Mittel zur Kartographischen Generalisierung", Kartographische Nachrichten, 14, 4, pp. 117-121.

W. Tobler, 1964a, "A Polynomial Representation of Michigan Population", Papers, Michigan Academy of Science, Arts, and Letters, XLIX, pp. 445-452.

W. Tobler, 1964b, An Experiment in the Computer Generalization of Maps, Technical Report #1, Office of Naval Research task No. 389-137.

A. Wiin-Nielsen, 1965, "Smoothing and Filtering of Two-Dimensional Fields", in A. Dingle and C. Young, Computer Applications in the Atmospheric Sciences, Ann Arbor, University of Michigan (NSF G-25204), pp. 214-221.

E. Whittaker, and G. Robinson, 1944, The Calculus of Observations, 4th ed. London, Blackie and Son, pp. 285-316.

Notes on the Analysis of Geographical Distributions

W. R. Tobler
Department of Geography
The University of Michigan
Ann Arbor

ABSTRACT

Theory implies the existence of a geographical periodicity in the arrangement of cities and other orders of service facilities, and to some extent ever predicts the frequencies of these geographical cycles. Even the casual cross-country tourist recognizes this geographical repetitiveness of events. Appropriate methods of investigation for the analysis of periodicities in geographical phenomena can be organized about the concept of a spatial series.

Paper delivered at the Institute on Emerging Concepts and Methods in Urban and Regional Analysis, East Lansing, 16 July 1965.

Notes on the Analysis of Geographical Distributions

Increasing attention is being devoted to the solution of regional and metropolitan problems. A common feature, recognized by the individuals responsible for such studies, is that the processes involved take place in a geographical context. A result is that vast amounts of geographical information are being accumulated. The convenient storage and retrieval of this information is being made possible by the introduction of large computing centers. A great deal of effort is being expended in this retooling for an increased informational capacity. At the same time effort is being devoted to consideration of methods of analysis appropriate to the expanding storage, retrieval, and processing capabilities. Some progress in this direction has been made in the field of geographical (spatial, regional, or areal) information processing. One of the encouraging developments has been the recent rise in the amount of economic, social, and cultural information identified in a locatively useful manner, as for example, the recent provision on the part of the U.S. Bureau of the Census of latitude and longitude coordinates for population enumerations. This allows direct computer analysis in greater volume and at less cost than previously employed methods and eliminates the need to rely heavily on nomographic approximations obtained from maps, the geographical version of the graph. At the same time the identification of data by geographical coordinates allows the automatic preparation of maps for those types of visual pattern recognition and speculative hypothesis-generation operations which the human can perform far better than current computers. A merging of the advantages of both formulations, the analytical and graphical, is represented by "sketch-pad" computing facilities.

Given the complexity of many geographical problems, and the fact that we would hope to be able to solve at least some of these in automatic real-time environments in the future, it is clear that this convenient locative identification is a step in the direction of the geographical information processing capacity required to cope with the magnitudes of data becoming available from "information centers" or from such recently developed geographical information collection systems as earth satellites. The megalopolis region may well find it to be economically beneficial to invest $4,000,000 for an orbiting (that is, stationary) sensor as a part of a larger information processing system.

There are many modes of organization for the analysis of geographical distributions. One which is common in the physical sciences but which is rarely applied in the social sciences considers space to be continuous (not compartmentalized into "regions"). By analogy with time series analysis, one can define a subject called geographical (or spatial) series analysis. Almost every textbook on economic statistics has a section on time series analysis; none has a section on spatial series analysis. There are many texts and monographs on time series. Again, there are none on geographical series. For each of the methods of analyzing time series data it seems possible to consider the equivalent for a geographical series. A first problem is that the mathematical extension often is difficult. Secondly, it is at least as critical that a geographical interpretation be supplied for the suggested manipulations. Generally this is simpler than would be imagined. Central place theory, for example, clearly implies that there exists a geographical periodicity in the arrangement of cities and other orders of service facilities, and to some extent even predicts the frequency of these cycles. It is apparent that methods of spectral (or harmonic) analysis would be of great assistance in the empirical verification of this important theory.

Table I is a presentation of some temporal data, arranged in alphabetical order. The first step in the analysis of this information might be to rearrange the table into its correct temporal sequence, and this is

TABLE II **

Anhwei	30	54
Chekiang	23	39
Chinghai	2	278
Fukien	13	48
Heilung-Kiang	12	179
Honan	44	65
Hopei	36	82
Hunan	33	81
Hupei	28	72
Kansu	13	167
Kiangsi	17	64
Kinagsu	41	41
Kirin	11	72
Kwangsi	20	85
Kwantunk	35	89
Kweichow	15	67
Lisoning	19	58
Shansi	14	61
Shantung	49	59
Shensi	16	76
Sinkiang	5	635
Szechwan	63	219
Yunnan	18	168

** China, population in millions and area in thousands of square miles.

TABLE I *

April	20
August	115
December	2
February	11
January	4
June	112
July	121
March	14
May	50
November	5
October	43
September	90

* Calcutta, average rainfall in tenths of inches.

rather easily done. Table II is a presentation of some geographical data, and the reader is asked to perform the comparable operation; arrange the data in its correct spatial sequence. The geographical ordering is the more difficult. Virtually everybody is familiar with the temporal sequence of the partitioning of the year into months, but few people would know the correct spatial sequence of the political partitioning of even as important a country as China. Another difficulty arises from the fact that the geographical sequence is quite irregular. Possibly one reason that geographical information is so frequently given in alphabetical order is that the printer would find a "geographical table" (with every entry in its correct spatial position) too irregular and with too many blank spaces for his symbol positioning techniques.

Related to the foregoing considerations is the disparity in size of intervals. For temporal data one would hardly collect information in a manner such that the first interval spans 30 days, the next 60 days, the next 25 seconds, the next 2 days, then 15 minutes, and so on. Instead one uses intervals of equal, or nearly equal, size. A similar strategy has been proposed (Hägerstrand, 1955) for geographical data collection but is rarely found.

A next step (recommended by most textbooks) in the analysis of temporal data is the preparation of a graph or histogram. Typically the data are assigned to the midpoint of the interval and represented with an ordinate proportional to the observations. A smooth curve is then drawn through the resulting points. Comparable procedures are employed for geographical distributions, with the results shown as a perspective diagram or in a plan view. The contour map of smooth level curves is of course the two-dimensional equivalent of the smooth curve on the ordinary graph. A difficulty arises here in choosing the center of the spatial interval, which has a shape as well as a size, and is therefore more complicated than in the case of temporal intervals. The two-dimensional interpolation required to obtain smooth level observations is also more complicated than the one-dimensional interpolation necessary for the temporal graphs.

From this simple introduction we can continue the analogy with time series in a straight forward manner. A continuing complication is the spacing of the observations, however. Several situations may arise: observations are obtained at irregular intervals (either by the sampling technique employed, or by the data collection agency, or because of the inherent nature of the events), or observations are obtained or assigned (by some interpolation technique) to regular grid points. These grid points may be either a square or a trigonal lattice on a plane, or polar coordinates, or an adaptation of latitude and longitude on a sphere or spheroid. It suffices here to consider only plane phenomena, and there is great temptation to consider only square grids since the mathematics are simplified. The greatest differences are between the square grid and the irregular scatter of observations. It is generally a simple matter to adapt the techniques to a trigonal (or hexagonal) grid, even on a computer, once they have been developed for a square grid. Data obtained by reference to polar coordinates can be converted to rectangular coordinates and then analyzed as are the irregularly scattered observations, though this is a rather unnatural procedure.

A simple smoothing technique often applied to time series data is the moving average. Here each observation is averaged with its neighbors, which may or may not be weighted. In the geographical case, and for a square grid, this may involve 5,9,13,17,21,25...points, or more typically, the k^{th} weighted average involves $(2k + 1)^2$ points. This can be written as

$$\overline{Z}_{ij,k} = \frac{\sum_{m=-k}^{m=k} \sum_{n=-k}^{n=k} W_{m,n} Z_{i+m,j+n}}{\sum_{m=-k}^{m=k} \sum_{n=-k}^{n=k} W_{m,n}}$$

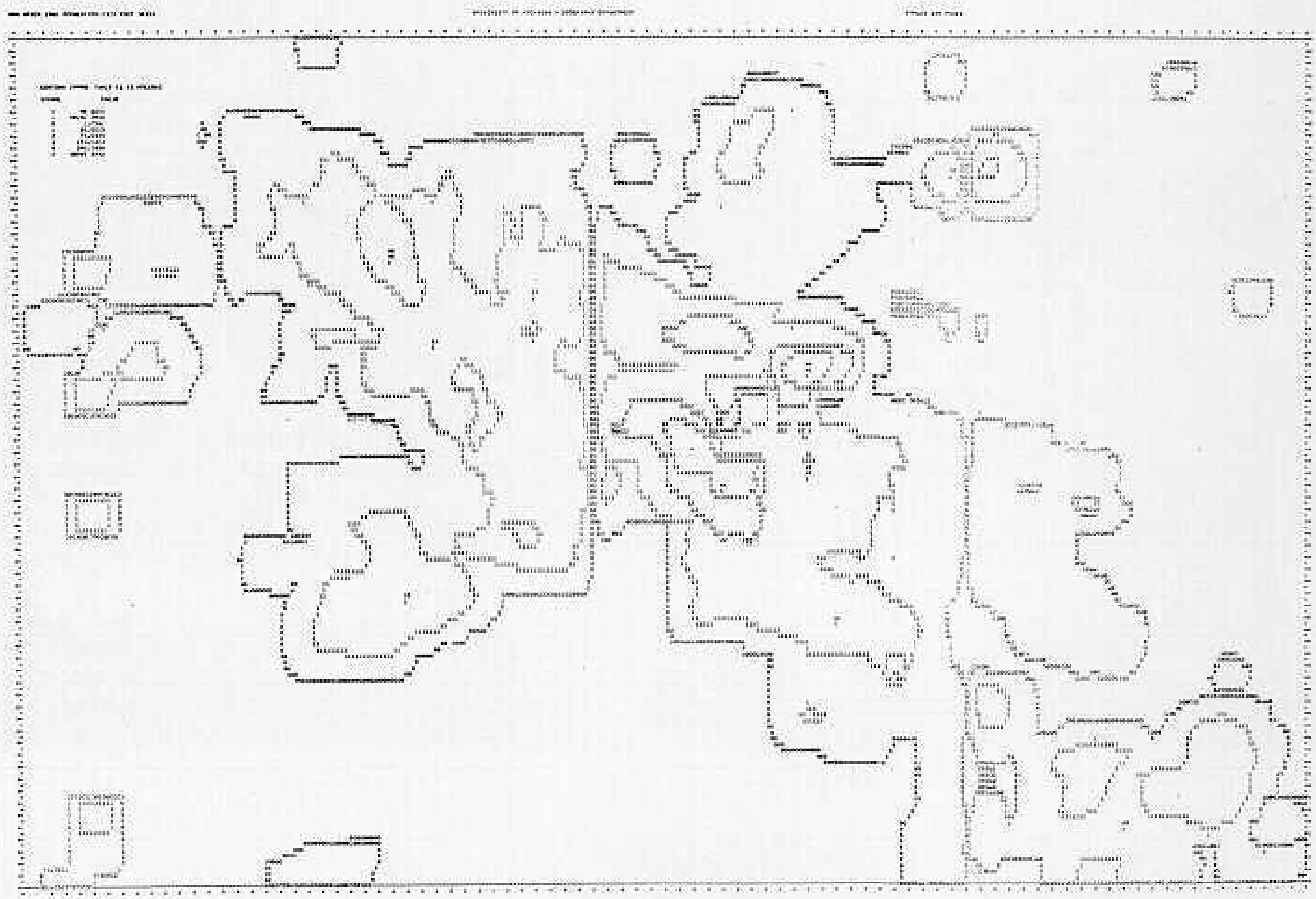
where W is an appropriate weighting factor. This operation is repeated for all grid points, except boundary points. For irregularly scattered observations a moving average might be established by averaging nearest (weighted or unweighted) neighbors, which is again rather simple.

As an example of the geographical use of a moving average, Mr. Yuill, a graduate student in the Department of Geography, recently compiled population figures for the city of Ann Arbor on a 300 foot grid. The population of each city block (or fraction thereof) was assigned to the grid point(s) which fell on that block. Grid points which landed in the street were assigned a value of zero, which caused extreme fluctuations in the data. A moving average was employed to smooth these observations enabling the computer to produce a population density contour map. (Another strategy of course would have been to use a floating template of larger than grid size in the initial determination of the population densities).

The moving average given above is a local or neighborhood operator (in this case a smoothing) in which k specifies the region considered to be local. As this neighborhood is enlarged the operator tends to become more and more global (the mathematical concepts of local and global are vaguely akin to the geographical concepts of site and situation). Other local operators may be of interest. For example,

$$\sigma^2_{ij,k} = \frac{\sum_{m=-k}^{m=k} \sum_{n=-k}^{n=k} (\overline{z}_{ij,k} - z_{i+m,j+n})^2}{(2k+1)^2}$$

will define a local variance, from which a contour map might be produced. This is of some theoretical interest since the variance may be regarded as the integral of the power spectrum over all frequencies. The local variance may also be regarded as a measure of texture (on aerial photographs) or a measure of roughness (in topographical situations). The local range, known from geographical studies of terrain as the "relative relief", can be defined



in a similar manner. Numerous additional local operators of interest can be defined.

Two asides may be appropriate here. One concerns the development and construction of a "parallel" computer at the University of Illinois. In contrast to most serial computers, which would have to apply the foregoing operations to all points in sequence, the parallel computer can apply the operations to all points simultaneously, with a resultant reduction in cost. The second point is of more theoretical interest and concerns the boundaries. It is clear that the smoothing operation cannot work in the vicinity of the edges of the region of observations. This is true of several of the operators discussed. The implication is that the region of observations should exceed the region of interest. For example, an agency concerned with the state of Iowa should also have observations from a strip completely surrounding Iowa.

In the analysis of temporal phenomena one is often concerned with the rates of change. A similar technique is available for geographical phenomena. This might, for example, be the rate of change of land values per mile. Clearly this depends on the direction it is natural to consider the maximum rate of change per unit distance (the gradient, orthogonal to the contours). For gridded data the gradient at the ij^{th} point can be calculated from

$$|\Delta Z| = \left(\left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2 \right)^{1/2}$$

where

$$\frac{\partial z}{\partial x} = \frac{Z_{i,j+1} - Z_{i,j-1}}{2\Delta x}$$

$$\frac{\partial z}{\partial y} = \frac{Z_{i+1,j} - Z_{i-1,j}}{2\Delta y}$$

This can again be contoured or can be used to compute the rate of change of the rate of change. This second spatial derivative, along with discriminatory analysis, information theory, and several other techniques, is useful in boundary delineation. In addition to the gradient, it may be of interest to calculate the directional derivative, as for example, the rate of change per mile of the population density from the central business district.

A fundamental question in the analysis of temporal data is whether or not there is a trend. This generally involves fitting a straight line to the graph of the observations, though a curvilinear trend is occasionally employed. Similar techniques can be employed for geographical phenomena. For example, the trend of population density (persons per square mile) in Michigan is

$$P = 6314 - 68.7*\text{latitude N} - 36.9*\text{longitude W},$$

or, in words, the population declines towards the northwest, which is not a particularly astounding result, though useful. For example, knowledge concerning the trend can be employed to adjust for population density in a rather simple manner, and the departures from the trend (residuals or anomolies) can be examined in greater detail. The technique of trend analysis, as employed in practice, is basically a linear or curvilinear least-squares curve-fitting procedure (multiple regression with geographical coordinates as the independant variables).

As another example of the utility of this technique consider the Von Thünen model of agricultural land use. The statement of the model is:

a) Commodities are chosen to maximize income.

b) $I = M-P-T$,

where

I = income

M_i = market price of the i^{th} commodity

P_i = production costs

T_i = transportation costs from farm to market

This model has been programmed for computer solution by Mr. A. Schwarz, a sophomore student at the University of Michigan. As formulated, the computer program allows many market locations and commodities. Transportation costs are linear functions of the distance from the market, and production costs depend only on the commodity. An obvious improvement in the program would be to allow production costs to be a function of the productiveness (fertility, etc.) of each location. A simple procedure is to enter the production costs as geographical trend equations. A climatic gradient could easily be represented as an equation similar (perhaps non-linear) to that given for population density in Michigan. Simple sinusoids could provide a first approximation to rolling topography, and so on. The final restrictions are basically limitations on data availability. The curve fitting technique can be extended to more and more complicated equations so that one may arrive at a "complete" mathematical equation which represents the observations in question to a high degree. Such "numerical maps" are employed in several fields. An alternate, but related direction is the fitting of theoretically derived curves, such as the S-shaped curve in demographic studies or the exponential decay function for the distribution of population within cities.

A practical application of these techniques might be the greater acceptance (and availability) of regional and metropolitan data which have been "adjusted" for regional trends. The net effect of this would be a reduction of regional complexity toward the "homogeneous plain" concept now employed in theoretical work of some elegance (this is perhaps somewhat comparable to the meteorologists "reduction to sea level", an effort to have comparable data despite the vagaries of topography). Presumably such adjustments for geographical trends would simplify some aspects of the current complexity of the environment which in turn might aid in the solution of numerous important practical problems of development and planning.

The economist often considers that events at time t may be related to events at some previous time t-k. It is not unreasonable to expect that events at one location (ij) may be related to events at some other geographical location (i+m, j+n). We are thus lead to a consideration of serial (or auto-)

correlation of each observation with, for example, its neighbor immediately to the east, i.e.

$$r\ (Z_{ij},\ Z_{i+1,j})$$

or with the neighbor immediately north

$$r\ (Z_{ij},\ Z_{i,j+1})$$

or with the neighbor two removed to the east

$$r\ (Z_{ij},\ Z_{i+2,j})$$

and so on. Consider all possible such lagged correlations. Now plot a contour map of the correlation coefficients as a function of the spatial lags in the two directions (i.e., $r = r(m,n)$, where $m = \pm 1, \ldots, n = \pm 1, \ldots$). "Peaks" on this correlogram indicate similarities of events, and the symmetry (or lack of symmetry) about the origin of the contoured correlation coefficients relates to the isotropicity of the space. This is somewhat difficult to interpret, but an appropriate question might be: Does the (economic, cultural, social) landscape repeat itself, and if so, how often and in which direction? As another aside, attention is drawn to the recent development of optical computers which perform this correlation operation (and several of the operations to be discussed shortly) extremely rapidly, and which involve no complicated computer program in the conventional sense.

For scattered and irregular observations the serial correlation procedure might be attempted in several ways. The most obvious is to correlate each observation with its nearest neighbor, perhaps adjusting for their distance apart. Obviously the process is symmetrical, which reduces the number of samples available. A second approach might be to correlate only with nearest neighbors within some sector. This should bring out directional tendancies more clearly than the previous method. Finally, correlation of all neighbors within some distance increment could be attempted. This involves a variable number of observations at each stage of the process, but seems appropriate for geographical situations in which some phenomenon is spreading from central places (city, or university, etc.).

It is often of interest to compare two sets of time series data with each other (e.g., stock market prices and consumer price index; or tree rings with weather records, etc.). This is known as cross-correlation. The geographical equivalent is the comparison of two maps, of either the same region or of different regions. A caveat should be entered here. Geographical patterns may be considered to consist of points, lines, areas, intensities, or flows. This analysis, if correct, would lead to 25 distinct types of intercomparisons. Among these might be, for example, comparison of the road pattern of the U.S. (lines) with the pattern of cities (points), or with the pattern of railroads (lines), or with the pattern of "depressed" areas (areas), or with the pattern of population densities (intensities), and so on. Most of the statistical techniques now available allow only comparison of like entities. Comparisons of pairs of point patterns, for example, has recently been shown to be possible via ordinary correlation and regression techniques, but extended to the domain of complex numbers. Similarly a two-dimensional version of rank correlation has been devised for the comparison of areal patterns, and a tensor correlation has been suggested for comparisons of intensity patterns. The cross-correlation technique, as employed in time-series analysis, really applies only to sets of observations which are considered (sampled) continuous functions. The geographical equivalent might be two maps showing, say, rainfall intensities and per acre value of agricultural land. The most common technique is to superimpose the two maps and calculate the immediate correspondences, that is, with both sets of data (Z and W) on a grid, calculate $r = r\ (Z_{ij}, W_{ij})$ for all ij.

The two dimensional cross-correlation goes one step further. Like the autocorrelation technique it shifts one of the maps around and produces a correlogram giving the correlation coefficients as a function of the spatial lags in two directions, viz

$$r = r(m,n) = r\ (Z_{ij}, W_{i+m,j+n}).$$

As a geographical example, the highest incidence of irritation from air polution might be downwind from the highest intensity of automobile traffic. Another application might be to attempt to find the "best" positioning of two maps relative to each other in order to "compare" two different cities. Comparisons of topographical map sheets with a file of "standard" physiographic types could also be performed via the cross-correlation technique.

The previous comments regarding autocorrelation for irregularly spaced observations also appear to apply to cross-correlation functions. The usual cross-correlation technique for gridded data involves only translations (no reflections or rotations) and therefore has several limitations, as workers in the field of pattern recognition have become aware: Letters which are upside down cannot be recognized, for example, and there may be geographical situations of interest which would not be detected because of this limitation. The use of two-dimensional convolutions may be required, and the statistical validity of cross-correlations must be evaluated by use of the relatively complicated coherence measures.

Seasonal trends, or business cycles, are often of interest in the analysis of economic time series. One approach to these phenomena is to adjust the data by taking out these fluctuations. For cyclical events the use, and empirical fitting, of Fourier series is common. Computer programs are now available for the fitting, to geographical distributions, of so-called double Fourier series. These will yield descriptive equations involving trigonometric functions. The next step requires some rather delicate analyses to decide whether periodicities actually exist or whether the data simply contain quasi-periodicities, or only apparent periodicities, or none of these. The mathematical details become rather complicated here, but the autocorrelation function can be shown to be related, via the Fourier transform, to the power spectra of the "waves". The interpretation, of course, is that central place theory does suggest a geographical periodicity in the arrangement of cities, stores, hospitals, and so on. This can be seen if one imagines and considers the highly regular and repeating pattern of the population density contours implied by Christaller's original statement of central place theory. In consequence, the search for geographical periodicities is not simply a vacuous empirical or mathematical exercise. A complicating factor, of course, is the geometry of geography, which is warped by transportation facilities and other real conditions. As a remark, the analysis of data for periodicities seems to be practical only when the observations are available (so collected or assigned) at uniform intervals. Choice of the appropriate grid interval appears to be related to the frequency expected, and the effects of aliasing must be considered.

Assume that central place periodicities could be estimated (at some level of statistical confidence) from empirical observations of, say, the population distribution in the United States. Aside from the obvious expectations regarding wave lengths and phases implied by the theory, additional validity tests are available. One can consider the population to consist of two sectors, urban and rural. These then should correspond rather well to a distinction between central place and non-central place populations. Consequently one would expect that subtraction from the total population of that portion of the population which could be attributed to the geographical cycles should leave as a remainder the "rural" population. Since the U.S. Bureau of the Census distinguishes between urban and rural population the empirical information is available for the test.

In summary, it appears that valid interpretations can be attached to geographical extensions of the methods of analysis developed for the study of time series.

Select References

J. L. Alward, "Notes on Spatial Discrimination", Engineering Summer Conference on Advanced Infrared Technology, lecture notes, Ann Arbor, June 1965.

R. Bachi, "Standard Distance Measures and Related Methods for Spatial Analysis", *Papers*, Regional Science Assn., X(1962), pp. 83-132.

M. S. Bartlett, "The Spectral Analysis of Two-Dimensional Point Processes", *Biometrika*, Vol. 51, Parts 3 and 4, Dec. 1964, pp. 299-311.

B. J. L. Berry, and A. Pred, *Central Place Studies*, Philadelphia, Regional Science Research Institute, 1961.

R. A. Bryson, and J. A. Dutton, "The Variance Spectra of Certain Natural Series", *Quantitative Geography*, W. Garrison (ed.), forthcoming.

Jon Cunnyngham, "The Spectral Analysis of Economic Time Series", Washington D.C., U.S. Bureau of the Census Working Paper #14, 1963.

L. J. Cutrona, E. N. Leith, C. J. Palermo, "Optical Data Processing and Filtering Systems," *IRE Trans. Inform. Theory*, Vol. IT-6, No. 3, June 1960, p. 391 et seq.

O. D. Duncan, R. P. Cuzzort, and B. Duncan, *Statistical Geography: Problems in Analyzing Areal Data*, The Free Press of Glencoe, Illinois, 1961.

E. S. Epstein, and J. A. Leese, "Application of Two-Dimensional Spectral Analysis to the Quantification of Satellite Cloud Photographs", Ann Arbor, University of Michigan (UM 05068-1-F), 1963.

T. Hägerstrand, "Statistika Primaruppgifter, Flygkartering Och 'Data Processing' Maskiner: Ett Kombinerings projekt", *Meddelanden Fran Lunds Geografiska Institution*, Nr 344, Lund, University of Lund, 1955, pp. 233-255.

J. W. Harbaugh, and F. W. Preston, "Fourier Series Analysis in Geology", *Computers and Computer Applications in Mining and Exploration*, J. Dotson and W. Peters (eds.), Tucson, University of Arizona, 1965.

J. L. Holloway, "Smoothing and Filtering of Time Series and Space Fields," *Advances in Geophysics*, Vol. 4, New York, Academic Press, 1958.

IBM Corporation, "Numerical Surface Techniques and Contour Map Plotting", E. 20-0117-0, undated (1964?).

W. C. Krumbein, "Trend Surface Analysis of Contour-Type Maps with Irregualr Control-Point Spacing", *Journal of Geophysical Research*, 67, 7(1959), pp. 823-834.

______________, and F. A. Graybill, *An Introduction to Statistical Models in Geology*, New York, McGraw-Hill, 1965.

Y. W. Lee, *Statistical Theory of Communication*, New York, J. Wiley, 1960, 501 pp.

M. Masuamya, "Correlation between Tensor Quantities", *Proceedings*, Physico-Math. Soc. Japan, 3rd Series, 21 (1939), pp. 638-647.

H. H. McCarty, and N. E. Salisbury, "Visual Comparison of Isopleth Maps as a Means of Determining Correlations Between Spatially Distributed Phenomena", Iowa City, Department of Geography, State University of Iowa, 1961.

R. L. Miller, and J. S. Kahn, Statistical Analysis in the Geological Sciences, New York, J. Wiley, 1962.

R. F. Minnick, "A Method for the Measurement of Areal Correspondence", Papers, Michigan Academy of Science, Arts, and Letters, XLIX (1964), pp. 333-344.

W. D. Montgomery, and P. W. Brome, "Spatial Filtering", Journal of the Optical Society of America, 52, 11 (1962), pp. 1259-1276.

H. A. Panofsky, and G. W. Brier, Some Applications of Statistics to Meteorology, University Park, Pennsylvania State University, 1958.

W. J. Pierson, ed., "The Directional Spectrum of a Wind Generated Sea as Determined from Data Obtained by the Stereo Wave Observation Project", Meteorological Papers, New York University, Vol. 2, No. 6.

A. H. Robinson, Mapping the Correspondence of Isarithmic Maps", Annals, Assn. Am. Geographers, 47 (1957), pp. 379-391.

A. Rosenfeld, and J. Pfaltz, Sequential Operations in Digital Picture Processing, TR-65-18, Computer Science Center, University of Maryland, 1965.

S. M. Simpson, Jr., "Least Squares Polynomial Fitting to Gravitation Data and Density Plotting by Digital Computers", Geophysics, XIX (1954), pp. 250-257.

P. Switzer, C. M. Mohr, and R. E. Heitman, "Statistical Analyses of Ocean Terrain and Contour Plotting Procedures," Cambridge, A. D. Little (AD 601538), 1964.

E. N. Thomas, "The Stability of Distance - Population - Size Relationships for Iowa Towns from 1900-1950," IGU Symposium in Urban Geography, K. Norborg (ed.), Lund, Gleerup, 1962.

E. N. Thomas, "Maps of Residuals from Regression: Their Characteristics and Uses in Geographic Research", Iowa City, Department of Geography, State University of Iowa, 1960.

W. R. Tobler, "A Polynomial Representation of Michigan Population", Papers, Michigan Academy of Science, Arts, and Letters, XLIX, 1964, pp. 445-452.

____________, "Computation of the Correspondence of Geographical Patterns", Ann Arbor, 1964 Regional Science Association Meeting.

J. W. Tukey, and R. B. Blackman, The Measurement of Power Spectra, New York, Dover, 1959.

N. Wiener, Extrapolation, Interpolation and Smoothing of Stationary Time Series, Cambridge, MIT Press, 1949.

A. Winder, and C. Loda, Introduction to Acoustical Space-Time Information Processing, ONR report ACR-63, Washington, GPO, 1962.

FOREWORD

Peter Gould presented a most stimulating lecture to the Michigan Inter-University Community of Mathematical Geographers seminar in Brighton, on the evening of the thirteenth of April, 1966. We succeeded in convincing Professor Gould that he should prepare these materials for the larger discussion paper audience, and are pleased to present the results.

The paper unquestionably falls within the field of cultural geography; not the cultural geography of the unique or esoteric, but a cultural geography which is concerned with the structuring of evidence in a manner which is consistent with the scientific criteria of generality and parsimony. Examining numerous questionnaires Gould finds that personal idiosyncracies, the residuals from the components analyses, cannot account for more than fifty percent of the variability contained in the data; in most cases they account for considerably less than fifty percent. People have a great deal in common. This suggests that empirically valid general models of geographical perception are feasible. One of the models implied in the paper relates migration to gradients of perception surfaces. It is then interesting to speculate on the degree to which perception and behavior are related.

The greatest compliment which any of our readers can accord the author will be to build upon the foundation provided. Will you be the one to take up this challenge?

ON MENTAL MAPS

Peter R. Gould
The Pennsylvania State University

> "Can geography be mixed up with psychology . . . ?"
>
> in Luigi Barzini, The Italians (New York: Bantam Books, Inc., 1965), p. 58.

At the start of this limited enquiry into the mental images that men have of geographic space I offer you no theories or even explicit hypotheses. On the contrary, you will find only unstructured, intuitive hunches, and interpretations pushed, in many cases, beyond the limits that the data allow. Of these speculations, for, indeed, this is what they are, many will undoubtedly be wrong. They will have served their purpose, and more, if they stimulate others to replace them with better notions. As usual, what we really need are more penetrating questions, but before these can be asked we must record what we do and do not know. The boundary of ignorance is not very far away, but it seems only sensible to stake it out before we try to cross it. We know so very little about the spatial images, the mental maps, that are in the minds of men. We know even less about how they are formed, the degree to which they are unique or general, and the way they impinge upon, and are reflected in, the decisions that men make. As human geographers reach out across traditional disciplinary boundaries to the other social and behavioral sciences, it is increasingly apparent that the truly satisfying explanation they seek is going to come from emphasizing the human as much as the geography. We may, perhaps, define our subject as essentially that which tries to understand the spatial aspects of Man's behavior.[1] If we grant that spatial behavior is our concern, then the mental images that men hold of the space around them may provide a key to some of the structures, patterns and processes of Man's work on the face of the earth. The emphasis upon the conditional tense is quite deliberate, and is only partly a result of intellectual cowardice and a general propensity to broadcast *caveats* in lieu of signing academic insurance policies. The other reason is that we really do not know whether mental maps are relevant to our problems. But the suspicion that they are is strong, and at the very least it seems worthwhile making some tentative probes along these lines.

IMAGES AND DECISIONS

The human landscape, in reality, or abstracted and modelled as a map, is nothing more, but equally nothing less, than the spatial expression of the decisions of men. As we examine even the most apparently superficial spatial patterns and processes that are a reflection of these decisions, we quickly become aware of the extreme complexity that underlies them, the myriad of variables that compete for attention and the way in which these form interlocking and convoluted structures that are numbing in their difficulty. Many of the decisions that men make seem to be related, at least in part, to the way in which they perceive the space around them and to the differential evaluations they place upon various portions of it. For the moment this is a bald and unsupported assertion, but it seems reasonable that the manner in which men view their spatial matrix impinges upon and affects their judgements to some degree. For example, men decide to migrate, not on a

---

[1] For an interesting, and rather similar view in history see: Lee Benson, "Causation and the American Civil War: Two Appraisals", *History and Theory*, Vol. 1, 1961, p. 163.

regular surface of equal opportunity or desirability, but in a world often perceived in an extreme, differential manner.[2] Men decide to grow crops and raise animals for their sustenance, not in an arbitrary way, but in part according to their particular views of the space around them. Men decide to locate their industries and business activities, and with more and more "footloose" industries coming onto the scene we are finding that traditional location factors are declining in importance. Törnqvist, for example, has indicated the virtual irrelevance of distance for some industries in Sweden,[3] and in this country Harris's classic paper indicated the vast area in the American Manufacturing Belt that lies around the point of minimum transport cost with only slightly higher access costs to the market.[4] Even traditional Weberian analysis discloses a basic characteristic of many extremum problems, with a large area of only slightly higher aggregate cost around the minimum point. What Rufus Isaacs has termed the "principle of flat laxity" seems to be operating extensively in geographic space,[5] and takes on new meaning as we become critically aware of what satisficing behavior in a spatial context implies. Thus, in view of the decline in importance of the more traditional location factors, might not the decision to locate be increasingly related to the image an area has in the

---

[2]Julian Wolpert, "Distance and Directional Bias in Interurban Migratory Streams", paper presented to the annual meeting of the Association of American Geographers, Columbus, Ohio, 1965.

[3]Gunnar Törnqvist, Aktiv Lokaliseringspolitik (Stockholm: Iduns Trycker iaktiebolag Esselte ab, 1963), pp. 215-292. See also his "Transport Costs as a Location Factor for Manufacturing Industry", Lund Studies in Geography, Series C. General and Mathematical Geography, No. 2. 1962.

[4]Chauncy Harris, "The Market as a Factor in the Localization of Industry in the United States", Annals of the Association of American Geographers, Vol. 44, December 1954.

[5]Rufus Isaacs, Differential Games (New York: John Wiley and Sons, Inc., 1965).

the minds of a few key people?[6] More and more the quaternary industries, the research and development companies, look to the scenic and recreational facilities, cultural assets and intellectual resources of an area. Snow and mountains are not essential to certain well-known electronic companies, but New Hampshire and Colorado are undoubtedly grateful for their physiographic and climatological inheritance! Similarly, large universities in pleasant surroundings are the locational loadstones for the research and development consultants. It is not difficult to think of many other examples where the maps that are carried in men's heads might be relevant in quite crucial ways.

WHAT DO WE KNOW ABOUT MENTAL MAPS?

Man's view of geographic space is extremely varied, and the views of individual men are always in part unique. Entering into the particular outlook of a particular man are a host of experiences, prejudices and desires, some shared widely with others, some quite specific to the individual. The Northerner is reluctant to be assigned by his company to the South, for he holds to a mental picture that is part of his northern cultural inheritance -- an inheritance absorbed in childhood, and reinforced by his daily sources of information. The townsman, comfortable and safe amidst the roar of traffic and bustle of urban life, is reluctant to live in the green peace of the country, which he associates with the stillness of bucolic decay. The New England family, suddenly presented with greater economic opportunity amongst the tall trees of Oregon, decides to stay with the known view and the familiar friends, for "Oregon is such a long way from civilization"! Thus, the political, social, cultural and economic values held by a man blend into an overall

---

[6] Andrew Wilson, "The Impact of Climate on Industrial Growth: Tucson, Arizona: A Case Study", Chapter 17 in W. R. Derrick Sewell (ed.), Human Dimensions of Weather Modification (Chicago: University of Chicago Department of Geography Research Series, No. 105, 1966), p. 249.

image about the space around him, an image whose components may be particular to him or held in common by many.

It hardly seems necessary to add that we know very little about these spatial images in the minds of men. While a concern for the mental maps of geographic space is nothing new, the literature is extremely sparse. We only have a very small number of examples where they are discussed at all, usually as interesting, but definitely peripheral points in larger investigations. For example, Tobler explicitly raised the question of the mental images that people have of their environment,[7] but his basic concern was for the mental transformations of distance that people make. Lowenthal and Prince have discussed the attitudes of a people towards the visual landscape,[8] while the former, in a synthesis that has yet to be equalled, has examined the relevance of the psychological literature in this area.[9] A number of other geographers have focussed upon the perception of environmental hazard and the spatial implications that such images have for locational decisions.[10] In political geography, only Herman seems to have moved beyond fuzzy speculation to investigate truly the changing values and attitudes of a people towards the national space.[11] Occasionally maps such as the "New Yorker's View of the

---

[7] Waldo Tobler, _Map Transformations of Geographic Space,_ Ph.D. thesis, Department of Geography, University of Washington, 1961, pp. 111-113.

[8] David Lowenthal and Hugh Prince, "English Landscape Tastes", _The Geographical Review_, Vol. LV, No. 2, April 1965, pp. 186-222.

[9] David Lowenthal, "Geography, Experience, and Imagination: Towards a Geographical Epistemology", _Annals of the Association of American Geographers_, Vol. 51, No. 3, September 1961, pp. 241-260.

[10] Ian Burton and Robert Kates, "The Perception of Natural Hazards in Resource Management", _Natural Resources Journal_, Vol. 3, No. 3, January 1964, pp. 412-441.

[11] Theodore Herman, "Group Values Towards the National Space: The Case of China , _The Geographical Review_, Vol. XLIX, No.2, April 1959, pp. 164-182.

United States" appear, but their humorous context actually obscures the fact that such cartograms of mental images can be extremely illuminating if properly used. Getis has shown how shape distortions can focus the eye upon a particular portion of the map,[12] and Mackay had a map in the late fifties showing Canada through French-Canadian eyes. Unfortunately, it was never published.[13] In a related field, only Lynch, as an urban planner truly concerned with the city as the home of man, has systematically investigated the differential images of the urban landscape, and in his most imaginative series of maps we have our only notions of these mental pictures.[14]

The literature in other fields is equally sparse. The psychologists, in their concern for "Perception", have barely touched upon the investigation of mental pictures of geographic space, for many of their efforts have concentrated upon the physics and physiology of the senses, often within highly controlled laboratory conditions. For example, the space with which Sandström was concerned was far removed in scale from geographical space,[15] and his insights on disorientation and loss of criteria for making locational judgements under formal experimental situations is hard to carry over in any meaningful sense to world scales. Even Hull's work, though a source of extremely stimulating analogy, does not deal with the larger

---

[12] In William Bunge, Theoretical Geography (Lund: C.W.K. Gleerup, 1962), Lund Studies in Geography, Series C. General and Mathematical Geography, No. 1, p. 43.

[13] Forrest Pitts, private communication.

[14] Kevin Lynch, The Image of the City (Cambridge: MIT Press, 1961).

[15] Carl Sandstrom, Orientation in the Present Space (Stockholm: Almqvist and Wiksell, 1951), pp. 138-147.

space of the earth's surface.[16] Nor did the swing of some psychologists to the Gestalt outlook produce a shift in the focus of their concern, and the discussions of space by this particular school do not really deal with the larger area of the earth's surface which is the geographer's realm. Though the pioneering work of the child psychologists Piaget and Inhelder is directly concerned with the way in which children learn about space, the world around them, and geometrical and topological concepts,[17] it does not deal with the essentially geographic images that children hold or the way they learn about them. Only Trowbridge's paper on imaginary maps, written half a century ago,[18] specifically raises the question of the spatial images people carry around in their heads. Unfortunately, this line of investigation was never followed up, and his paper represents the solitary gold nugget at the bottom of the psychological pan. The rest is a residue of vaguely structured insight that hardly rewards the effort of panning it out from the ground material.

In other areas the prospect is equally bleak. The mythical space of Cassirer, though treated as a mental construct, seems less than useful to the geographer,[19] and the Weltanschauungen, or world views that have been discussed by other philosophers may be splendid flowers of many hues, but they are

---

[16]Clark L. Hull, A Behavior System: An Introduction to Behavior Theory Concerning the Individual Organism (New York: John Wiley and Sons, Inc., 1964), pp. 215-274.

[17]Jean Piaget and Barbel Inhelder, The Child's Conception of Space (London: Routledge and Paul, 1956); Jean Piaget, Barbel Inhelder and Alina Szeminska, The Child's Conception of Geometry (New York: Harper Torchbooks, 1964); Jean Piaget, The Child's Conception of the World (Patterson, N.J.: Littlefield, Adams and Co., 1963).

[18]C. C. Trowbridge, "On Fundamental Methods of Orientation and Imaginary Maps", Science, Vol. 38, No. 990, 1913, pp. 888-897.

[19]Ernst Cassirer, The Philosophy of Symbolic Forms: Volume II: Mythical Thought (New Haven: Yale University Press, 1955), pp. 83-94.

difficult to transplant into the hard earth with which the geographer deals.

THE QUESTION OF UNIQUENESS AND GENERALITY OF VIEWPOINT

Because the total experiences of individual men are unique it might seem, at first glance, that they perceive the world around them in quite distinct, totally individualistic ways. But, if this were really so, it would be impossible to say anything of general, and, therefore, of scientific worth about their spatial perception. Though this statement may sound almost tautological, it does raise, by jarring our commonsense experience, the notion that the views of men are not, in fact, totally disparate. We may disagree with some about the desirability or undesirability, the beauty or ugliness, of a particular place, but we can be almost certain of finding someone whose "view from the bridge" closely parallels our own. Perhaps, then, this is the key: a portion of our viewpoint is quite particular to ourselves, while another part is shared, or held in common, with many of our fellows.

Given some information about the preferences of a group of people for various portions of an area, we require a way of separating out the general or shared portion of their perception from that which is quite specific to them individually. Putting it another way, we would like to partition the total variation in space preferences for a given sample of people into those portions that indicate general or common viewpoints, and those that represent unique portions that may be assigned to individuals themselves. It is for this reason that the problem has been approached through principal components analysis.[20]

---

[20] Many standard works are now available. A good introduction to basic notions is Raymond Cattell, "Factor Analysis: An Introduction to Essentials, I. The Purpose and Underlying Models", Biometrics, March 1965, pp. 190-215, and " . . . II. The Role of Factor Analysis in Research", Biometrics, June 1965, pp. 405-435. Full reference works include H. H. Harman, Modern Factor Analysis (Chicago: Chicago University Press, 1960), and Paul Horst, Factor Analysis of Data Matrices (New York: Holt, Rinehart and Winston, Inc., 1965).

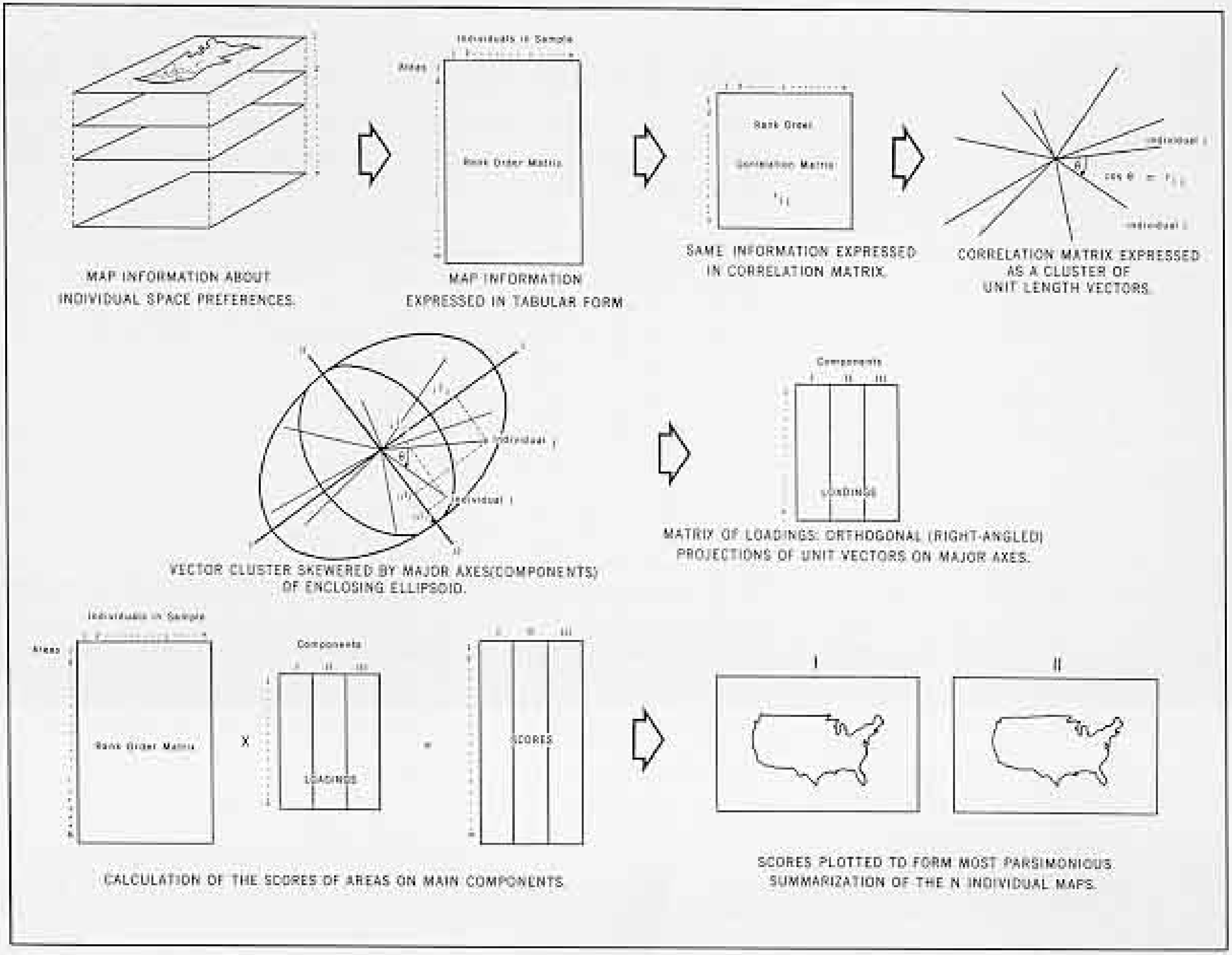
In all the examples that follow, people were asked to provide rank order listings of their preferences for various areas. The question was posed in the context of residential desirability with "all other things being equal". For example, in the United States, students were asked to imagine themselves married and settling down with a family with complete freedom of location according to their own particular views as to what was desirable. The question was similarly posed in Europe and Africa, with only minor modifications to adapt it to local circumstances. Thus the basic data consisted of a matrix whose rows represented areas (states in the U.S.A., countries in Europe, and administrative districts in Africa), while the columns represented people. Each column contained the rank order values that a particular person had assigned to places in the rows, so that in a crude sense each person became a variable upon which the residential desirability of a place was measured.

Clearly, if two people held very similar views their rank order lists would match quite closely. Thus, the whole basic data matrix may be summarized by a smaller matrix of rank correlations (Figure 1). It is upon such matrices that the principal components analyses are performed to break out the underlying structures of space preferences in terms of a smaller number of dimensions or components. By definition, such dimensions are unrelated to one another, because we are imposing an orthogonal structure upon the data,[21] and they may be thought of as independent scales upon which the areas have particular values or component scores. Thus, to summarize, the maps are constructed from scores on an orthogonal principal components structure, which is merely used descriptively

---

[21]Rotations can be performed, but the standardized criteria seem a bit simple-minded for this problem. What we really need is a physical structure in a two or three dimensional space which is attached, via a computer, to an oscilloscope displaying the map with a contoured "perception surface". Actual rotation of the structure by the investigator would almost instantaneously produce the new surface interpolated from the new scores based on the new loadings. Our "simple structure" criterion would then become truly spatial, which, after all, is what it should be.

Figure 1: Steps in the Construction of the Component Maps



as a statistical summary device, to see if we can break apart and simplify the structure underlying the views and values men hold and place upon geographic space.

THE PERCEPTION OF RESIDENTIAL DESIRABILITY IN THE UNITED STATES

Apart from turning their attention temporarily on particular places during times of physical or human crisis, perhaps most people in our highly mobile society perceive and think about the geographic space called the United States in terms of residential desirability. At the state universities of California (Berkeley), Minnesota, Pennsylvania and Alabama,[22] students in beginning geography courses were asked to provide rank order lists of the forty-eight contiguous states in terms of their own, quite personal preferences.[23]

Two problems were recognized. First, the state units were quite gross, and any analysis must be made at an extremely general, macro-level. While it might be better to make the mesh of perception a finer one, either by using county units, or gridding the map with 100 mile squares, such a notion can be quickly dismissed when we realize that people would have to rank literally thousands of areas defined in terms quite unlike those with which they were familiar. At least states, while gross areal units, are familiar objects

[22] I would like to thank Professors Alan Pred, University of California, Berkeley, Philip Porter, University of Minnesota, and Eugene Wilson, University of Alabama, for distributing the questionnaires to their students.

[23] Sample sizes varied from about twenty-five to fifty. While the lower bound constitutes quite a small sample, repeated samples of this size taken at the Pennsylvania State University have indicated an extremely high degree of consistency in the preferences and in the appearance of the final maps. I would like to thank Messers Marich, Bigelow and Knighton, graduate students in geography at the Pennsylvania State University, for providing me with their seminar exercise results.

with quasi-collective images. Secondly, most people faced with the problem of ranking items in order of preference have some immediate, and usually quite strong, likes and dislikes, but there may be a large number of items "in the middle" to which they are more or less indifferent. Thus, instructions were given that where difficulty was experienced in assigning a rank order to a state, that it should be matched against the others in succession with the question in mind "If I _had_ to choose, which would I prefer?".[24] It should be recognized, though, that the middle area rankings will be less valid, although there is little reason to suspect any systematic bias and we can probably regard the effect of indifference quite legitimately as random noise injected into the data.

The View From California

The surface of perception derived from the first dimension (Figure 2),[25] may be considered a general, overall view of the residential desirability of the United States as seen from California. A ridge of high desirability extends along the entire west coast, with the highest peak of the whole perception surface in

---

[24]In later work, still on-going, people have been asked to include and rank in their lists Neutral Points. As an individual compiles a list of states, starting with the most desirable, his initial feelings are positive -- he would like to live in the areas he most prefers. Further on down the list, however, his feelings become indifferent, at which point he injects the Neutral Point, assigning it the next rank in the list. Thus the Neutral Point provides a base point for any subsequent scales derived from a principal components analysis. See Harold Gulliksen,"Intercultural Studies of Attitudes", in Harold Gulliksen (ed.), _Contributions to Mathematical Psychology_ (New York: Holt, Rinehart and Winston, Inc., 1964), pp. 62-108. See too his contribution "The Structure of Individual Differences in Optimality Judgements", and Ledyard Tucker, "Systematic Differences Between Individuals in Perceptual Judgements", both contained in Maynard W. Shelley and Glenn Bryan (eds.) _Human Judgements and Optimality_ (New York: John Wiley and Sons, Inc., 1964).

[25]Factor scores have been transformed to percentage values relative to the highest score to make comparisons easier. These values were plotted at the approximate center of population of each state, and isolines added to form a three-dimensional "surface of perception" -- a useful, if somewhat pretentious notion. This is sheer cartographic license, used without apology, to heighten the visual effect and to provide the concept of gradient.

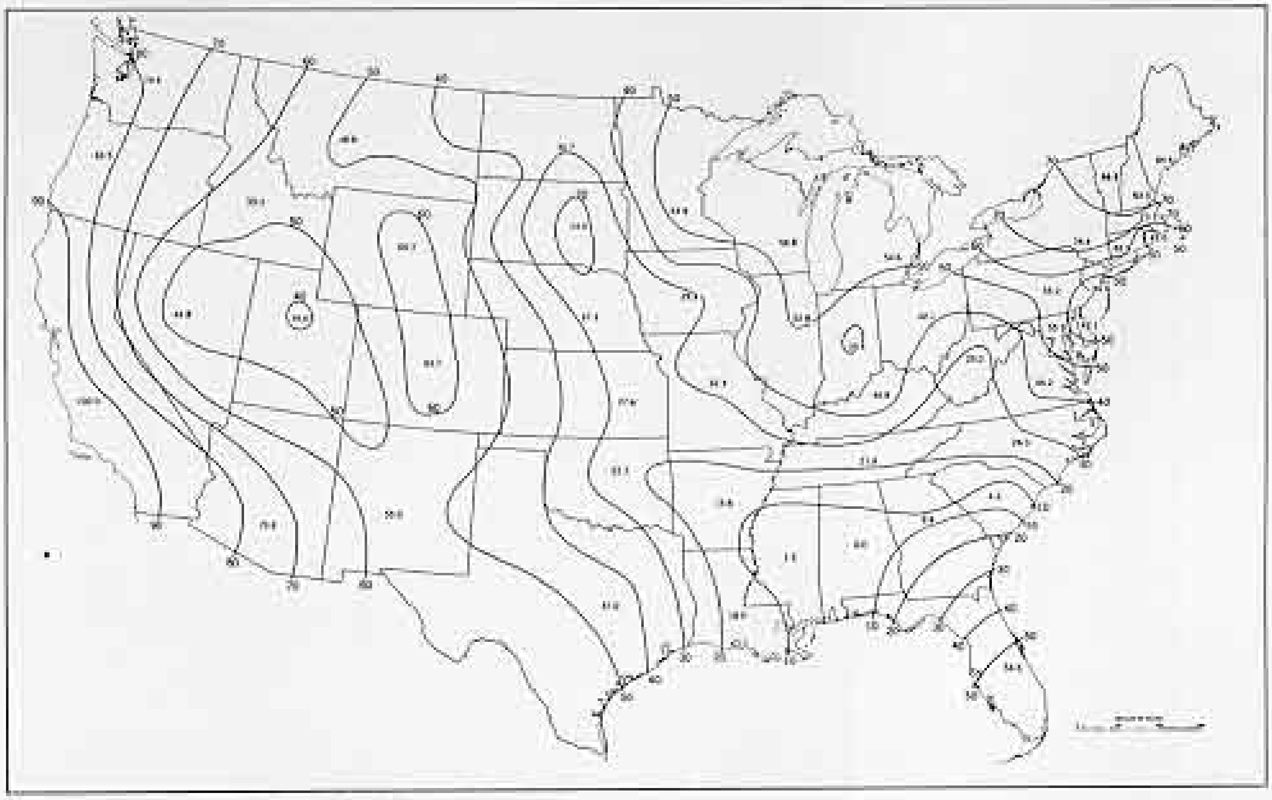


Figure 2: The View From California: The First Dimension

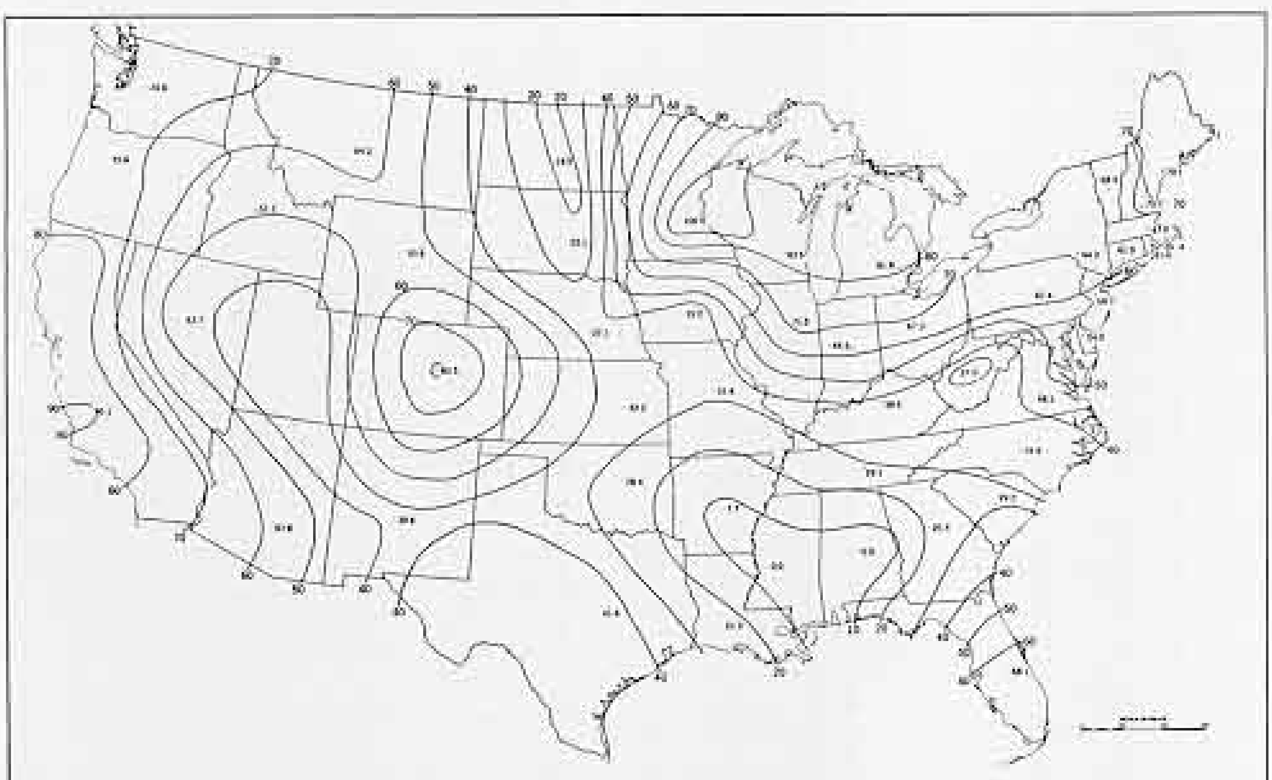


Figure 3: The View From Minnesota: The First Dimension

California itself. However, the gradient to the east is very steep and there is a clear "Perceptual" as well as "Great" Basin with a low point in Utah. Overall, as the view moves eastwards, there is a steady decline in desirability to the Great Plains, with the exception of a local peak in Colorado, and a quite definite "sinkhole" in South Dakota. However, upon reaching the ninety-fifth meridian, the general east-west trend of the perception surface changes radically, for the overall orientation shifts by ninety degrees and a very clear discrimination is made in a north-south direction between the Midwest and the Northeast, where the surface begins to rise once more, and the South, which forms the lowest perceptual trough of the entire surface. Alabama, Mississippi, Georgia and South Carolina, with their images of civil and social unrest, are the last places in the country for California's students. Only Louisiana is perceived as a slightly more desirable place of residence, but even here the gilding is somewhat tarnished and worn. Florida escapes the general Southern trend, and while Californians are not prepared to place this state as high as some other groups, possibly because of an old rivalry for the title of America's premier place in the sun, the gradient is steep from the low point in Alabama. To the north, the surface trends upwards all the way to New England. Noticeable, however, is the way West Virginia distorts the even march of the iso-percepts, possibly because of the recent emphasis upon the problems of poverty in Appalachia and the high social awareness usually ascribed to California's students. Thus, much of Kentucky may be in a similar economic plight, but her image and her value on this first, general dimension of perception is bolstered to almost twice that of her well-publicized eastern neighbor. Perhaps white-fenced, bluegrass pastures and sour-mash bourbon with pretensions to spiritous greatness, familiar themes in many advertisements, have conveyed an image brighter than the purely economic facts over much of the state would warrant! From Pennsylvania north-east to New England the rise is very rapid, reflecting an image that appears to include more than the bright lights of Megalopolis. The view of the northeastern "cultural hearth" seems to carry visions of the mountains and lakes of Vermont and New Hampshire and the quiet, rock-strewn coasts of Maine.

The View From Minnesota

From Minnesota the view of the United States is almost the same as that from California (Figure 3). True, the highest peak on the perception surface has shifted to the point from which the perception took place, but the high west coast ridge is still very much in evidence, together with the steep gradient to the Utah perceptual basin, the rise to the Colorado high and the fall to the Dakota sinkhole. Indeed, the Minnesotan seems all too aware of his western neighbor, and is most reluctant to trade his "land of sky-blue waters" for the dry flat dreariness a few miles to the west. The steep decline in the surface to the flat country of Iowa reinforces an impression of spatial chauvinism, although to the east Wisconsin appears quite acceptable for residential purposes. Once again, the general west to east trend shifts ninety degrees around the one hundredth meridian, and a low trough, centered in Mississippi and Alabama, blankets the South. Florida is again an exception. Northwards the iso-percepts rise, twisted by the West Virginia (Appalachia) distortion, and the Midwestern and New England states bask in residential acceptability with almost uniform values in the sixties.

The View From Pennsylvania

Perhaps more cosmopolitan than the Minnesotans, or set on edge by their rural location at the point of maximum inaccessibility, students at the Pennsylvania State University are seduced by the Californian siren to form the only sample in which the perception point is not the most preferred (Figure 4). Otherwise the surface is almost identical in general form to the Californian and Minnesotan examples. Once again, the perceptual ridges and basins of the West are repeated, and the South is the lowest, or least desired part of the country.

Alabama: Different Values and Different Views

Apart from local peaks of desirability, the three northern and western viewpoints are virtually identical. It is almost as though we had the same perception surface drawn for all three upon a rubber sheet and could reproduce the exact surface by moving a tennis ball beneath it to form, or reinforce, the local point. However, the view of the college student from Alabama, while sharing one or two points of similarity in the West, is generally quite different (Figure 5). The high peak of the surface is centered at the point of perception once again, but whereas the previous three samples tended to lump the "South" into a single low trough, Southerners appear to perceive the area with a high degree of spatial discrimination. Most noticeable is the very steep gradient down to Mississippi in the west: there seems to be little love lost between the two states that Northerners tend to place together, possibly because of the extreme violence associated with the civil rights movement in that state. A fairly high degree of discrimination is also apparent between South Carolina and North Carolina. While California's politically and socially concerned students did tend to single out the former, generally the difference in desirability perceived by the northern students between these two states was not marked. Southerners, on the other hand, with a more intimate knowledge of this area, place North Carolina on par with gentlemanly Virginia and Kentucky, while assigning the southern neighbor the next-to-lowest score in the whole South. With the effect of the Civil War still being handed down in the minds of men one hundred years later, the surface falls away rapidly to the northeast, with a very steep gradient along the Mason-Dixon line. North of this historic divide, there is considerable perceptual homogeneity with all the Yankees lumped together in Southern minds, just as the Northern viewpoints blotted out and homogenized possible differences in the South. Such disparagement is shared by the Midwest and the whole set of northern states as far as Washington and Oregon on the Pacific. Northern, remote and blizzard-ridden, with only eye-straining horizons of waving wheat, the Dakotas are the last place in America for the college-bound Alabaman! In the West, only the Californian ridge and the Colorado peak appear, though less

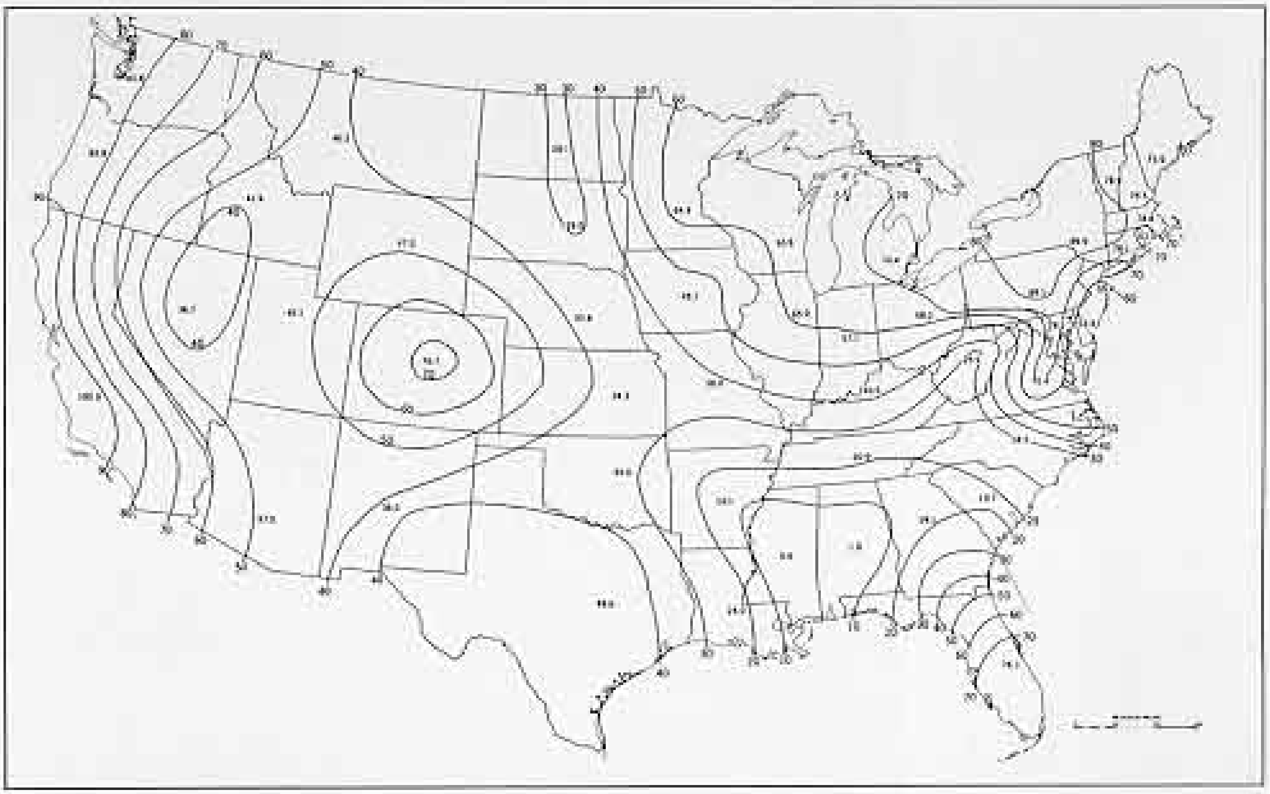


Figure 4: The View From Pennsylvania: The First Dimension

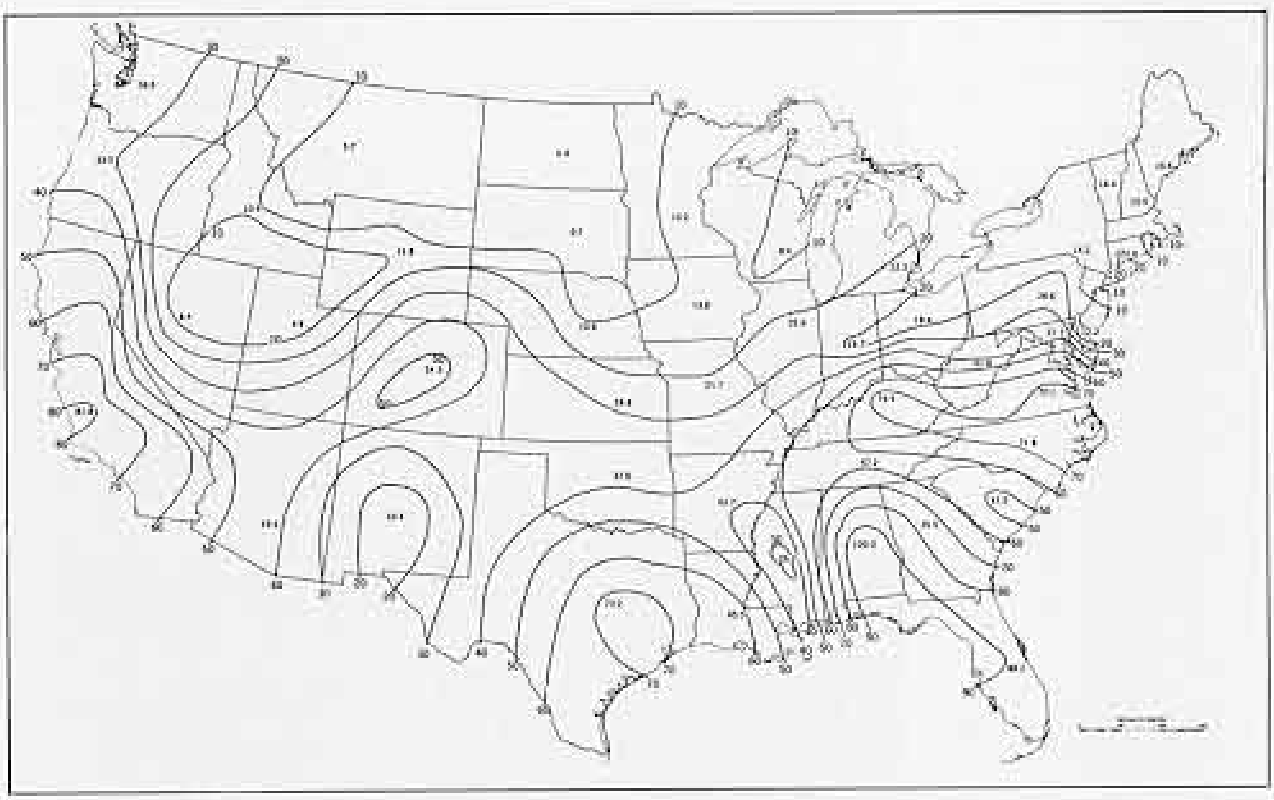


Figure 5: The View From Alabama: The First Dimension

strongly than in the other examples. It is interesting to speculate upon the reason for the sinkhole in New Mexico, compared with the high value of Texas (still in the South), and Arizona (Goldwater country in the election year, shortly after which the sample was taken). What mental image does the word "Mexico" conjure up for the white Southerner that so reduces the desirability of this state far below its neighbors?

QUESTIONS AND SPECULATIONS

The remarkable degree of similarity in the mental maps held by the Californian, Minnesotan and Pennsylvanian samples raises the question of the degree of perceptual homogeneity within the groups themselves. The proportion of the total variation explained in each case by the first component does not vary radically (Table I), but the differences, if they do not simply arise from sampling fluctuations, are intriguing. Pennsylvania possesses the highest degree of perceptual homogeneity, closely followed by Minnesota, while Californian students appear to agree the least

TABLE I

State	Percentage of Variance Extracted		
	I	II	III
Pennsylvania	46	16	6
Minnesota	41	15	7
California	36	15	9
Alabama	28	13	9

about places of residence. Perhaps this is because the University of California at Berkeley draws upon a more heterogeneous population for its student body, while the state of California itself receives large numbers of migrants from other states in the Union. Surprisingly, the sample with the least homogeneity of outlook is Alabama, and we might postulate a sort of spatial schizophrenia, for while the most dominant portion clings tightly to Alabama, and shares a mental map that discloses all the century-old cliches about "Yankeeism", the rest are split in their views with little agreement between their mental images.[26] We shall meet such heterogeneity again when we examine the mental maps of French students.

Other questions also emerge from an examination of the maps based upon the first component scores. One of these is the effect of size on the ordering process. While it has been postulated that in viewing a map our minds act as a high pass filter, so that small-scale features are accentuated,[27] it is worth noting that Rhode Island is consistently lower in overall score than its New England neighbors. Does it really have the image of a less pleasant place to live than nearby Massachusetts and Connecticut? Or is it, in fact, so small that people tend to forget about it, and in partially overlooking it assign a lower rank to it than it might otherwise receive?

Another source of possible bias could result from the propensity for people to group together things that are spatially contiguous.[28] It would be difficult to design a watertight experiment to get at this effect if it existed, but one approach might be as follows. Identical maps of the United States could be given to two groups (one the sample, the other a control), and data obtained on the space preferences. At a later time, the experiment could be repeated with the control group getting the same map as on the first run, while the sample group would receive a "map" showing the outlines of the states located in a random, and non-contiguous fashion. The successive

---

[26]Since the sample was confined to white students it would be highly desirable to examine the mental maps of negro students in the same area. The viewpoints might not be identical.

[27]J. Leith Holloway, "Smoothing and Filtering of Time Series and Space Fields", Advances in Geophysics, Vol. 4, 1958, pp. 386-387.

[28]Julian E. Hochberg, Perception (Englewood Cliffs: Prentice-Hall, Inc., 1964), p. 86.

canonical correlations between the control groups' data sets should be very high, while the same correlations between the sample groups' contiguous and non-contiguous sets should be significantly lower if a contiguity effect is operating. The inferential question of the significance of the difference between two successive canonical correlation analyses, even if assumptions of normality are regarded as plausible, seems worth pushing.

Finally, the question of indifference must be raised once again. By requiring each person to insert and rank a Neutral Point in his list of preferences we may obtain at least some notion of the severity of the problem. In the one example presently available for Pennsylvanian students,[29] the overall score of the Neutral Point on the first component places it in the twenty-third rank. Thus nearly one half of the states are generally perceived as positively desirable to live in; a remarkable comment upon the spatial mobility of the present college population. Indeed, the level at which a Neutral Point appears on the overall scale represented by the first dimension may be considered as a measure of the degree of parochialism of the sample group. A Neutral Point high on the first dimension, indicating that people are positively inclined to only a few states, might measure either a high degree of parochialism or a high level of discrimination, depending upon one's own attitudes towards such value-loaded words. We might hypothesize that the degree of parochialism, as measured by the position of the Neutral Point, would be directly related to the average age of the group, and the degree of social and cultural isolation experienced by the people in it.

---

[29] I would like to thank Mr. Bruce Marich, NDEA Fellow in Geography at the Pennsylvania State University, for allowing me to use the results of his paper "Space Speaks: Some Metaphysical Roots of Spatial Decomposition", p. 15. In the West, California, Washington and Texas were above the Neutral Point; in the South only Florida was so perceived; while all the states east of Wisconsin and north of the Ohio River were seen as positively desirable.

THE RISE AND FALL OF A HYPOTHESIS

Our assumption that indifference exists in the rank orders as random noise makes the interpretation of further components somewhat hazardous. Nevertheless, the remarkable consistency (Table I) in the proportion of the variance explained by these second scales invites interpretation.

Remembering that the dimensions or scales that we impose in a principal components analysis are orthogonal, and, therefore, unrelated to one another, we might expect that successive maps of perception surfaces should illustrate quite independent concepts about the mental images that men have of geographic space. The view from California on the second dimension (Figure 6), immediately suggests that underlying the overall, general surface (Figure 2), there is another surface, quite independent of the first, that is strongly related to distance away from the point of perception. In fact, apart from a rather awkward contretemps around Oregon and Washington, and small distorting pockets in the Ohio, Kentucky and West Virginia areas, the correlation of the scores with raw, crow-flying distance from California is remarkably high ($r_s = .90$). Thus, there appears to be some strong evidence that a distance component is present in mental maps, and that mental images of the differential desirability of geographic space cannot be meaningfully represented on a single, general scale.[30]

The second perception surface for Pennsylvania (Figure 7), bolsters the idea that a distance component is present. Once again the west coast states prove to be exceptions, destroying the fairly regular march of the isopercepts

---

[30] Implying that the rank of the correlation matrix is greater than one. For readers unfamiliar with the geometry of such a notion, there is an intriguing physical model in a geographical context in Robert E. Blackith, "Morphometrics", Chapter 9 in Talbot H. Waterman and Harold Morowitz (eds.), Theoretical and Mathematical Biology (New York: Blaisdell Publishing Company, 1965), p. 236.

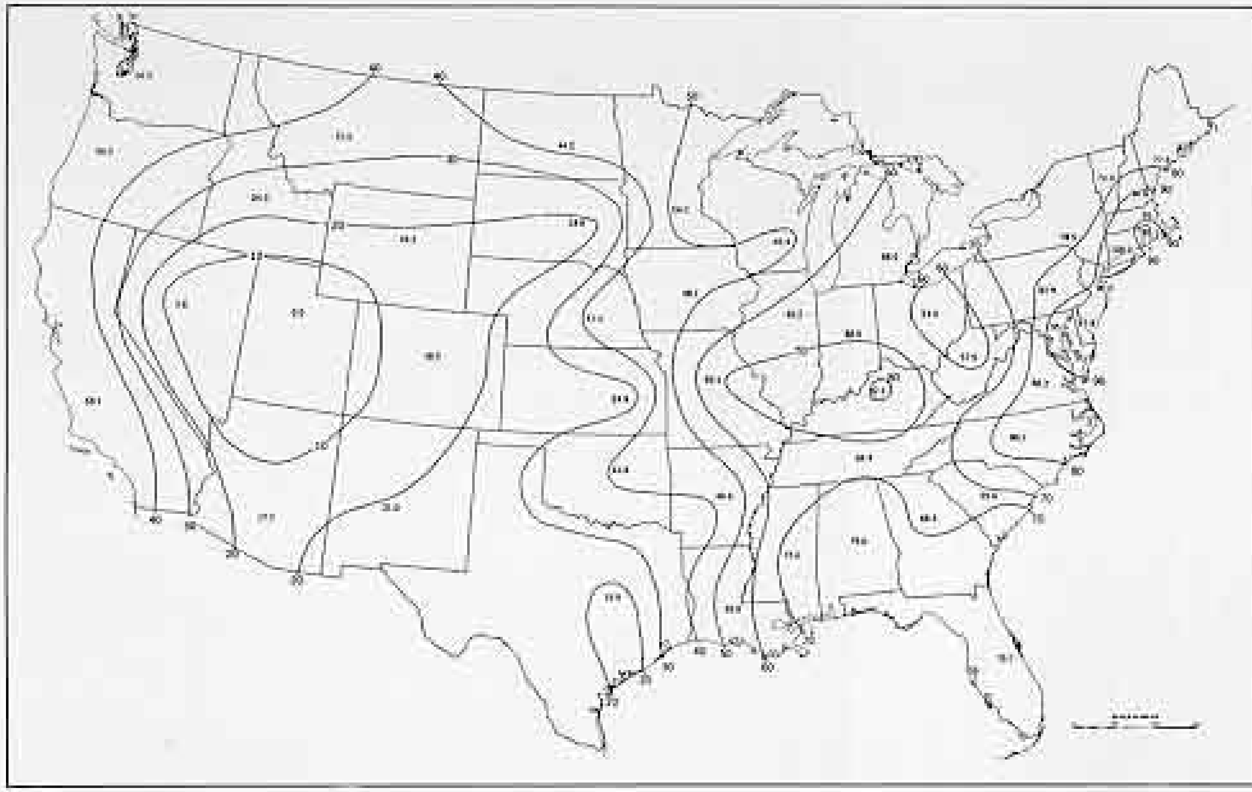


Figure 6: The View From California: The Second Dimension

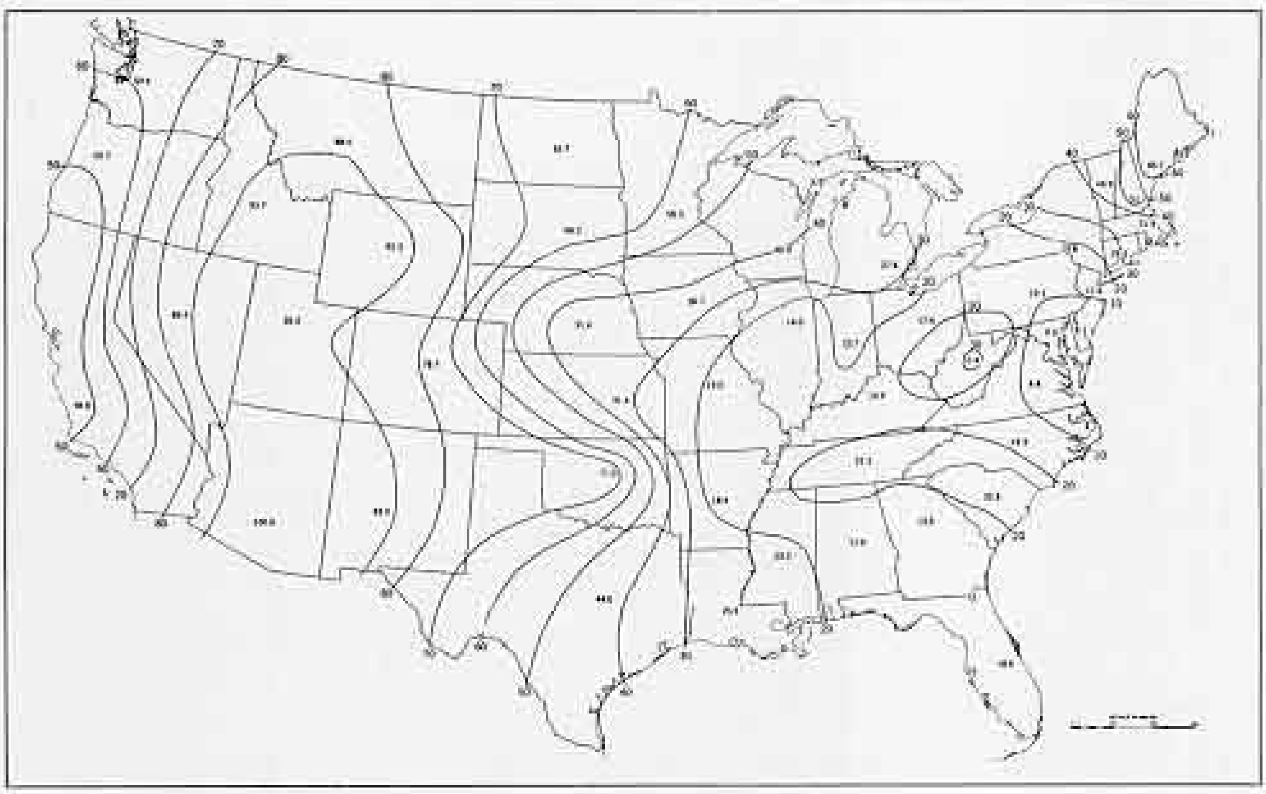


Figure 7: The View From Pennsylvania: The Second Dimension

with distance away from Pennsylvania, but even so the correlation with raw distance is significantly high ($r_s = .64$). New England, too, distorts the distance effect and lowers the overall relationship, but the scores vary with such regularity that one has the feeling that if the map were drawn upon a rubber sheet mere stretching of the space, rather than tearing or inverting it, could raise the correlation significantly. In other words, a spatial transformation, with some rather interesting psychological implications, could disclose a much stronger distance effect than the one crudely indicated by the actual association between component scores and straight geographic proximity.[31]

The hypothesis that a distance effect is another dimension to the mental map appears tenable so far. Unfortunately, it is exploded by the next example: Minnesota, true to form, and ever willing to demonstrate geographical inconsistencies, does not fit! If the scores of states on the second component of Minnesota are related to distance (Figure 8), the correlation is virtually zero, and no amount of stretching of a rubberized surface would appear to make this dimension conform to our distance hypothesis.

But another question now appears worth considering. In all three of the second component maps (Figures 6-8), there appears to be some propensity for the isolines to run generally north and south. In the case of California, the surface is high in the west, dipping to low values over the Great Basin and Mountain states, and rising again slowly at the Great Plains to another high, though convoluted plateau in the east. Similarly, the second perception surface of Minnesota is moderately high in the west, dips to a low trough and rises again with considerable regularity eastwards from the Great Plains only to dip again finally in the New England area. Pennsylvania's second surface also displays the "east-west" effect to a marked degree, although the shape appears

---

[31] Other samples, taken by participants in a seminar at the Pennsylvania State University, confirm the general "east-west" effect of the second component scores.

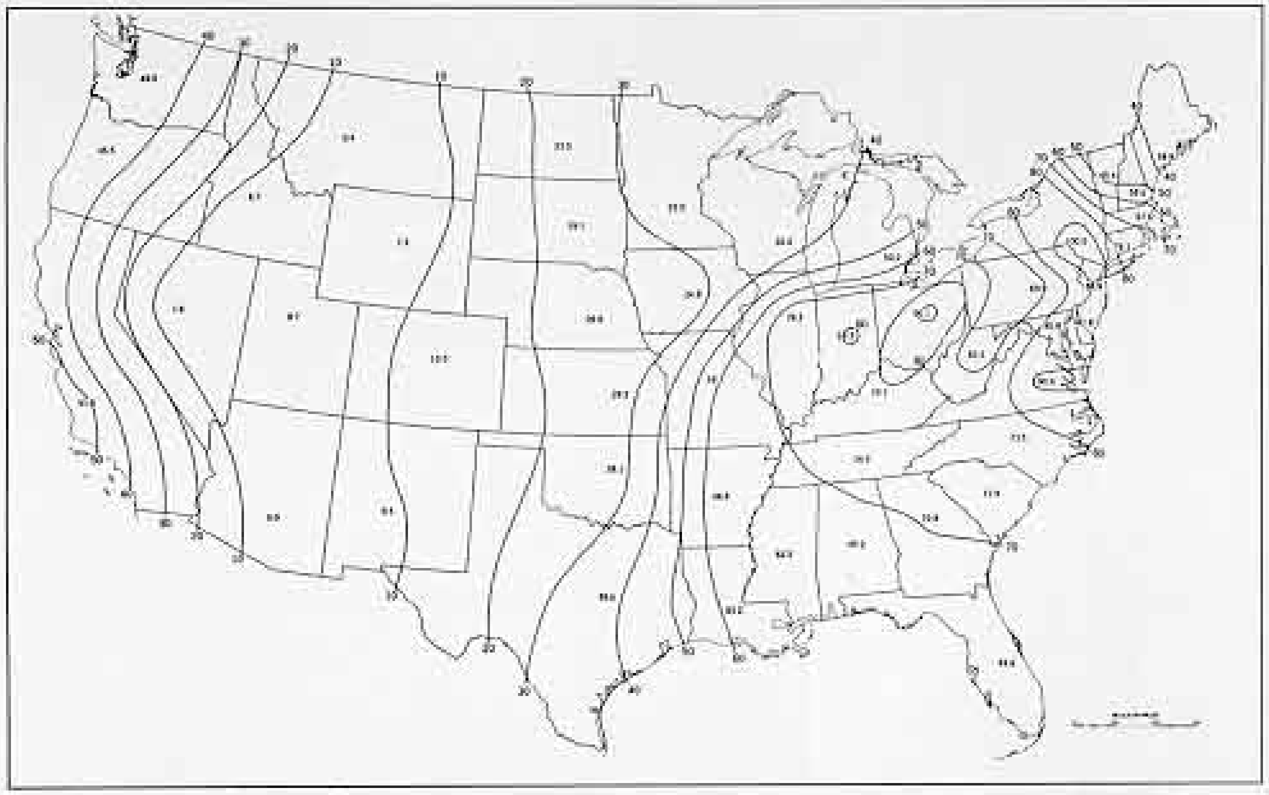


Figure 8: The View From Minnesota: The Second Dimension

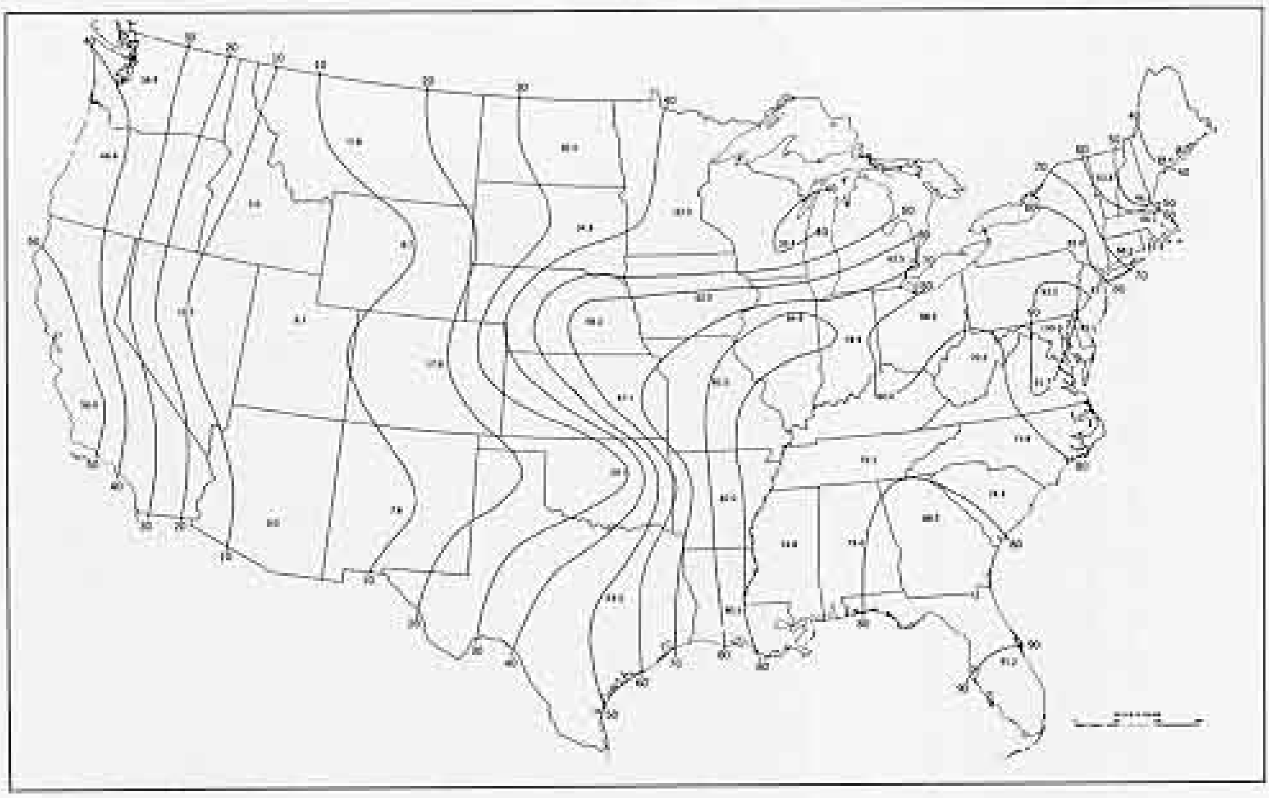


Figure 9: The View From Pennsylvania: The Inverted Surface on the Second Dimension

to be the negative image of California and Minnesota. However, the reversal of the surface raises the question of the mathematical structure of the component model which has been imposed upon the data. If we remember that the loadings of the unit person vectors upon the second component are the elements of the corresponding normalized eigenvector weighted by the square root of the eigenvalue (namely, the length of the second axis of the ellipsoid), then the positive and negative signs may be reversed without changing (1) the variance accounted for by the dimension or (2) the position of the eigenvector in the hyperspace. In other words, the structure of the model is quite unchanged by consistently reversing the positive and negative signs of the loadings.[32] When this is done, the Pennsylvanian surface corresponds closely to those of California and Minnesota, with a high western ridge and a low trough over the Basin and Mountain states which rises to the eastern plateau to dip, finally, over New England. We might hypothesize, therefore, that there is a second dimension to the mental maps that illustrates a propensity to perceive and evaluate the geographic space of the United States in a fairly consistent east-west direction. Thus, when the point of perception is on the "edge" of the rectangular space, such as in California and to a lesser extent in Pennsylvania, an apparent, though spurious distance effect appears reasonably tenable. Only when the perception point moves to the center of the space, as in the case of Minnesota, must the distance hypothesis be discarded to be replaced by an east-west interpretation that is much less intellectually satisfying because if is difficult to relate to any other insights we have. We do know that east-westness in travel produces distinct psychological and physiological

---

[32] Comments upon this argument would be greatly appreciated since I have seen nothing in the literature referring to such a change that seems quite arbitrary until the geometrical structure is examined closely.

effects compared to north-south movement,[33] but this observation provides little confirmation that we are on the right track and the hypothesis awaits much more study with larger and more numerous samples. For example, if one looks at the map of Alabama (Figure 10), through reasonably charitable eyes some confirmation of the east-west hypothesis is obtained. Moving west to east, the high Pacific ridge dips to a north-south trending trough at the 110 meridian followed by a rise in the eastern part of the country. But severe local anomalies, such as Arkansas, occur and the surface appears equivocal in its ability to support or deny the rather clear configurations of the other three examples.

EUROPE LOOKS AT EUROPE

In these days of newly emerging nations in Asia and Africa, we are apt to forget that it is really Europe that has experienced the most shattering political change of the twentieth century. Smashed by the First World War, and beset afterwards by violent revolution, catastrophic inflation and economic depression, the stable and ordered continent of the early years is now but an historical curiosity to three generations. Laid waste by yet another world war, in which civilian casualties alone were counted in the millions, and divided afterwards by ideologies that confront one another across a line drawn through her very heart, Europe is still in the process of settling down to a new, though hopefully a dynamic and moving equilibrium.

Faced for the first time with two powers larger than any that they have experienced before, the countries of Europe are seriously considering a degree of cooperation and unification that would have been unthinkable a few decades ago.

---

[33] G. T. Hauty and T. Adams, Phase Shifts of the Human Circadian System and Performance Deficit During the Periods of Transition, Part I East-West Flight, Part II West-East Flight and Part III North-South Flight (Washington, D.C.: Federal Aviation Agency, Office of Aviation Medicine, 1965).

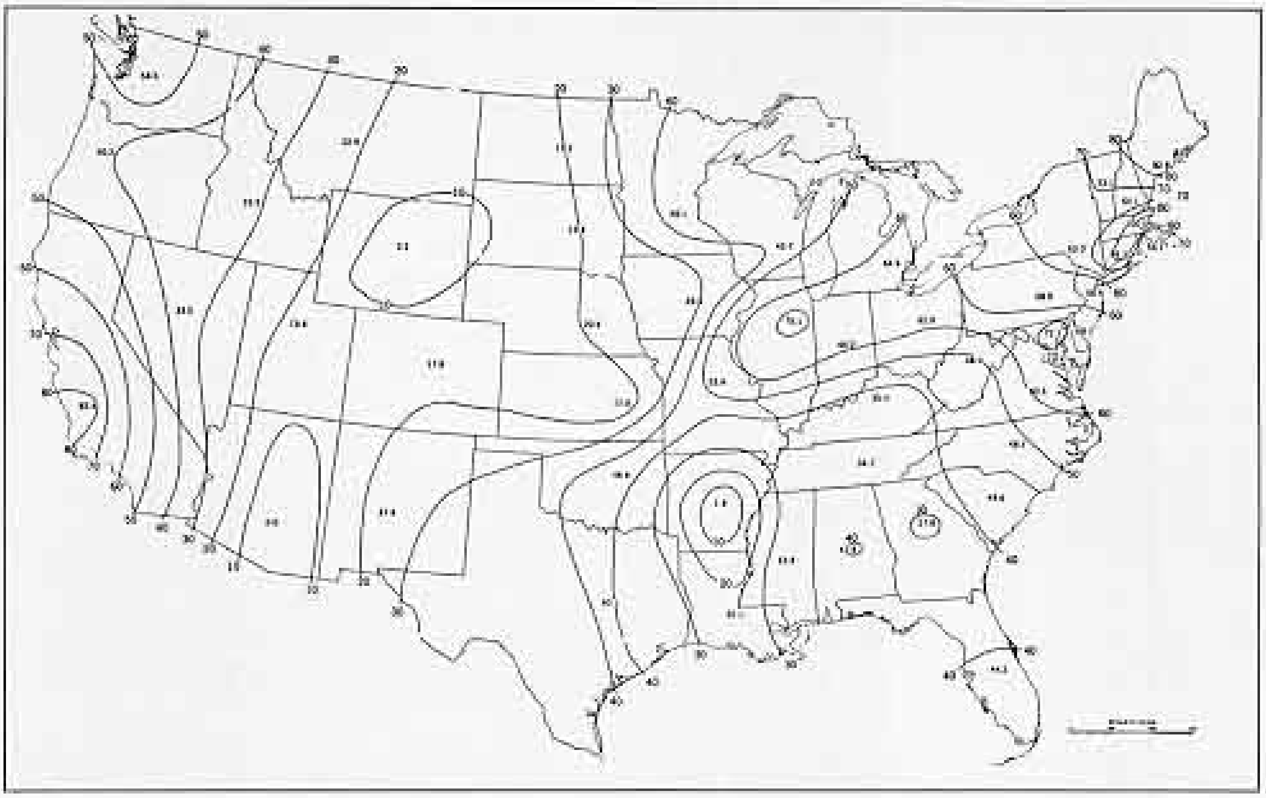


Figure 10: The View From Alabama: The Second Dimension

Military alliances, steel communities and joint atomic facilities, common aeronautical projects and Common Markets -- all are signs of an overall process of drawing together. True, the process is not altogether smooth, and there are backward steps as well as forward gains, but as economic cooperation leads towards some degree of political integration, it is pertinent to enquire about the mental images and preferences that these people have of other countries in the larger European community. For in the last analysis, the barriers to unification are mental, and where a government is ultimately responsible to an electorate the mental images that a people have are reflected to some degree in the policies towards other nations.

It would be foolish to pretend that the space preferences for residential desirability by university students are indicative of the overall mental images held by one nation for another. Nevertheless, by posing the question in terms of residential choice, we may record the mental maps of a small, but important group in the post-war generation -- a generation, let it not be forgotten, that has known only an uneasy peace broken by sporadic outbursts that could all to easily have triggered the Armageddon. Such a generation is likely to give thoughtful answers.[34]

Europe From Different Viewpoints

The sample from Sweden displays the highest degree of perceptual homogeneity with seventy-two percent of the variance collapsing upon the first component (Figure 11). On this general, overall scale, Sweden is by far the most preferred, followed by Switzerland, Norway, Denmark and the United Kingdom in a tight cluster. Other West European, democratic nations appear next, such as France, West Germany, Italy and the Benelux countries, followed by those with dictatorial regimes such as Spain and Portugal. The Eastern bloc is scored low, with East Germany and Albania per-

---

[34] I would like to record my thanks to Professors Olsen, Garner, Manshard, Pecora and Juillard, geographers at the universities of Uppsala, Bristol, Giessen, Rome and Strasbourg respectively, for distributing questionnaries to their students for me.

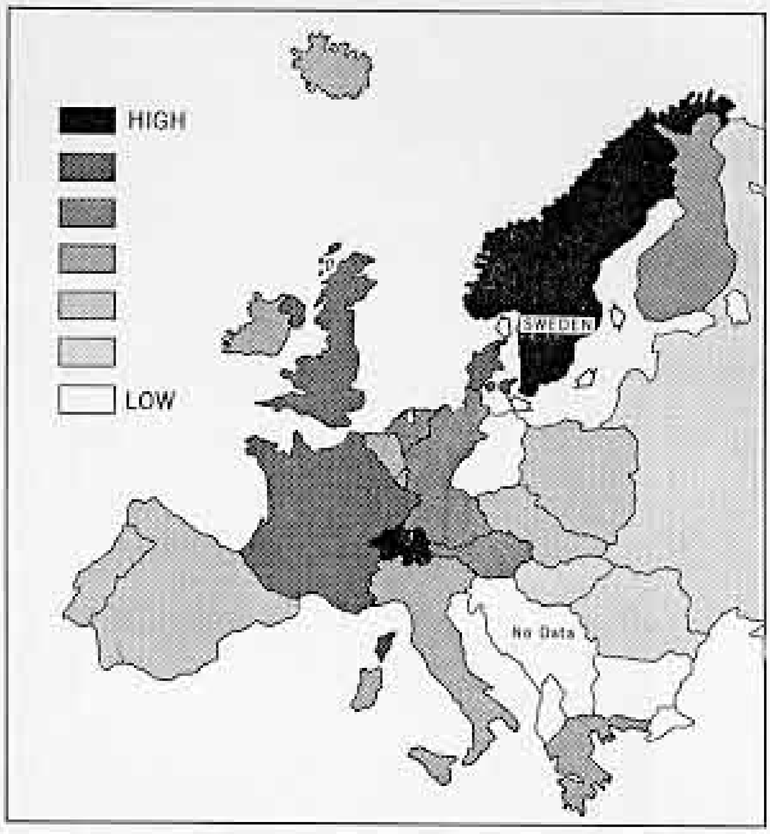


Figure 11: Europe Viewed From Sweden:
The First Dimension

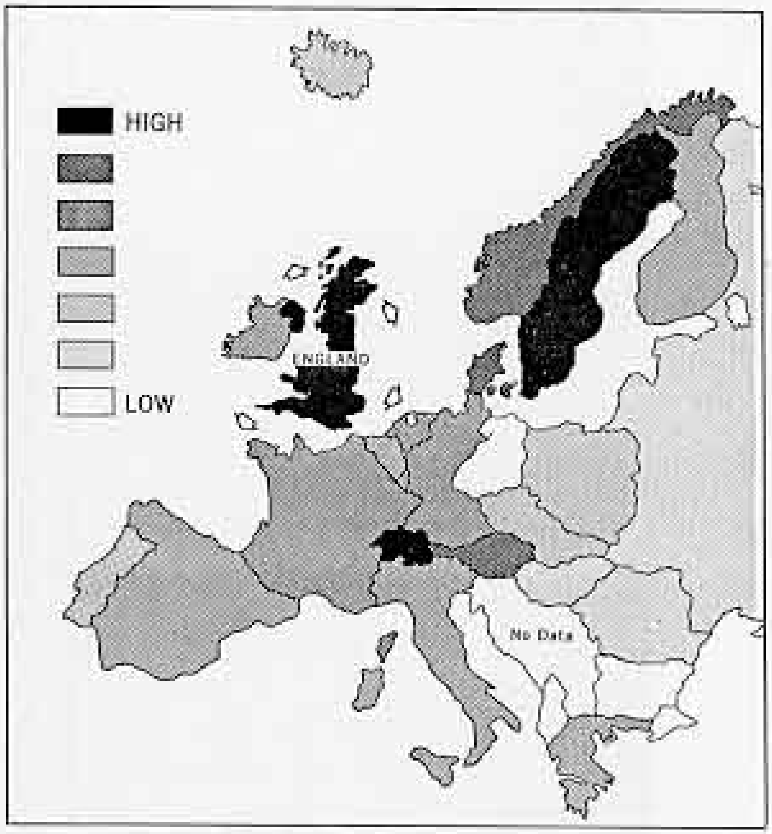


Figure 12: Europe Viewed From England;
The First Dimension

ceived as the least desirable of all.

While not quite as homogeneous in outlook as Sweden (Table II), the overall mental map of the English university students is remarkably similar to that of their contemporaries in Sweden (Figure 12). Indeed, with the exception of England's view of Finland, and Sweden's view of Eire, students in these countries seem to

TABLE II

Country	Percentage of Variance Extracted	
	I	II
Sweden	72	6
United Kingdom	66	8
Germany	56	10
Italy	55	8
France	45	15

view Europe through almost identical eyes. Both groups display a strong preference for western, democratic nations and tend to shun authoritarian governments able to exert a high degree of coercion upon their citizens. The contrast between east and west is virtually identical in every respect. Portugal scores slightly higher on the English map, indicative perhaps of the long historical ties between the two countries, while Iceland is placed somewhat lower than on the Swedish scale.

The view from Germany (Figure 13), reinforces a notion of cultural affinity that was displayed in a less severe form by Sweden's preference for the Scandanavian countries. After West Germany itself, the german-speaking countries of Switzerland and Austria are preferred over all others, with such linguistically similar nations as Sweden, Denmark and the Netherlands following closely. Significantly, Belgium, Luxembourg and France are preferred next, and the United Kingdom scores considerably lower on the German map than does Germany on that of the United Kingdom's! On the whole, however, Germany shares the viewpoint of England and Sweden for the West, and

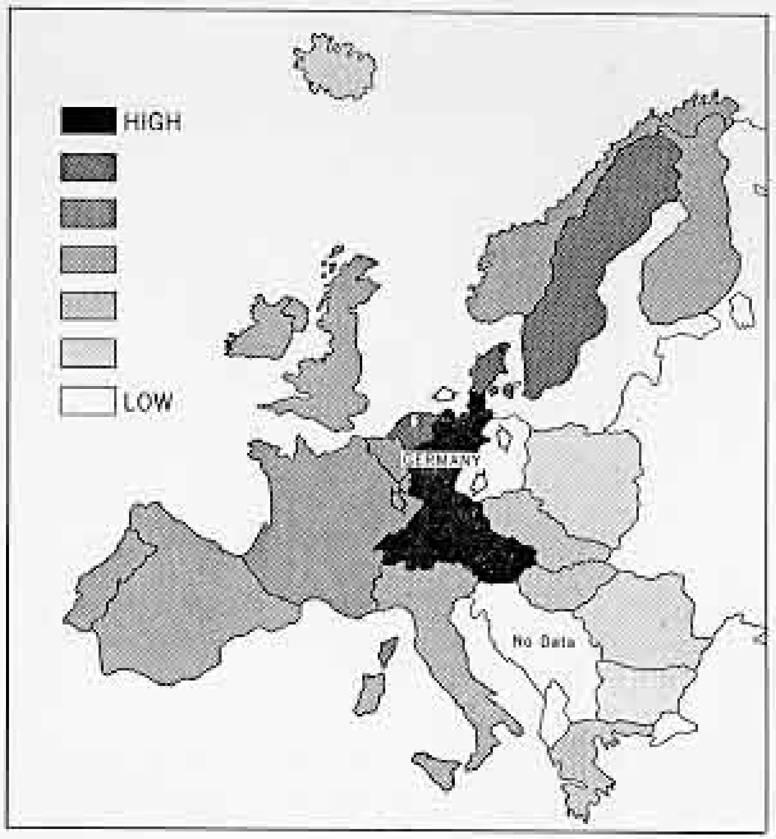


Figure 13: Europe Viewed From Germany:
The First Dimension

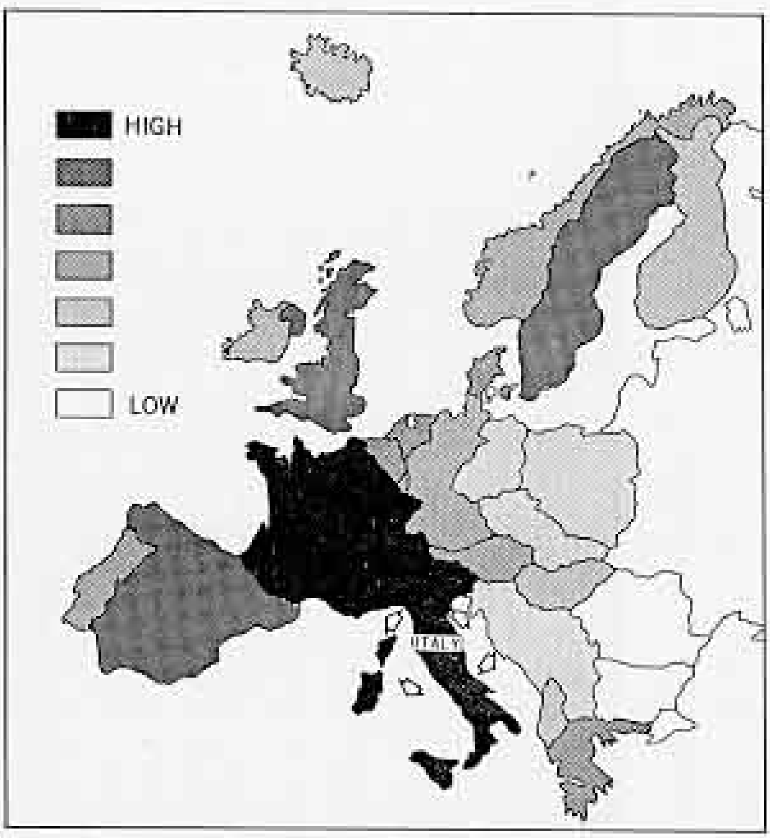


Figure 14: Europe Viewed From Italy:
The First Dimension

the Iron Curtain is again a most vivid divide.

Similar to Germany in the degree of perceptual homogeneity (Table II), but displaying a mental map that is just as distinctively individualistic (Figure 14), Italian students have a view of Europe that places the Catholic and Latin countries of France and Spain much higher than any other. Greece, too, with its sunshine and blue waters, is perceived much more favorably, while Western Germany is given a lower rank in the Italian sample than by any other group. Nearby Albania, normally the last country in Europe where students would like to live, receives a higher rank than usual, close behind Finland and Iceland and well ahead of most of the Eastern bloc.

France possesses the least homogeneous of viewpoints, although the dominant one is very close to those of Sweden and the United Kingdom (Figure 15). However, there is a second view from France (Figure 16), producing a larger second dimension (15%) than any other, which places the U.S.S.R. and Yugoslavia on par with the homeland and eschews Scandanavia, the Low Countries and perfidious Albion!

<u>Overall Views of Europe</u>

We can consider the scores of countries upon the main dimensions extracted by the principal components analyses as values on new variables that are a result of linear combinations of the original rank order lists. These variables display some interesting summary interrelationships. For example, while the dominant views of all five samples are closely related (Table III), Italy meshes less strongly than the other four, agreeing least of all with Germany.

TABLE III

Correlations Of Component I Scores

Country pair	Similarity of Viewpoint
Sweden-U.K.	.94
U.K.-Germany	.90
Germany-Sweden	.90
Sweden-France	.89
France-U.K.	.88
Italy-France	.81
Germany-France	.79
Italy-U.K.	.73
Italy-Sweden	.69
Italy-Germany	.62

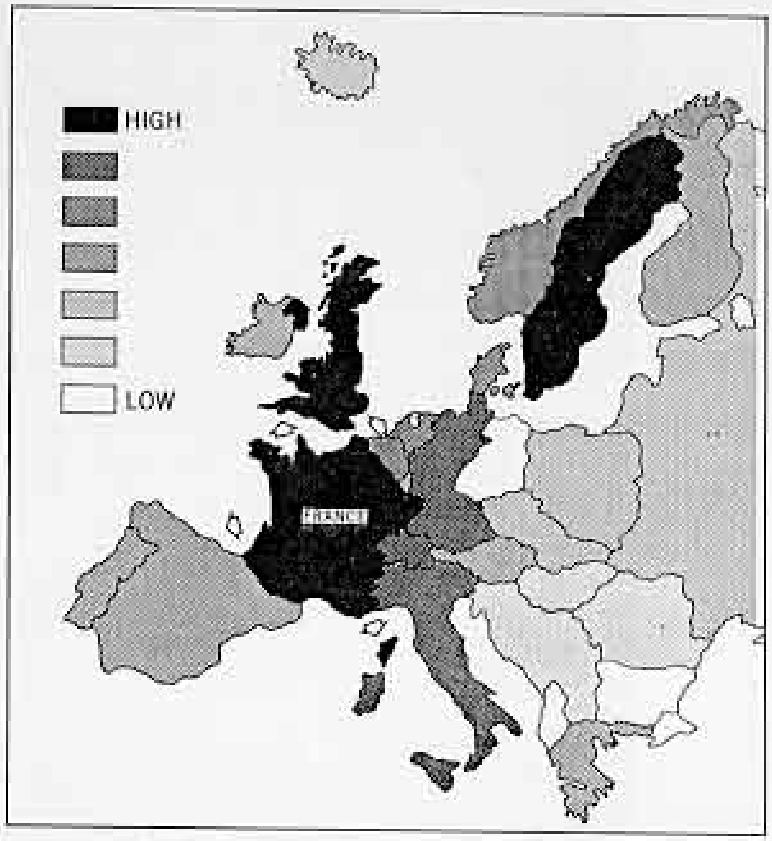


Figure 15: Europe Viewed From France:
The First Dimension

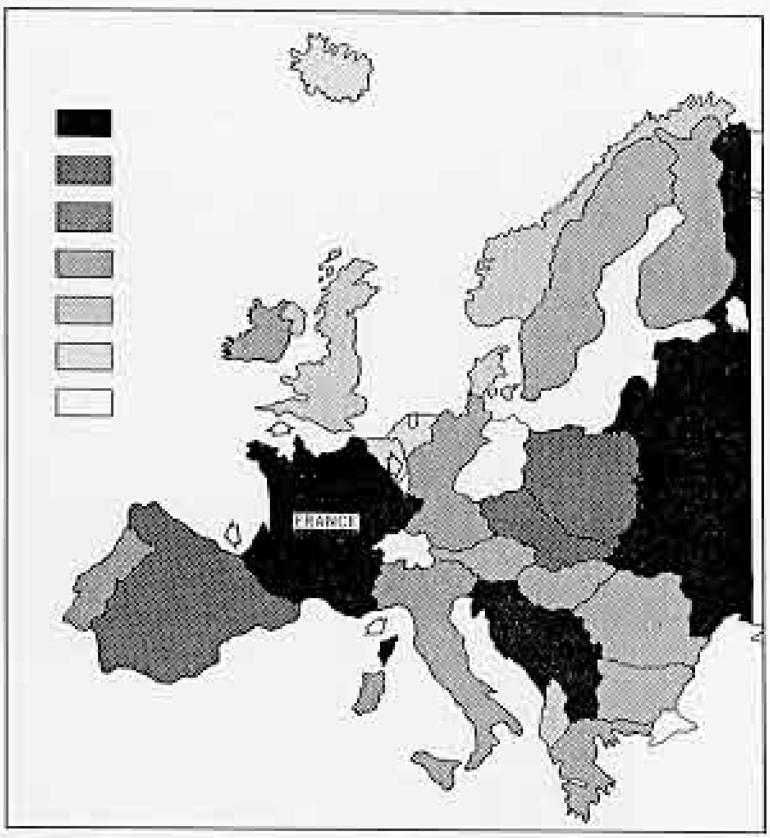


Figure 16: Europe Viewed From France:
The Second Dimension

Given the problem of indifference in ranking, the second components, which do not produce uniform maps, must be used with caution. Relationships between the scores on this second group of artificial variables (Table IV), are not always strong, but the directions of the signs are intriguing: all are positive with the exception of Germany. Whatever the second viewpoint from Germany represents, it relates slightly to the second from Italy, and inversely with all the others. Favorably

TABLE IV

Correlations Of Component II Scores

Country pair	Similarity of Viewpoints
Sweden-U.K.	.60
U.K.-France	.46
Italy-Sweden	.32
France-Sweden	.22
Italy-France	.14
Italy-U.K.	.09
Germany-U.K.	-.06
Germany-Sweden	-.14
Germany-France	-.17
Germany-Italy	.16

evaluating the surrounding neighbors (Figure 17), including Fascist Spain, this dimension is inversely related to the first components of Sweden and the United Kingdom, components that displayed such strong preferences for democratic institutions (Table V). Similarly, Italy's second dimension relates inversely to Sweden I and the United Kingdom I, although all these should be interpreted with caution since

TABLE V

Correlations Across Components

Country pair	Similarity of Viewpoints
Sweden I-Germany II	-.23
U.K. I-Germany II	-.13
Sweden I-Italy II	-.52
U.K. I-Italy II	-.43

both Germany II and Italy II have high variation remaining on the diagonal of the

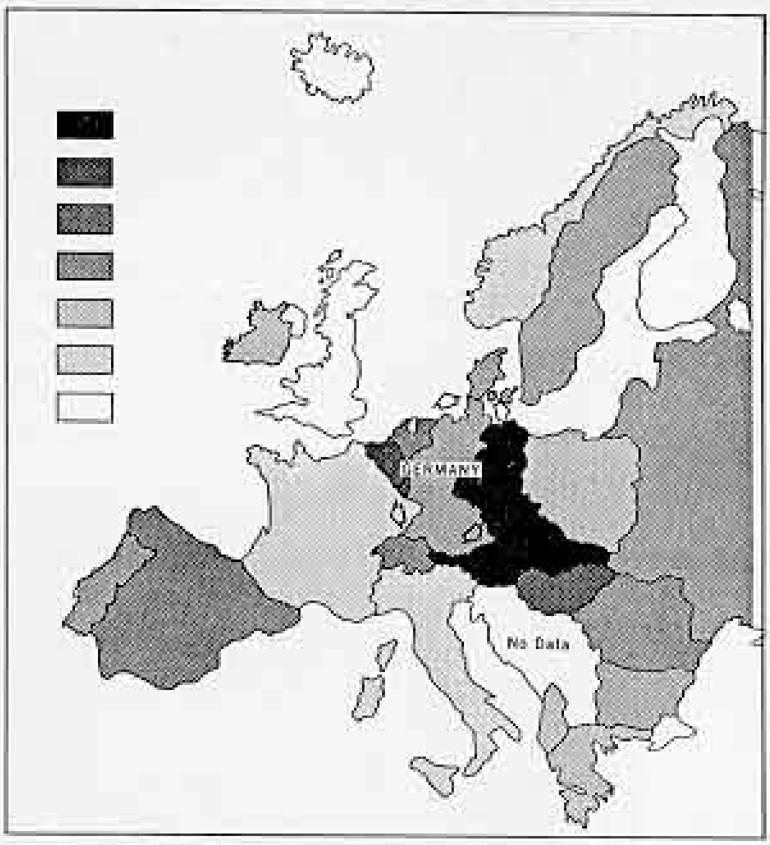


Figure 17: Europe Viewed From Germany:
The Second Dimension

residual matrix, indicating that low communalities would be appropriate and that the specificity of their variation is high.

A second, or higher principal components analysis provides us with an overall scale for all the samples (Figure 18).[35] Summarizing forty-five percent of the variation, it may be interpreted as a complex scale that seems to reflect both the level of the standard of living and the degree of political authoritarianism. Highest are Switzerland and Sweden, both havens of neutrality with high standards of living and liberal democratic governments. This cluster is followed by most of the democratic socialist countries of Western Europe with high living standards and predominantly protestant church affiliation. The remainder of Western Europe scatters down the scale, with dictatorial regimes such as Spain and Portugal coming last. A fairly wide gap separates the Eastern bloc, led by Poland which is perceived as the most liberal of this group. East Germany, Bulgaria and Albania are the least desired countries for residential purposes.

GOVERNMENT ASSIGNMENT: THE VIEW OF AN AFRICAN ELITE

In most of the countries of Africa the number of university students and graduates is still extremely small. With the exception of a few in business and private law practice, most of them enter government service in a variety of diplomatic, administrative and teaching positions where the pace of advancement can be very rapid indeed. The result is that positions of considerable responsibility are often held by men and women only a few years away from their graduation. When such decision-making power is in the hands of a fairly small elite group, the view they hold of their own nation becomes of more than passing interest. For example,

---

[35] Once again, comments on this analysis would be welcomed. Usually higher order analyses are performed upon components which have been rotated according to some criterion that results in obliqueness. Thus the dimensions are no longer orthogonal, and a more parsimonious space and expression can be obtained. Here we are closer to a components analysis of canonical variates resulting from five, separate analyses.

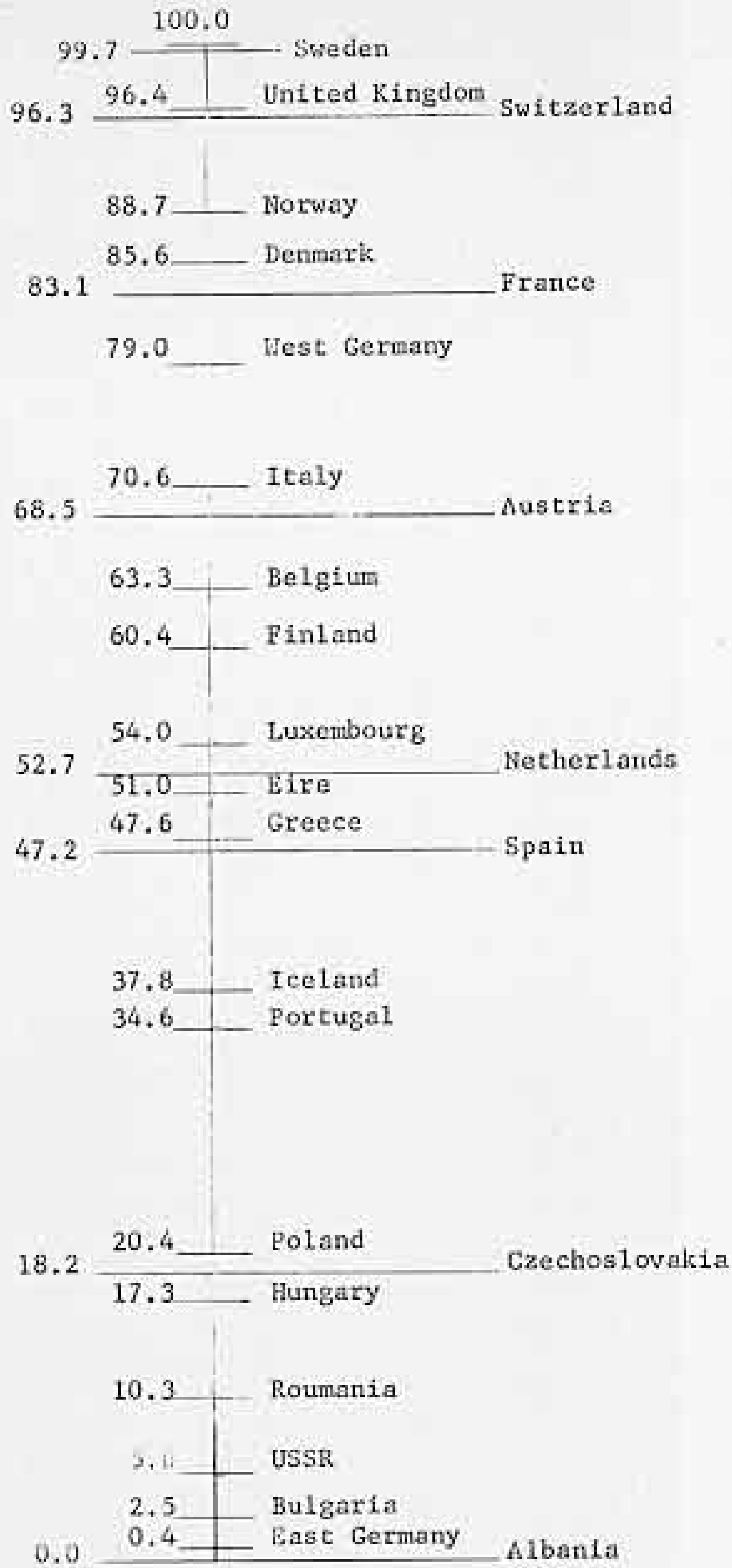


Figure 18: Space Preferences in Europe: The Overall View

investment must be assigned spatially, as well as sectorally, and the overall differential images held by such a group may influence the allocation of the meagre development funds that are under such intense competition.

Ghana

In sampling the mental maps of Ghanaian and Nigerian students,[36] the question was posed in terms of the residential desirability of districts given complete freedom of choice of assignment in some type of government service. In Ghana (Figure 19), there is a very high degree of agreement of spatial perception, with 67.4% of the total variation collapsing upon the first component. Generally the coastal districts containing the major urban centers all have high scores, and the preferences decline fairly regularly away from the southern "core" of the country towards the north. On successive components the drop in the explained variation is marked. After the variation in the overall mental map is extracted, the remaining components appear to be small regional effects that contrast one traditional area against another. For example, component II, extracting 7.5% of the variation (Figure 20), contrasts vividly the eastern Ewe area with the core area of Ashanti, and such an interpretation is bolstered by the close matching of contrasting signs in the factor loadings with places of birth and residence of the students. The third dimension highlights in a similar fashion a core area centered on the town of Koforidua north of Accra.

The degree of spatial regularity in the scores of the overall mental map (Figure 19), suggests that fairly simple perception surfaces could describe the national pattern with some accuracy and so allow us to separate broad trends from local anomalies. If the perception scores are related to their geographical

---

[36] I would like to thank Professors Amano Boateng and George Benneh of the University of Ghana, and Professor Akin Mabogunje of the University of Nigeria at Ibadan, for distributing the questionnaires and maps to their students. Samples from Uganda and Tanzania are still in the process of completion.

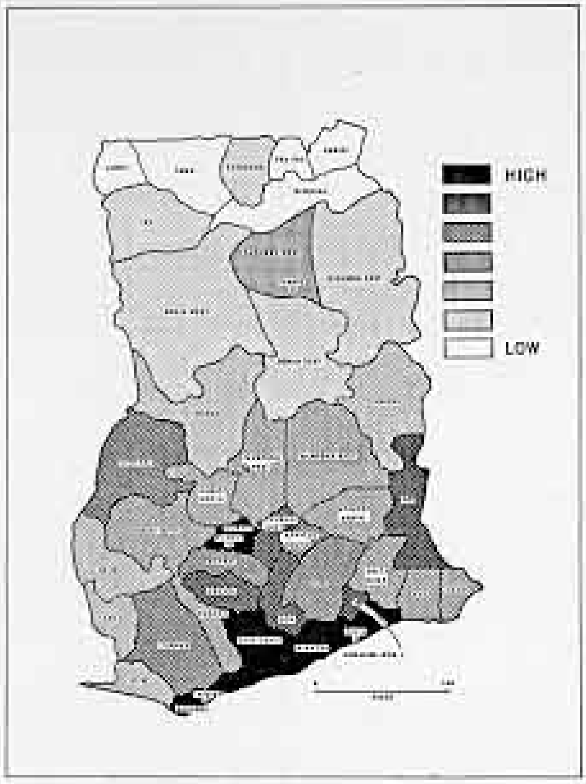


Figure 19: The View From Ghana:
The Most General Viewpoint

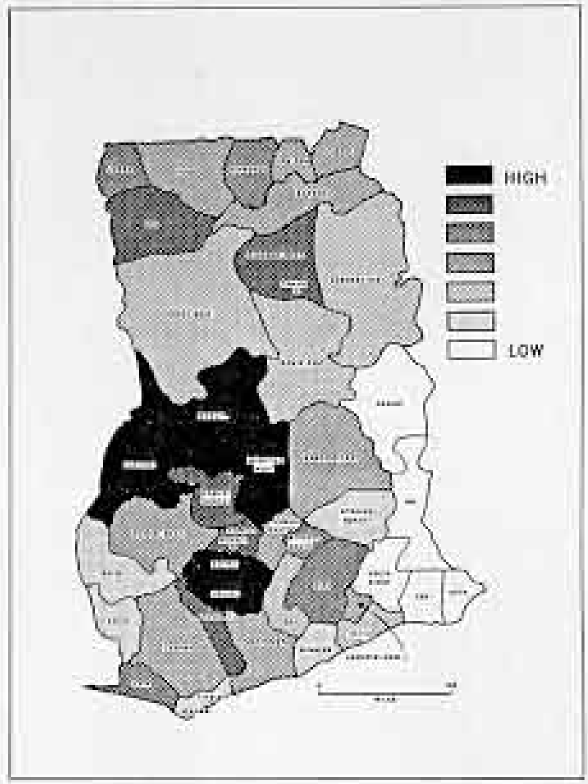


Figure 20: The View From Ghana:
Regional Contrasts

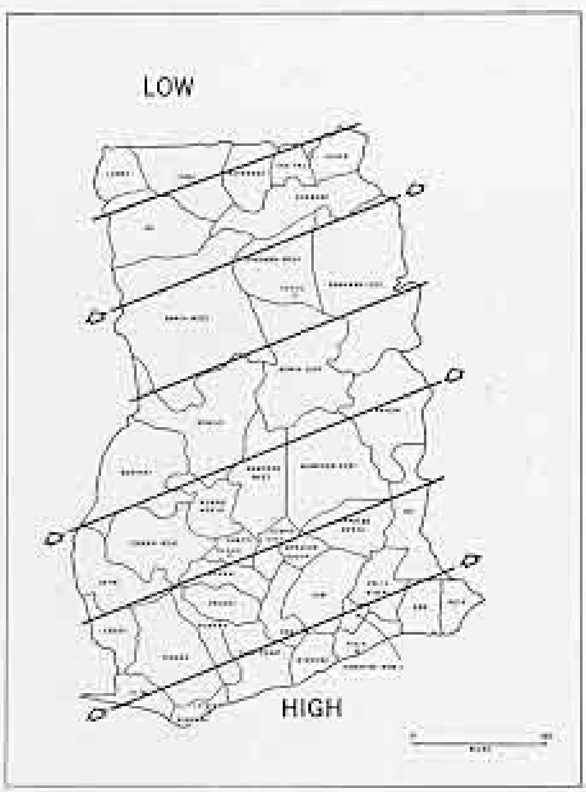


Figure 21: Ghana: The Linear
Perception Surface

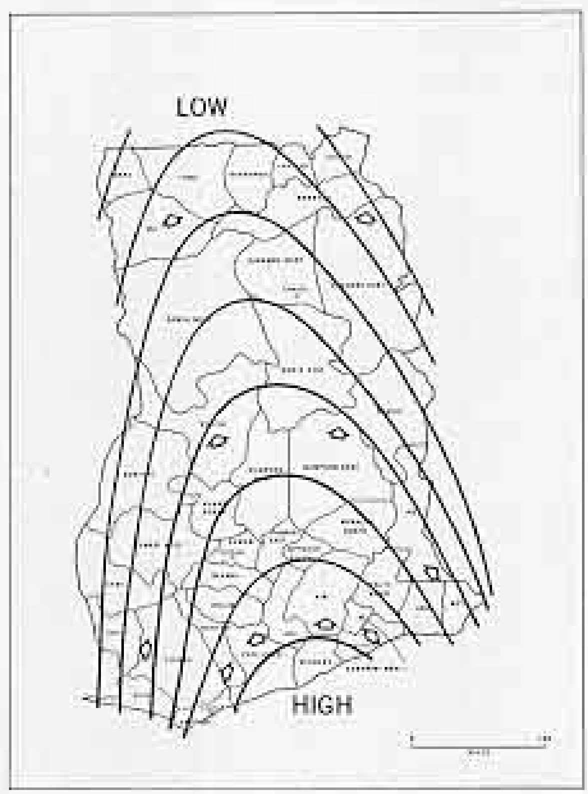


Figure 22: Ghana: The Quadratic
Perception Surface

coordinates, the simplest surface (Figure 21), which in terms of sums of squares accounts for 64% of the variation, is a plane tilted in a northwesterly direction whose southern edge almost parallels the coastline. However, at the expense of only slightly greater complexity, the quadratic surface is a much better description, accounting for an additional 15% of the variation (Figure 22). It appears to be the "best" in the sense that it combines parsimony of equational terms with an ability to account for perception scores by their geographical locations. For example, adding the cubic terms only raises the sums of squares by 3%, from 79% to 82%.

Having estimated the overall trend surface in an objective fashion, the anomalies may now be interpreted (Figure 23).[37] Here the power of the urban centers to cast bright images in the minds of men is shown all too clearly. Accra the capital of Ghana, Kumasi the center of Ashanti and the principal inland town, Ho the administrative center of the Volta Region, and Tamale the center of the north all have images far brighter than the overall trend would predict. Even Kusasi, Navrongo and Wa in the far north are above the surface, for they contain the main towns and administrative centers. Of particular interest is the image that the Sunyani district holds, due not only to the presence of a main town, but because this is an area that is exploding economically. Here traditional authority and modern higher education blend to sparkplug an area whose dynamism is clearly perceived so that it appears much more attractive than its peripheral position away from the core would suggest. Similarly Enchi in the southwest is higher than one would predict due to recent economic developments in the area. It, too, partakes of the pioneer glow, perhaps because of the increased information people have about these areas from newspapers and news reports.

[37]The notion that spatial relationships may be illuminated by contrasting opposite anomalies finds an intriguing expression in a rather different context, see F. C. Happold, Religious Faith and Twentieth Century Man (Harmondsworth: Penguin Books, Ltd., 1966), p. 46.

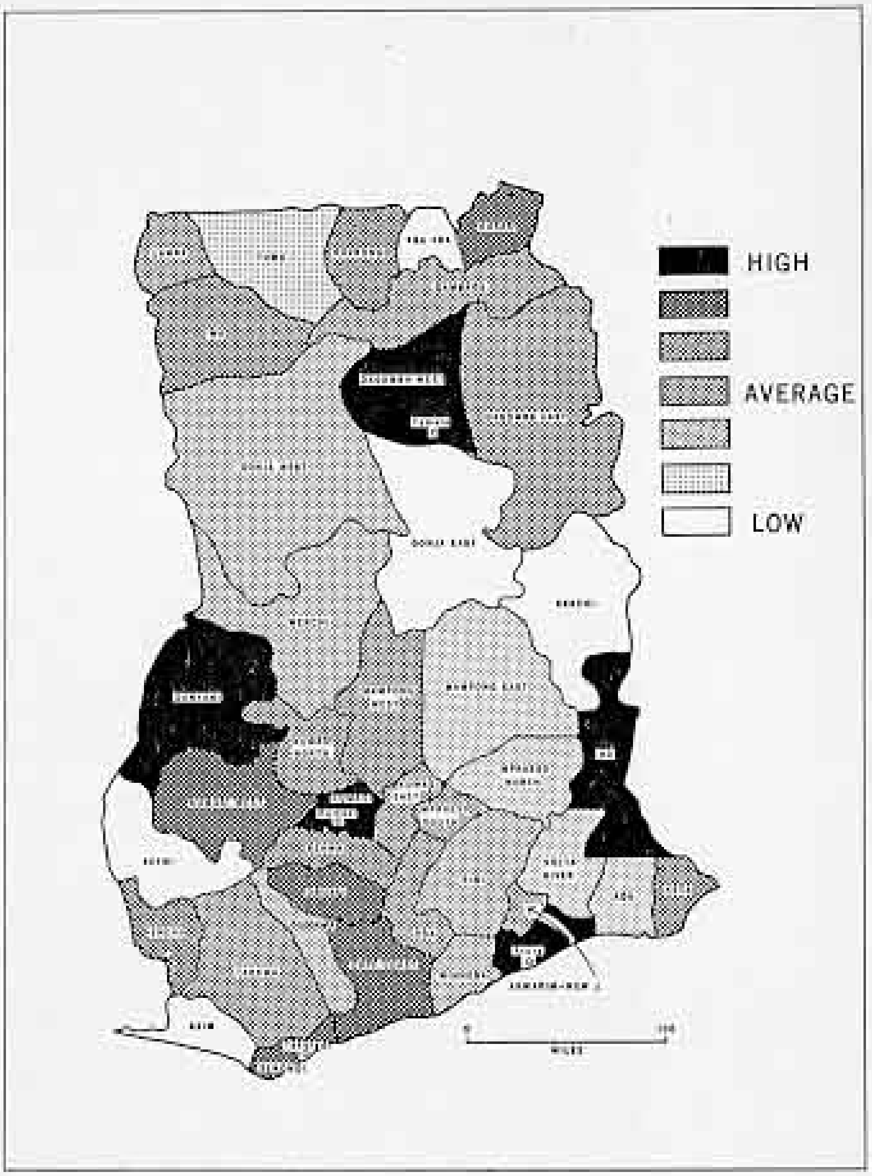


Figure 23: Ghana: Residuals from the Quadratic Perception Surface

The negative residuals, those areas perceived as even less desirable than the overall trend would suggest, lie mainly in a zone trending from the southeast to the northwest. Many lie in the Barren Middle Zone, an area that for environmental (tsetse fly and low, irregular rainfall), and historical reasons (devastated by slave raiding from the north and south), appear especially unattractive. Indeed, it is only now that population is trickling back to the area,[38] settling along the major north-south roads whose original purpose was to provide administrative and economic links between the poorer north and the bountiful south.

Nigeria

The overall viewpoint of university students in Nigeria is only slightly less homogeneous (62.5%) than that of Ghanaian students. For those who are aware of some of the less optimistic political prognostications that have been made in recent years, such agreement will come as a pleasant surprise. While Nigeria is assumed to face divisive forces along traditional regional lines between the North, Southeast and Southwest, the strong agreement in the space preferences of university students has considerable implications for forging national unity. Despite different backgrounds and home residences, nearly all the sample loaded highly on this first, general dimension indicating that each agreed to a marked degree about this overall mental map (Figure 24). The perception surface, however, is not so simply described as in the case of Ghana. Two distinct cores appear; one a band stretching over most of the southern portion of the country from the Yoruba east to the Ibo west, the other a perceptual peak in the north centered on the Jos, Zaria and Kano districts. While the overall linear trend is clearly from the southwest to the northeast, the variation accounted for by even the quadratic

---

[38] David Grove, Population Patterns: Their Impact on Regional Planning (Kumasi: The Kwame Nkrumah University of Science and Technology, 1963), pp. 13-14 and map p. 47.

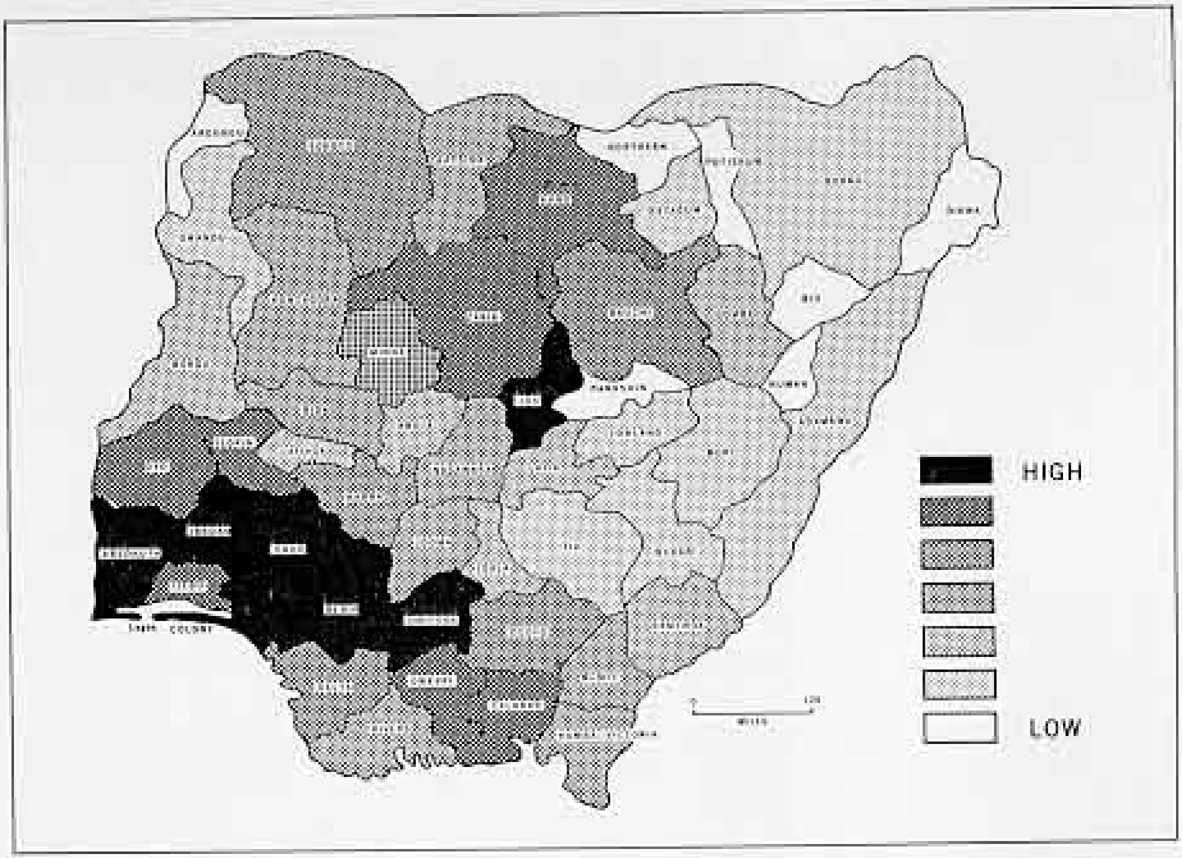


Figure 24: The View From Nigeria: The Most General Viewpoint

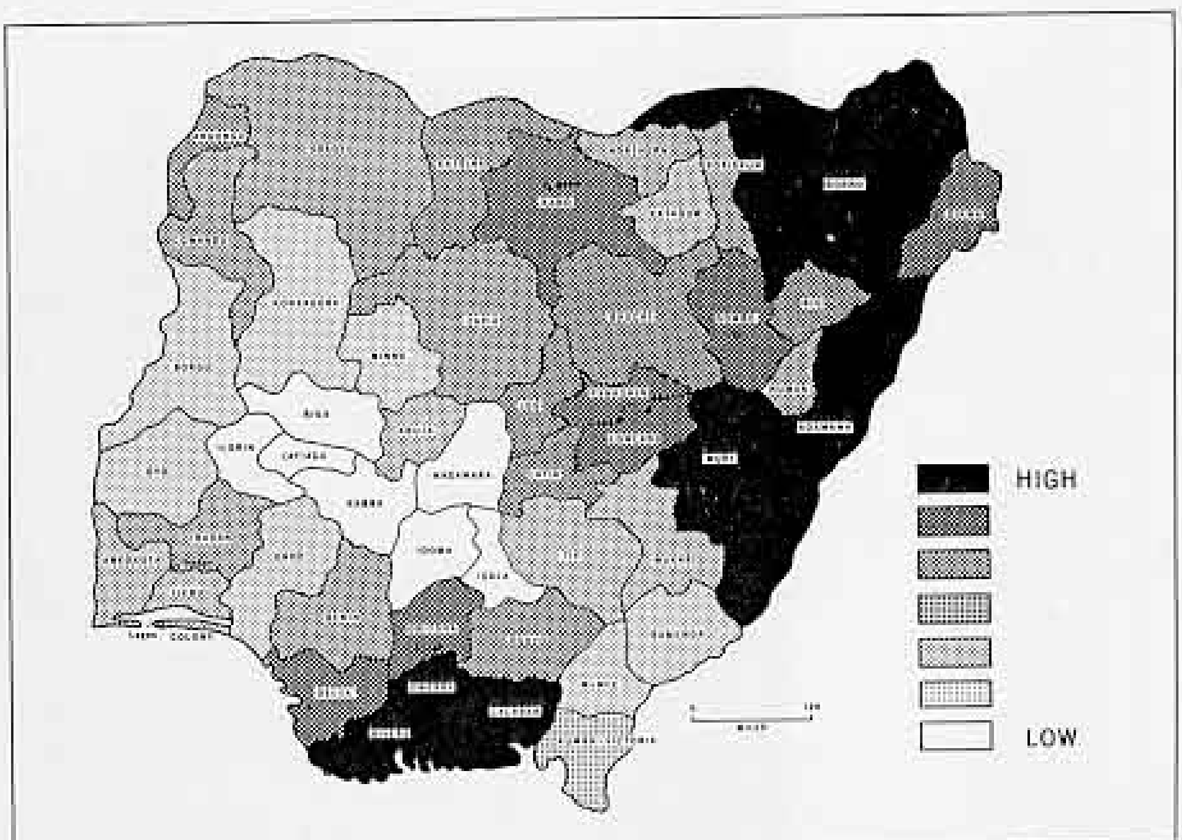


Figure 25: The View From Nigeria: Regional Contrasts

surface (52%) is considerably lower than in the case of Ghana. One is reminded vividly of Whitten's examination of the multiple intrusion hypothesis in the Lacorne granitic massif,[39] where two cubic surfaces were required to raise the sums of squares significantly. Here there is little doubt that a similar division of the area along the Barren Middle Zone, an area that is quite marked on the mental map, would provide a much more accurate description of the overall surface. Nigeria, at least in the minds of this small elite group, seems to be divided into two, quite desirable parts -- the Northern core and the Southern band composed of a blend of the Eastern, Western and Benin regions.

Away from the Northern core perceptual scores drop sharply, and much of the northeastern part of the country is not regarded at all favorably. Generally there is some element of "peripheralism" as in the Ghanaian map, and it is striking the way in which both of the perception surfaces closely reflect the pattern of road density.[40] Thus, we may have some confirmation that road density is a useful surrogate measure not only of accessibility, but those aspects of the modernization and development process that seem to be closely related to this slippery, but useful concept.

The second component (Figure 25), like the subsidiary dimensions of Ghana, accounts for only 5% of the variation and is a scale upon which small

---

[39] E. H. T. Whitten, A Surface-Fitting Program Suitable for Testing Geological Models which Involve Areally-distributed Data, ONR, Geography Branch, Technical Report No. 2, Contract Nonr 1228(26), discussed in R. J. Chorley and Peter Haggett, "Trend Surface Mapping in Geographical Research", Transactions and Papers of the Institute of British Geographers, Publication No. 37, 1965, pp. 56-57.

[40] See, for example, Peter Gould, The Development of the Transportation Pattern in Ghana (Evanston: Northwestern University Department of Geography Research Series, No. 5, 1960), p. 109; and Edward Taaffe, Richard Morrill and Peter Gould, "Transport Expansion in Underdeveloped Countries", The Geographical Review, Vol. 53, No. 4, October 1963, pp. 512 and 515, in which reference is also made to the peripheral border areas away from the main cores.

regional effects are contrasted. Compared to the strength of the first dimension, these do not appear to be strong or important, and they support the idea that for Nigerian university students petty regionalism has long been put aside.

ON THE RECONSTRUCTION OF PERCEPTION SURFACES

On the assumption that people's actions in an area may be partially related to their perception of the space and the differential evaluations they place upon various portions of it, it is possible that by working backwards we can make some rough reconstructions of the mental images held by men long ago. For example, in a recent study of the process of historical settlement in western New York State just after the revolution,[41] trend surface analyses up to the cubic were carried out in which dates of first settlement in an area (time dimension) were related to the geographical locations (two space dimensions).[42]

We might consider the even march of the isochrones defining the simplest, or linear surface (Figure 26), as indicative of the waves of settlers that might have moved across the country from east to west if the area had been perceived as a uniform transport surface completely isotropic in all the opportunities it presented to settlers at that time. However, such an assumption is obviously not tenable. Roads and tracks were beginning to lace the area at this time, making it easier to travel in some directions than others, and the information people had about different portions of the space

---

[41] I would like to thank Mr. Gary Fuller, NDEA Fellow in Geography, The Pennsylvania State University, for giving me permission to use the maps and information from his study "Western New York: A Culture Hearth?", June, 1966, to which reference should be made for a much fuller [sic] discussion than is possible here.

[42] The idea comes from: Peter Haggett and Richard Chorley, "Trend Surface Mapping . . .", op. cit., p. 64.

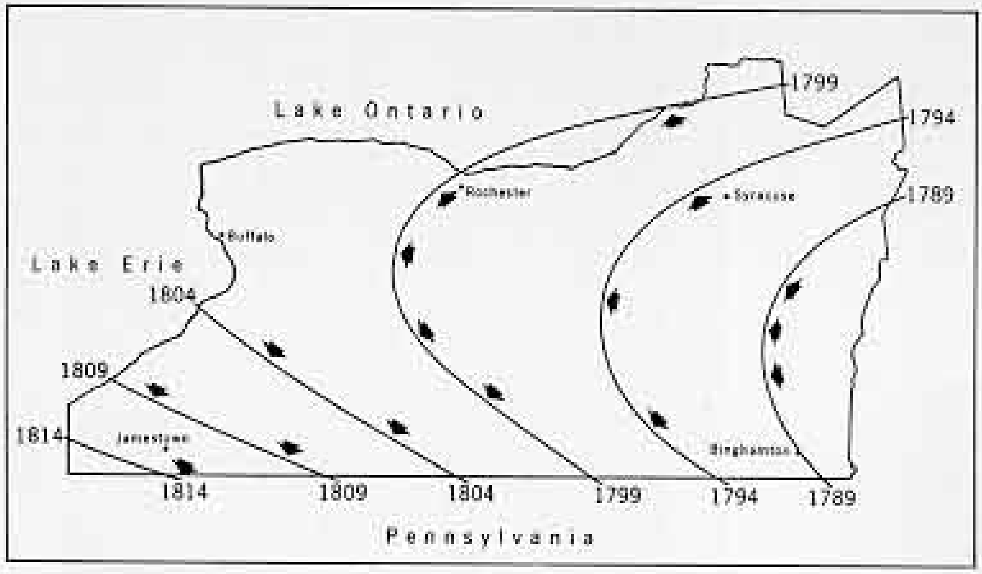


Figure 28: Western New York State: The Quadratic Surface

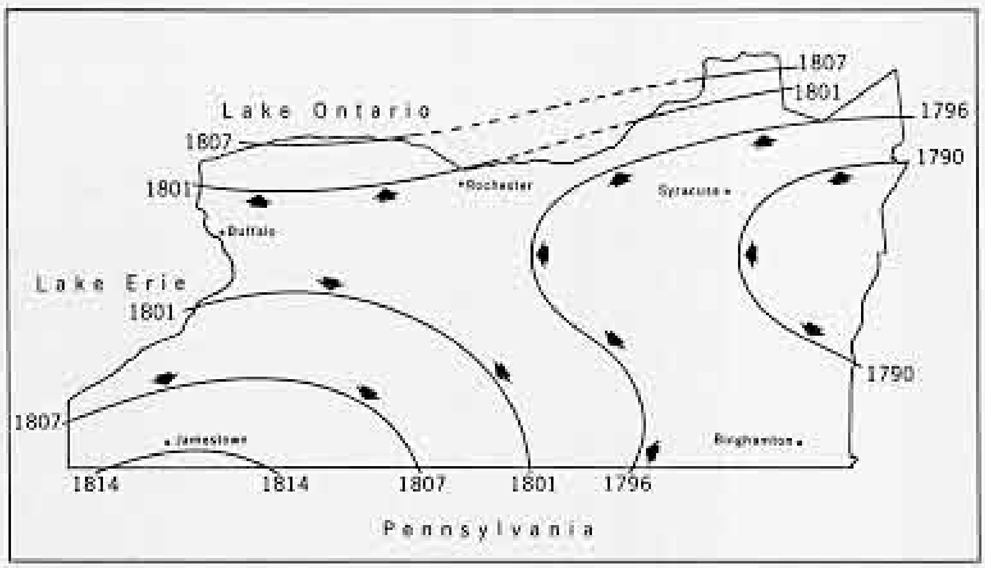


Figure 29: Western New York State: The Cubic Surface

varied and was strengthened by differential feedback processes. While analogies may be dangerous, I agree wholeheartedly with Bauer that they " . . . may play two roles: the scientific role of developing generalized knowledge and the practical role of illuminating other events".[43] Thus, in the same way a submarine valley can distort an evenly spaced wave train (Figure 27)[44] so we might think of the underlying surface of perception distorting the even waves of settlement over the land. Fitting the quadratic surface (Figure 28), which represents the next level of accurate description gained at the least expense of complexity,[45] provides us with some notion of the ease of travel in certain directions, the information flowing back to the points of origin, and the way opportunities were perceived by the people at the time. The even settlement waves are pulled along the main route to the west, and the lakes to the north and south of this main corridor are marked. Describing the time and space relationships with the next most complex surface, the cubic (Figure 29), indicates even more strongly the way in which the Lake Ontario plain was perceived as a less desirable area for settlement, for the time gradient is extremely steep to the north as the settlers by-passed it in their push westwards along the Lake Erie corridor to the New opportunities in Ohio. This was also an area of military activity where towns were frequently raided by the British in the early years of the nineteenth century.[46] Similarly, the southwestern corner forms a pocket of late settlement in an area of rougher terrain that was filled in after the initial waves of settlers had pushed into

---

[43]Quoted by Bruce Mazlish, The Railway and the Space Program: An Exploration in Historical Analogy (Cambridge: MIT Press, 1965), p. xiii.

[44]Blair Kinsman, Wind Waves: Their Generation and Propogation on the Ocean Surface (Englewood Cliffs: Prentice Hall Inc., 1965), p. 19.

[45]Sums of squares increased from 36% to 49%.

[46]Fuller, op. cit., p. 11.

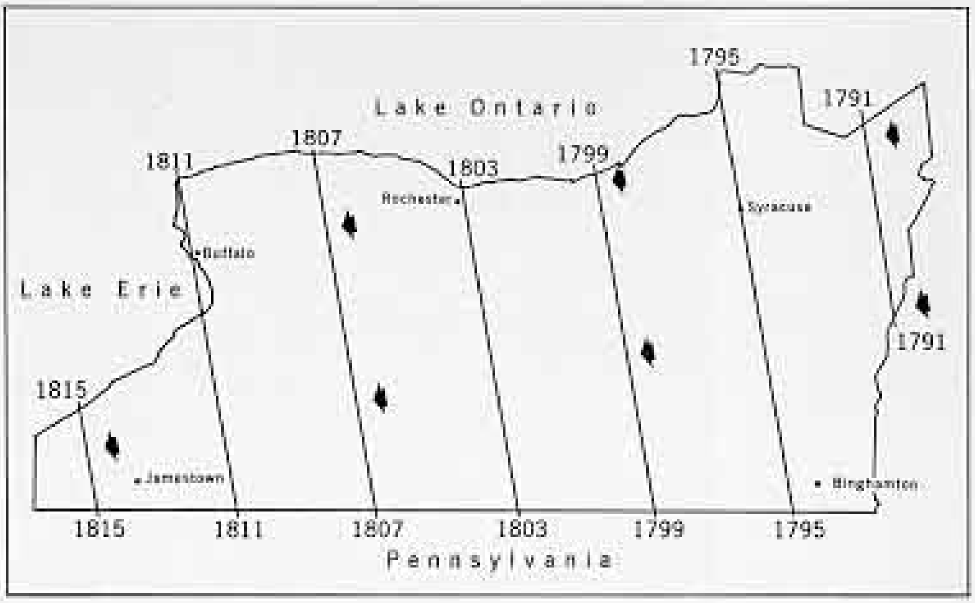


Figure 26: Pioneer Settlement Waves in Western New York State

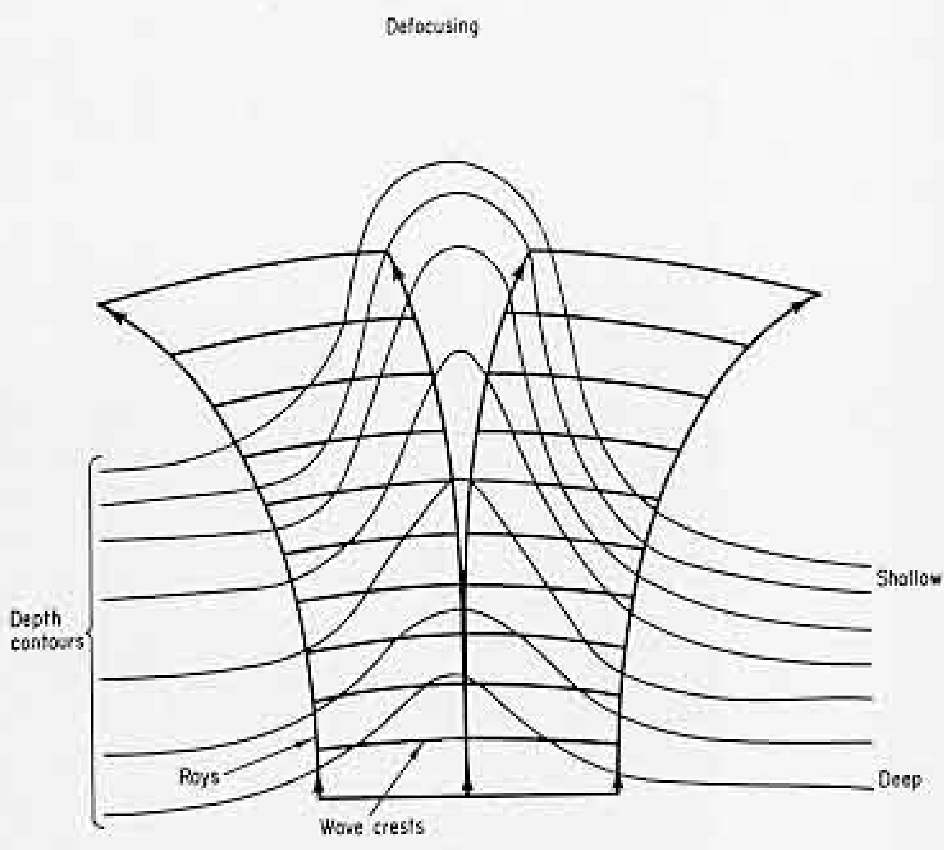


Figure 27: The Refraction of a Wave Train over a Submarine Valley. From Blair Kinsman, WIND WAVES: THEIR GENERATION AND PROPAGATION ON THE OCEAN SURFACE, (C) 1965. Reprinted by permission of Prentice-Hall Inc., Englewood Cliffs, New Jersey, and the author.

the new lands of the west.[47]

To separate out the broad regional regularities and trends from the smaller local effects has always been a challenging task for the geographer. Where a dynamic spatial process such as pioneer settlement is going on, the use of trend surface analysis, combining space and time, may not only allow us to achieve such a goal in an efficient and objective manner, but obtain, in addition, some insight into the mental images that men held at the time. If such a notion is valid, we may be able to examine the way in which such mental maps change through time, and so trace the line of inheritance for these images. Perhaps a series of careful content analyses may allow us to observe which areas maintain their brightness in the minds of men, and which are quickly tarnished as new opportunities, new technologies and new values change the very matrix in which they are evaluated and perceived.[48]

SOME IMPLICATIONS OF MENTAL MAPS

What are some of the implications of these mental maps, and what lines of further investigation seem worth pursuing? I hope the examples have indicated that there may be an area of enquiry here that is not only geographically intriguing, but one that smears the line between pure and applied research. For perhaps the most obvious implications lie in the broad area of planning, whether this is undertaken by governments or individuals. Many locational decisions in industry are going to be influenced by the mental maps of a few, key people. We

---

[47]Maps of residuals highlight the areas that were perceived as particularly attractive or repellent, see Fuller, op. cit., p. 14.

[48]Useful references include: Richard Budd and Robert Thorp, An Introduction to Content Analysis (Iowa City: State University of Iowa Press, 1963); Robert North, Ole Holsti, M. George Zaninovich and Dina Zinnes, Content Analysis (Evanston: Northwestern University Press, 1963); while imaginative applications include David McClelland, The Achieving Society (New York: VanNostrand and Co. Inc., 1961) and Richard Merritt, "Systems and the Disintegration of Empires", General Systems Yearbook, Vol. VII, 1962, pp. 91-103.

can see this in the choices of many footloose industries in this country, while in England the image of the southeast is becoming a source of continuing frustration for planners trying to disperse new factories away from the London "magnet" to relieve congestion and to pump-prime other areas that are in need of additional employment opportunities. Even the channel tunnel, which will simply bolster the locational advantage of the southeast, is receiving criticism on the grounds that it will reinforce the pull of the area.[49]

In much of the underdeveloped world, the allocation of social investment is still of critical concern as many countries try to forge the basic infrastructure of transport, education, sanitation and health facilities. Are the areas that are already "mentally bright" going to receive a large share because they are prominent already in the minds of men? Would an awareness and self-knowledge of this tendency have any beneficial influence? The stricture "Unto them that hath shall be given" seems to describe the basic features of a system of allocation with strong feedback features to produce the agglomerations and clusters of goods and people that are the main feature of the urban revolution.

There are also some implications for administrative planning. In the African countries particularly, the mental maps closely corresponded to the accessible, modernized areas illuminated by the bright lights of the cities and towns. Yet one of the great needs in most of these countries is to get people, particularly teachers of all kinds, into the "bush" areas that are so disparagingly viewed. Are there not some implications here for incentive allowances that might be inversely related to the perceptual scores that various areas receive? Of course, this is not a problem unique to Africa. Salaries for teachers in Alaska are incentively inspired beyond the difference in the cost of living, and the Soviet Union is using

---

[49] Anon., "Under and Over", Manchester Guardian Weekly, July 7, 1966, p. 8.

very high incentive pay to lure her people into the new and dynamic lands of Siberia.[50]

In the area of migration, too, mental maps may shed some light on the gross and long-term movements of people. Thomlinson, for example,[51] after trying to estimate the effect of many variables on migration in the United States, comments upon the high residual variation of the Pacific states. Interestingly enough, they are all part of a prominent ridge of desirability that is consistent across all the mental maps of the students sampled. Similarly, the areas of marked migrational loss in this country, the Great Plains in particular, are low troughs and sinkholes. The implications for depressed areas are obvious, and in some of the backward pockets of Appalachia it would be useful to know about the mental maps of the young and the old.[52] In England, work is currently proceeding on the mental maps of pupils about to leave school in the hope that they will shed some light on the migrational streams of young people that are causing such concern to regional planners.[53]

At the more academic level, the mental maps raise the question of the geographical implications of the informational flows to which people are subjected. Many writers, across a range of disciplines and concerns, have commented

---

[50]Some investigations are proceeding in New Guinea on the mental maps of district officers for administrative assignments.

[51]Ralph Thomlinson, "A Model for Migrational Analysis", Journal of the American Statistical Association, Vol. 56, No. 295, September 1961, pp. 675-689.

[52]Two thousand questionnaries were recently obtained from high school students by Mr. Robert Ziegenfus, Department of Geography, The Pennsylvania State University, on this topic.

[53]In cooperation with Mr. Peter Haggett, the author is receiving returns from thirty schools widely scattered throughout England, Wales and Scotland. The results will be used in the 1966 Madingley Lectures at Cambridge University, and will be reported upon at the NSF-sponsored symposium on Advances in Cultural Geography, Columbus, Ohio, November, 1966.

upon the way in which viewpoints are molded by the available information. As Herbert Simon notes in a critique of some common cliches:

> Does a man live for months or years in a particular position in an organization, exposed to some streams of communication, shielded from others, without the most profound effects upon what he knows, believes, attends to, hopes, wishes, emphasizes, fears and proposes?[54]

What are the flows of information that form and mold the surfaces of mental maps? All other things being equal, do they change in content and intensity as one moves up the ladder of central places to the critical nodes of connectivity in an inter-urban network? St Paul's migration to Rome may well have been influenced by his mental map of the geographic space that comprised his "world". Surely, by his demonstrated awareness of the relationships between location, information and space, modern geographers can claim him as one of their own? After all, they claim the best of everything else!

In Western New York, it was noted that differential flows of information may have had a profound influence upon the rate and direction of pioneer settlement. At a later time, and a little further west, Cochran has described the psychological effect of the railway in altering geographic horizons,[55] and the way the "big city" newspapers raised the information level of the rural population and altered their consciousness of time and space.[56]

Finally, there is the question of the information available to one generation, and the way it is filtered through the minds of the last. To what extent do we

---

[54]Herbert Simon, Administrative Behavior (New York: The Free Press, 1965), p. xv.

[55]Thomas C. Cochran, "The Social Impact of the Railway", in Bruce Mazlish (ed.), The Railroad and the Space Program: An Exploration in Historical Analogy (Cambridge: MIT Press, 1965), p. 177.

[56]Ibid., p. 178.

inherit our mental maps? It would be interesting to sample the geographic images in successive generations to see what significant changes existed between them. To what extent, for example, is the bright image of the Jos district in Nigeria due to it being the traditional "local leave" resort for the European population in colonial times? How closely would the mental maps of district officers in the 1920's and 1930's match those of the present? As Happold notes in a different context:

> One cannot stress too strongly the extent to which our world view has been conditioned by our mental history and development.[57]

Using the notion of the positions of neutral points as indices of parochialism, we may be able to measure, crudely to be sure, but <u>measure</u> nevertheless,[58] the changes in mental images from one generation to the next. As Mencius noted more than two millenia ago:

> By weighing, we know what things are light, and what heavy. By measuring, we know what things are long, and what short. The relations of all things may thus be determined, and it is of the greatest importance to measure the motions of the mind. I beg your Majesty to measure it.[59]

---

[57] Happold, <u>op</u>. <u>cit</u>., p. 40.

[58] Louis Guttman, "The Nonmetric Breakthrough for the Behavioral Sciences", invited address to the Automatic Data Processing Conference of the Information Processing Association of Israel, January 5-6, 1966.

[59] Mencius, circa 335 B.C., quoted in Truman Kelley, <u>Essential Traits of Mental Life</u> (Cambridge: Harvard University Press, 1935).

On the Length of Empirical Curves

by

Julian Perkal

Translated from

Julian Perkal, "O Dlugosci Krzywych Empirycznych", Zastosowania Matematyki, III, 3-4 (1958), pp. 257-286.

by

R. Jackowski

under the direction of

Professor W. R. Tobler
Department of Geography
University of Michigan
Ann Arbor, 1965

In this paper, I shall point out difficulties encountered in the measurement of the length of empirical curves, define a new method for approximating lengths of order ϵ, and describe a longimeter used for the measurement of such lengths.

1. We will consider arcs of curved lines on a plane (henceforth simply arcs) as pictures of homeomorphical sections. By the length of an arc we understand the limit of the sum of segments of length inscribed in this arc, by letting the segments go to zero. Such a limit (finite or infinite) always exists. Arcs of finite length we will call rectified. By the distance between arcs A and B we take a small number r, such than an arbitrary point on either arc A or B is not further than r from some (nearest) point on the other arc (See Fig. 1). The distance between two arcs is zero if and only if these arcs coincide.

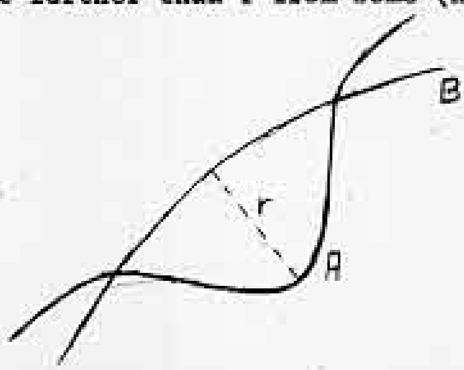


Fig. 1

If we know the manner in which the arc was constructed (for example if we know the equation of the arc) we can attempt to state whether the arc is rectifiable or not. If it is rectifiable we can ask for the length of the arc. If however, we do not know the manner in which the arc was constructed then we cannot assume that the arc is rectifiable.

This is the situation with respect to empirical curves. The circumference of leaves, the length of the seacoast, or of the edges of sharp razors are examples of arcs which may not be rectifiable. H. Steinhaus points out (Ref 6), that sharp razors observed with the naked eye have the shape of a straight line, but the same razor observed through a magnifying glass has an entirely different shape because of small notches; the same razor observed under a microscope again has an entirely different shape, depending on the structure of the steel. Figure 2 shows the edge of a leaf (Peucedanum sp.) observed (a) with the naked eye, (b) magnified ten times, (c) 100 times, and (d) 1000 times (these illustrations were kindly provided at my request by Z. Hejnowicz from the Academy of Anatomy and Cytological Plants at the University of Wroclaw). The true appearance of the edges of leaves and razors is not known, and it is not known whether the edges are arcs which are rectifiable.

Can these small deviations, perceptible only through a microscope, actually give rise to a real difference in the length of the arcs? The answer is yes. The length of an arc is not a continuous functional. Arbitrarily near a given arc A we can draw another arc B (saw-toothed, as in Fig. 3) whose length is considerably greater than that of arc A. Thus a nearly rectilinear sharp razor observed with the naked eye appears as a great deal longer and more complicated curve when observed under the microscope. And if we observe the razor under progressively stronger microscopes we will see a progressively more complicated curve, and consequently one of greater length. In fact we cannot even talk of the true and final shape of the curve. This is why we cannot talk of the true length of a razor. The area of a flat surface bounded by a curve behaves differently. The area of a plane is a continuous functional. If two curves lie near each other, than there is little difference in the areas bounded by the curves.

There exists a naive method of measuring the length of empirical curves. Lay a thread along the length of the arc, and then measure the length of the straightened thread. In practice however, it is impossible to complete the first of these instructions. No matter how we lay the thread, when we look through a microscope we will find that there are places where the thread deviated from the arc. A similar result holds for all other methods of measuring the length of empirical curves, whether it be leading a tracing wheel

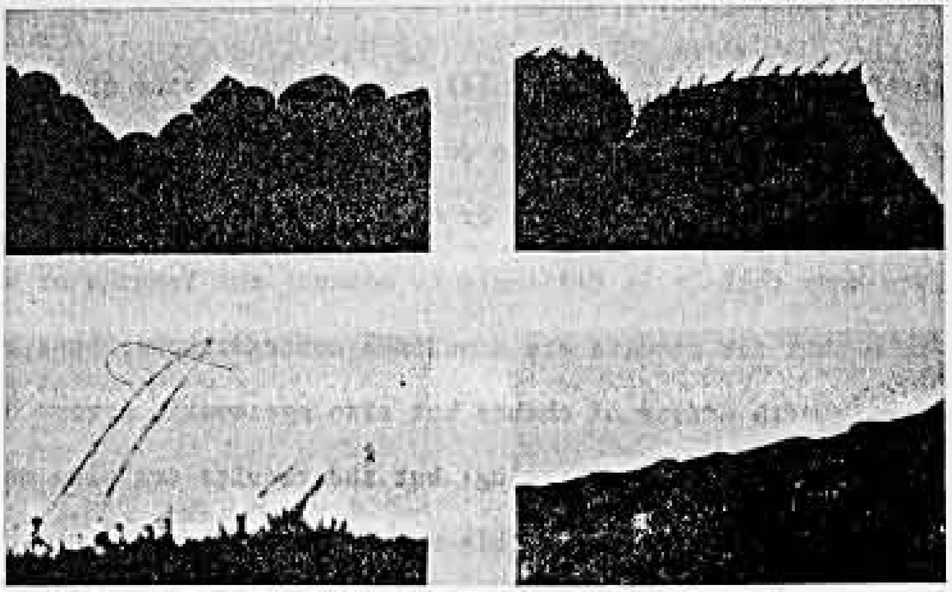


Figure 2
a b
c d

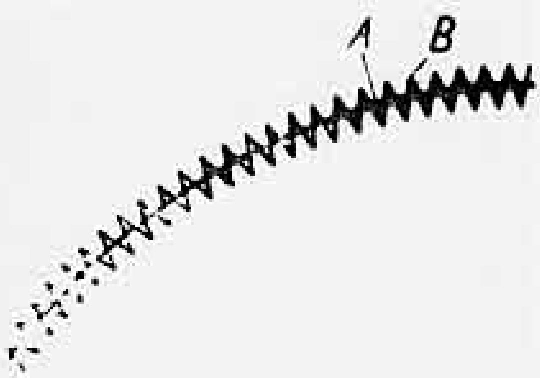


Figure 3

along the length of the arc, or incremental measurements using dividers. They must fail, for as I mentioned earlier we do not know how to ascertain if the measured empirical curve is rectifiable or not.

Naturalists know that it is difficult to measure the lengths of empirical curves; they know that the results are sometimes contradictory; that they are encumbered not only with errors of chance but also systematic errors depending on the method and on the person measuring; but the results are also encumbered with important general errors. This was first realized by geographers; the Viennese geographer A. Penck in 1894 presented the length of the seacoast between two points on the peninsula of Istrii (Ref 1) measured on various differently generalized maps. These are the results:

Map No.	scale	length in km
1	1:1 1,1600000,000	105
2	1: 3,700,000	132
3	1: 1,500,000	157.6
4	1: 750,000	199.5
5	1: 75,000	233.8

Note that the scale and the degree of generalization both effect the divergence of the results. But the degree of generalization should not be used. If map 4 is photographically enlarged ten times, we would have a map of the same scale as map 5, but the generalization thus obtained would be the same as generalized map 4, for this reason it would be expected that on this map the length of the measured section of the coast would be 199.5 and not 233.8 km. This would be so because the photographically enlarged map 4 would not bring out any new featuresathat are apparent on the more suitable map 5.

In the presence of such difficulties and divergencies the question arose: Is there,iin general, any sense in talking about the length of empirical curves? It appears that a new notion of length should be developed, which should also approximate the classical, but should be accepted in lieu of the classical method. Such a new length should serve for all empirical curves and should be simple for measurement.

H. Steinhaus has contributed in this direction (Ref 5 and other works cited there) and gives methods for measuring and comparing lengths of order n. These notions are based on results due to Crofton. Crofton's theorem relates the length of a curve lying in the original plane with a set contained in Crofton's plane. Take a curve lying on the first plane and a series of straight lines cutting this curve. Each one of these lines becomes a point of Crofton's plane. Now a measure of this set of points is equal to the length of the curve. However, points describing straight lines intersecting the measured curve twice must be counted twice, and in general, points describing straight lines intersecting our curve n-times must be counted n-times. Crofton's theorem should be expressed thus: the series K_i is a set of points of Crofton's plane, which describe the straight lines of the first plane intersecting the measured curve at least i-times. The length of the arc is the sum of the infinite series

$$(1) \qquad \sum_{i=1}^{\infty} K_i$$

2. Steinhaus' known longimeter is based on this equation. This is a sheet of transparent paper with parallel lines. This sheet is placed on the curve to be measured and the number of points that intersect the curve with the straight longimeter are counted. The number of these points of intersection are approximately proportional to the length of the curve. Many-fold repetition of this measurement increases the accuracy in that the error of chance becomes smaller.

But this longimeter, like all other equipment used for measuring classical lengths, is not suitable for measuring the length of empirical curves since it is not known whether the pertinent series is convergent.

The method presented permitted H. Steinhaus to introduce a new concept of length, length of order n. H. Steinhaus calls the length of order n the sum of the first n expressions of the series (1). A measurement of this length of order n occurs with the aid of the Steinhaus longimeter just as in the classical measurement of length, with the one difference that for every straight longimeter no more than n points of the intersection of the line with the measured curve are counted. If the straight lines intersect the curve less than n-times, or n times, then all of the points of intersection of the line and curve are counted. If however any line intersects the curve more than n times, it is accepted that this line intersects the measured curve only n times. In agreement with the accepted program, the length of n^{th} order is approximately the classical length in the sense that it tends towards the classical length when n increases to infinity. The length of order n serves for all empirical curves and is easy to measure. It would appear therefore that the problem is solved and that the length of order n instead of classical length can be introduced conventionally. The concept of the length of order n solves many questions pertaining to the measurements of empirical curves. This approximated length is useful in general for the comparison of the length of curved lines on geographical maps or on several maps of various generalizations. It can be admitted that in applying the length of the indicated order, for example the 5th, Penck would not have obtained such variations in the measurement of the length of the seacoast of Istrii. There are however natural problems for which the aforementioned length of order n is insufficient.

The arguments are as follows:

(1^{st}) A difficulty with the method is that there is a slightly weak connection between the length of order n and the classical length. Naturalists want to measure the true length, the so called classical length of the empirical curve, and want to believe that such a length exists. It seems that it is easier for the naturalist to digest generalization of a curve rather than to introduce a new abstract measurement, such as the length of order n. It is known that the length of order 1 gives the length of rectilinear classical curves of first degree and the lengths of order 2 give the lengths of convex classical curves. It is not known however for what class of curves of length of order n is the classical length when n>2. It is further known that by increasing n the length of order n increases and approaches the classical length. It is not known what is the connection between n and a suitable length of order n, the difference between classical length and the length of order n. If the naturalists would obtain an approximated length of measurement he would want to visualize a curve approximating the measured (generalized) one whose classical length would be equal to the length of order n of the measured arc.

(2^{nd}) Every method of approximating measured length must embrace a convention. In H. Steinhaus' method the number n is conventional in that it is the order of length. This integer is not dependent on customary units of length (centimers or inches). But a difficulty with this convention is the lack of correspondence between the arbitrary n and the suitable measurements of classes of curves whose length, classical and of order n, are equal. In certain cases it would appear to be more convenient to use a convention connected with suitably approximated measurements and which would define classes of curves for which the approximated length would be equivalent to the classical.

(3^{rd}) The length of order n, as in classical length, is a discontinuous functional which could lead to the aforementioned paradoxes of length. In measuring the length of order 10, for example, for intricate empirical curves it may occur that points of intersection of the arc with the longimeter line would appear under the microscope not as one point but as greater number of points of intersection. The length of order 10 is therefore dependent on whether the curve is observed with the naked eye or with a microscope or on the generalization of the curve. It could occur, that every point of intersection of the empirical curve with the lines of the longimeter is actually such a collection of points (in practice we could not prove this otherwise). In the meanwhile misgivings would arise that the length of order n would be simply n times length of order 1.

3. I will now describe a concept of approximate length of order ϵ, where ϵ is a real number. The length of order ϵ will be a continuous function of the curve and will depend on the number ϵ in a continuous manner. The length of order ϵ does not have the fault I previously indicated for the length of order n. A more comprehensive description of this concept can be found in my paper (Ref 3).

In the paper (Ref 2) I defined the collection A (X); that is, the ϵ-halo of the arc X, as the collection of all points on the plane not more than ϵ distant from the arc X.

$$(2) \qquad A_\epsilon(X) = \mathop{E}_{x}\left[(x,X) \leq \epsilon\right]$$

where E is Lebesgue's symbol, and (x,X) indicates the distance of the point x from the arc X; that is, the distance of point x from the nearest point on arc X.

Consider the area $a_\epsilon(X)$ of $A_\epsilon(X)$. Figure 4 shows the ϵ-halo of the arc X. As is evident it consists of a belt whose width is 2ϵ enclosing the arc X plus two semicircles of radius ϵ.

The length of order ϵ of the arc X I indicate with the symbol $L_\epsilon(X)$ and define by the following formula:

$$(3) \qquad L_\epsilon(X) = \frac{a_\epsilon(X) - \pi\epsilon^2}{2\epsilon}$$

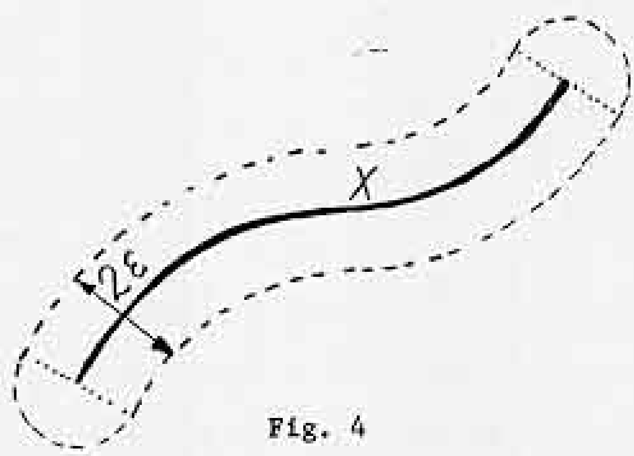


Fig. 4

For the arc shown in Figure 4 the length of order ϵ is therefore the area enclosing the arc X (the area of the ϵ-halo) minus the area of the two semicircles, divided by the width of this belt. If the arc X were a straight line, this belt would be a rectangle of width 2ε and of a length equal to the length of arc X. The area of the belt will not change if the arc X is subject to small variations. For straight arcs the definition (3) is intuitive and the length of order ε is equal to the normal length.

It is evident (see Ref 3) that the length of order ε is a continuous function of the arc; that is, if arc X lies close to arc Y, then the length of order ε of these arcs differs only slightly. This length is a diminishing and continuous function with respect to a changing ε, for example, if ε diminishes the length of order ε increases and the reverse. By this, if the change in ε is insignificant, the length of order ε either will not be subject to change, or the change will likewise be insignificant.

In an earlier investigation (Ref 4) I wrote on ε-convex sets. An arc is ε-convex, if a circle of diameter ε could fit on both sides of this arc. In other words, an arc is ε-convex, if every point on it has a radius of curvature of not less than ½ε. For curves of 2ε convexity (with the ends separated by at least 2ε) the length of order ε is equal to the classical length. For other curves the length of order ε can be equal to or smaller than classical.

Figure 5 represents an ε-convex set (ε = 5mm) and an arc X which is not 2ε-convex. The thin continuous line traces the ε-halo of the arc X. As

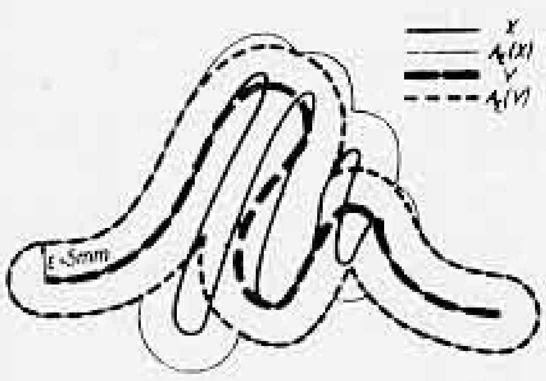


Figure 5

formula (3) shows, every arc, whose $\mathcal{E}$-halo is equal to $A_{\mathcal{E}}(X)$, has the same length of order $\mathcal{E}$ as arc X. In general, if it would be possible to draw the arc Y as $2\mathcal{E}$-convex and with $\mathcal{E}$-halo equal to $A_{\mathcal{E}}(X)$, the classical length of arc Y would be equal to its length of order $\mathcal{E}$ and equal to the length of order $\mathcal{E}$ of the arc X. Furthermore, we would be able to show that arc Y approximates the given arc X (lying in $A_{\mathcal{E}}(X)$), whose classical length would then be the length of order $\mathcal{E}$ of arc X. Unfortunately, such an arc Y in general does not exist. We can only inscribe the $2\mathcal{E}$-convex arc Y in such a manner that one has $A_{\mathcal{E}}(Y) \subset A_{\mathcal{E}}(X)$. Such an arc Y with its $\mathcal{E}$-halo is drawn in Figure 5 as a broken line. In the meanwhile it is known that $L_{\mathcal{E}}(X) \geqslant L_{\mathcal{E}}(Y) = L(Y)$, where the last symbol represents the classical length. We therefore know how to draw arc Y approximately the same as arc X, with a classical length less than $L_{\mathcal{E}}(X)$ with a difference depending on $\mathcal{E}$. It is easy to estimate this difference; that portion of the collection $A_{\mathcal{E}}(X)$ that is not covered by the $\mathcal{E}$-halo of arc Y is divided by $2\mathcal{E}$, or $\left[a_{\mathcal{E}}(X) - a_{\mathcal{E}}(Y)\right] /2\mathcal{E}$.

The length of order $\mathcal{E}$ includes the conventional parameter $\mathcal{E}$ (of known length) and indicates a difference between the classical length of rectifiable arcs and the length of order $\mathcal{E}$. I wrote above that arcs with small curvature (radius of curvature not less than $\mathcal{E}$) have as the length of order $\mathcal{E}$ the classical length. It can be seen that the length of order $\mathcal{E}$ measures accurately an arc of very small curvature and other arcs in an approximate matter; consequently the error of approximation is larger the greater the curvature of the measured arc. For unrectifiable arcs this error is infinitely large, since the classical length of such arcs is infinitely large, but the length of order $\mathcal{E}$ is finite. From this it is evident that the actual error, or the difference between classical length and length of order $\mathcal{E}$ should not be a standard of accuracy of measurement, for the error depends not so much on $\mathcal{E}$ as on the arc. For unrectifiable arcs the error is infinite (independent of $\mathcal{E}$) and therefore for empirical arcs is indeterminate. It can be agreed that the standard of precision will be the ratio = /L (X) (H. Steinhaus' suggestion). When decreases, the length of order (the denominator) will increase, and the ratio will also decrease. It can be conventionally required that in measuring the length of the arc X, a length of such order be used so that the ratio is not greater than 10% or 5%. The number can be chosen in another, more natural manner, by considering which small curves of the arc can be discarded (generalized) without harm to the problem under consideration. If we decide to disregard curves of radius less than , then the length of order will really be the length of such curves. We will return to this problem again after discussing generalization.

The length of order has properties which should be required of empirical length. It is a continuous function, and completely removes the paradox of length on the consideration that a very long arc may still be found near a short arc. For 2 -convex arcs the length of order is equal to the classical. For other arcs it is smaller than the classical length but increases in a continuous manner when we decrease . If an arc is rectifiable, then by decreasing to zero the length of order can in a sense be treated as an approximation of classical length, and, for a given measured arc, another arc can be constructed, whose classical length is similar (the error can be estimated) to the length order of the measured arc. The parameter of length , can be fitted to the arc, or to the problem, in a natural manner. This is related to the generalization of the arc. In the following portion of the paper I will describe a simple tool for measurement of length of order .

4. Figure 4 presents an arc X and its -halo. In formula (3), in order to designate the order length of this arc, the area a (X) of the -halo of the arc X must be measured. This can be accomplished with a so-called point planimeter (see Ref 7); this is a plane with a network of points arranged

regularly, as for example on a square lattice (see Figure 6) with points with coordinates (am,an) where m and n are whole numbers and a is an arbitrarily assumed real number (constant for the entire problem); or a triangular lattice with sides b (see Figure 7). Every such regular arrangement of points corresponds to a plane figure whose centers are points of the figure. The figures are sketched in Figure 6 and 7. In the case of rectangular points these are squares with areas a^2, and in the triangular coordinates they are hexagonal with area $\frac{1}{2}\sqrt{3b^2}$. The point planimeter is a network of points regularly distributed on a sheet of transparent paper. This sheet is placed on the set to be measured (Fig. 8) and the number of points falling within the set are counted. The area of the set is equal to the expected number of these points increased by the area of the pertinent figure. The area can be estimated with the aid of arithmetic means from a number of random applications of the planimeter to the figure. From the Jarniaka-Steinhaus' theorem the error with respect to the measurement carried out with the point planimeter is directly proportional to the length of the arc bordering the measured feature and inversely proportional to the area of this feature (for which the unity is a). If we decrease a q-times, then the length of the arc increases q-fold and the area of the feature q^2-times. In view of this the error with respect to the measurement decreases q-times. The denser is the network of rectangular points (for triangular network likewise), the more efficient is the planimeter, therefore the more accurate is the measurement carried out with the planimeter.

The measurement of area is carried out in the following manner: the point planimeter (rectangular with side a, triangular with side b) is placed k-times on the feature and the number of points falling in the area are counted. Denote the first application as n_1, the second n_2, ... the k^{th} as n_k. The area of the set is

$$\frac{a^2}{k}\sum_{i=1}^{k} n_i \qquad \text{(rectangular network)}$$

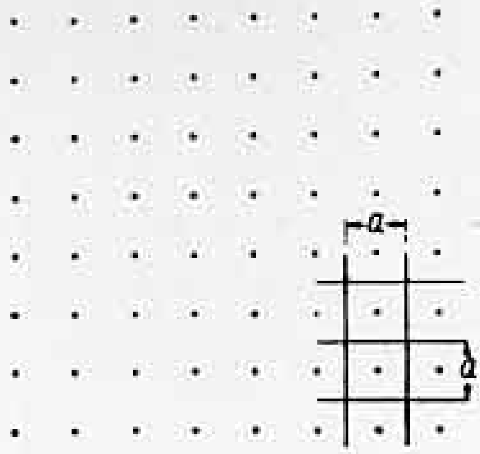


Figure 6

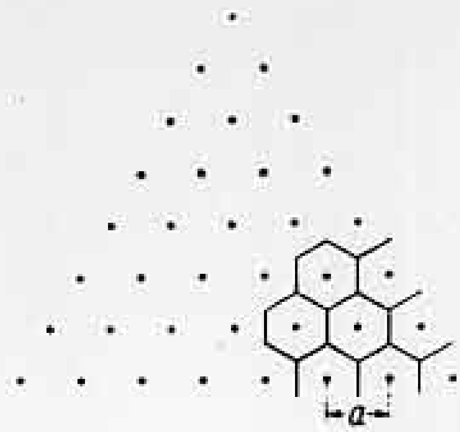


Figure 7

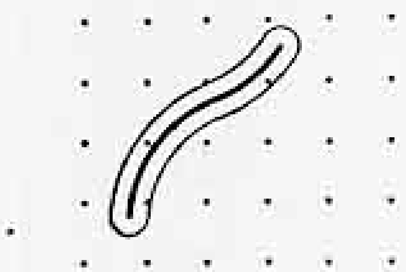


Figure 8

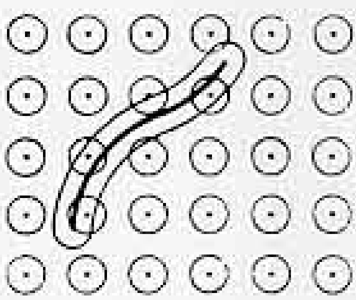


Figure 9

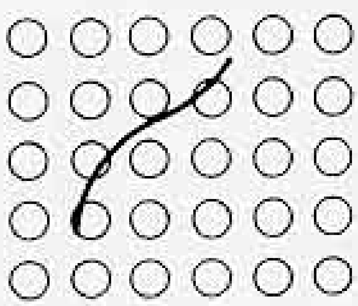


Figure 10

$$\text{or} \quad \frac{b^2\sqrt{3}}{2k}\sum_{i=1}^{k} n_i \qquad \text{(triangular network)}$$

From formula (2) we conclude (Fig. 9) that the point P falls in $A_\varepsilon(X)$ if and only if a circle of radius ε centered at point P encounters the arc X. This permits us to measure the area $a_\varepsilon(X)$ from a drawing on which only the arc X is traced (Fig. 10). It will be satisfactory to replace points with circles of radius ε on the planimeter. The drawing at the end of the paper gives such an arrangement based on a rectangular point planimeter with sides of the lattice a = 10 mm, that is, for ε = 5 mm. The second drawing gives a planimeter based on a triangular network with sides b = 8 mm, or for ε = 4 mm. With these drawings we use a light table and can count the number of circles encountered by the arc X. This incidence number increased by a^2 for the rectangular arrangement of circles and by $\frac{1}{2}\sqrt{3b^2}$ for the triangular, will be equal to $a_\varepsilon(X)$, the area of ε-halo of the arc X. For better estimation of the number of circles cutting the arc X, the measurement should be repeated k times. Summing the number of circles encountering the arc X for all k measurements and using the earlier mentioned formula we obtain from equation (3) the following:

$$\text{(4)} \qquad L_\varepsilon(X) = -\frac{\pi}{2}\varepsilon - \frac{a^2}{2k\varepsilon}\sum_{i=1}^{k} n_i \qquad \text{(rectangular lattice)}$$

$$\text{and} \qquad L_\varepsilon(X) = -\frac{\pi}{2}\varepsilon + \frac{b^2\sqrt{3}}{4k\varepsilon}\sum_{i=1}^{k} n_i \qquad \text{(triangular lattice)}$$

We add to equation (4) constants a, b, and k for the various ε's so that the tool will be most convenient and effective. We will call our tool for short an ε-longimeter (either rectangular or triangular). Because the planimeter with a denser number of points in the network is more effective it follows that a denser layout of circles in the longimeter will also be more effective. The longimeter in Figure 10 is less effective then in the larger drawings, but the circles would overlap and the use of such a longimeter would be inconvenient. The most convenient longimeter is one with circles touching each other, with $a = 2\mathcal{E}$ in the case of a square lattice, and with $b = 2\mathcal{E}$ in the case of a triangular one.

In the case of the square longimeter, the constant coefficients in equation (4) are simplified if they take the form $a^2/2k\mathcal{E} = 2\mathcal{E}/k$. This occurs if $k = 2\mathcal{E}$, where $\mathcal{E}$ is expressed in suitable units of length (for instance in millimeters) and the results will be in the same units. Formula (4) then becomes:

$$(5) \qquad L_{\mathcal{E}}(X) = -\tfrac{1}{2}\pi\mathcal{E} + \sum_{i=1}^{k} n_i$$

Thus, for example, on the large drawing the longimeter for $\mathcal{E} = 5$ mm is shown. The side of the lattice comes out as $a = 2\mathcal{E} = 10$ mm, and the number of repetitions of the measurement is $k = 2\mathcal{E} = 10$. The subtrahend $\tfrac{1}{2}\pi\mathcal{E} = 7.854 \approx 8$. The method for the measurement of the length of order $\mathcal{E} = 5$ mm of the arc X is very simple:

We lay the longimeter at random 10 times on the arc X and count the (grand total) number of circles falling on the arc X. Eight is subtracted from this sum to obtain the length of 5 mm order of the arc X in millimeters. The numbers a and k can be obtained similarly for arbitrary $\mathcal{E}$'s. Such numbers together with the subtrahend $\tfrac{1}{2}\pi\mathcal{E}$ for various $\mathcal{E}$'s are presented in Table 1:

TABLE I

$\mathcal{E}$	1	1.5	2	2.5	3	4	5	6	8	10	15	20	30	50
a	2	3	4	5	6	8	10	12	16	20	30	40	60	100
k	2	3	4	5	6	8	10	12	16	20	30	40	60	100
$\tfrac{1}{2}\pi\mathcal{E}$	2	2	3	4	5	6	8	9	13	16	24	31	47	79

Each column of this table with ε expressed in arbitrary units of length, permits the construction of suitable ε-longimeters. Thus, for instance, the longimeter in the large drawing is made with the aid of the column whose initial number is 5.

In the case of the triangular lattice the coefficient in equation (4) is $b^2\sqrt{3}/4k\varepsilon$; after substitution for $b = 2\varepsilon$ it will be equal to $\frac{\varepsilon\sqrt{3}}{k}$. It will be equal to unity for $k = \varepsilon\sqrt{3}$, which in the view that k (the number of times the measurements are made) must be a natural number, makes it possible to construct an equally suitable longimeter only for certain values of ε. For example, the ε-longimeter in the case of the triangular coordinates shown on the drawing has circles of radius $\varepsilon = 4.05$ mm. This is a number chosen so that $4.05\sqrt{3} \approx 7 = k$. The subtrahend $\frac{1}{2}\pi\varepsilon$ here is equal to $6.37 \approx 6$. The rule to use this longimeter is: we obtain the length of order $\varepsilon = 4.05$ mm for the arc X by sevenfold (random) application of the longimeter to the arc X, and subtract 6 from the total number of circles hit by the arc. Similarly, we can construct other sets of numbers for ε-longimeters using the triangular lattice. They are shown in Table 2.

TABLE 2

ε	1.15	1.73	2.31	2.89	4.05	5.78	8.67	11.5	14.4	28.9	57.8
b	2.30	3.46	4.62	5.78	8.10	11.6	17.3	23.0	28.8	57.8	115.6
k	2	3	4	5	7	10	15	20	25	50	100
$\frac{1}{2}\pi\varepsilon$	2	3	4	5	6	9	14	18	23	45	91

In the equations used for longimeters I recommend k-fold random applications to the arcs. In some cases it is advisable to use systematic applications of the longimeter instead of random. With this aim, a k-pointed star can be drawn (on the rectangular longimeter 10 points, on triangular longimeter 7 points) in the center of the longimeter. On the arc a section should be traced and one of its ends marked. Then the longimeter should be applied to the arc to be measured so that the center of the star falls on the marked point, and the sections of the arc traced on the points of the star. I-measurements of I-pointed stars should be applied to the section. In this case, it is not necessary to write down the several quantities of circles of the longimeter falling on the arc in the Ith application of the longimeter. It is sufficient to obtain the number at one counting, and after finishing k measurements (after going around the whole star) the proper subtrahend should be deducted. The length of order ε of the measured arc is then available immediately.

5. The above described measurement with the ε-longimeter gives the length of order ε of the arc X. If the arc X is 2ε-convex and has ends at least 2ε apart, then the length of order ε is equal to the classical length. If however the arc X is not 2ε convex, that is, has a radius of curvature at some place of less than ε, the length of order ε is a new quantity which we conventionally accept as the length. Because for empirical curves we cannot confirm, as mentioned earlier, whether they are rectifiable or not, the more so it cannot be confirmed whether they are 2ε-convex. It therefore should be noted that for all empirical curves the length of order ε is a new conventional measurement.

The shape of empirical curves can be very capricious. In an earlier paper (Ref 4) I brought attention to the need of simplifying empirical sets and proposed the use of ε-convexity. We will apply this to the simplification of the geometrical shape of an arc. Usually empirical curves are of interest as boundaries between two areas, as for example, the seashore, leaves and their background, and razors and their background. If we were interested in a curve which is not such a boundary we could artifically introduce, for example, a circle connecting the ends of the arc and two areas would be created which our arc divides.

Let A and B be two areas divided by the arc X (Fig. 11). In paper (4) I designated the 2ε-convexity of an area A by $C_{2\varepsilon}(A)$, or the smallest set of 2ε-convexity enclosed in area A. This is called an ε-generalization of area A. The edge of this set, or rather that portion of it which is enclosed in the closed area B will be designated by $G_\varepsilon(A/B)$ and will be called the ε-generalization of the edge of area A in area B:

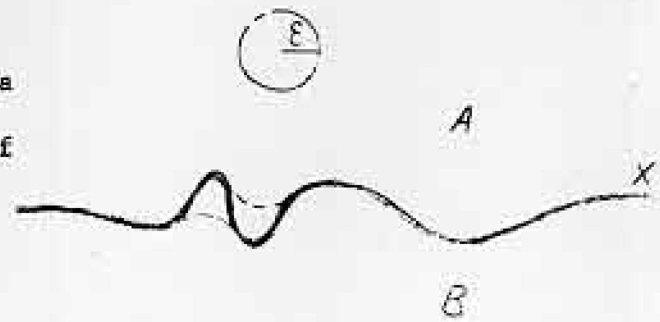


Fig. 11

$$(6) \qquad G_\varepsilon(A/B) = BF_r\left[C_{2\varepsilon}(A)\right].$$

The arc $G_\varepsilon(A/B)$ is shown in Figure 11 by a broken line. It can be formed in the following manner. Along the arc X we will roll (on the B side) a circle with a radius of ε. The envelope of all positions of the circle consists of two branches; that branch of the ε-generalized edge of area A in area B will be that which lies nearest the arc X. In the case where area A is 2ε-convex, $G_\varepsilon(A/B)$ coincides with arc X.

Analogously by $G_\varepsilon(B/A)$ is meant the ε-generalized edge of area B in area A. This is obtained by running a circle of radius ε along arc X on the A side. The arc X is traced with a heavy line in Figure 12, and arcs $G_\varepsilon(A/B)$ and $G_\varepsilon(B/A)$ by thin lines. Both area A and B are 2ε-convex at the same time, if and only if the arc X is 2ε-convex; then arcs $G_\varepsilon(A/B)$ and $G_\varepsilon(B/A)$ coincide with arc X. If only one of the areas A or B is 2ε-convex, then one of the arcs $G_\varepsilon(A/B)$ or $G_\varepsilon(B/A)$ coincides with the arc X, and the other is different. If finally neither of the areas A or B is 2ε-convex then all three arcs differ. In the last two cases, that is if the arc X is not 2ε-convex, then between arcs $G_\varepsilon(A/B)$ and $G_\varepsilon(B/A)$ there exists a two dimensional set, the area $C_{2\varepsilon}(X)$. This area is called the ε-generalized edge of areas A and B:

$$G_\varepsilon(A,B) = C_{2\varepsilon}(X).$$

This is the natural interpretation of these definitions. Imagine that area A is the ocean and area B the land and let us consider the boundary between the land and the sea (Figure 12). We will limit ourselves to an instantaneous moment, that is, we will not consider the changes caused by tides and waves. A series of sufficiently accurate photographs are taken at the chosen moment. But then the boundary should distinguish the individual grains of sand from the water around these grains, and, with greater accuracy the curve will become even more complicated. Thus the following arc will be the ε-generalized

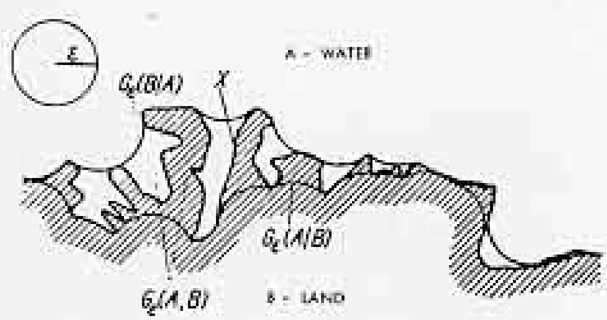


Figure 12

edge of the land in the sea: we will place a plane circle or radius on the sea, floating on top of the water. We will try to swim as close as possible to the land with the circle, so that at every moment at least one point on the circle is in contact with land and so that the circle is at all times on the surface of the water. Bays narrower than 2 (that is, in those in which our circle will not fit) will be counted as part of the 2 -convex land the line along which our circle will run will be the envelope of this circle under all conditions, that is, it will be the -generalized edge of land in the sea.

Similarly, if we permit the circle to move as near as possible to the sea but in such a way that it is entirely on land and does not enter any peninsulas narrower than 2ε, they will be counted as 2ε-convexity of the sea. The line formed by the movement of this circle will be the ε-generalized edge of the ocean in the land. Note that these two generalized curves do not coincide. The curve $G_\varepsilon(B/A)$, the -generalized edge of the land in the ocean, will leave some narrow bays on the side of the land and water on the side of the sea. On the other hand, arc $G_\varepsilon(A/B)$, leaves narrow peninsulas on the side of the sea and dry land on the landward side. An ε-generalized edge of land and sea will be contained between these two curves; G (A,B) will not be a line but a domain of certain added area, an area consisting of narrow bays of the oceans and peninsulas of the land. Similarly small lakes with sea water and islands of continental land can appear there. However, in this area, we will not find a single piece of land or sea which could contain a circle with a radius of .

The length of order ε of two curves can be measured with the aid of the ε-longimeter described in paragraph (4): the ε-generalized edge of area A in area B and ε-generalized edge of area B in area A can be measured without tracing these edges. This means that the measurement of curves G (A/B) and G (B/A), or only one of these curves, can be carried out having only one drawing of the curve X dividing areas A and B (or these features in nature) and an -longimeter. For this we need in addition a separate circle with a radius of (the same as one of the circles in the -longimeter) cut out of pasteboard.

In measuring the length of order of the arc X, the circles of the -longimeter falling on arc X are counted. We will differentiate between circles attainable from area A, attainable from area B, and not attainable. We will call circles of the -longimeter attainable from area A if we can fit in area A additional circles cut from pasteboard so that they have points in common with the circles of the longimeter under consideration. Circles of the longimeter not attainable either in area A or B we call unattainable. In Figure 13 is shown a curve X dividing area A and B and the -longimeter placed on curve X. Circle a of this longimeter is attainable equally from area A as from B. Circle b is

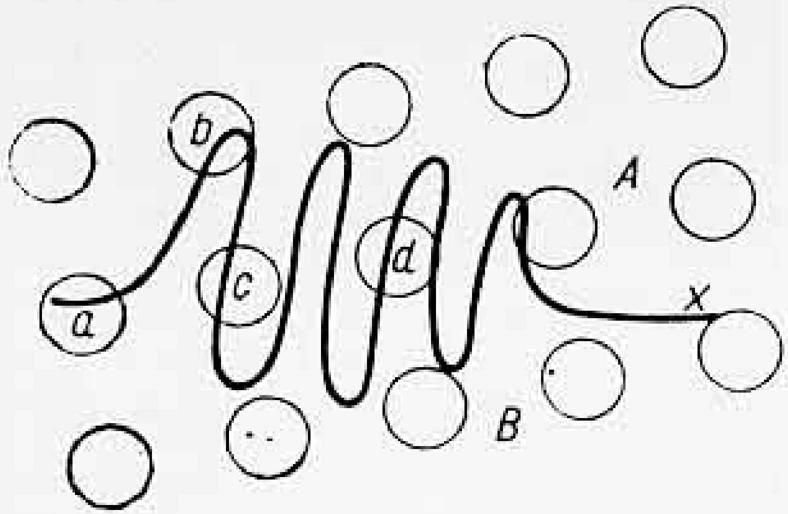


Figure 13

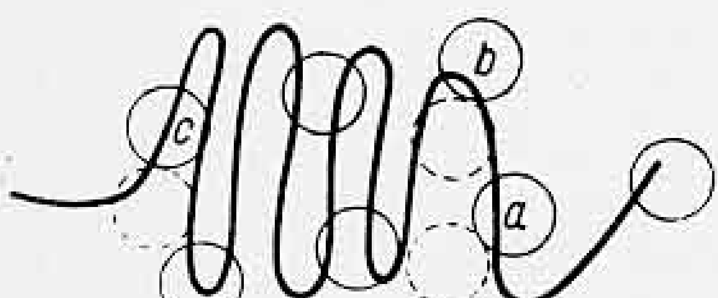


Figure 14

attainable only from A, circle c only in B, and circle d is unattainable. Occasionally we can ascertain without additional circles if the circles are or are not attainable in some area. In certain cases we have to use additional circles (from pasteboard) thus making the measurement more difficult. Examples of simple and difficult determinations are presented in Figure 14 (in the latter case, the additional circles are shown as dotted circles). For the circles of the ε-longimeter to fall on the arc $G_\varepsilon(A/B)$, it is necessary and sufficient for the circle to fall on the arc X and to be attainable from area B. Actually, if the circle of the longimeter falls on the arc $G_\varepsilon(A/B)$ then it cannot lie completely either in area A, or in area B. Therefore it is concluded that it must fall on the arc X dividing areas A and B. If furthermore the circle falls on the curve $G_\varepsilon(A/B)$, then the points of the set $C_{2\varepsilon}(B)$ are enclosed; the points of this set have the property that a circle of radius can be added to every such point that lies within area B. Therefore, the circles of the longimeter falling on curve $G_\varepsilon(A/B)$ are attainable from area B. Conversely, if the circle of the longimeter falls on the arc X and is attainable from area B, this includes points in series $C_{2\varepsilon}(B)$ and points of area A, therefore it falls on the edge of area $C_{2\varepsilon}(B)$ or curve $G_\varepsilon(A/B)$.

Analogously, for a circle of the ε-longimeter to fall on the curve $G_\varepsilon(B/A)$ it is necessary and sufficient that the circle fall on curve X and that it be attainable from area A.

From this results the following method of measuring the length order ε of the ε-generalized edge of area A in area B; we apply the ε-longimeter to the curve X and count the number of circles of the longimeter falling on curve X and at the same time attainable from area B. The measurement is repeated k times (the number k from Table 1 or 2 should be recorded on the longimeter) and from the combined quantity of counted circles subtract the pertinent coefficient $\frac{1}{2}\pi\varepsilon$ (also given on the longimeter). In an analogous manner the length of order ε of the curve $G_\varepsilon(B/A)$ can be measured. It should be noted that curves $G_\varepsilon(A/B)$ and $G_\varepsilon(B/A)$ do not have to be 2ε-convex and in view of this their length of order ε does not have to be a classical length.

The number of circles of the ε-longimeter falling on the arc X, but attainable neither from area A or B, in a certain manner indicates the size of the ε-generalized edge of areas A and B. A circle is unattainable if it lies entirely within the set $G_{\varepsilon}(A,B)$, or if the center of this circle lies within the ε-core (see Ref 2) of the set $G_{\varepsilon}(A,B)$. The expected number of unattainable circles falling on the arc X is an indication of the size of the ε-generalized edge of areas A and B (the boundary between the land and sea, consisting of narrow bays and peninsulas). Just because this number is zero does not necessarily mean that the ε-generalized edge of areas A and B has an area equal to zero; this means only that this edge is narrow, and that the ε-core of this edge has an area equal to zero.

6. The circles of the ε-longimeter are numbered 1, 2, ... , r and the event Z_i is considered as being dependent on placing the i-th circle on the arc X, when using the longimeter. To x_i we assign a random variable (which can have the value one) when the event Z_i is approached, and a value of 0 when the opposite event approaches, that is, when the i-th circle does not fall on the arc X. By $x = x_1 + x_2 + \cdots + x_r$, we mean the number of circles of the longimeter which fall on the arc X in one application (in paragraph 4 the variable x is indicated by $\sum n_i$).

If the occurrences of Z_i were independent of each other and if the probability P of the occurrence of Z_i were independent of i, the random variable x would have a binomial distribution with an average rp, equal to the anticipated number of circles of the ε-longimeter falling on the arc X, with Bernoulli's variation rpq (where q=1-p). But the events Z_i are not independent but are precisely correlated. For that matter, if one circle of the longimeter falls on the arc X, then the probability that neighboring circles of the longimeter also fall on this arc is greater than the probability that a random circle of the longimeter will fall on the arc X. The correlation between the events Z_i depends on the shape of the arc X; for some it is greater and for others less. This positive correlation increases the variance of the random variable x, making it greater than normal.

On the other hand, the probability of the occurrence of Z_i is dependent on i. By the application of the longimeter to the arc X the center of the longimeter is utilized and for that reason the circles lying in the center of the tool have a greater possibility of falling on the arc X than the circles lying on the edges of the tool. This brings about a decrease in the variance of the random variable x (to less than normal) and the size of this change depends on the person making the measurement; that is, on whether he applies the center or the edge of the longimeter to the arc X.

These two influences to a certain extent cancel, and one can with a certain amount of approximation maintain that the random variable x has a normal variance rpq. The variance of the random variable N which is the summation over k of the independent random variable x (the combined number of circles falling on the arc in k independent applications of the longimeter) is therefore approximately equal to krpq, and with a large longimeter, where r is great in comparison to the anticipated value of variable x, q is close to unity, and the variance in the random variable N is equal to krp/N. From this we can draw the conclusion that one should expect that the variance in the length of order ε of the arc X is of the order of the measured length, and the average quadratic deviation is on the order of the root of this length. Therefore the greater the length (of order ε) of the arc, the greater the precision (percentwise) in measuring with the longimeter. If, for example, the length of order ε of some curve is equal to 20, then it is expected that the probable average error of this length is of the order of 4 to 5, that is, 20 to 25%. If the length is equal to about 100, then the probable average error is on the order of 10, that is, about 10%. If, as I mentioned at the end of paragraph 3, ε is taken so that the ratio $\eta = \varepsilon / L_\varepsilon(X)$ is not greater than 10%, then in view of the fact that $\varepsilon \geqslant 1$ and $L_\varepsilon(X) \geqslant 10$, the probable average error would become not greater than 30% of the length of order ε.

In practice measurement with the longimeter proved to be far more precise than would appear from the above considerations. Evidently the influence of the different rolls played by the central and the extreme circles of the longimeter to diminish the variance is far stronger than the influence among the events Z_i. The result of this is that in practice the variance is remarkably less than normal.

Together with laboratory assistants of the Mathematical Institute, J. Dobrowolske, A. Huskowski, and M. Kusiatkow, we measured several curves with the longimeter (the measurements were repeated a number of times, in one case 200 times). The aim of these measurements was the examination of the systematic error in practice, the random measurement of different curves with different longimeters, and the determination of the time required to accomplish the measurements.

For the investigation of the systematic errors three persons; A, B, and C applied at random the ε-longimeter (ε = 3 mm) 60 times on arc 2 (Figure 15) and counted the number of circles of the longimeter that hit the arc. In Table 3 are presented in a series of columns the number of circles falling on arc, the average, the variation, and the average quadratic deviation of these numbers. As can be seen, there are no systematic differences between the average results obtained by

TABLE 3

Number of circles hitting arc 2	Number of applications per person		
	A	B	C
12	-	1	2
13	2	5	8
14	9	13	24
15	29	12	15
16	15	23	10
17	4	4	1
18	-	1	-
19	1	1	-
Average of the count of circles	15.23	15.20	14.43
Variation	1.02	1.73	1.15
Average deviation of the quadratic mean	0.13	0.17	0.14

persons A and B. However, there is a real difference between A and C (Student's criterion gives the likelihood of a hypothesis saying that there is no difference as being equal to 0.0001). The difference in the results obtained between B and C is equally real. In the variance, the results obtained by B differ greatly (at a reliability level of 1%) from those of A; the variances obtained by A and C do not really differ.

The result of this is that one might worry about the systematic errors encountered in the results of measuring with the longimeter. In our case the extreme difference was 0.80, or 5 to 6%. It should be realized from the systematic differences in errors, that these differences could be 20% of the errors.

I next took up random errors. From what I have written previously the

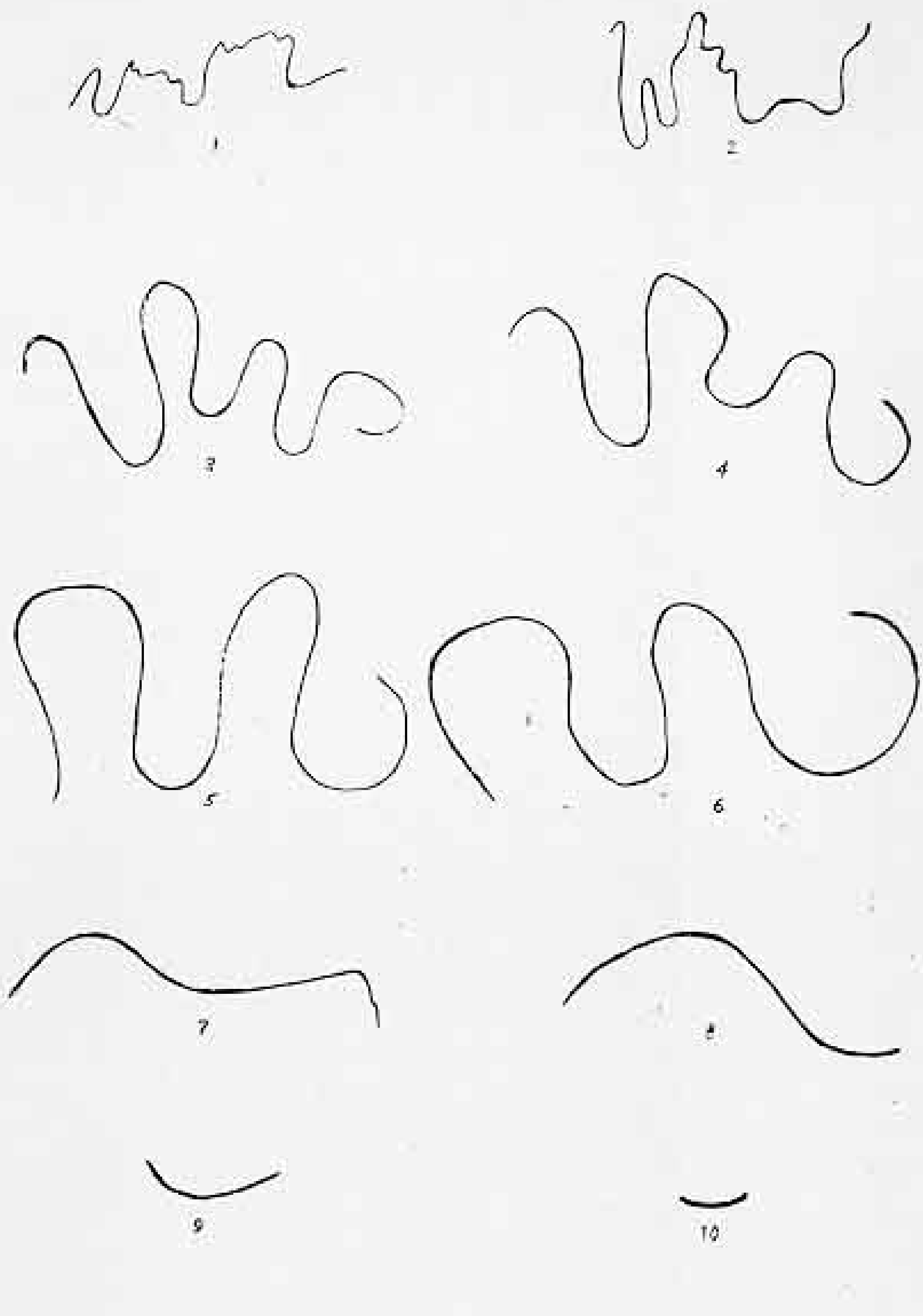


Fig. 15

results show that the random error depends to a large extent on the person making the measurements; for person A the error was 20% less than that for person B. The variation in the results of one application of the longimeter we indicate by S^2. The length of order 3 mm we obtain using the sum of six measurements; consequently the variation in the length of order 3 mm will be the sum of the variation of six random independent variables of variation S^2, or $6S^2$ (a constant coefficient does not affect the variation).

The average length of order 3 mm of curve 2 obtained by persons A, B, and C are presented in Table 4. The average quadratic deviation of these lengths calculated from 6 applications and the same deviations expressed in percent of the length of the arc are also given. In the beginning of the paragraph, we calculated that in the case of normal variation the average quadratic deviation should be in the order of $\sqrt{90} \approx 9.5$, therefore the empirical variation of length is actually subnormal.

TABLE 4

	A	B	C
Average length of the order of 3 mm of arc 2	86.7	86.5	81.9
Average quadratic deviation of the length obtained from 6 applications	2.47	3.22	2.62
The same deviations in %	2.9	3.7	3.2

The random error of measurement (average quadratic deviation) can be estimated on the basis of one measurement made up from 6 applications of a longimeter with $\epsilon = 3$ mm. The error will actually be estimated with less precision, but for comparative purposes it will be sufficient.

Ten curves are presented in Figure 15. Double measurements of length were made on each of these, of order 3 mm, 5 mm, and 8 mm with the appropriate longimeter (square lattics). The results are presented in Table 5. From the table we can investigate how the lengths get smaller as the order of the arc increases (ε = 3, 5, 8 mm), for various curvatures and for large (curves 1 and 2) and small (No. 8) curves. Curves 9 and 10 were particularly short. In the last

TABLE 5

LENGTH \ ARC		1	2	3	4	5	6	7	8	9	10
$L_3(X)$	A	69	87	145	135	165	158	64	58	26	10
	B	67	86	145	138	175	166	65	63	20	10
$L_5(X)$	A	61	79	132	132	168	163	65	56	20	11
	B	54	75	121	128	174	171	61	60	20	13
$L_8(X)$	A	56	63	111	107	156	159	66	57	21	11
	B	51	69	106	111	158	162	66	58	20	9
$L(X)$		76	105	151	138	176	171	65	58	22	10

line are given the lengths L(X) of these curves, measured by an incremental method. This method is based on counting small steps made along the length of the curve with point dividers (the span of the points is adjusted with a screw). The length of the steps are determined by dividing 100 steps along a straight line. As can be seen, arcs of sufficiently small curvature have the length of order ε very close to the length obtained by the incremental method. Letters A and B represent the persons performing the measurements. The results from both persons are given in order to see the difference between two measurements (each consisting of 6 applications) performed by the different persons.

As can be seen from the table (Table 5), the first three arcs have lengths of order 3 mm which are smaller than those lengths measured incrementally. The lengths of order 5 mm of these curves are smaller than the lengths of order 3 mm, and still smaller are the lengths of 8 mm order. The following curve (4) has a length of 3 mm order equal to the length established by the stepwise method. On the other hand, its length of 5 mm order is somewhat smaller and the length of 8 mm order is considerably smaller. The next two curves (5 and 6) have lengths of 3 mm and 5 mm order equal to the lengths measured stepwise, but the length of 8 mm order is smaller. Finally the last four curves 7, 8, 9, and 10 have the lengths of 3 mm, 5 mm and 8 mm order equal to the stepwise measured lengths.

Curves 9 and 10 are especially short. The ratio $\eta = \varepsilon/L(X)$ for these curves approaches 75% (when ε = 8 mm). The remaining curves have lengths exceeding 60 mm and for these the ratio η is decidedly smaller, although they occasionally exceed 10%, which earlier we had accepted as the upper limit. Despite this the random errors inherent in the results of measurements are not large. Table 6 consists of ratios η_ε and the indices of variation W_ε, that is, the percent error in the length of order ε. The number of applications of the longimeter necessary for one measurement of length of the order ε are given in the column titled k . Actually if the entire measurement were repeated n times (that is the longimeter were applied nk times) the average error of length of order ε obtained from these n measurements would become $\sqrt{n}$ times smaller.

TABLE 6

K	ARC	1	2	3	4	5	6	7	8	9	10
6	3	4.4	3.5	2.1	2.2	1.8	1.9	4.7	5.0	13.0	30.0
	W_3	3.5	2.8	1.7	1.7	3.3	1.7	3.7	3.9	6.3	12.2
10	5	8.2	6.3	3.8	3.8	3.0	3.1	7.7	8.9	25.0	45.0
	W_5	3.6	3.6	3.6	2.1	3.1	1.8	2.2	5.8	9.5	15.5
16	8	14.3	13.	7.1	7.5	5.1	5.0	12.1	14.0	38.1	72.8
	W_8	7.9	5.7	3.0	4.2	3.6	2.1	5.5	6.0	13.2	18.2

As can be seen, in cases where the ratio η does not exceed 10%, the error of measurement does not exceed 6% and in only two cases does it exceed 4%. It is further seen that when ε increases, in general, the error becomes greater, and η increases simultaneously.

In Table 7 are presented the lengths x of the ε-generalized edges of areas which are separated by the plane curves of Figure 15. We will call P the area lying above the arc X, and Q below this arc. The length of the arc $G_\varepsilon(Q/P)$ we will call $\overline{L}_\varepsilon$ and the length of arc $G_\varepsilon(P/Q) = \underline{L}_\varepsilon$. Each one of these lengths were measured by two persons, A and B. Each measurement consisted of k applications of the longimeter. These measurements contain greater errors, both random and systematic. Person A had results systematically greater than those of person B

TABLE 7

ORDER OF LENGTH / ARC		1	2	7	8	9	10
$\overline{L}_3$	A	50	73	59	56	24	9
	B	37	62	59	56	22	11
$\underline{L}_3$	A	54	81	62	57	23	11
	B	34	73	60	57	22	11
$\overline{L}_5$	A	39	63	67	58	24	13
	B	42	54	57	54	22	12
$\underline{L}_5$	A	52	65	65	57	22	12
	B	58	56	59	52	18	14
$\overline{L}_8$	A	42	49	66	61	19	10
	B	41	46	58	56	23	10
$\underline{L}_8$	A	46	67	65	61	20	10
	B	42	52	57	56	19	9

by a factor of 4. After eliminating these systematic errors it is found that the random error is established at about 5% of the measured length.

The lengths of curves 1, 2, and 6 measured by Steinhaus' line-longimeter are included in Table 8 for comparison. This table presents the averages obtained for fifty applications of the longimeter for each result. These averages contain random errors more or less of the order resulting from single measurements with the ϵ-longimeter. For example, the average error (from 50 applications of Steinhaus' longimeter) of the length of order 1 of arc 1 is established at 4.5% and for the order 8 it is established at 3.3% of the length of the measured arc.

TABLE 8

ORDER OF LENGTH / ARC	1	2	3	4	5	6	7	8
1	44	62	74	78	79	80	80	81
2	47	71	90	100	108	110	111	111
6	77	126	153	163	168	168	168	168

In conclusion, I will give the time required to accomplish the measurement of the various lengths. For the measurement of length of order (ϵ = 3, 5, or 8 mm) of any arc in Figure 15, two to four minutes are required. Such a measurement (consisting of 6, 10 or 16 applications of the longimeter) results in an error of about 4%. To obtain the length of order n of such an arc with the same precision by Steinhaus' method would require 50 applications which would take 15 minutes or four times longer than in required by the ϵ-longimeter. The incremental method of measurement (a double measurement and double marking of the length with steps) requires more than 10 minutes. The precision of this method depends on the shape of the curve. For arc of small curvature (as in arc 6) the stepwise method gives very precise results (the error is established at about 0.3% of the length), but for arcs of greater curvature the precision decreases decidedly (for example, in arc 1 the error is established at about 4%).

Literature cited:

(1) A. Penck,"Morphologie der Erdoberfläche,"Stuttgart 1894.

(2) J. Perkal,"O epsylon aureolach", Roczniki PTM, seria I, Prace Matematyczne (w druku).

(3) J. Perkal,"On the ϵ-length,"Bull. Acad. Pol. Sc. Cl. III, Vol. IV (1956), str. 399-403.

(4) J. Perkal,"Sur les ensembles epsilon convexes,"Colloquium Mathematicum 4 (1956), str. 1-10.

(5) H. Steinhaus,"Length, shape and area,"Colloquium Mathematicum (1954), str. 1-13.

(6) H. Steinhaus, "O dlugości krzywych empirycznych i jej pomiarze, zwlaszoza w geografis", Sprawozdania Wroclawskiego Towarzystwa Naukowego 4 (1949), Dodatek 5.

(7) H. Steinhaus, "O mierzeniu pól plaskich," Przeglad Matematyczno-Fizyczny 2 (1924), str. 1-6.

(8) H. Steinhaus, "Sur un théorème de H. V. Jarnik", Colloquium Mathematicum 1 (1947), str. 1-5.

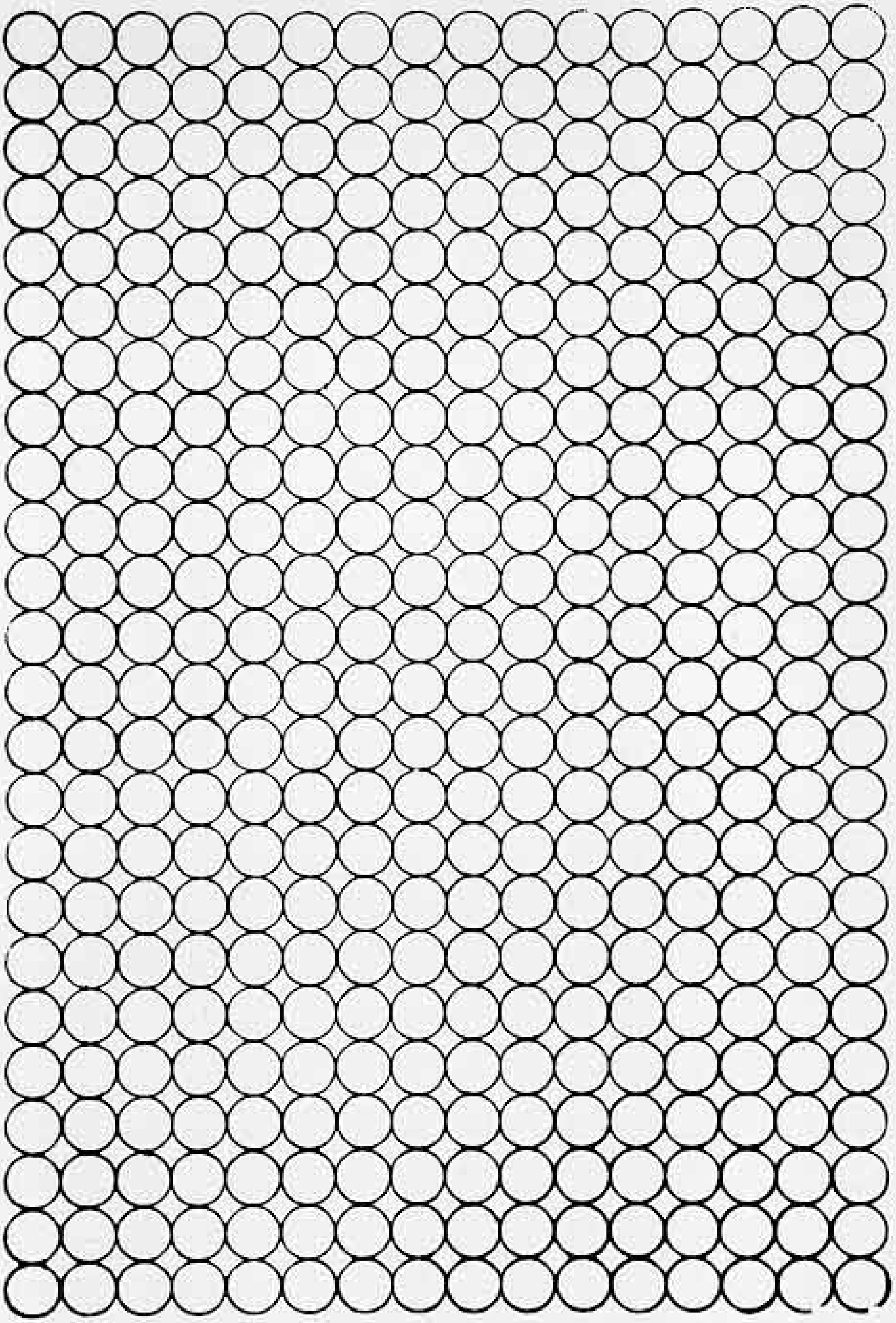


Figure 16

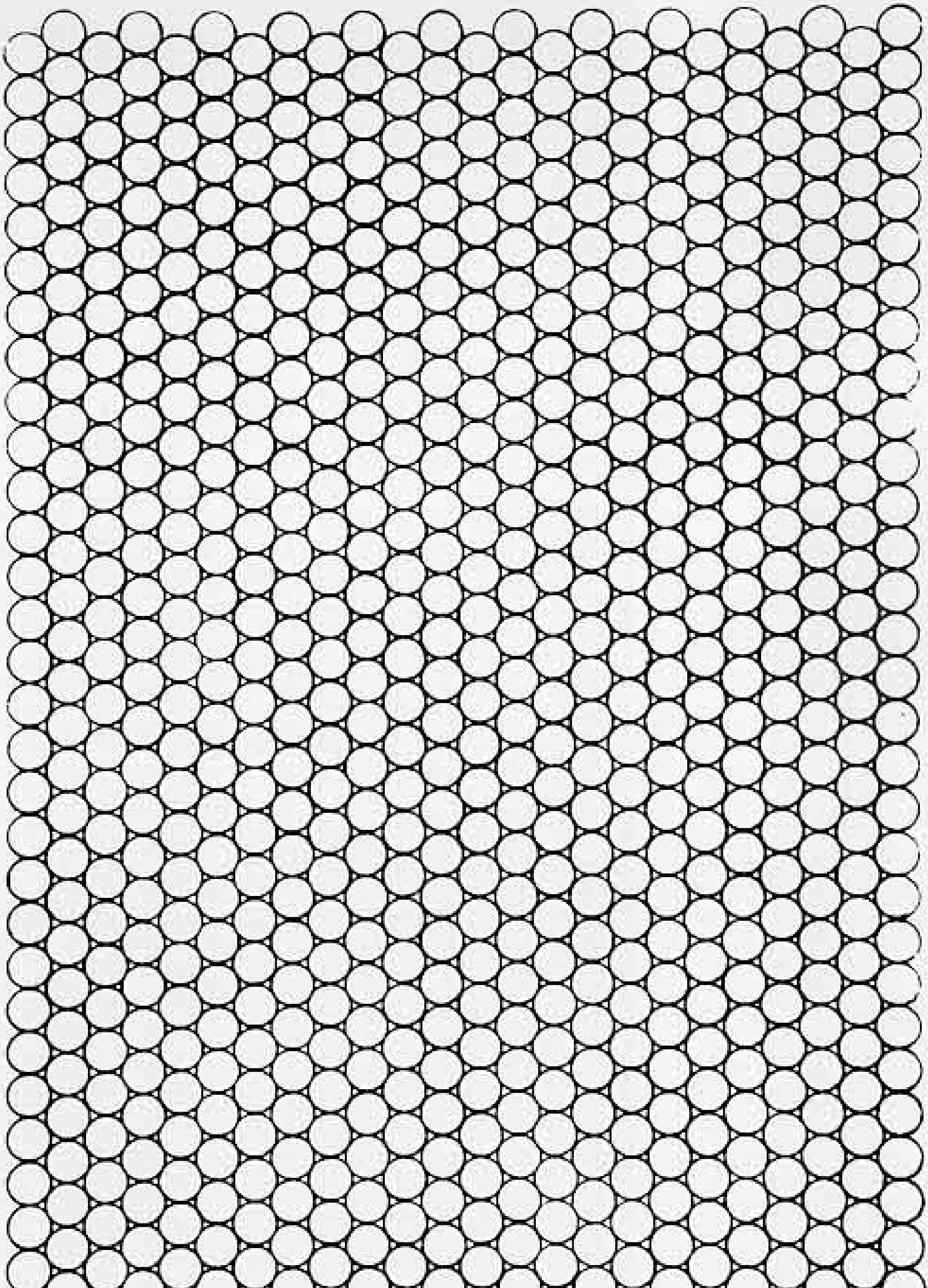


Figure 17

An Attempt at Objective Generalization

by

Julian Perkal

Translated from

Julian Perkal, "Proba obiektywnej generalizacji," Geodezja i Kartografia, Tom VII, Zeszyt 2 (1958), pp. 130-142.

by W. Jackowski

under the direction of

Professor W. R. Tobler
Department of Geography
University of Michigan
Ann Arbor, 1965

Cartographic transformations, employed to represent the surface of the earth on maps, are of two simple types; map projections and generalization. Map projections are obtained by an objective mathematical operation. The subject of this paper is the second cartographic transformation-map generalization.

Consider first the subject to be transformed, a portion of the earth's surface. The areas to be mapped contain geographical features of various types; water, forests, land lying between 200 and 250 miles above sea level, or an area belonging to a particular governmental unit. A first important problem encountered in mapping these areas is the following. The area to be mapped is determined only if we know exactly which points belong to the area, and which do not; that is, if the line bounding the area is exactly determined. In certain cases this is simple. For example, as when the boundary between countries is a series of straight line segments connecting a specified set of points. Here the boundary consists of straight lines. This type of boundary presents no problems. It is simple to draw at small scale, and simple to generalize (the segments remaining straight lines), and the lengths of the lines are easily measured. This type of line boundary will be referred to as a conventional line, since they are created by convention and do not actually exist.

A different type of area is bounded by natural features, not by convention. For example, an island consists of an area not covered by water which is surrounded by an area which is covered by water. I do not wish to consider difficulties such as high and low tides, or the movement of waves or other phenomena which affect the status of a particular point such that it at one moment belongs to the island and then to the sea. We have sufficient difficulty in deciding the area of the island at any instant of time. Let us assume that we are able to freeze the island momentarily, as in photography (restrictions on the photographic process are presented later).

Let us now attempt to determine the area of the island mentally, that is, to contemplate the difficulties encountered in this procedure. We must decide on the dividing line between the island and the sea. It is simple to do approximately. We designate the shoreline as something more or less irregular. On closer scrutiny, however, this process is unsatisfactory. Small creeks are mistakenly retained as part of the island and small peninsulas are assigned to the ocean. If we wished to correct our shoreline empirically, by drawing an actual line with a stick, this would not help. It would not be successful since smaller misidentifications will remain, for example, stones or pebbles (or coarser stones) from the water. Many stones would be surrounded by water, forming very small islands along the shore. Many small pools of sea water are formed on the surface of the island, thereby forming small lakes. To these should be added numerous small peninsulas, bays, straits, and isthmuses. Even if one were able to cope with these situations, separating the islands from the sea, the solution would not be absolute. For, by suitable observation through a magnifying glass or through a microscope, new complications would appear in the form of individual grains of sand or particules. (The limitations of photographs should be recognized, since lines on photographs, observed with a microscope, may have an entirely different shape or form than these same lines in nature, as observed under a microscope.)

A line separating two empirical regions will be called an empirical line. There is no point in debating its true nature, since its shape depends on the precision with which the observations are made. The problem of outlining such areas can be solved only by approximation. We can only approximate the boundary lines. With what precision whall we do this? What does it mean when we approximate an empirical line with a given precision?

Leaving the question unanswered temporarily, we will go to yet another difficulty. How can one define the length of an empirical line? As illustrated by the foregoing example, the length of such a line depends on the precision of observation. The more precise the measurement of the boundary line, the greater will be its length. It is not even known if greater increases in the precision of the length of the empirical line would increase to some limit, or whether the increase would be unbounded (lines of this type are termed unrectifiable in mathematics). Furthermore, there is no point in contemplating whether an empirical line has a finite length. H.Steinhaus (6) has examined the so-called paradox of the length of an empirical line. It makes no sense to speak of the real length of an empirical line, or of the real length of a line which represents an empirical lines.

Having described the nature of the main difficulties, I will pass on to the matter of generalization. I will differentiate between actual and symbolic generalization. By actual generalization is meant a simplification of empirical lines, or their picture on a mpa, which bound an area. By symbolic generalization is meant the substitution on a map by conventional symbols for areas, as in the substitution of a circle for the area of a city, or a conventional line for the area of a road or a river.

Generalization is a subjective action. Every cartographer has his own method of simplifying or reducing the boundary between areas. Two generalizations to the same reduced size of a single map made by two different cartographers may differ considerably. The authors of cartographic handbooks leave the reader considerable leeway in generalizing the map. In his book, G. N. Liodt writes (2, p.358) "The task of the cartographer consists of the following: In the drawing of shore lines of oceans and large lakes the irregularities in the contour lines are drawn in agreement with the scale of the map. But in reductions to scale of small details, a cartographer should comprehend all of the peculiarities necessary to describe the type of shore and place them on a map at even the smallest scale. In generalizing networks of rivers, first the small river channels, small streams, and small islands are reduced, then the larger bends, larger streams, and so on."

Some consider generalization an even more subjective transformation.

For example, L.S. Garajewskaja in Cartography (1) devotes over forty pages of text to problems of generalization and proposes, for instance, the replacement of many small bends in a river with a few larger bends having the same apparent characteristics as the initial river bends.

In what follows, I intend to present certain geometrical manipulations developed previously (4), and to propose that cartographers use them for their generalizations. The method which I will describe transforms every region into a generalized area, and the transformation is unique. Generalization using this convention is therefore objective. I will concede that with this objective generalization I will encounter serious difficulties for it will depend singularly on the shape of the area, and does not satisfy geomorphological considerations (as desired by cartographers, e.g., Garajewskaja).

The degree of generalization will be defined by a real number ϵ , which represents the length of a segment. Figure One shows a circle[1] of diameter ϵ and a region, D. Some of the points (e.g., the point p) included within this region have the property that there exists a circle of diameter ϵ which lies entirely within D and which contains the point. There, however, are also points (e.g., the point q) in D, such that no circle of diameter ϵ which is in D can contain the point. The set of all

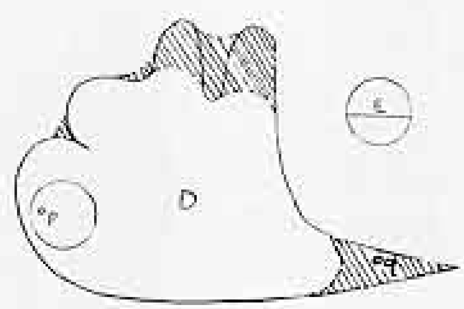


Fig. 1

points, such as p, having the property that they are contained within circles of diameter ϵ which can be completely included in the region D, I shall call an ϵ-generalization of the region D and will assign the symbol $G_\epsilon(D)$ to this set of points. The boundary of this set of points does not coincide with the original boundary of D. The regions of divergence have been shaded in Figure

One by a dashed line.

The ϵ-generalization of the area D can be visualized as follows. Within the area D move a circle of diameter ϵ in such a manner that the circle always lies completely inside the area D; it is never outside the area D. Positions covered by the circle by this movement form the ϵ-generalization of the area D.

The method of ϵ-generalization results from this definition. First, cut out a circle of diameter ϵ, then using this circle try to touch all points of the area of D. It may occur that the circle cannot touch or cover certain points, such as the point q. It will therefore stop in contact with at least two points of the border of D. With a bow compass with center in $G_\epsilon(D)$ draw a circle connecting these points of tangency, thereby excluding the inaccessible portions of D. In practice I employed (for Figures 3,5, and 7) a drop bow compass with a radius of $\epsilon/2$.

The area of the plane which remains after deletion of D is called the complement of D and is designated by D'. The complement D' of D is also an area and consequently we can also speak of an ϵ-generalization of this new external area. The edge of the ϵ-generalized complement of $G_\epsilon(D)$ is indicated by a dotted line in Figure One. It also departs from the edge of D. Thus, in Figure One, we have inside the inner (continuous and dashed) line the ϵ-generalization of area D, or on the outside of the outer (continuous and dotted) line the ϵ-generalization of the complement of the area D. Between these two lines is a collection of points: a composite of continuous lines (which bound the meeting of the ϵ-generalizations of areas D and D') and of two-dimensional areas (contained between the dashed, dotted, and continuous lines). This collection, shaded in Figure One, I will refer to as the - generalized edge of the areas D and its complement D'.

I propose these geometrical operations for objective cartographic generalizations. Figure Two shows a section of a sea coast.[2] The map is divided into a region of land and a region of ocean. All of the places lying in the ocean which are contained in a circle of 2 cm. diameter which fits into the ocean area we will call the 2 cm. generalized area of the ocean (Figure 3c). In other words, the 2 cm. generalized area of the ocean consists of those areas in which circles of 2 cm. diameter can be placed without touching the land. Actually part of the ocean remains beyond the 2 cm. generalized ocean area. These are the lakes, bays, and straits into which a 2 cm. circle will not fit.

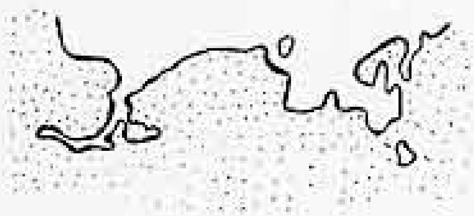


Fig. 2

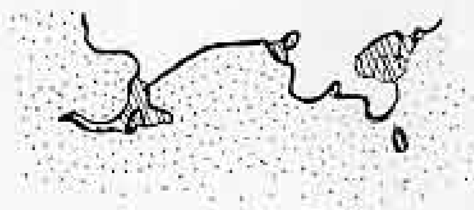


Fig. 3a

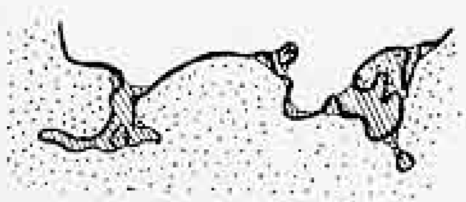


Fig. 3b

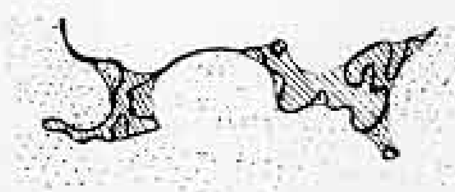


Fig. 3c

If 2 cm. on the map actually represents 200 m., then the 2 cm. generalized ocean represents that portion of the ocean into which a raft 200 m. in diameter can float.

A 2 cm. generalization of the land areas (Figure 3c) will consist of those places on land which can be covered by 2 cm. circles on the map without touching the ocean. This area will not include all of the land; it will not include small islands, peninsulas, and isthmuses so narrow that circles 2 cm. in diameter will not fit inside them. The area that remains between the 2 cm. generalized ocean and the 2 cm. generalized land we will call the 2 cm. generalized shore. This consists of small (since 2 cm. circles will not fit inside them) lakes, bays, straits, islands, peninsulas, and isthmuses.

This border has a real significance of its own. As previously indicated, it is not possible to construct a suitable line to represent the boundary between two areas, for example, between ocean and land. Increased precision in the observation of the line would increase the complication. These complications are in fact the small lakes, bays, straits, islands, peninsulas, and isthmuses. Operationally (by measurement or on photographs) we must define the real coast so that we do not take into account these six complicating entities if they are somewhat too small; that is, smaller than some arbitrary number. Thus, for instance, peninsulas 4 meters wide will be taken into account, but narrow half-meter bays (crevices between rocks) we will not consider in our observations. But if a certain area consists entirely of these small elements not worthy of consideration, to which domain should these elements be assigned? They can be counted neither as a part of the land not as a part of the ocean. There is a third indistinguishable (at some level of observation) area which is a mixture of land and sea which can be called the edge of the land and of the ocean. This area should be distinguished by color on a map of the coast, for example, by a suitably heavy black line (which hides the edge or coastal area).

Figure Three presents generalizations of the areas shown in Figure Two for different ϵ, namely 5 mm., 10mm., 20 mm. It is seen that the ϵ-generalized edge consistently becomes broader.

Figure Four shows built-up and vacant areas (of the city of Wroclaw). Figure Five shows ϵ-generalizations of these areas and their ϵ-generalized edges for $\epsilon = 1$ cm., 3 cm., and 6 cm. We note that the 6 cm. generali-

Fig. 5

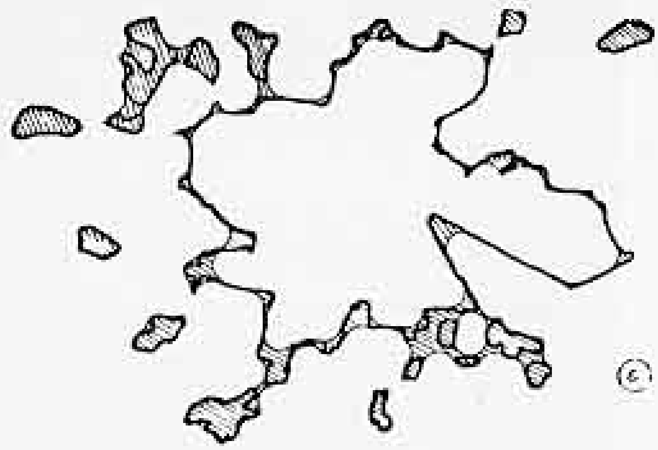


Fig. 5a

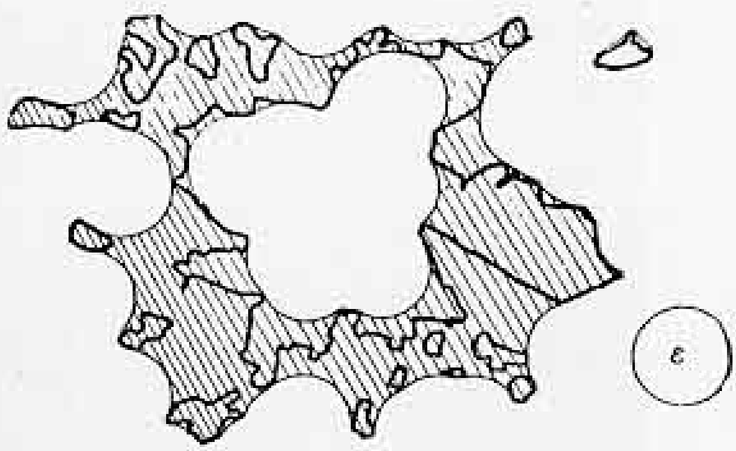


Fig. 5b

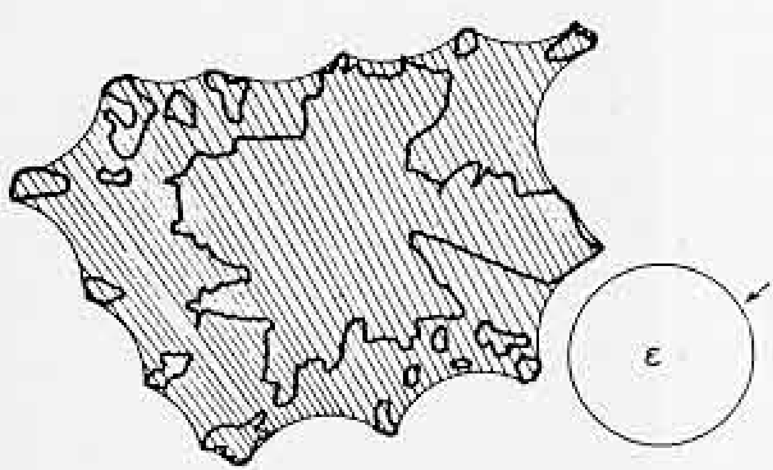


Fig. 5c

zation causes the city to disappear completely, leaving only the vacant area and the edge area. This situation may also arise while generalizing other small or narrow regions, such as islands, lakes, rivers, or roads. The disappearance of an inner area at a given level of generalization can be taken as a reason for omitting the given element from a map, or for a representation by stylized symbols (e.g., for cities, roads, or rivers). When an area is to be shown on a map, the disappearance of the inner area may be an indication that the element should be replaced by a cartographic convention, with the elements shown symbolically.

The selection of the number ϵ is arbitrary but can be justified on several grounds, including technical considerations. Let us assume that we have a map of some area at a scale of 1:25,000 and we wish to redraw this map at a scale of 1:500,000. We will have to perform a suitable generalization of the initial map, and choose a suitable ϵ for this purpose. A requirement that all of the smallest features on the new map should be visible with the naked eye seems justifiable. That is, none of the lines on the map should be thicker than about 0.5 mm. (or some other number, e.g., 1 mm.). It follows that the initial map, at a scale of 1:25,000 or twenty times as large, should be generalized in such a manner that no lines remain which are wider than 10 mm. (0.5 X 20). Accordingly, the ϵ chosen for generalization will be 10 mm., which guarantees that only areas which will fit a circle of 10 mm. diameter will be included on the generalized map (at 1:25,000), and consequently only areas which will fit a circle of 0.5 mm. diameter will appear on the map at a scale of 1:500,000.

Figure Six shows a section of the Oder River traced from an old German map (Messtischblatt 2828, Preussische Landesaufnahme of 1886) of Wroclaw at a scale of 1:25,000. I have generalized this figure for various values of ϵ and have reduced it photographically so that, after reduction, $\epsilon = 0.5$ or 1 mm. Figure Seven shows these generalizations and their reductions to a suitable

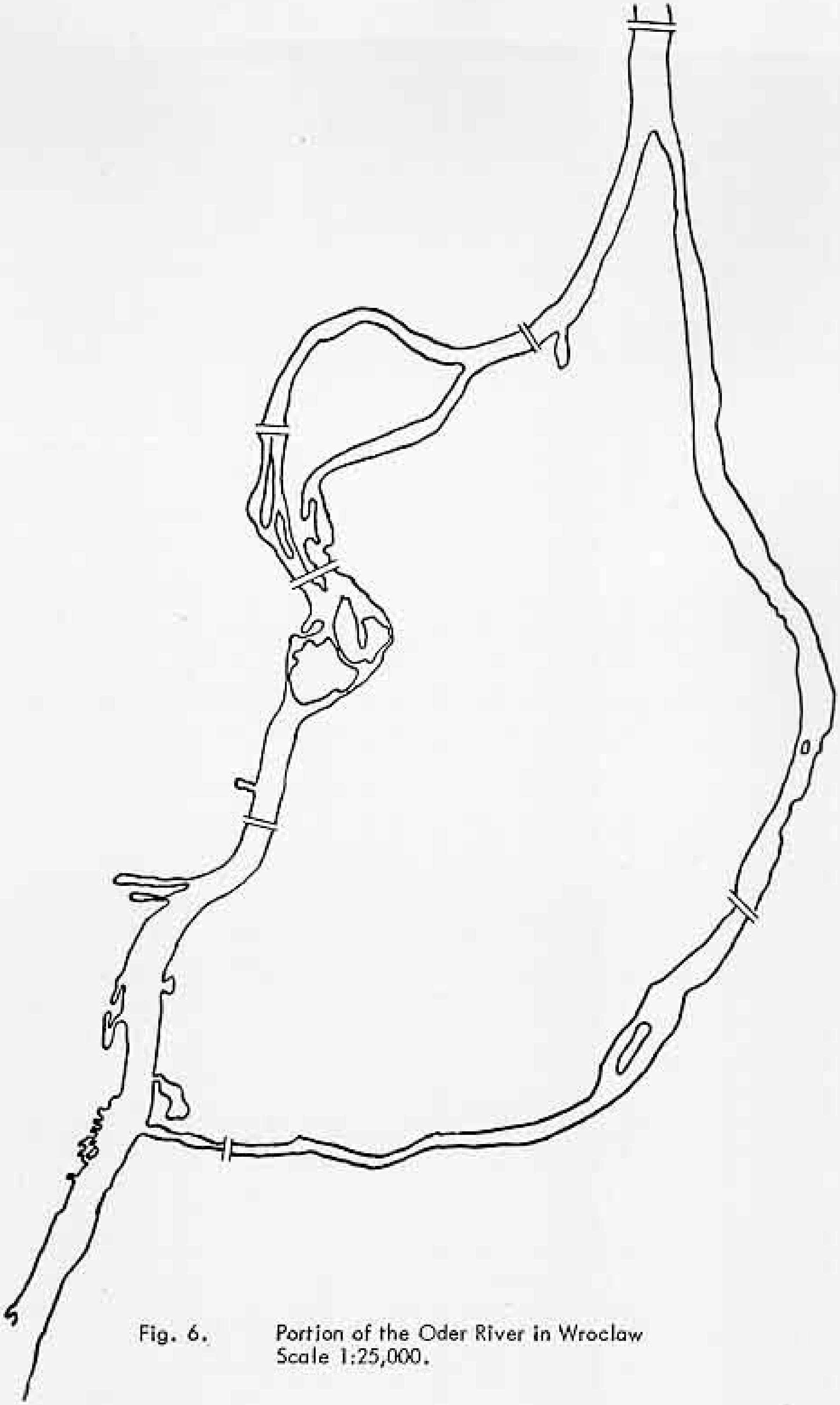


Fig. 6. Portion of the Oder River in Wroclaw
Scale 1:25,000.

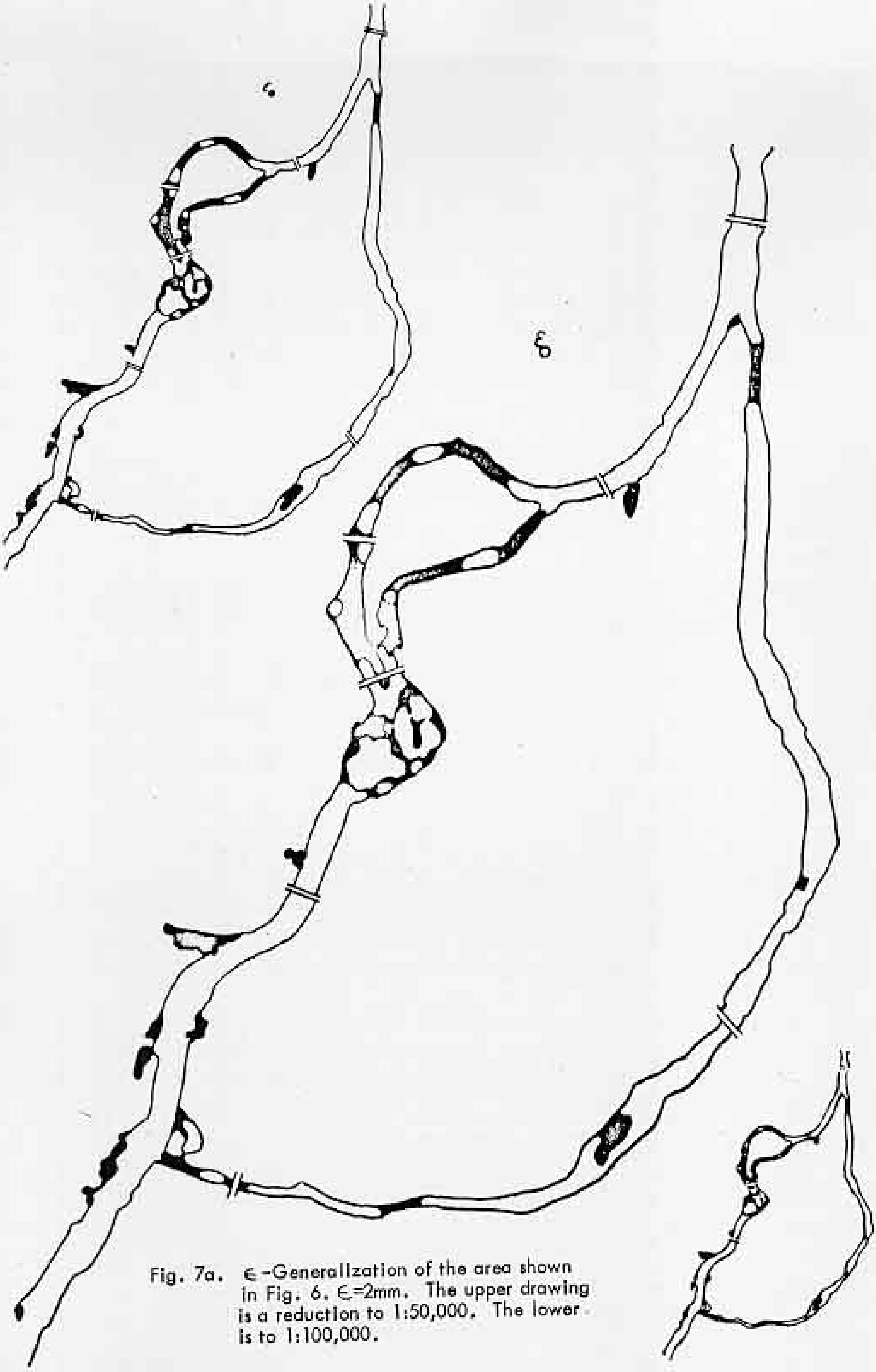


Fig. 7a. ϵ-Generalization of the area shown in Fig. 6. ϵ=2mm. The upper drawing is a reduction to 1:50,000. The lower is to 1:100,000.

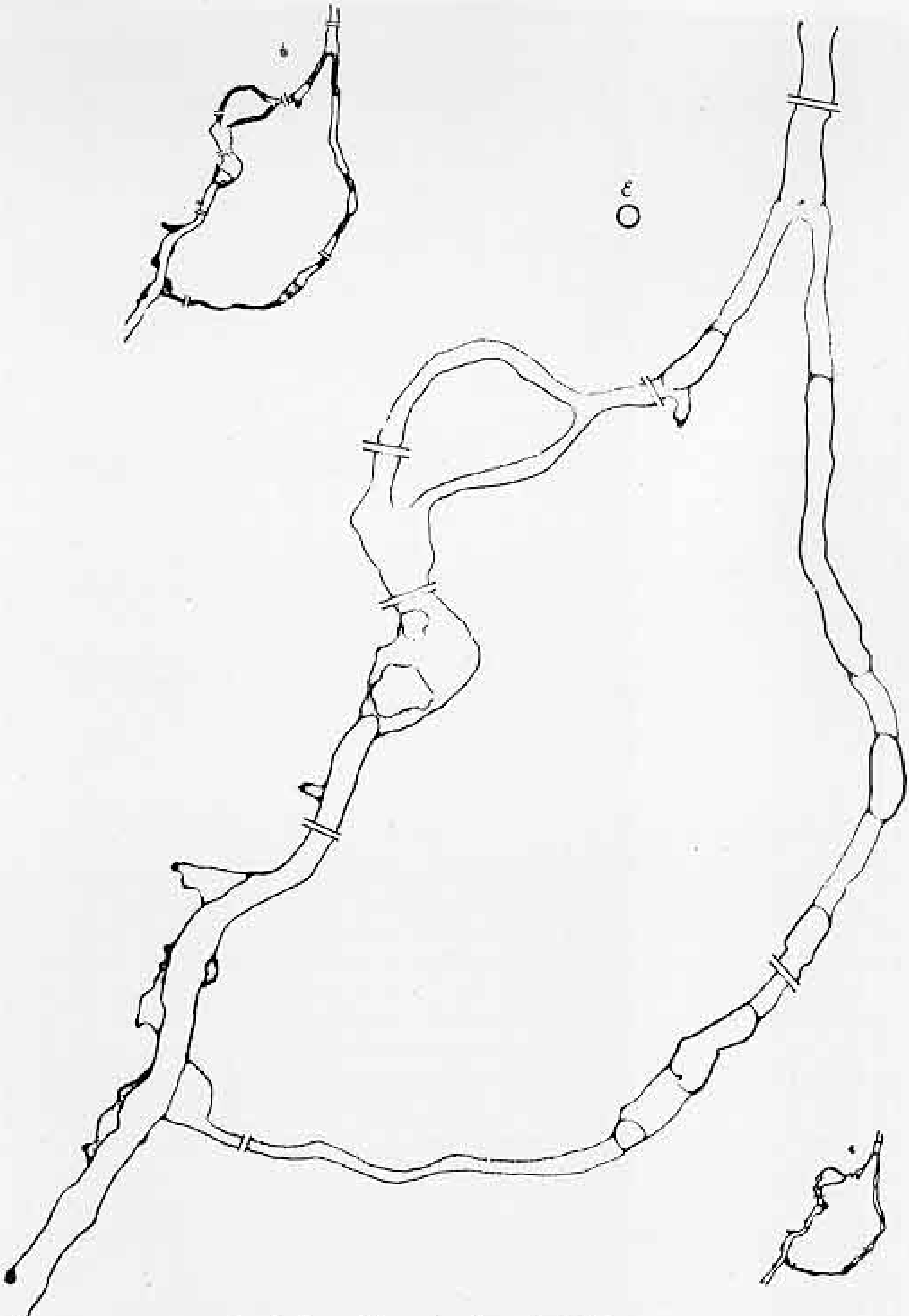


Fig. 7b. Є=4mm. Upper drawing reduced to 1:100,000. Lower one to 1:200,000.

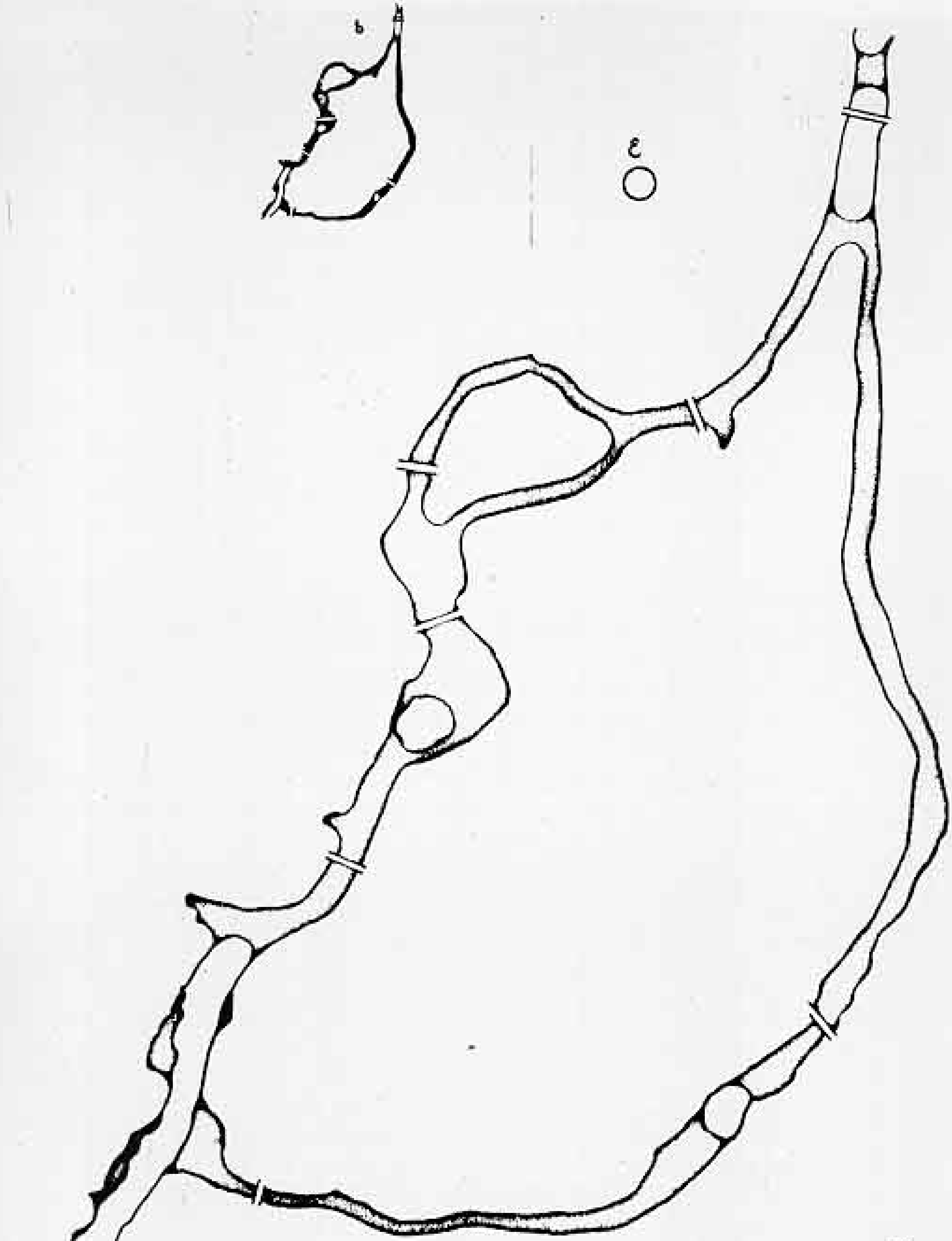


Fig. 7c. ε=6mm. Reductions to 1:150,000 and 1:300,000.

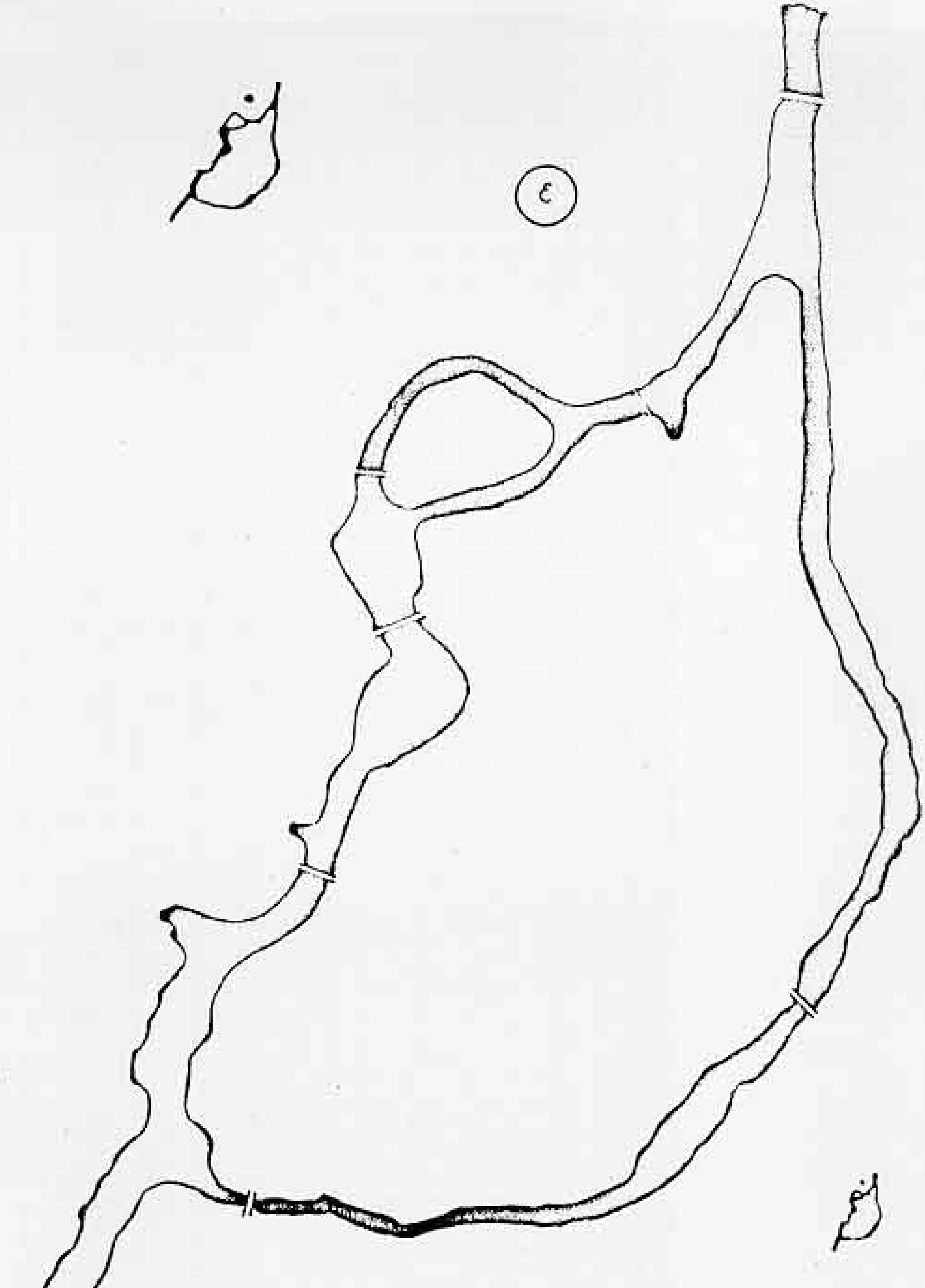


Fig. 7d. ∈=10mm. Reductions to 1:250,000 and 1:500,000.

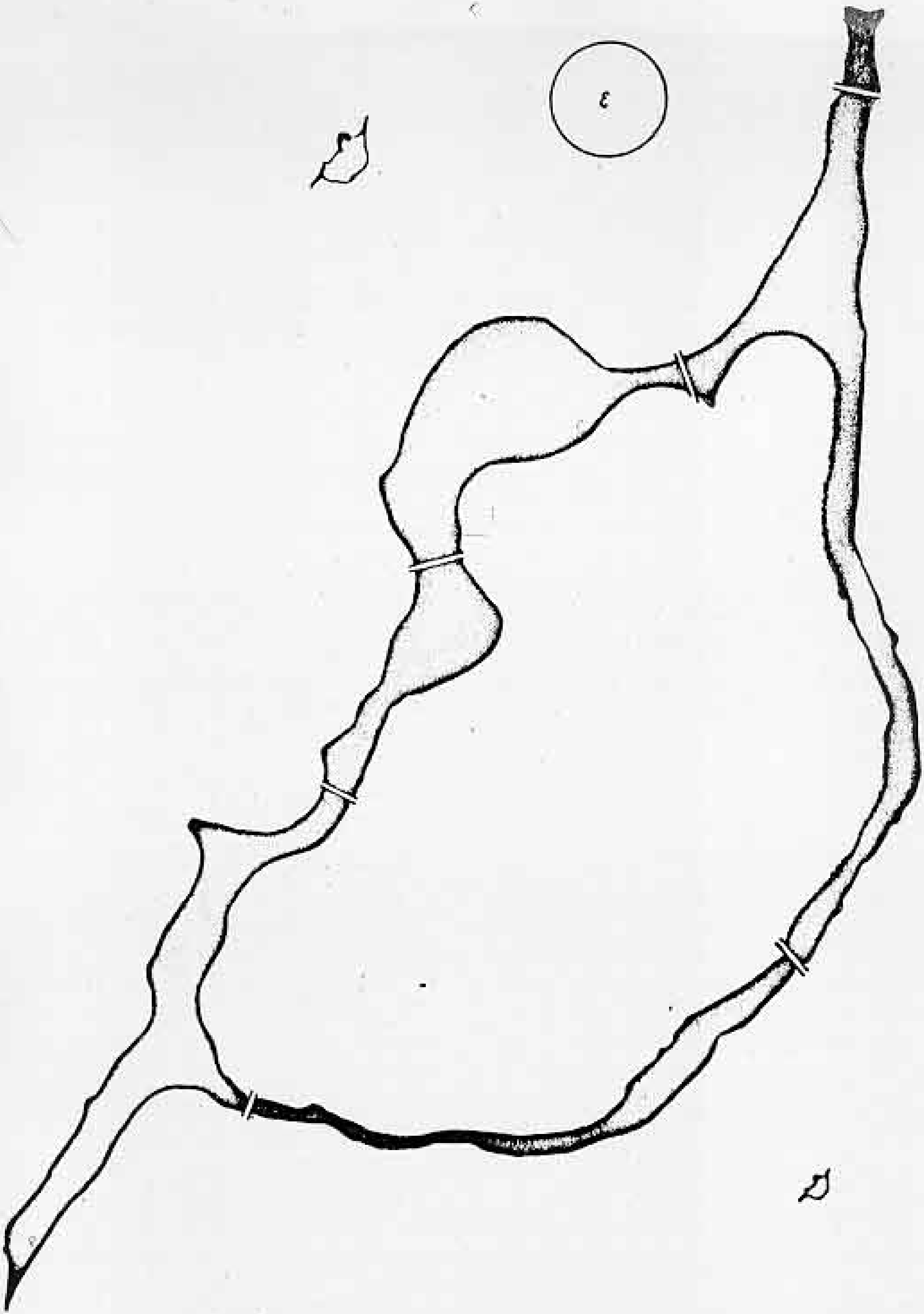


Fig. 7e. $\in$ =20mm. Reductions to 1:500,000 and 1:1,000,000.

scale. The generalized edges (shaded in the previous figures) have been blackened in Figure Seven.

The notion of ϵ-generalization described had its origin in the notions of ϵ-convexity, which I presented in an earlier paper (4). It is related to the problem of the measurement of length, to which I have devoted some effort (5). As mentioned at the beginning of this paper, it serves no purpose to contemplate the real length of an empirical line, since we do not know in general if this line can be rectified. We can, however, speak of length approximated to a given degree. For this purpose we can use one of several known methods (for example, the stepwise method for length ϵ, or the method of H. Steinhaus). The most correct method appears, however, to be my method of the measurement of length of order ϵ.

Figure Eight shows a longimeter of 3 mm. order; that is, a device with 3 mm. intervals for measurements of length. A sheet such as this (on transparent paper) should be lain at random six times over the curve to be measured. Each time count the number of circles that are touched or intersected by the curve being measured. For example, in Figure Nine, the curve touches or intersects fifteen circles. Subtract five from the sum of the number of circles intersected by the curve in six tries, and the resulting number is the 3 mm. order length of the curve in millimeters. The length of order ϵ is appropriate for ϵ-generalization, in the sense that if this does not change after ϵ-generalization, then the length of order ϵ is the real length (and there actually exists a finite length for this curve). The ϵ order length of an arbitrary edge of two areas is equal to the length of the $\epsilon\eta$-generalization of that edge (this length will be the length of a line which has been thickened in places), for an arbitrary number $\eta \leq \epsilon$. This property permits the comparison of the length of different lines on maps generalized differently.

These methods are not the only ones given in the reference cited above (5). It appears that with the help of the ϵ-longimeter one can measure the ϵ-length of a curve in a figure, separating the ϵ-generalized edge of an area from the ϵ-generalized areas. These kinds of measurements can be useful in determining the length of sea coasts, river basins, and transportation routes.

Fig. 8

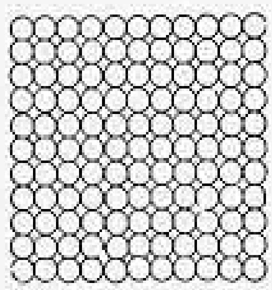
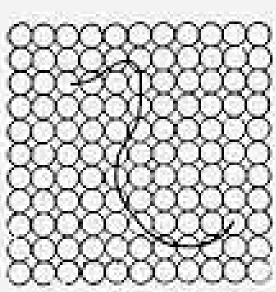


Fig. 9



Footnotes:

(1) In my previous papers on this subject (see references), I defined $\in$ as the radius, not the diameter, of the circle. Therefore, $\in$ in this paper is $2\in$ in the previous papers.

(2) Figures 2, 3, 4, and 5 have been reduced by such an amount as to make the numbers illegible.

Literature:

1. L.S. Garajewskaja, Kartografija, II izdanije, Izdatielstwo gieodeziczeskoj literatury, Moskwa 1955.

2. G.N. Liodt, Kartowiedienie, II izd., Gos. Ucz. Ped. Izd. Min. Prosw. RSFSR, Moskwa 1948.

3. A. Lomnicki, Kartografia matematyczna, II wyd. PWN, Warszawa 1956.

4. J. Perkal, "Sur les ensembles $\in$-convexes", Colloquium Mathematicum IV, 1, str. 1-10, Wroclaw 1956.

5. J. Perkal, "O dlugoscikrzywych empirycznych," Zastosowania Matematyki (w druku).

6. H. Steinhaus, "Length, shape and area," Colloquium Mathematicum III, Wroclaw 1954, str. 1-13.

FOREWORD

Does the coastline have a length? Like Lewis Richardson's "Does the wind possess a velocity?" "This question, at first sight foolish, improves on acquaintance" (1926, Proc. Roy. Soc., A, 110, p. 709). As Nystuen points out in the accompanying paper, these questions should not be thought of as theoretical curiosities. Close examination instead leads to very useful results. Richardson, in his paper, abandons the notion of limits in favor of nearest neighbor relations. Steinhaus and Perkal involke epsilon neighborhoods to overcome the paradox of length. Nystuen adopts this latter strategy to examine geographical boundaries.

Ratzel would have appreciated Nystuen's paper, with its clear definition of the area of an edge, for he once wrote "Der Grenzraum ist das Wirkliche, die Grenzlinie die Abstraktion davon." The well known literature of political goegraphy contains numerous boundary studies, to which Nystuen's paper now adds concepts and operational procedures which should have great impact. Even the cartographer should take note because it is now clear that generalizing the inside of a boundary differs from generalizing the outside of the same boundary.

One of the justifications for an organization such as the Michigan Inter-University Community of Mathematical Geographers is the exchange of ideas which it fosters. Professor Bunge brought the works of Steinhaus and Perkal to the attention of Professor Nystuen, who characteristically developed numerous and insightful geographical interpretations of the materials. Nystuen's first presentations on this topic were to the Community at its meeting place in Brighton during the spring of 1965. He subsequently gave a short paper (reproduced here) on the subject at a Regional Science Association meeting. A more recent paper, which includes further valuable geographical generalizations and extensions to problems of international concern, will be published in the Papers of the Peace Research Society (Volume VII).

To assist the reader, we have included translations from the Polish of two of Perkal's original papers. These also refer to the publications in French and English which Professor Nystuen had available. We were distressed to learn of Perkal's demise during the summer of 1966 and regret that he was unable to become familiar with these applications of his work.

W. R. Tobler

Regional Science Association
Annual Meeting
Philadelphia, Pennsylvania
November, 1965

EFFECTS OF BOUNDARY SHAPE

and the

CONCEPT OF LOCAL CONVEXITY

by

John D. Nystuen

The University of Michigan

Many spatial processes depend upon the shape of the partitions created by their boundary patterns. If the boundary shape is changed the process itself is changed, in fact, the very existence of the process may depend on the boundary shape.

In this paper I report on a concept which I believe is useful in analyzing the interplay between boundaries and spatial processes. This notion may be called _local convexity_ or perhaps, _convex-in-the-small_.

It seems to me three classes of subject matter are affected by boundary shapes. In one case the boundary affects the processes in the domains on one or both sides of the boundary. In the second case the boundary affects processes which are crossing it. And finally, I recognize a class of processes or spatial elements which exist only at boundaries. The main purpose of this paper is to suggest some measurement concepts which will be of aid in analysis of all three of these types of spatial processes.

Boundary Effects in Spatial Processes.

Some spatial domains have clearly defined boundaries. The boundary between land and sea is often well identified. At least it is well identified at certain scales. As an aside, I note that one is hard pressed to find the land-sea boundary to centimeter accuracy even in an instant in time such as would be available in analyzing an air photograph. Leaving that question for the moment, we can agree that at accuracies of one-tenth a kilometer, we can usually know where the land ends and the sea begins.

Sea processes are confined by this land-sea boundary and are affected by it. Consider the tides in the Bay of Fundy. The funnel shape of the bay causes a steep face on the tidal bore and creates at this place some of the highest tides in the world. Clearly here is a boundary shape which affects the spatial process it bounds.

The dimension of a spatial boundary is one less than the domain it bounds. For example, the tidal forces in the Bay of Fundy operate on the volume of water contained in the bay. The boundary shape which is important in determining the height of the tide, is the shape of the two-dimensional surface of the bottom and sides of the bay. Other processes can be considered to operate in essentially two-dimensional domains. They are bounded by one-dimensional edges. An example is the wave pattern on the surface of the bay. The wave pattern is sensitive to the shape of the one-dimensional shore line.

The first step in any analysis of the effect of shape is to determine the relevant dimension of the domain in which the spatial process operates. Many geographically significant processes can be considered two-dimensional. Line boundaries are the critical constraints to these processes. For simplicity, most of my remarks in this paper are concerned with two-dimensional domains and line boundaries.

A second class of boundary effects involves transfer processes. In processes or activities crossing a boundary there is often observed a transfer impedance which acts to reduce the effectiveness of communication across the boundary. Consider the customer contact of a supermarket across a river in a city. Figure I shows a sample of home places of customers to a supermarket in Ann Arbor. The river impedes contact across it. In part, this is geometric. There are only two bridges close enough to be of effect in the supermarket's trading area. Actually, people on the opposite side of the river must take a dog-leg route to the bridges in traveling to the store. In addition because the barrier is impermeable to automobile and pedestrian traffic except at the bridges, congestion occurs at the bridges which further reduces drawing power beyond the river. In analyzing this type of situation we need to be concerned with several general problems. The function of the barriers in terms of such questions as whether it is permeable all along its length or only at certain points is one problem. The second problem is geometric in the sense of the need to analyze the effects of the configuration and relative position of the barrier. Furthermore there is a trade-off implied between costs of additional crossings versus costs of congestion or queuing at a few points. Boundary function and boundary shape clearly affect transfer problems. Notice in passing that in order to analyze the function of a boundary, the activity involved in the crossing must be identified. The river would have a completely different role if the process under consideration was the spread of polluted air from nearby factories. Boundaries must be functionally defined.

Boundary shapes create another situation of great interest. Consider the coastline again. This boundary is no doubt the most important on earth. It is actually the line dividing three volumetric domains: the ocean mass, the ocean of air, and the land mass. Three volumes can meet at most in a line. Scientists postulate that life itself started at this interface.

There are many phenomena which exist only at the edge of a domain or in a region which we may call the boundary zone. Coast lines are in some places smooth and possessing very small average curvatures while at other places they may be

FIGURE 1

A Portion of Ann Arbor, Michigan

Sample of Supermarket Customers' Homeplaces, 1964

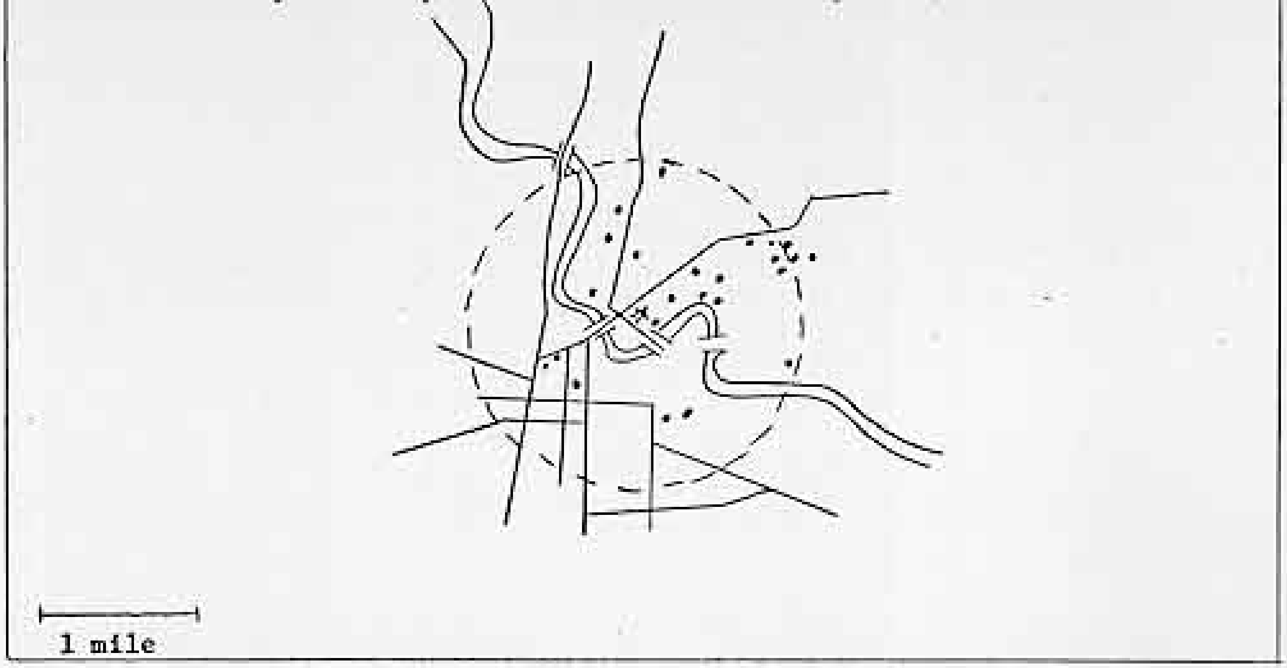
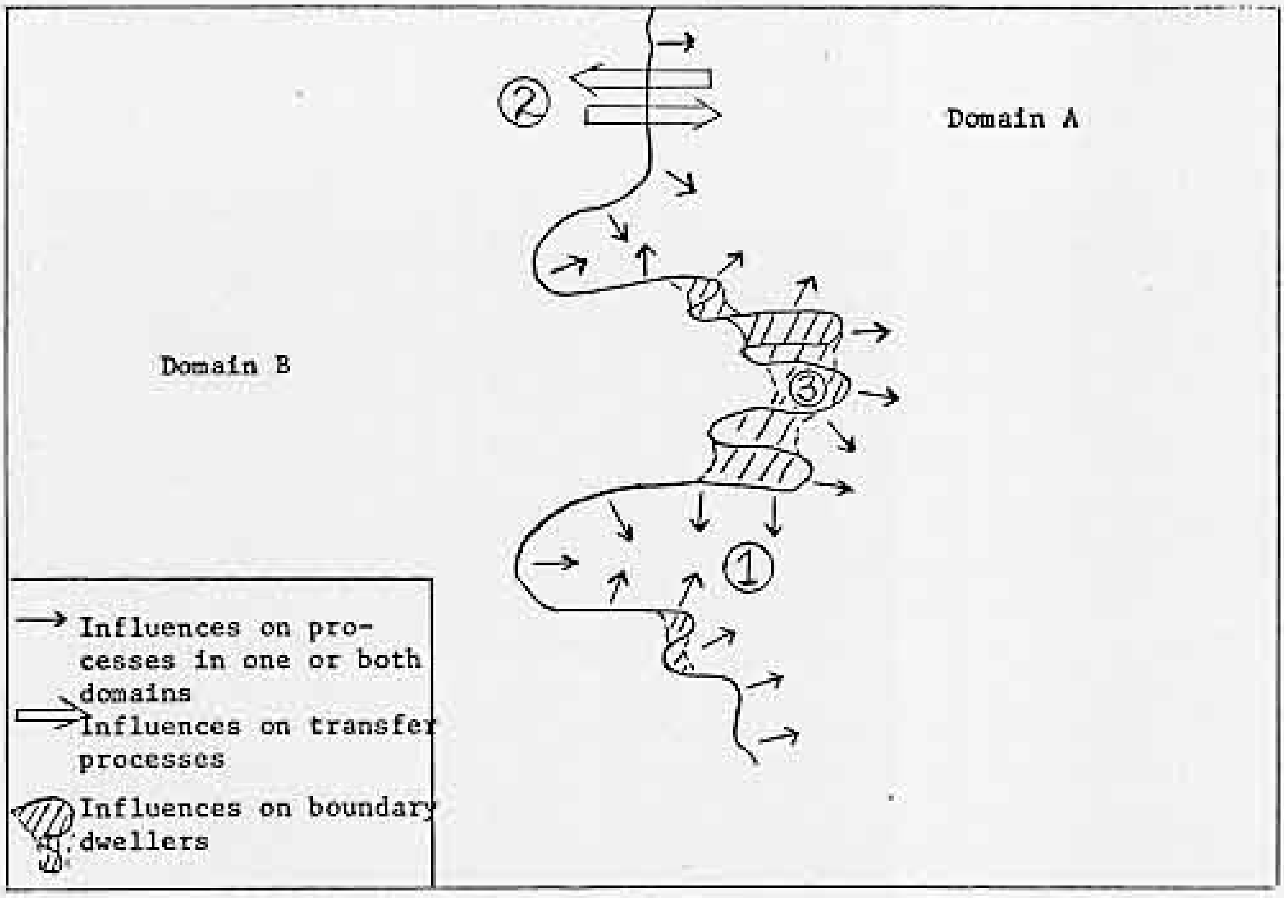


FIGURE 2

Three Classes of Boundary Influences:



deeply embayed and assume a sinuous form. This difference in shape makes for a difference in width of the boundary zone. The smooth coast has a narrow boundary zone and there is little room for boundary dwellers. The sinuous coast defines a boundary zone which is broader and many boundary creatures and elements exist. Clams are boundary dwellers. They are found most often in embayed areas where they are protected from the effects of the open sea. They also benefit by nearness to the domains of air and land even though they cannot exist in those domains. I propose to define boundary zones as a first step in analyzing those activities which are found confined to these zones. On the seacoast neither the sea nor the land dominates; each has its influence modified by the presence of the other.

I have purposely considered several types of phenomena at different geographic scales. I want to emphasize the unity of these topics insofar as they deal with boundary functions and shapes. I am convinced that there is an underlying unity in the spatial analysis. In order to make any progress in the analysis, specific abstract properties of boundary shapes and functions must be defined and measured. Analogy between widely different real world subject matters is a very dangerous procedure unless the phenomena are reduced in a crucible of abstract reasoning to the essential properties only. No unwanted transfer of meaning can be allowed. The essential properties we seek here are the ones with implications for spatial processes. We need clear definitions and measures of these properties.

Boundary Properties.

Several properties of boundaries may be specified. Clearly the dimensions of boundaries and the domains they contain are important characteristics. Boundaries have functions as well as dimensions. Some boundaries turn back elements that reach them. These are reflecting barriers. Their effect is to change the direction of movement but not to appreciably diminish the energy involved in the movement. Other barriers may absorb the elements that touch them thereby reducing the energy level of the process they confine. These are absorbing barriers. Such barriers reduce the level of energy of the process they bound. The Berlin Wall is an absorbing barrier. A third class of barrier is permeable. Part of the process passes through the boundary and part is either turned back or absorbed. These are permeable barriers. Obviously, a single boundary may involve to some degree all three of these functions. Subclasses of these three functions are also useful to consider.

It is clearly meaningless to talk about a boundary without specifying the activities involved in interaction with it. The same boundary will have different functions for different activities.

A major problem in analysis of the effects of boundaries is to devise operational definitions of boundaries which will allow systematic measure of the effects of a variety of boundary shapes and functions.

A reasonable assumption with regard to transfers from one domain to another is that transfers will depend on the permeability per unit length of a linear boundary or per unit area for a two-dimensional boundary. In cases with constant unit permeability the total transfer potential will be directly proportional to the length of the boundary line or, in the two-dimensional boundary case, to the total surface area.

Given the assumption above, the two ways to affect total transfers are to change permeability or to change the length of the boundary. If permeability is constant, different shapes of the same sized domains will have different transfer potentials. For minimum transfer, the circle is the optimum shape under these assumptions. The optimum shape for maximum interaction is not obvious. A shape could be compact with a very frilly edge resulting in a long perimeter. Another could be deeply indented such that all points within the domain are within some minimum distance to points outside the domain and yet the perimeter need not be exceedingly long.

Intuitively, measures of the length of a boundary and possibly its curvature would seem to be useful in analyzing effectiveness of boundary shapes. For example, the ratio of the length of a boundary to the area of the domain it encloses might be one such measure. The problem reduces to one of measuring length and possibly average curvature.

The length of a rectifiable arc and curvature at a point are familiar mathematical concepts. A rectifiable arc is one in which the tangent to the curve is continuous at all points on the segment. The following theorem emphasizes this points: (Courant, p.277)

> Every curve $y = f(x)$ for which the derivative $f'(x)$ is continuous, is a rectifiable curve, and its length between $x = a$ and $x = b$ $(b \geq a)$ is given by the formula
>
> $$S(a,b) = \int_a^b (1-y'^2)^{1/2}\, dx$$
>
> where $$y'^2 = \left(\frac{dy}{dx}\right)^2$$

This definition introduces a serious measurement problem. If the length of boundary is to be used in an index of effectiveness of shape, then the length of the line should be at least measurable in theory. Unfortunately most empirical lines have many points of discontinuity on them. The first derivative does not exist at all points on the curve. On deductive grounds it is not at all clear that the length of a boundary can be measured with theoretical rigor.

The Paradox of Length.

Steinhaus (1954) has pointed out the paradox of the length of empirical lines. Empirical lines are generally not rectifiable. The more accurately an empirical line is measured the longer it gets. The series of lengths obtained by repeated measures with finer and finer instruments does not ordinarily converge to a finite value. In general, the magnitude of the increments of change between successive measures does not vary systematically.

The length of a river or ridge line or other empirical line may be approximated by summing straight line segments between points on the line. Longer and longer lengths are obtained as the points are chosen closer and closer together. The same is true of the perimeter of a leaf. The finely serrated edge may be approximated at various scales but the finely measured lengths are always longer than those that are roughly measured. Even if a microscope were employed the length would continue to grow until at molecular level the length would approach the infinite.

The paradox is not to be confused with the fact that all physical quantities are subject to errors of measure. The problem remains regardless of the level of accuracy. A finer measure will always be longer. Nor should this problem be thought of as a theoretical curiosity. No agreement as to the length of an empirical line can be expected unless the scale of measurement is given. This is rarely done. One often hears, for example, recreation promoters who claim their region has hundreds of miles of lake front or fishing streams. No mention of how such values were arrived at is given. Nor can diplomats agree that the Polish-German border is so many hundred kilometers long, and so forth. Table 1. shows the extent of disagreement regarding the length of a mutual boundary between various nations. All such measures have little meaning.

Table 1. *

Disagreement on Length of Land
Frontier between Selected Nations

Land-frontier between	Kilometers as stated by	
	the former country	the latter country
Spain and Portugal	987	1214
Netherlands and Belgium	380	449.5
USSR and Finland	1590	1566
USSR and Romania	742	812
USSR and Latvia	269	351
Estonia and Latvia	356	375
Yugoslavia and Greece	262.1	236.6

* From Lewis F. Richardson, "The Problem of Contiguity" General Systems Yearbook v.6 (1961) p.169.

On the Epsilon Length.

Julian Perkal (1956b) has suggested a method of measuring the length of an empirical line. He proposes to define empirical length as a new concept which is analogous but not identical to the abstract length of a rectifiable arc. The notion depends on the idea of an ε-neighborhood of a curve. The ε-neighborhood may be defined as follows. Let capital letters represent sets of points and small letters represent individual points. The ε-neighborhood of a curve X is the set of all points on the plane for which the distance from the curve is not greater than ε, where ε is a fixed small distance. In set notation:

$$(1) \qquad A_\varepsilon(X) = E_X[d(p,X) \leq \varepsilon]$$

where $d(p,X)$ is the distance from a point p to the nearest element of the set X.

The set $A_\varepsilon(X)$ may be regarded as a function of ε and X and is monotonic increasing and continuous with respect to both arguments. The set $A_\varepsilon(X)$ may be regarded as defining the area $a_\varepsilon(X)$ which is therefore also a continuous, monotonic increasing function of ε and X.

Perkal (1956b) modified a definition of length of a rectifiable curve X given by H. Minkowski which is:

$$(2) \qquad L(X) = \lim_{\varepsilon \to 0} \frac{a_\varepsilon(X)}{2\varepsilon}$$

When this definition is used without passing to the limit, the ε-neighborhood of X includes a strip along the curve and two semi-circular areas around the end point of the curve to be measured. Not passing to the limit, the area required for definition of length is the strip along both sides of the line less the end areas and the definition becomes

$$(3) \qquad L_\varepsilon(X) = \frac{a_\varepsilon(X) - \pi\varepsilon^2}{2\varepsilon}$$

This is analogous to finding the length of a rectangle by dividing the area by its width. It may be seen by expression (3) that the length of a line is directly proportional to the area of the ε-neighborhood and inversely proportional to the size of the ε chosen. (See Figure III.)

Epsilon - Convex Curves.

Let the curve X be called ε-convex if we can draw at any point on the curve tangent circles with diameter ε, on both sides of the curve, having the point of contact as the only point in common with the curve. The epsilon length of a rectifiable curve is always less than the abstract length and approaches that length from below as ε gets smaller. Thus, using this definition we avoid the run-away length with closer and closer approximations.

The ratio ε/L characterizes the approximation that ε-length holds to ordinary length. For a suitable approximation ε should be chosen to make this ratio small, for example, one-tenth or less.

Measures of the Epsilon - Length of a Line.

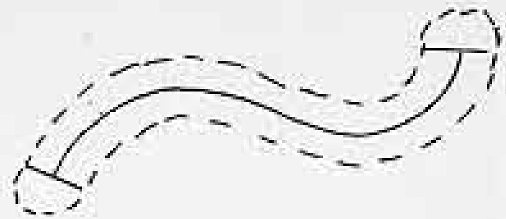
Steinhaus (1960) describes a method of measuring an area by using a lattice of points contained in the area as an approximation. For a square lattice in which a is the horizontal and vertical distance between points, the area of a region is proportional to n/k (a^2) where n is the number of points in the region, and k is the number of trials or placing of the lattice on the region. For a triangular lattice in which b is the distance between points, the formula for the approximation to an area is $\frac{n\sqrt{3}}{2k} b^2$

Substituting these equations for the area value in equation (3) yields

$$(4) \qquad L_\varepsilon(X) = \begin{cases} \frac{na^2}{2k\varepsilon} - \frac{\pi}{2}\varepsilon & \text{for a square lattice} \\ \frac{nb^2\sqrt{3}}{4k\varepsilon} - \frac{\pi}{2}\varepsilon & \text{for a triangular lattice} \end{cases}$$

The size of the lattice used for measuring the area is arbitrary and may be

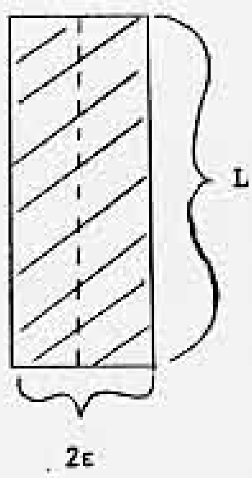
FIGURE 3



a ε(x) equals the area within ε distance of the curve. The ε-length of the curve is this area less the two half circles at the endpoints divided by the width of the strip, that is, by 2ε.

$$(3) \qquad L_\varepsilon(x) = \frac{a_\varepsilon(x) - \pi\varepsilon^2}{2\varepsilon}$$

This is analagous to finding the length of a rectangle by dividing its area by its width. $L = \frac{A}{W}$ where L = length, A is area, W is width (2ε).



chosen to be a function of the ε. Perkal was very clever at adjusting these constants so that they cancelled one another out and yielded a very simple formula for the measure of the ε-length of a line using the lattice point approximation. By setting $a = 2\varepsilon = k$ for the square lattice; and by setting $b = 2\varepsilon$ and $k = \varepsilon\sqrt{3}$ for the triangular lattice (k is rounded to the nearest integer) the following formula results

$$(5) \qquad L_{\varepsilon}(X) = n - \frac{\pi}{2}\varepsilon \begin{cases} a = 2\varepsilon = k & \text{for square lattice} \\ b = 2\varepsilon,\ k = \varepsilon\sqrt{3} & \text{for triangular lattice} \end{cases}$$

The ε-length of a line may be measured directly using this formula and the appropriate-sized lattice with ε-neighborhood circles drawn around each lattice point. A template with the lattice of circles may be used. The lattice is placed on the curve to be measured and each circle containing any segment of the curve is counted. Repeat the process k times with random placement of the lattice. If the curve is closed the correction factor $\pi/2$ (ε) is left off. If ε is in millimeters, the ε-length is in millimeters. Convert this value to actual ε-length by the scale of the map or photograph used.

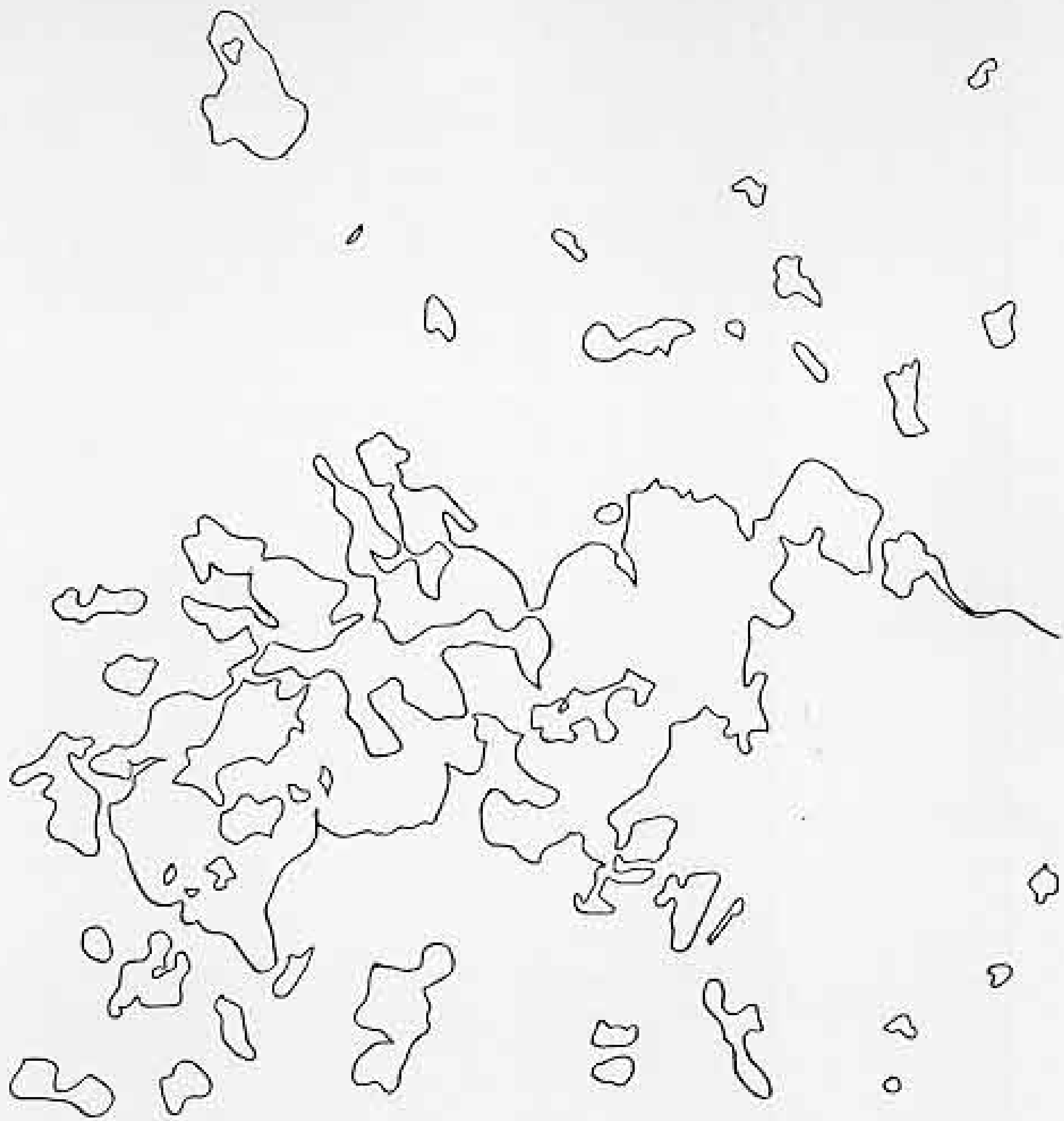
Figures IV and V and Tables 2 and 3 show the results of an application of these measures for Lake Minnetonka and a portion of the East coast of the United States. As expected, the smaller the ε, the longer the ε-length.

Mechanically Defined Epsilon-Values.

Some immediately practical consequences result from adoption of the notion of the ε-length of a line. Consistent and comparable measures of length can be made for line phenomena by suitable choice of ε.

Each map measuring device, such as a hand held map measurer, the plotting head of an automatic plotter, and so forth, have a minimum turning radius as a consequence of the mechanical properties of the measuring device. This turning radius is the ε-value of the instrument. Lengths measured by the device depend upon this value. Naturally different lengths of the same boundary will result from using different

FIGURE 4



Lake Minnetonka

1 0 1 2

miles

FIGURE 5

Portion of the East Coast of the United States

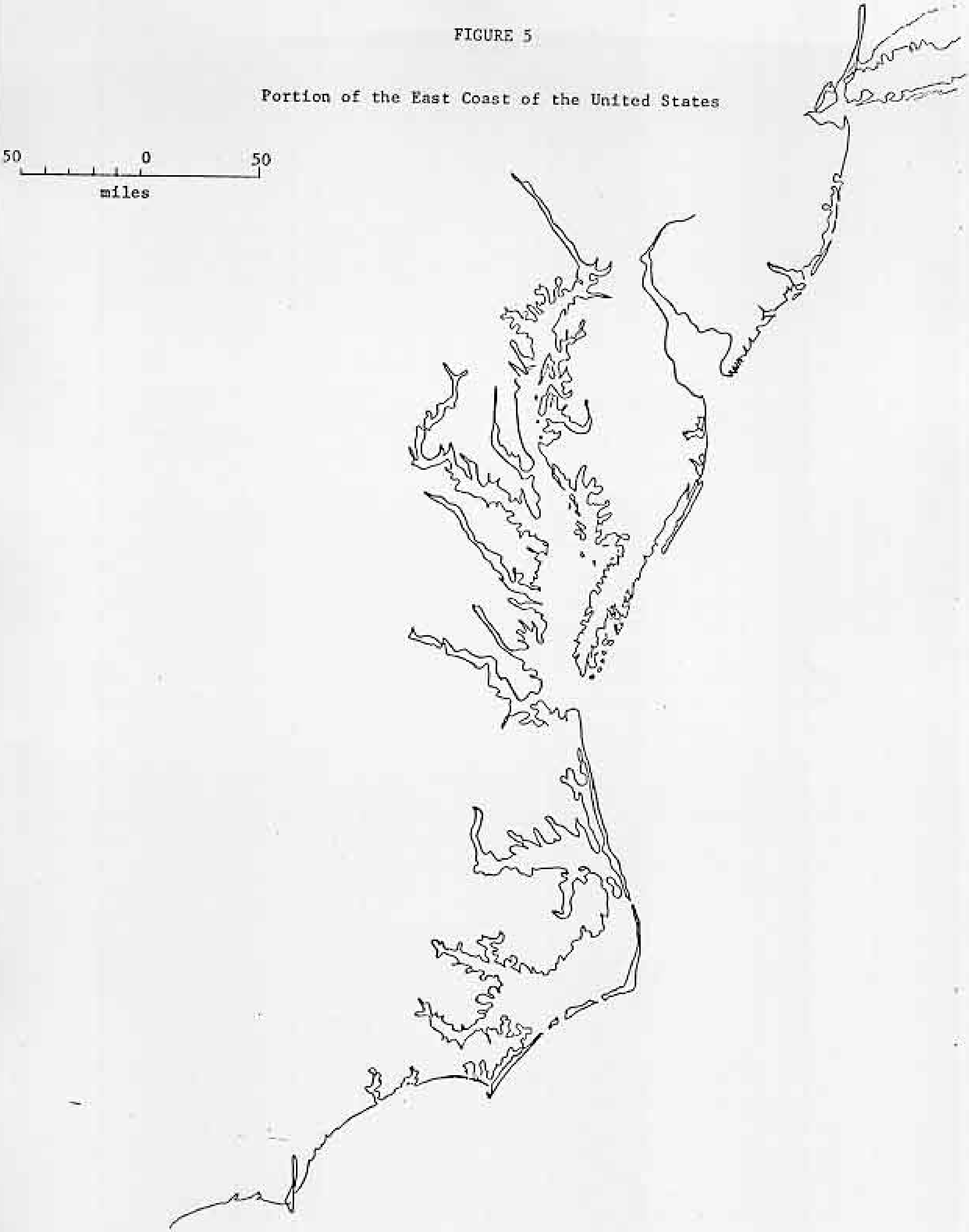


TABLE II

Measure of the ε-length of a Line Using Lattice Point Approximations

Lake Minnetonka
map scale 1:126,720

Triangular Lattice

trials

K	n
1	92
2	100
3	104
4	103
5	96
6	97
7	100

$\sum n = 692 - 6 = 686$ mm

$\varepsilon = 4$ mm
4x 126,720 .507 km
or .32 miles
~ 1700 feet

Total length of lake boundary (main lake only)

686 x 126,720 86.93 km
or 54.0 miles

Square Lattice

trials

K	n
1	62
2	63
3	63
4	58
5	62
6	64
7	65
8	61
9	63
10	59

$\sum n = 619 - 8 = 610$ mm

$\epsilon = 5$ mm
5 x 126,720 .634 km
or .39 miles
~ 2060 feet

Total length of lake boundary (main lake only)
610 x 126,720 77.3 km
or 48.0 miles

TABLE III

Atlantic Coast Approximate
ε-length Using Triangular Lattice

map scale 1:3,168,000

trials	
K	n
1	163
2	153
3	160
4	164
5	155
6	168
7	170

$\sum n = 1123 - 6 = 1117$ mm

$\varepsilon = 4$ mm
$4 \times 3{,}168{,}000 = 12.67$ km
or 7.9 miles

Total length of portion of coast shown,
(islands not included)

$1117 \times 3{,}168{,}000$ 3538×10^6 mm
3538.7 km
or $2199 \pm 8 = 2200$ miles

instruments or different scales of map or photographs of the object to be measured. No true length exists and if each measure by several methods are carefully done, there is no basis for choosing the best value independent of the subject matter under study. That is, the researcher must decide if there is a functionally significant minimum radius involved in the objects under observation. If such a radius can be identified the ϵ may be consciously specified by the researcher rather than being arbitrarily established by mechanical limits of the methods of observation.

Other means of empirical observation are subject to the same restraints. The grain size of photographs, the resolving power of lens and other remote sensing devices such as infra-red sensors establish minimum ϵ for measurement and all length depends upon this ϵ. Clearly, if length is an important variable in the analysis, the researcher, not his machines, must specify the most suitable ϵ to be used.

For some purposes a consistent ϵ is all that is needed. Perkal (1958a) suggests a method using the ϵ length concept in which consistent generalization of a map can be achieved when constructing a small scale map. What is required in this case is not the absolute ϵ-radius of the objects studied but rather a technique for consistent generalization of all parts of the map. Tobler (1964) has found using the mechanical ϵ of graphic plotters yields satisfactory results in map generalization.

Functionally Defined Epsilon-Neighborhoods.

The concept of the ϵ-length of a line provides an operational definition which may be used to measure empirical lines such as seacoasts and rivers. The length of a seacoast varies depending upon the purposes involved. For example, for any stretch of the Atlantic seacoast, the length is longer from the point of view of somebody in a rowboat than for somebody in an ocean liner. The difference depends upon the turning radius of the vehicles.

The length of this coast is also different from the point of view of railroads. Another factor is involved here, namely, the boundary of the ϵ-neighborhood may be thought of as the trace of points nearest the coastline left by the edge of a circle of ε-radius which is rolled along the coastline on both sides. Because of the shape of the capes and bays, the length of this trace will not be the same on each side of the seacoast. Ships swing wide to avoid the capes but trains swing inland to avoid bays and estuaries. The coastline, then, appears to have an inside length and an outside length for purposes of movement along its perimeter depending upon the technical requirements of the vehicle used. The statement may be made more general. The boundaries of the ε-neighborhood of a line are ordinarily not equal; the one-side length equals the other-side length only if the line is ε-convex. I will return to this peculiar point later.

Clearly, one method of defining a functionally significant ε-radius has to do with the turning radius of vehicles. This radius depends upon technical consideration of mass and energy and of uncertainties in steering including both navigational needs and shoulder and head room requirements at operational speeds. The length of barriers, therefore, are different for purposes of an expressway as compared to a highway; for very high speed ground vehicles, such as proposed for the Northeast Corridor route, as compared with a typical railroad; for vehicular travel as compared with walking and so forth. Distances between points in a region vary by mode of travel, in part, as a consequence of these technical reasons.

Transportation characteristics are one way a functional ε-radius may be specified for a study. Another, somewhat more subtle but probably more important, functional meaning can be applied to the ε-neighborhood. The ε-neighborhood of a point is the area of a circle of ε-radius. Most objects under study in spatial systems have internal spatial requirements. Take, for example, various land uses in a metropolitan region. Farms, factories, residential sub-divisions, individual residential lots, and so forth have typical internal dimensions within some range of variability. These site requirements must be taken into account in defining the spatial patterns of the land uses in the region. Suppose, for example, you are planning to identify the shape of the urbanized area of the region and to do this you are to employ a remote sensing device with a resolution power of one square foot. Let us say this instrument can identify soft ground from hard surfaces. The decision is made that hard-surfaces are urbanized area and soft ground are areas of vegetation and non-urban land. The mechanical ε-neighborhood of the instrument covers one square foot and it assigns each square foot to one of the two classes depending on the proportion of hard and soft within the unit area.

If the entire metropolitan region were then monitored and plotted on a map, the urbanized areas would not be found. The lawns of residential plots would have been classified as non-urbanized area, country roads and farm buildings would have been classified as urban. The problem is that the concept urbanized area contains an implied ε-neighborhood, which depends on the site requirements of the various activities found in metropolitan areas. If an ε-neighborhood of the instrument had been made larger so that it was equivalent to the average-sized residential lot and if the typical proportion of soft to hard surface for the residential site requirements were specified; the urbanized zone produced by these automatic means would be in much better correspondence to the theoretical concept of an urbanized area.

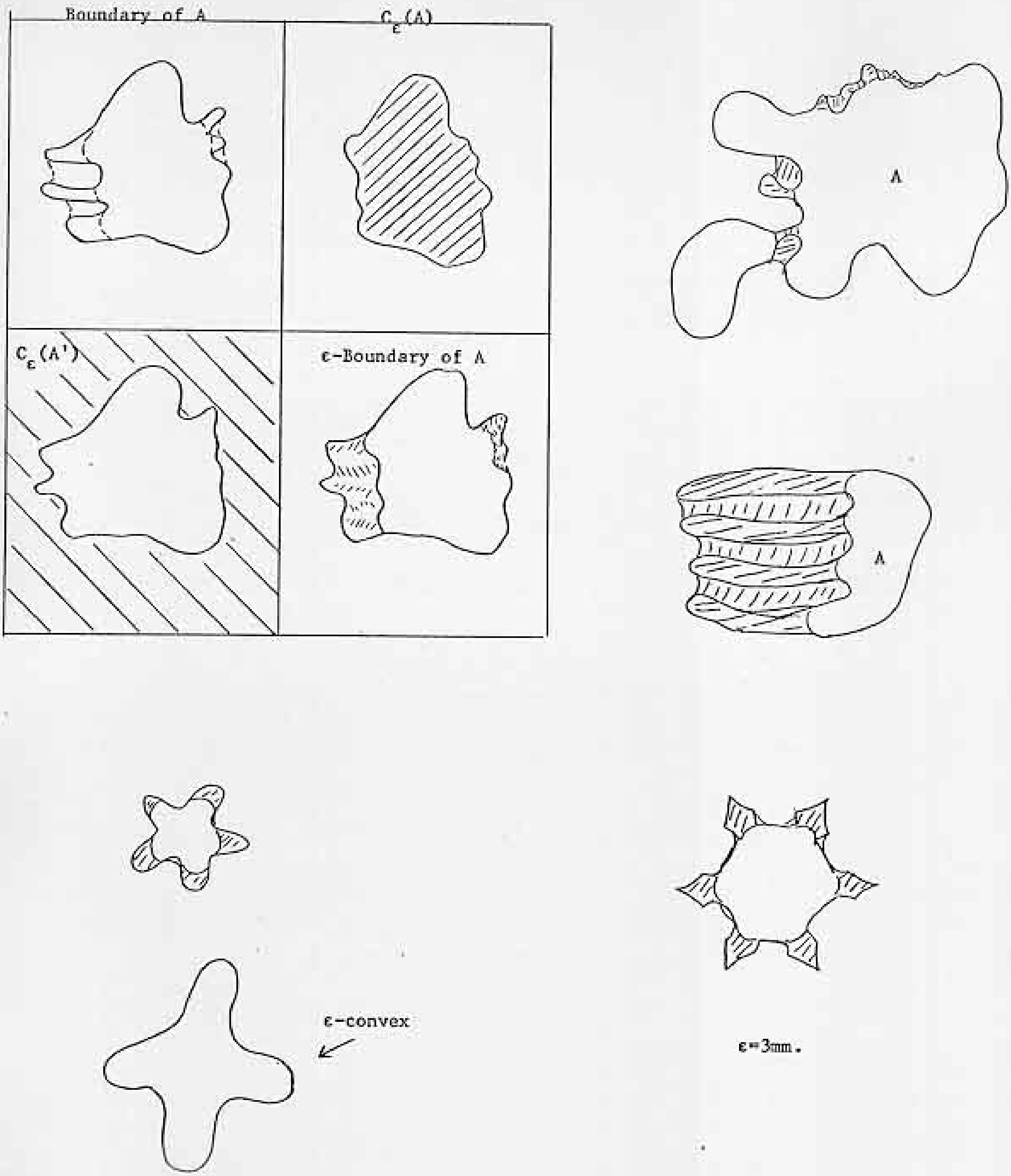
These observations provide insight into the dichotomy of site versus situation commonly found useful in spatial analysis. Variations within an ε-neighborhood are site variables and are specified by indices of texture, roughness, intensities, etc. Variations between ε-neighborhoods are situational or locational variables and these variables define shape, relative position, length, etc.

The Boundary Zone.

The idea of an ε-neighborhood of a point leads to an operational definition of boundary zone. Consider domain A surrounded by an irregular boundary. The ε-convex set of A is defined in the following manner. Let p be a member of a set of points contained within but not on a circle of radius ε, which in turn is wholly within the domain D. Then the points p, in the collection of all such circles define the open ε-convex set of A. Denote this set by $C_\varepsilon(A)$. All the points of A need not be in $C_\varepsilon(A)$. Those points are excluded that are in places where the boundary of A curves back on itself with less than 2ε distance across the loop. Because this set is open the boundary points of A are also excluded.

FIGURE 6

Boundary Zones Defined by ε-Convex Edge of Domain A and its Complement A'



The complement of A is A'. A $C_\varepsilon(A')$ exists but is equal to the complement of $C_\varepsilon(A)$, i.e., $C_\varepsilon(A)'$ only when the boundary is an ε-convex curve as defined previously. If the boundary is not an ε-convex curve, a space or collection of points is left between $C_\varepsilon(A)$ and $C_\varepsilon(A')$ which is the <u>ε-edge</u> or <u>ε-boundary zone</u>. This edge is reduced to the boundary line for ε-convex curves.

The boundary zone may be wide or narrow depending on the shape of the boundary and, of course, on the ε chosen. Figure 4 provides some examples. In some cases the ε-boundary zone contains more of one domain than its complement. Segments of either domain or both may be isolated as well by the boundary zone.

A boundary zone defined in this fashion suggests many ways in which the efficiency of a shape for various purposes might be analyzed. If the permeability of the boundary zone is less than the permeability of a unit area of the domain on either side, the most efficient locations for crossing the boundary would be in places where the boundary zone is the narrowest.

The relative strength of each domain in the boundary zone may be established by the proportion of each domain found there or perhaps by the relative difference in the length of the convex envelope on the one side of the edge zone compared to the other side.

Phenomena found to be confined to zones within a certain distance of a boundary may be designated as boundary dwellers and a functional ε for them would be indicated. It would remain, of course, to establish theories for each particular subject matter as for why certain activities are found at boundaries only.

Many other interesting topics are suggested by the ideas I have presented here. Central to Julian Perkal's ideas is the ε-neighborhood. It may be characterized as <u>site convexity</u> or <u>convex in the small</u>. I have great expectations for its applications to spatial analysis of all types. I think it will be particularly useful in studying the effects of boundary shapes on spatial processes, subjects which have received little attention to date, and which I have tried to introduce to you today.

References

1. R. Courant, 1937, Differential and Integral Calculus, Vol. I, 2nd Edition, Vol. II, 1936. New York: John Wiley and Sons.

2. Julian Perkal, 1956a, "Sur les ensembles epsilon convexes." Colloquium Mathematicum v. 4: 1-10.

3. __________, 1956b, "On the epsilon-length," Bulletin of the Polish Academy of Science, Cl III, v. IV #7:399-403.

4. __________, 1958a, "Proba Obiektywnej generalizacji," Geodezia i Kartografia, Tom VII, zeszyt 2.:130-142.

5. __________,1958b, "O Dtugosci Krzywych Empirycznych," Zastosowania Mathematyki, III, 3-4:258-283.

6. William Bunge, 1963, Theoretical Geography, Lund Studies in Geography, Series C, Mathematical and General, #1 (Lund: Gleerup).

7. Lewis F. Richardson, 1961, "The Problem of Contiguity," an appendix to Statistics of Deadly Quarrels (edited by Quincy Wright and C.C. Lienau) Pittsburgh, Boxwood Press, 373pp. Also in General Systems, Yearbook of the Society for General Systems, v. VI:139-187.

8. Hugo Steinhaus, 1954, "Length, Shape, and Area," Colloquium Mathematicum: 1-13.

9. __________, 1960, Mathematical Snapshots, Oxford University Press.

10. Waldo Tobler, 1963, "Geographic Area and Map Projections," The Geographical Review, v. 53:59-78.

11. __________, 1964, An Experiment in the Computer Generalization of Maps, Technical Report #1, Office of Naval Research, Task no. 389-137.

12. Edward L. Ullman, 1941, "A Theory of Location of Cities," American Journal of Sociology, v. 46. #6.

13. __________, 1956, "The Role of Transportation and the Bases for Interaction," in Man's Role in Changing the Face of the Earth, William Thomas, editor. University of Chicago Press:862-880.

14. Robert S. Yuill, 1964, "A Simulation Study of Barrier Effects in Spatial Diffusion Problems," Office of Naval Research, Geography Branch. Task no. 389-140, Contract NONR1228(33) Technical Report #1.

A METHOD FOR THE STEPWISE SEPARATION OF
SPATIAL TRENDS

E. Casetti R. K. Semple

The Ohio State University

The problem of separating trends occurs in very similar terms in the analysis of time series[1] and of space series. Trends are general tendencies persisting through time or through space that cause time or space series to increase or decrease "smoothly." In order to isolate them, the original series are split into components associated respectively with the trend and with the residual, such that the summation of their corresponding terms reproduces the original series.

Techniques of time series analysis can be and have been extended to space series. The interpolation of planes and surfaces to geologic series[2] closely parallels the fitting of lines and curves to time series. The only difference is that functions of two spatial coordinates (functions of two variables) are used in the first case, and functions of time (of one variable) in the second. Also, the mobile averaging can be easily extended to two dimensional cases, for instance, by

---

The authors gratefully acknowledge the constructive criticism by Professor W. R. Tobler, University of Michigan.

[1]See, for instance, H. T. Davis, The Analysis of Economic Time Series, Detan Printing Company, Colorado Springs, 1941.

[2]A Discussion and bibliography is contained in R. L. Miller and J. S. Kahn, Statistical Analysis in the Geological Sciences, Wiley, 1962, p. 394-.

replacing the value of each term of a space series by some average of the values of the terms within a given distance from it.[3]

The Method

The methods proposed in this paper are expecially suited to extract spatial trends,[4] but may be applied to time series as well. Basically, they are designed to give a quantitative expression to smooth changes functionally related to distance from points or lines of reference. In particular, these methods may be used in order to identify the smooth decline of the "influence" of an urban center, or of a cluster of urban centers with distance from a point of reference that may or may not coincide with downtown areas. The methods identify trends by means of optimal points of reference or origins. A point of reference is considered optimal when the correlation between some predetermined transformation of the distance of the series terms from it, and the values of the terms are not smaller than for any other admissible point of reference. In order to clarify the concepts involved, a simple time series example is now discussed. Suppose, for instance, in the time series below

Value	5	4	3	2	1	2	3	4	5
Time	t	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8

that each term of the series has a value and a time coordinate. If the

---

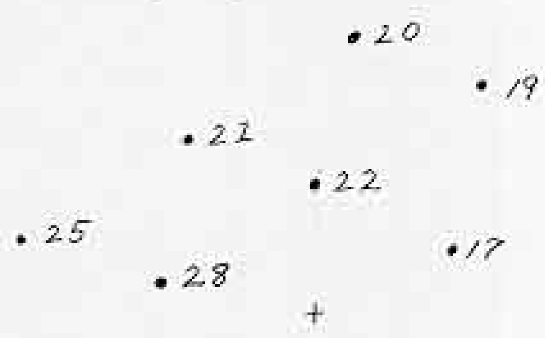
[3]See P. E. Potter, "The Petrology and Origin of the Lafayette Gravel," Part I: Minerology and Petrology, Journal of Geology, 63, (1955): pp. 1-35.

[4]The first application of the methods described in this paper can be found in R. K. Semple, A Quantitative Separation and Analysis of Spatial Trends in the Viability of Small Urban Centres in Southern Ontario. Unpublished M.A. Thesis, University of Toronto, 1966.

time coordinate t+4 is taken as origin, distances of the terms from it can be easily obtained. The correlation between these distances and the respective term values is 1. Clearly the correlation between term values and distance from any other origin would be less than 1, so that in this case, t+4 is the optimal origin. A regression, in which time distances from the optimal origin (or some transformation of them) is the independent variable and the values of the time series are the dependent variable, could be calculated and residuals obtained. In the example above the residuals are all zero. But suppose that this was not the case. Then the procedure could be repeated on the residual time series, and a new optimal origin determined, new residuals obtained, and so on, iteratively. These iterations would be interrupted when the residuals are random or smaller than a threshold. The difference of the initial time series minus the last residual time series constitutes a trend component.

For extracting spatial trends, a two-dimensional generalization of the procedure above can be applied. The concept of optimal origin is easily extended to space series. For example, suppose the series identified by the following diagram:

Figure 1



The circles indicate the location of the terms of the space series and the numbers adjacent to them are their values. The cross is an arbitrary "origin." The distance between the cross and each of the circles (or some function of this distance) can be calculated for each "term" and correlated with the values of the terms. An origin is optimal if the correlation coefficient associated with it is not smaller than the one associated with any other admissible origin.

The procedure for extracting spatial trends involves the following steps:

1. An optimal origin is obtained.
2. The values of the space series terms are regressed against the distance of the locations of the terms from the optimal origin. The values predicted by the regression identify a trend associated with the optimal origin.
3. The residuals, that is the difference between the original values and the predicted values, form a new space series that can be again separated into a trend and a new residual. To this effect, a new optimal origin is located, such that the correlation between the distances from it to the location of the terms of the space series, and the residuals from the first regression is maximal.
4. A new regression can be calculated and new residuals obtained.

The procedure can be repeated again and again. Each time it yields an optimal origin, a space series incorporating the portion of the trend associated with the optimal origin used, and a space series consisting of residuals. The iterations can be stopped when the residual space series is random.

The variance of the original space series can be decomposed into the variance of the trend component associated with the successive optimal origins plus the variance of the residual. A rule of thumb for deciding when to stop the extraction of new trend components from the residual may be based on the proportion of the total variance in the last (nth) trend component. A threshold is set and trend components are extracted only when their variance is greater than the threshold.

The Scope of the Technique

This procedure for the extraction of spatial trends is eminently applicable when empirical observations or theoretical considerations suggest that a given spatial variable tends to have values the higher (or lower), the closer they are to given locations. However, it can be shown that the procedure is also effective for trends related to straight or curved lines. Suppose, for instance, the space series in the following figure:

Figure 2

27 0	25 0	26 0	27 0	25 0	26 0
30 0	29 0	31 0	30 0	30 0	29 0
26 0	27 0	25 0	25 0	26 0	27 0

High values occur along the central "ridge," and decline on both sides of it. The values of the variable decline with distance from a line rather than from a point, and therefore the trend can be associated with a line. Suppose the procedure discussed above is applied. One of the points along the central ridge would be chosen as first optimal origin, and the linear regression of the values of the series against a transformation of the distances from the optimal origin would be calculated. Clearly, the residuals will tend to be larger on the portions of the central ridge away from the first optimal origin. Therefore, the second optimal origin would be located on the ridge, away from the first origin. A similar reasoning can show that the procedure will tend to locate, along the central ridge, a sequence of points that will capture the trend.

The same will happen for other kinds of trends. Each optimal origin will specify further portions of them until purely random residuals remain.

When the value of a variable in a space series tends to decline (or to increase) with distance from certain "poles" the procedure will yield optimal origins that will tend to coincide with these poles. In this case the optimal origins have a substantive meaning. If instead the spatial trend is related to lines rather than points, the optimal origins taken separately do not have meaning other than a formal one.

Similar considerations apply to the trend components associated with individual optimal origins. Whenever these origins correspond to poles their trend components identify the areas where the influence of the poles is felt, and it can be illuminating to map them.

Instead, when line trends are extracted, only the sum of the trend components associated with all the optimal origins has a substantive meaning.

The method proposed can be useful to geographers in different respects. It decomposes space series into two components containing respectively smooth variations of variables over space, and residuals. These two layers can be investigated, for instance, by multiple regressions, in order to analyse separately the causative factors related to trends and those related to residuals from trends.

Furthermore, within some substantive contexts, the optimal origins and the trend components associated with them may identify cores of regions, and areas in which the influence of these cores is felt.

The spatial trends to which this procedure is suited involve a non-linear decline with distance from a point or line of reference. The trends, therefore, can be visualized as mounds, ridges, or hills rather than as traingular cones. Suitable transformations of the variables can be used so that the influence of each optimal origin declines more than proportionally with distance.

The transformations suggested are 1) the logarithmic transformation of the values of the series terms, and 2) the reciprocal of the distance from the optimal origins--increased by one. If the first transformation is used, the relationship between trend and distance from optimal origins is obtained in the following form:

$$V_{ij} = \exp (a_1 - b_1 d_{1ij} + a_2 - b_{2ij} + \cdots + a_n - b_{nij})$$

Where V_{ij} is the value of the space series term with coordinates i and j, d_{kij} is the distance of the term with coordinates i and j from the kth optimal origin, and a_k and b_k are regression coefficients.

With the second transformation instead, the relationship between trend and distances is:

$$V_{ij} = a_1 + \frac{b_1}{d_{1ij} + 1} + a_2 + \frac{b_2}{d_{2ij} + 1} + \cdots + a_n + \frac{b_n}{d_{nij} + 1}$$

<u>An Empirical Application</u>

In the example that follows it was attempted to separate spatial trends from data on population growth in small urban centers in Southern

Ontario.[5] The computer program used is discussed in detail in Appendix A and only its basic structure is described here. The input of this program consists of coordinates and the values of the terms of a space series.

Optimal origins are searched for by superimposing over the study area successively finer grids. The differences of the largest and the smallest ordinate and abscissa values of the space series are calculated. The largest of the two differences is chosen as the length of a square study area. The sides of this square area are then divided into fifteen equal parts which identify a 15 by 15 grid. From each intersection of the grid the distance to every term of the space series is calculated, and the (transformed) distance is correlated with the values of the terms so that a 15 x 15 matrix of correlation coefficients is obtained. The largest correlation coefficient and its coordinates are identified. In order to determine more precisely the actual coordinates of the optimal origin, the four diagonal correlation coefficients adjacent to the one identified are used to define the limits of a second finer grid from which a 10 x 10 matrix of correlation coefficients is calculated. Again the largest correlation coefficient and its coordinates are identified. One final 8 x 8 matrix of very fine coefficients is then obtained, and the largest coefficient in it is assumed to indicate the actual point of highest correlation on the study area. This point is taken to be the optimal origin and the values of the space series are regressed against the transformed

[5]For the locations and identification of the towns in the study area see Table 1 and Map 1.

distances from the optimal origin to the locations of the space series terms. The variance explained by the distances is calculated and its ratio to the variance of the space series obtained.

The regression residuals are then calculated. They provide the new space series on which the steps above are repeated. The procedure is iterated until an optimal origin is obtained, such that the distances from it explain less then five per cent of the variance of the residual. The program is then terminated. The program outputs: 1) the coordinates of the optimal origins, 2) the residuals from the trends associated with each optimal origin and the final residuals, 3) the proportion of the variance of each residual accounted for by each component.

The program was used to analyse the population growth of incorporated towns of Southern Ontario with a 1951 population between one and five thousand. Different areas of the province appear to be characterized by different growth rates. Thus, the purpose of the experiment was to identify and separate these trends. The input consisted of geographical coordinates and of the percentage population change between the 1951-1961 census period (See Table 2). Five optimal origins were obtained which account for over 90 per cent of the population change's variance (See Table 3).

Growth rates tend to decline with distance from three optimal origins clustered in a "horseshoe" fashion around the western end of Lake Ontario and increase with distance from a two optimal origin located at Parry Sound and Woodbridge. (See Map 2).

A test of the effectiveness of the procedure was carried out by regressing the original space series values, the spatial trend values and the residuals from the trend against variables selected to represent the most likely "causes" of spatial trends. (See Map 3 for trend component). If these variables really express causes of regional trends of viability, and if the procedure proposed is effective in separating a trend component from a residual, it necessarily follows that:

$$R_T > R_V > R_R$$

where R_T, R_V and R_R are the multiple correlation coefficients of the regional variables and, respectively, of the trend component (R_T), the original space series value (R_V), and the residuals (R_R).

In other words, the variables will be more related to the trend measures than to the original space series which includes a residual in addition to the trend. Also the original variables will be more related to the original space series values than to the residuals that presumably do not contain trend elements.

In order to calculate the multiple correlation coefficients, a BIMED 29 stepwise multiple regression program was used. This program regresses independent variables, one at a time against a dependent variable until their added explained variance falls below a specified level. Correlation and regression coefficients are output.

The following variables were used: (See Table 4).

- X_1 township population density
- X_2 value of farm sales per farm
- X_3 township agricultural assessment in dollars per acre of land
- X_4 market potential measures
- X_5 improved acreage as a percentage of occupied

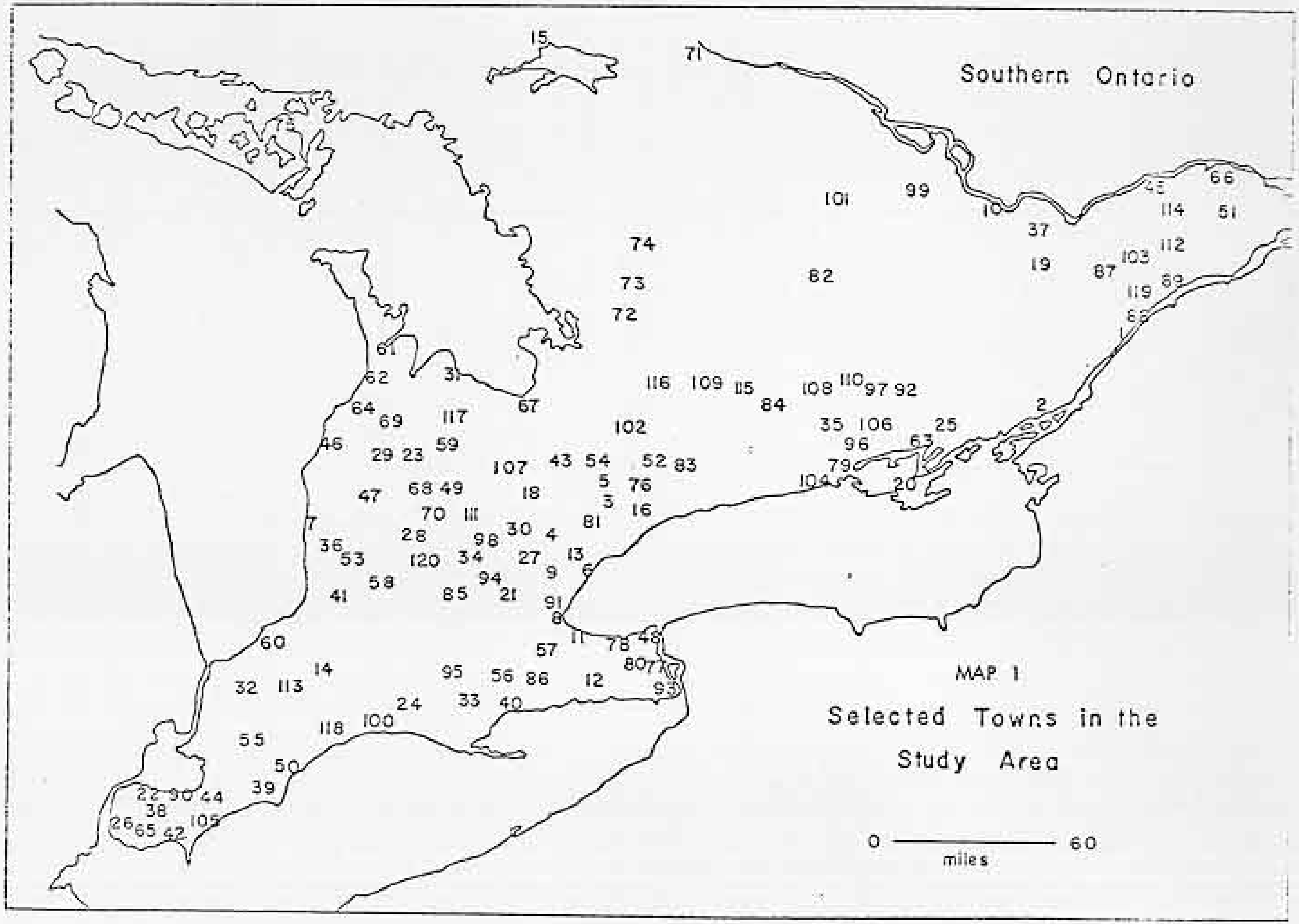
The market potential measures were obtained from the following,

$$MP_i = \sum_{j \neq i} a_i b_i \; / \; d_{ij} \qquad i,j = 1,2 \ldots n$$

Table 1

Selected Towns in the Study Area

1.	Prescott	41.	Exeter	81.	Woodbridge
2.	Gananoque	42.	Kingsville	82.	Bancroft
3.	Richmond Hill	43.	Alliston	83.	Port Perry
4.	Georgetown	44.	Tilbury	84.	Lakefield
5.	Aurora	45.	Rockland	85.	New Hamburg
6.	Port Credit	46.	Kincardine	86.	Hagersville
7.	Goderich	47.	Wingham	87.	Kemptville
8.	Stoney Creek	48.	Niagara	88.	Cardinal
9.	Milton	49.	Mount Forest	89.	Morrisburg
10.	Arnprior	50.	Ridgetown	90.	Belle River
11.	Grimsby	51.	Alexandria	91.	Waterdown
12.	Dunnville	52.	Usbridge	92.	Tweed
13.	Streetsville	53.	Seaforth	93.	Crystal Beach
14.	Strathroy	54.	Bradford	94.	Bridgeport
15.	Sturgeon Falls	55.	Dresden	95.	Norwich
16.	Markham	56.	Waterford	96.	Frankford
17.	Penetanguishene	57.	Caledonia	97.	Madoc
18.	Orangeville	58.	Mitchell	98.	Elora
19.	Carleton Place	59.	Durham	99.	Eganville
20.	Picton	60.	Forest	100.	Port Stanley
21.	Hespeler	61.	Wiarton	101.	Barry's Bay
22.	Tecumseh	62.	Southampton	102.	Sutton
23.	Hanover	63.	Deseronto	103.	Winchester
24.	Aylmer	64.	Port Elgin	104.	Colborne
25.	Napanee	65.	Harrow	105.	Wheatley
26.	Amherstburg	66.	Vankleek Hill	106.	Stirling
27.	Acton	67.	Stayner	107.	Shelburne
28.	Listowel	68.	Harriston	108.	Havelock
29.	Walkerton	69.	Chesley	109.	Fenlon Falls
30.	Fergus	70.	Palmerston	110.	Marmora
31.	Meaford	71.	Mattawa	111.	Arthur
32.	Petrolia	72.	Gravenhurst	112.	Chesterville
33.	Delhi	73.	Bracebridge	113.	Watford
34.	Elmira	74.	Huntsville	114.	Casselman
35.	Campbellford	75.	Little Current	115.	Bobcaygeon
36.	Clinton	76.	Stouffville	116.	Beaverton
37.	Almonte	77.	Chippawa	117.	Markdale
38.	Essex	78.	Beamsville	118.	West Lorne
39.	Blenheim	79.	Brighton	119.	Iroquois
40.	Port Dover	80.	Fonthill	120.	Milverton



MAP 1

Selected Towns in the Study Area

Table 2

LOCATION AND GROWTH OF SMALL TOWNS
IN SOUTHERN ONTARIO (1951-1961)a

Town	1	2	3
RICHMOND HILL	320	260	6600
STREETSVILLE	300	290	3437
STONEY CREEK	303	325	2144
GEORGETOWN	285	285	1983
MARKHAM	332	262	1674
AURORA	320	250	1618
MILTON	290	300	1297
PORT CREDIT	310	295	977
BANCROFT	428	248	960
STOUFFVILLE	335	253	881
GRIMSBY	312	328	856
CHIPPAWA	348	343	848
FONTHILL	328	338	646
BRADFORD	313	266	579
CRYSTAL BEACH	347	360	566
PRESCOTT	588	172	525
BEAMSVILLE	319	329	483
BRIDGEPORT	240	298	471
ALLISTON	292	236	451
ACTON	280	285	439
ORANGEVILLE	275	258	414
CLINTON	175	285	371
WOODBRIDGE	310	260	369
WATERDOWN	290	314	367
DELHI	245	361	362
AYLMER	211	366	351
WHEATLEY	106	427	334
KEMPTVILLE	579	145	317
STRATHROY	167	348	314
PORT PERRY	350	240	314
CALEDONIA	285	338	308
STAYNER	230	251	305
GODERICH	163	273	299
UXBRIDGE	342	240	297
BELLE RIVER	88	410	296
ROCKLAND	598	93	293
ELMIRA	243	288	289
NIAGARA	344	324	287
BLENHEIM	139	407	281

Table 2 (Page 2)

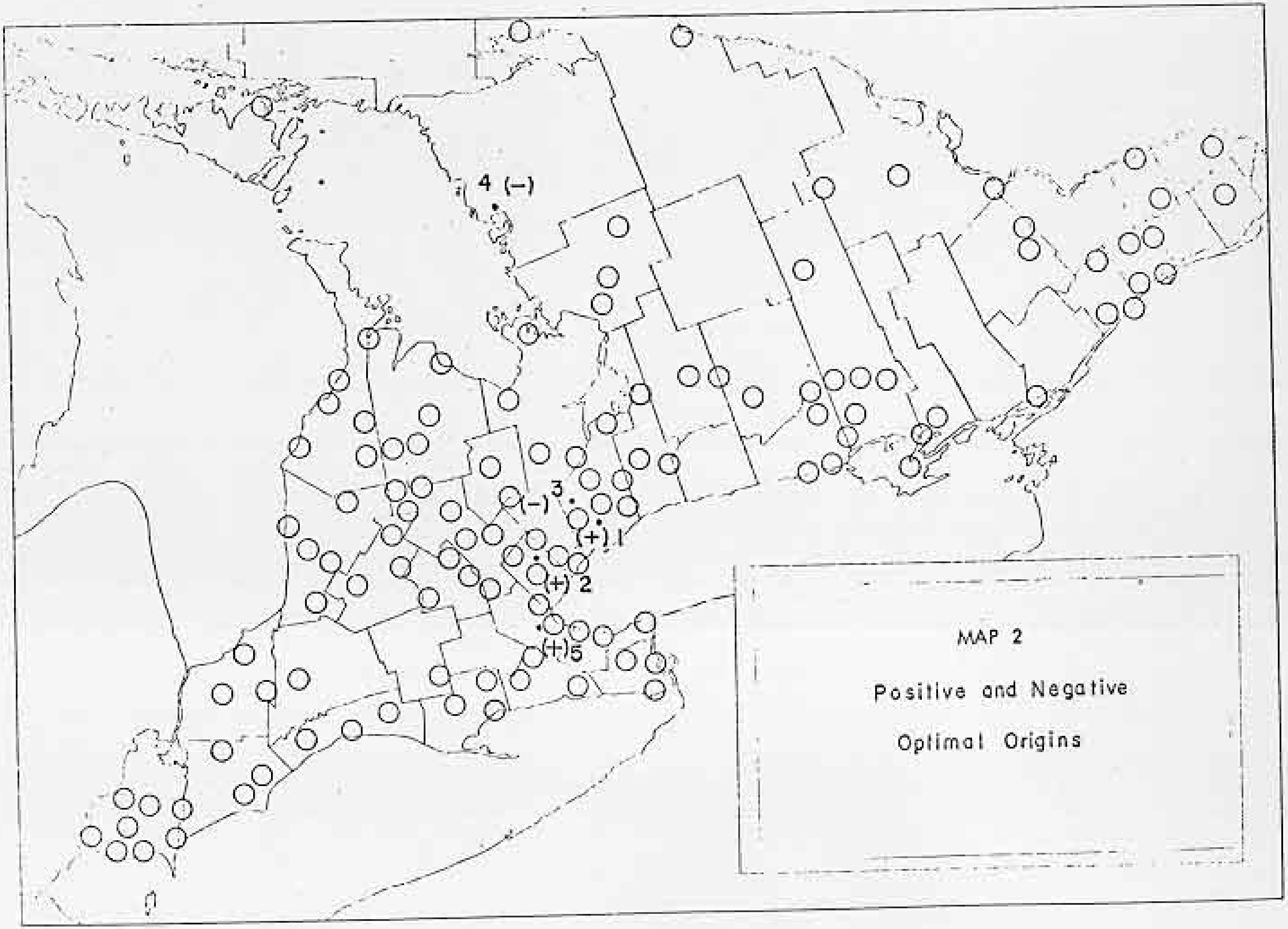
	1	2	3
WATERFORD	260	353	273
STURGEON FALLS	287	22	267
LAKEFIELD	401	208	267
TECUMSEH	075	405	263
NEW HAMBURG	232	310	262
SUTTON	327	220	259
PORT DOVER	266	366	256
ESSEX	080	390	251
ARNPRIOR	527	108	249
HANOVER	211	234	246
MARMORA	441	202	236
ALMONTE	540	128	223
AMHERSTBURG	059	427	223
BRIGHTON	439	244	222
FOREST	140	340	222
COLBORNE	429	249	206
MEAFORD	241	191	206
EXETER	178	311	196
STIRLING	451	219	195
PETROLIA	130	354	194
WINCHESTER	598	137	190
HAGERSVILLE	277	350	188
DURHAM	225	233	185
NORWICH	239	348	183
BARRY'S BAY	439	105	181
DESERONTO	486	228	181
WALKERTON	203	237	180
FRANKFORD	448	230	179
ALEXANDRIA	644	104	178
HARROW	072	434	176
VANKLEEK HILL	640	092	172
HESPELER	260	304	170
EGANVILLE	477	100	168
BEAVERTON	341	208	161
DUNNVILLE	309	355	157
LISTOWEL	215	275	154
NAPANEE	492	223	154
KINGSVILLE	085	435	150
TWEED	467	207	147
MOUNT FOREST	249	292	145
DRESDEN	127	382	143
CHESTERVILLE	606	135	141
MITCHELL	199	301	135
PICTON	480	155	132
FERGUS	254	278	131
TILBURY	108	412	130
ST. MARY'S	202	319	121
GANANOQUE	545	213	114
HAVELOCK	427	207	113
WINGHAM	192	259	105

Table 2 (Page 3)

	1	2	3
CASSELMAN	616	117	103
ARTHUR	245	265	102
ELORA	276	210	102
RIDGETOWN	147	396	101
WIARTON	205	178	94
LITTLE CURRENT	155	057	93
HARRISTON	221	259	92
BRACEBRIDGE	330	150	91
CARDINAL	597	165	91
MADOC	455	200	86
MARKDALE	236	220	82
PENETANGUISHENE	286	176	79
MATTAWA	371	027	78
WATFORD	153	350	77
CAMPVELLFORD	434	219	75
SOUTHAMPTON	189	201	69
SEAFORTH	184	292	65
KINCARDINE	169	231	63
MILVERTON	218	291	53
PORT ELGIN	188	207	47
SHELBURNE	267	242	46
IROQUOIS	601	160	46
FENELON FALLS	370	198	42
WEST LORNE	163	385	38
GRAVENHURST	326	162	24
CHESLEY	215	220	15
CARLETON PLACE	543	135	15
BOBCAYGEON	382	199	2
PALMERSTON	223	266	-12
MORRISBURG	610	152	-21
PORT STANLEY	195	377	-21
HUNTSVILLE	335	122	-30

(a) Note: -30 means that a town's population declined by -3.0 percent between 1951 and 1961.

1,2 Map coordinates
3 Percent population change 1951-1961



MAP 2

Positive and Negative Optimal Origins

Table 3

Town		Location of Optimal Origins Coordinates		Cumulative Explained Variance
		X	Y	
Richmond Hill	(+)	322	260	68.8%
Milton	(+)	296	291	85.1%
Woodbridge	(-)	311	263	88.5%
Parry Sound	(-)	257	123	89.5%
Stoney Creek	(+)	303	326	93.7%

(+) Decline in growth with distance from optimal origin
(-) Increase in growth with distance from optimal origin

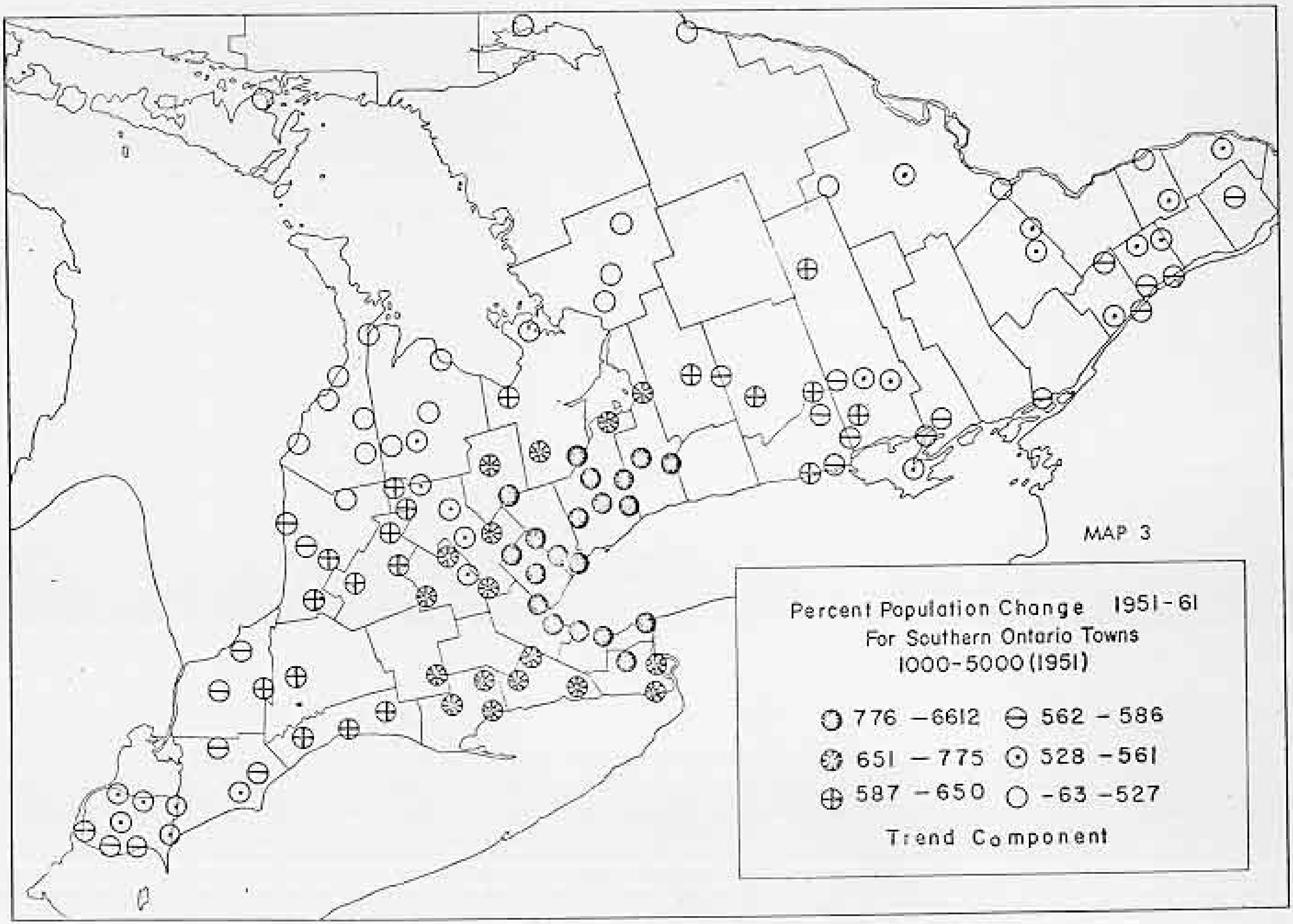

MAP 3
Percent Population Change 1951-61
For Southern Ontario Towns
1000-5000 (1951)
776 - 6612
562 - 586
651 - 775
528 - 561
587 - 650
-63 - 527
Trend Component

Table 4

Town	1	2	3	4	5	6
RICHMOND HILL	6600	2198	34192	272	6242	358
STREETSVILLE	3437	6770	7157	1455	3339	98
STONEY CREEK	2144	5997	21233	905	2138	6
GEORGETOWN	1938	1878	4148	528	999	984
MARKHAM	1674	2198	34192	275	1502	172
AURORA	1618	2809	4192	161	1397	221
MILTON	1297	4728	4148	1483	1160	137
PORT CREDIT	977	6770	7157	1455	1074	-97
BANCROFT	960	458	1224	8	230	730
STOUFFVILLE	881	2809	34192	215	1621	-175
GRIMSBY	856	4348	3160	527	657	199
CHIPPAWA	848	9270	3691	557	356	492
FONTHILL	646	1588	3691	210	433	213
BRADFORD	579	648	2366	68	515	64
CRYSTAL BEACH	566	1653	3691	412	319	247
PRESCOTT	525	672	272	98	165	360
BEAMSVILLE	483	1994	3160	400	457	26
BRIDGEPORT	471	1038	4685	47	176	-71
ALLISTON	451	722	2366	68	401	50
ACTON	439	1878	4148	128	791	-352
ORANGEVILLE	414	325	292	60	456	-42
CLINTON	371	600	656	55	208	163
WOODBRIDGE	369	2997	34192	391	504	-135
WATERDOWN	367	2455	21233	258	640	-273
DELHI	362	1032	1148	135	288	74
AYLMER	351	1027	1142	112	235	116
WHEATLEY	334	1000	927	108	180	154
KEMPTVILLE	317	201	272	33	192	125
STRATHROY	314	682	3793	63	223	91
PORT PERRY	314	533	5128	78	497	-183
CALEDONIA	308	599	663	179	405	-97
STAYNER	305	1008	2366	35	236	69
GODERICH	299	232	656	54	192	107
UXBRIDGE	297	346	5128	100	579	-282
BELLE RIVER	296	877	2541	101	176	120
ROCKLAND	293	710	223	42	157	136
ELMIRA	289	1038	4685	135	345	-56
NIAGARA	287	2323	3160	425	413	-126
BLENHEIM	281	723	927	110	188	93
WATERFORD	273	713	1148	91	322	-49

Table 4 (Page 2)

Town	1	2	3	4	5	6
STURGEON FALLS	267	1160	621	13	32	235
LAKEFIELD	267	1218	1438	136	235	32
TECUMSEH	263	1242	2541	130	202	61
NEW HAMBURG	262	808	4658	104	299	-37
SUTTON	259	1127	5128	166	366	-107
PORT DOVER	256	2655	1148	117	304	-48
ESSEX	251	1138	2541	71	174	77
ARNPRIOR	249	883	919	36	152	97
HANOVER	246	642	810	30	145	78
MARMORA	236	131	1224	5	212	24
ALMONTE	223	713	464	60	187	36
AMHERSTBURG	223	1329	2541	75	202	21
BRIGHTON	222	643	538	38	218	4
FOREST	222	871	1371	69	209	13
COLBORNE	206	462	539	50	253	-47
MEAFORD	206	525	810	22	-19	225
EXETER	196	551	656	63	224	-28
STIRLING	195	334	1224	28	220	-25
PETROLIA	194	513	1371	48	208	-14
WINCHESTER	190	601	201	52	161	29
HAGERSVILLE	188	450	663	83	349	-161
DURHAM	185	201	810	16	179	6
NORWICH	183	735	1339	24	293	-110
BARRY'S BAY	181	173	919	5	156	25
DESERONTO	181	368	1224	28	214	-33
WALKERTON	180	602	403	54	147	33
FRANKFORD	179	2287	1224	71	203	-24
ALEXANDRIA	178	301	205	31	192	-14
HARROW	176	1206	2541	133	175	1
VANKLEEK HILL	172	2457	346	47	191	-19
HESPELER	170	2533	4685	176	341	-271
EGANVILLE	168	138	919	17	170	-2
BEAVERTON	161	461	5128	80	285	-124
DUNNVILLE	157	490	663	120	353	-196
LISTOWEL	154	508	995	58	248	-94
NAPANEE	154	579	289	48	211	-57
KINGSVILLE	150	1636	2541	170	205	-55
TWEED	147	274	1224	24	183	-36
MOUNT FOREST	145	475	1986	37	378	-233
DRESDEN	143	871	927	85	210	-67
CHESTERVILLE	141	601	201	52	161	-20
MITCHELL	135	389	995	58	224	-89
PICTON	132	1384	230	61	187	-55
FERGUS	131	766	1986	55	392	-261
TILBURY	130	750	2541	84	180	-50

Table 4 (Page 3)

Town	1	2	3	4	5	6
ST. MARY'S	121	828	995	68	250	-129
GANANOQUE	114	761	566	78	202	-88
HAVELOCK	113	132	1438	24	221	-108
WINGHAM	105	723	652	111	-35	142
CASSELMAN	103	402	223	26	191	-88
ARTHUR	102	233	1986	31	316	-214
ELORA	102	766	1986	55	172	-70
RIDGETOWN	101	269	927	58	216	-115
WIARTON	94	155	403	44	-31	126
LITTLE CURRENT	93	206	60	4	73	30
HARRISTON	92	518	1986	44	235	-143
BRACEBRIDGE	91	573	282	63	22	69
CARDINAL	91	989	272	78	195	-104
MADOC	86	231	1224	16	182	-96
MARKDALE	82	209	810	24	148	-66
PENETANGUISHENE	79	2144	2366	111	-63	142
MATTAWA	78	568	621	5	76	2
WATFORD	77	535	1371	52	216	-139
CAMPBELLFORD	75	562	538	38	203	-128
SOUTHAMPTON	69	744	403	53	88	-19
SEAFORTH	65	353	656	53	220	-155
KINCARDINE	63	474	403	45	125	-62
MILVERTON	53	458	995	62	266	-216
PORT ELGIN	47	744	403	53	79	-32
SHELBURNE	46	237	292	32	322	-276
IROQUOIS	46	421	201	49	194	-173
FENELON FALLS	42	413	464	54	231	-189
WEST LORNE	38	418	1142	55	221	-183
GRAVENHURST	24	331	282	58	47	-23
CHESLEY	15	225	403	35	133	-118
CARLETON PLACE	15	394	464	29	188	-173
BOBCAYGEON	2	277	464	46	211	-209
PALMERSTON	-12	518	1986	44	280	-247
MORRISBURG	-21	495	201	50	194	-215
PORT STANLEY	-21	1033	1142	60	237	-258
HUNTSVILLE	-30	350	282	32	72	-22

1 Percent population change 1951-1961
2 Township population density
3 Market potential
4 Agricultural assessment
5 Regional component of population change
6 Residuals

where MP_i is the market potential of the ith county, d_{ij} is the distance from the ith county to the jth county, and a_i and b_i are, respectively, disposable income and the population of the ith county.

The results of the analysis are given in Table 5.

Clearly R_T is greater than R_V and R_V is greater than R_R as predicted. Also, the variables selected by the stepwise program for the prediction of the spatial trend (X_4 and X_3) are different from the ones selected for predicting the residuals from the trend (X_1). This implies that the components into which the original space series is split have altogether different characteristics, as it should be expected if the procedure proposed is effective in separating spatial trends and residuals.

Conclusion

The technique discussed in this paper is particularly suited to separating trends that can be expressed as functions of distances from points and lines of reference, and it can be used for the quantitative identification of cores of regions, growth poles and depressed areas. It may provide additional tools for the geographic research that aims at results in the form of precise and testable statements.

Table 5

RESULTS OF THE MULTIPLE REGRESSION AND CORRELATION ANALYSIS

$$y_1 = .053x_4 + .892x_3 + 85.6 \qquad R_T = .726$$

$$y_2 = .051x_4 + 1.05x_3 + 56.1 \qquad R_V = .693$$

$$y_3 = .032x_1 - 46.04 \qquad R_R = .248$$

y_1 = trend component

y_2 = percentage population change

y_3 = residuals from the spatial trend

x_1 = township population density

x_2 = township agricultural assessment

x_3 = market potential

APPENDIX

i) A Program for studying the relationships of Spatially Distributed Variables with Distances from points

ii) General Flowchart

iii) Program

A PROGRAM FOR STUDYING THE RELATIONSHIPS OF

SPATIALLY DISTRIBUTED VARIABLES

WITH DISTANCES FROM POINTS

LANGUAGE: Fortran IV

OUTPUT:

I. Initial Data
  a. The 'X' and 'Y' co-ordinates and the magnitude of the spatially distributed variable

II. Three correlation matrices
  a. An initial 15x15 matrix of correlation coefficients
    i. The location and magnitude of the largest correlation coefficient (absolute value)
    ii. The location and magnitude of the four diagonal correlation coefficients.
  b. A second 10x10 matrix of fine correlation coefficients
    i. The location and magnitude of the largest fine correlation coefficient
    ii. The location and magnitude of the four adjacent diagonal coefficients
  c. A third 8x8 matrix of very fine correlation coefficients
    i. The location and magnitude of the largest very fine correlation coefficient

III. The actual 'X' and 'Y' map co-ordinates of the point of highest correlation

IV. The residuals from the correlation analysis
  a. The residuals are printed eight to a row with the first residual corresponding to the first item of the initial data. The residual corresponding to the second item of data is found in row one, column two; and so on...

V. The explained variance due to the regression of the distance from the initial point of highest correlation to each item and the magnitude of each item.

VI. A second set of three correlation matrices
   a. The output is of the same form as sections two through five. This time, however, the explained variance is due to the regression of the distances from the second point of highest correlation to each item and the magnitude of the residuals for each item from the initial correlation.

VII. A third set of three correlation matrices
   a. Output continues until explained variance of a set of residuals on the previous set of residuals falls below five per cent.

RESTRICTIONS: This program handles up to 500 items and their co-ordinates.

CARD PREPARATION:

First Control Card

cols 1-5 1 if residual print-out is desired, 0 otherwise.
6-10 1 if reciprocal transformation plus one of distance required.
11-15 1 if log transformation of distance required.

Second Control Card

cols 1-3 number of items (variables)
4-40 variable format statement defining the format of the card input (The first three fields should be in A-conversion and are used for the variables' names). The variable format should begin with 1X. The subsequent fields can be numeric or blank.

For example:

a statement (1X, 3A6, 3F9.0) reserves column 1 for machine control, cols 2-19 for the variable name, and cols 20-28 for the 'X' co-ordinate, cols 29-37 for the 'Y' co-ordinate, and cols 38-46 for the magnitude of the variable.

Data cards: The data cards are punched in accordance with the variable format statement contained in the second control card.

DECK-MAKE-UP:

i. $JOB Monitor Control Card
ii. Fortran IV Source Program Cards
iii. $DATA Control Card
iv. First Control Card
v. Second Control Card
vi. Data Cards

Note: This program is designed to be read on input-tape 5 and write on output-tape 6. It can be adjusted to suit the local computer installation be changing the 5 and 6 on source program cards SPADI012 and SPADI013 respectively.

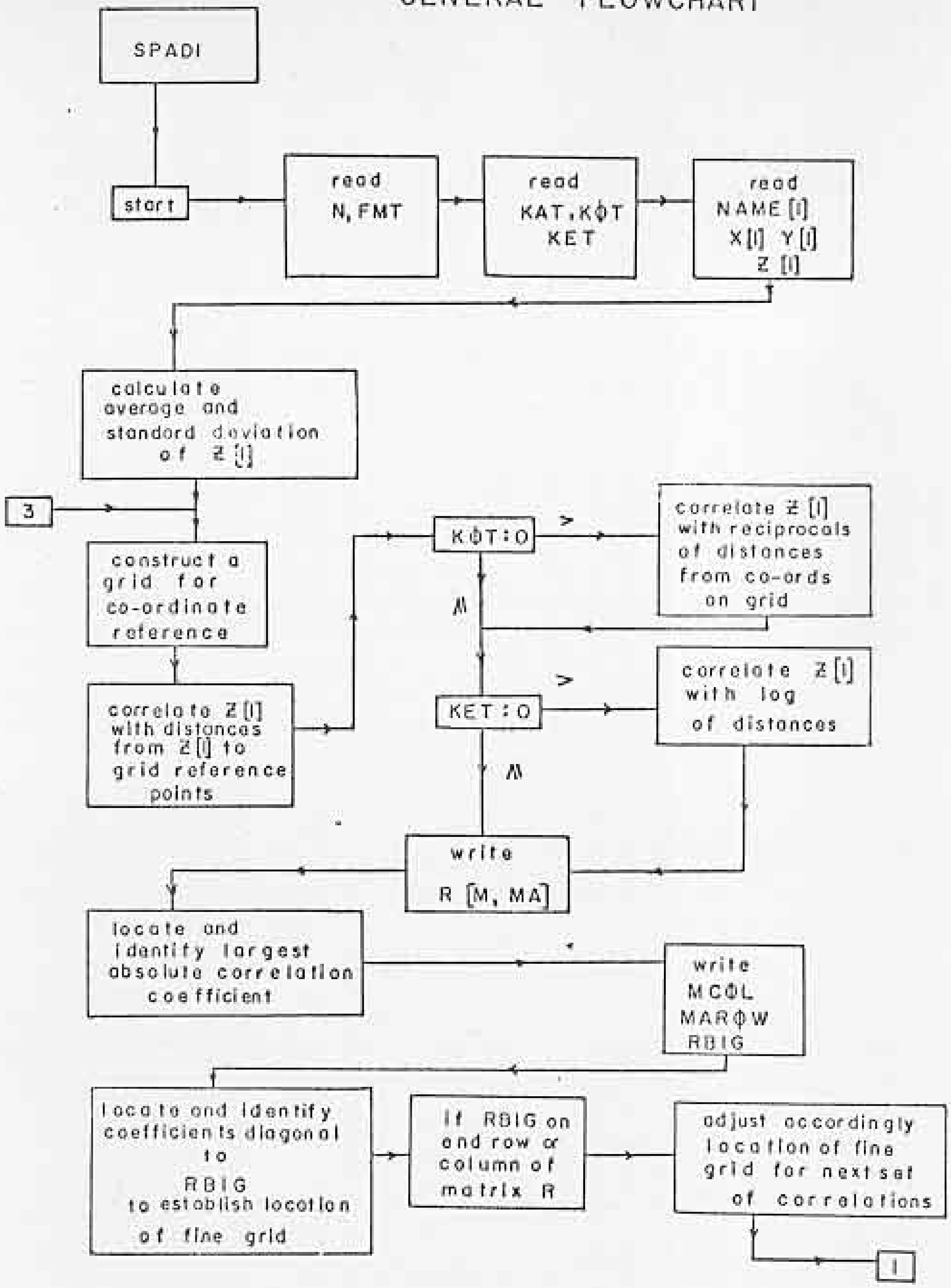
LIST OF SYMBOLS USED IN THE PROGRAM

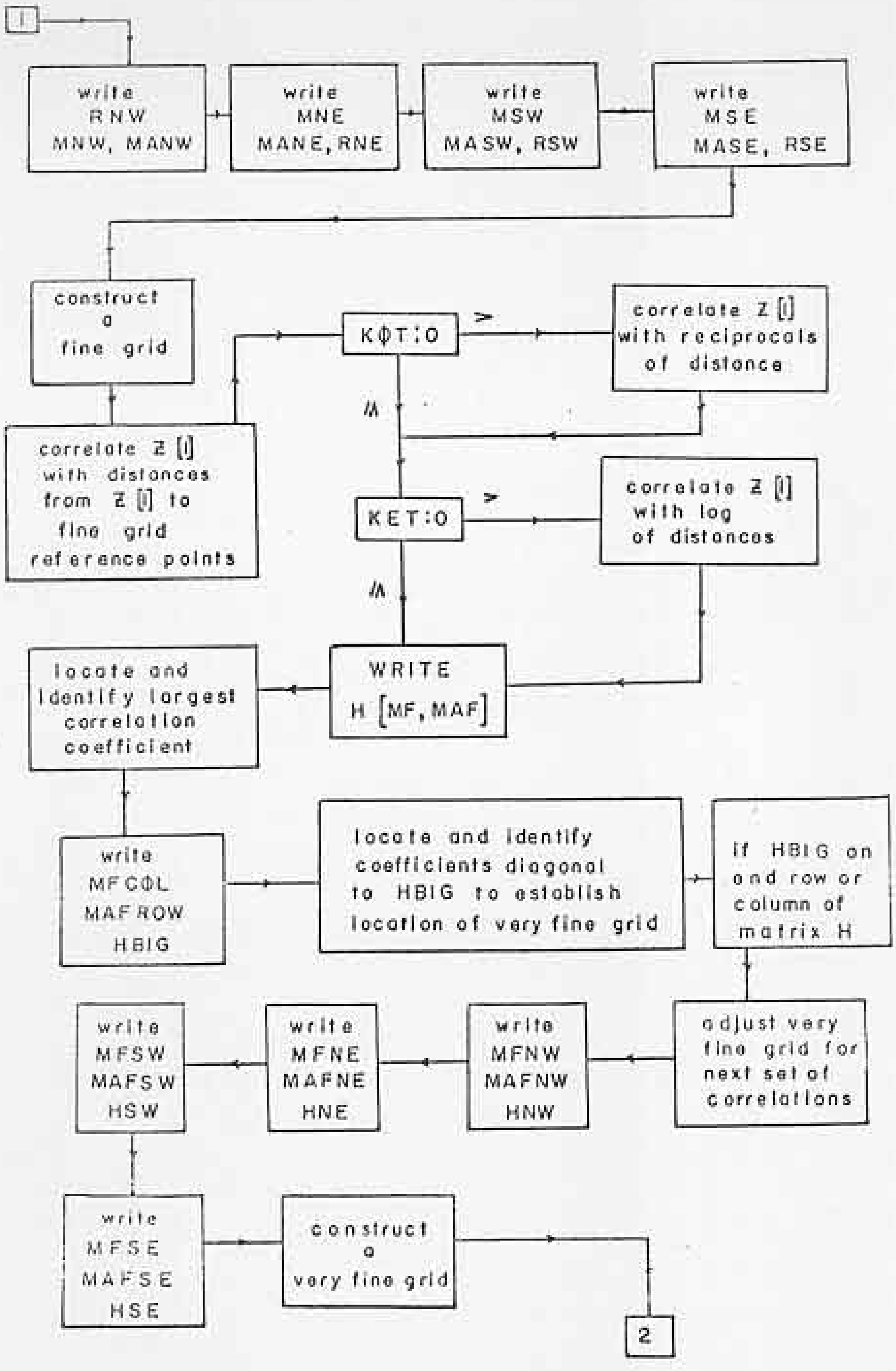
FLOW-CHART

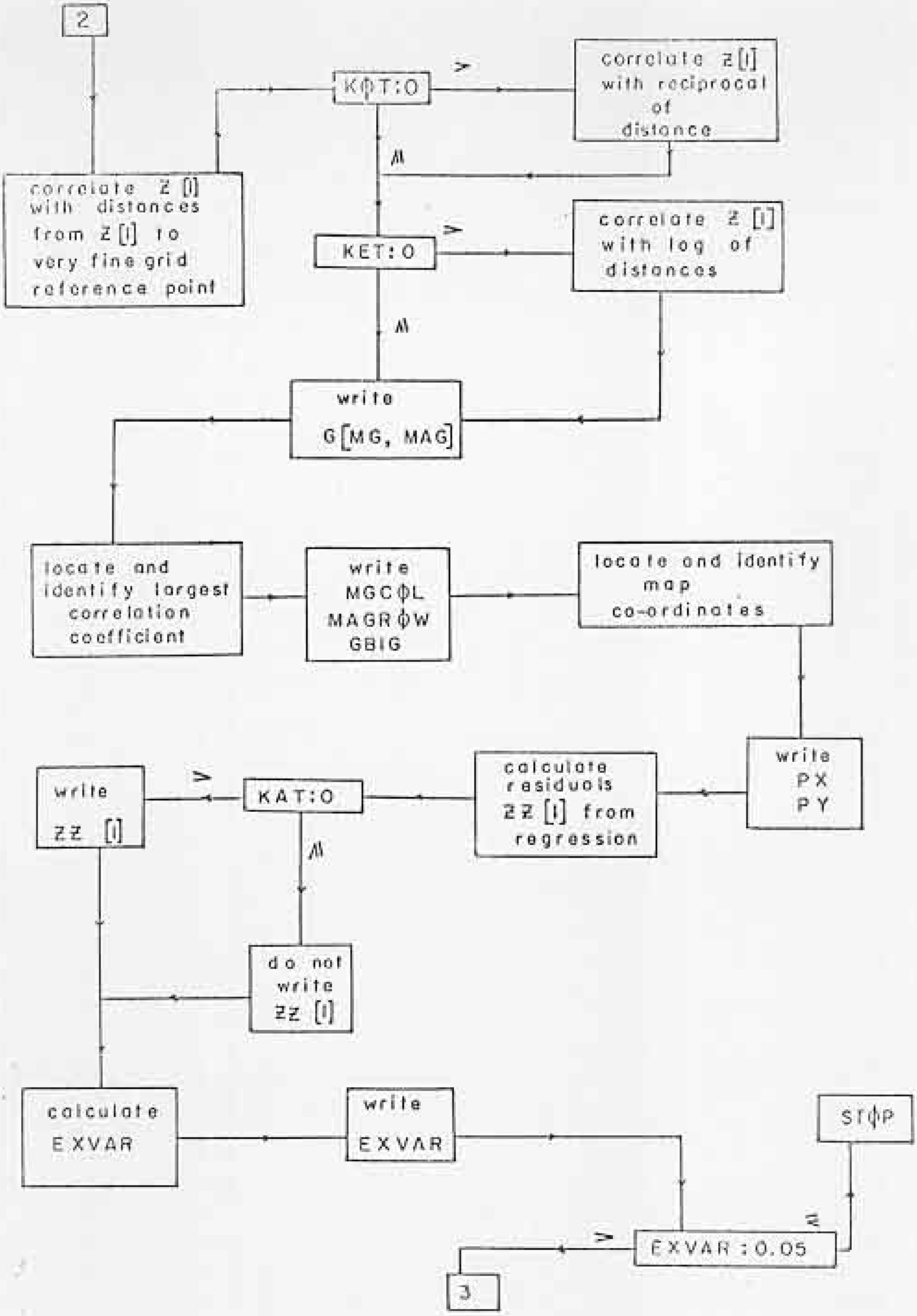
N	Number of spatially distributed variables
FMT	Variable format
KAT	Option for residual print-out
KOT	Option for reciprocal transformation
KET	Option for log transformations
NAME(I)	Name of variables
X(I)	X co-ordinate of each variable
Y(I)	Y co-ordinate of each variable
Z(I)	Value assigned to each variable
R(M,MA)	Initial matrix of correlation, M and MA equal 15
MCOL	The column of the large correlation coefficient
MAROW	The row of the large correlation coefficient
RBIG	The largest correlation coefficient (Absolute value)
MNW	Column to the left of RBIG
MANW	Row above RBIG
RNW	Upper left diagonal correlation coefficient to RBIG
MNE	Column to the right of RBIG
MANE	Row above RBIG
RNE	Upper right diagonal correlation coefficient to RBIG
MSW	Column to the left of RBIG
MASW	Row below RBIG
RSW	Lower left diagonal correlation coefficient to RBIG
MSE	Column to the right of RBIG
MASE	Row below RBIG

RSE	Lower right diagonal correlation coefficient to RBIG
H(MF,MAF)	Matrix of fine correlation coefficients, MF and MAF = 10
MFCOL	The column of the largest fine correlation coefficient
MAFROW	The row of the largest fine correlation coefficient
HBIG	The largest fine correlation coefficient
MFNW	Column to the left of HBIG
MAFNW	Row above HBIG
HNW	Upper left diagonal correlation coefficient to HBIG
MFNE	Column to the right of HBIG
HNE	Upper right diagonal correlation coefficient to HBIG
MFSW	Column to the left of HBIG
MAFSW	Row below HBIG
HSW	Lower left diagonal correlation coefficient to HBIG
MFSE	Column to the right of HBIG
MAFSE	Row below HBIG
HSE	Lower right diagonal correlation coefficient to HBIG
G(MG,MAG)	Matrix of very fine correlation coefficients MG equals MAG equals 8
MGCOL	Column of the largest very fine correlation coefficient
MAGROW	row of the largest very fine correlation coefficient
GBIG	The largest very fine correlation coefficient
PX	The 'X' co-ordinate of the point of highest correlation
PY	The 'Y' co-ordiante of the point of highest correlation
ZZ(I)	The residuals
EXVAR	Explained variance.

GENERAL FLOWCHART




1
write RNW MNW, MANW
write MNE MANE, RNE
write MSW MASW, RSW
write MSE MASE, RSE
construct a fine grid
correlate Z [I] with distances from Z [I] to fine grid reference points
KØT:O
>
correlate Z [I] with reciprocals of distance
≤
KET:O
>
correlate Z [I] with log of distances
≤
WRITE H [MF, MAF]
locate and identify largest correlation coefficient
write MFCØL MAFROW HBIG
locate and identify coefficients diagonal to HBIG to establish location of very fine grid
if HBIG on and row or column of matrix H
adjust very fine grid for next set of correlations
write MFNW MAFNW HNW
write MFNE MAFNE HNE
write MFSW MAFSW HSW
write MFSE MAFSE HSE
construct a very fine grid
2


2
correlate Z [I] with distances from Z [I] to very fine grid reference point
KØT:O
correlate Z [I] with reciprocal of distance
KET:O
correlate Z [I] with log of distances
write G [MG, MAG]
locate and identify largest correlation coefficient
write MGCØL MAGRØW GBIG
locate and identify map co-ordinates
write PX PY
calculate residuals ZZ [I] from regression
KAT:O
write ZZ [I]
do not write ZZ [I]
calculate EXVAR
write EXVAR
EXVAR :0.05
STØP
3

```
   95 CONTINUE
      WRITE(NOUT,3)
    3 FORMAT (1H1,5X,11H REGRESSION,5X,7H NUMBER,I3)
      WRITE (NOUT,6)
    6 FORMAT(16X,25H CORRELATION COEFFICIENTS)
      WRITE (NOUT,7) ((R(M,MA),MA=1,15),M=1,15)
    7 FORMAT(3X,15F8.4)
C
C     SECTION TO LOCATE AND IDENTIFY THE LARGEST CORRELATION COEFFICIENT
C
      RBIG=R(1,1)
      MCOL=1
      MAROW=1
      M=1
      MA=1
  100 IF(ABS(R(M,MA))-ABS(RBIG)) 102,102,101
  101 RBIG=R(M,MA)
      MCOL=M
      MAROW=MA
  102 IF(MA-15) 103,105,105
  103 MA=MA+1
      GO TO 100
  105 IF(M-15) 106,120,120
  106 M=M+1
  107 MA=1
      GO TO 100
  120 CONTINUE
      WRITE (NOUT,8)
    8 FORMAT(10X,7H COLUMN,10X,4H ROW,10X,6H VALUE)
      WRITE (NOUT,9) MCOL,MAROW,RBIG
    9 FORMAT(14X,I3,11X,I3,8X,F8.4)
C
C     SECTION TO LOCATE AND IDENTIFY THE CORRELATION COEFFICIENTS
C     DIAGONAL TO  RBIG
C
      MCEL=MCOL
      MAREW=MAROW
      IF (MCEL-1) 160,160,170
  160 MCEL=2
  170 CONTINUE
      IF(MAREW-1)180,180,190
  180 MAREW=2
  190 CONTINUE
      IF(MCEL-15) 210,200,200
  200 MCEL=14
  210 CONTINUE
      IF(MAREW-15) 230,220,220
  220 MAREW=14
  230 CONTINUE
      RNW=R(MCEL-1,MAREW-1)
      MNW=MCEL-1
      MANW=MAREW-1
      RNE=R(MCEL+1,MAREW-1)
      MNE=MCEL+1
      MANE=MAREW-1
      RSW=R(MCEL-1,MAREW+1)
      MSW=MCEL-1
      MASW=MAREW+1
      RSE=R(MCEL+1,MAREW+1)
      MSE=MCEL+1
      MASE=MAREW+1
  235 CONTINUE
      WRITE(NOUT,10)
   10 FORMAT(7X,4H MNW,7X,5H MANW,10X,4H RNW,7X,4H MNE,7X,5H MANE,10X,
     14H RNE)
      WRITE (NOUT,11) MNW,MANW,RNW,MNE,MANE,RNE
      WRITE (NOUT,12)
   12 FORMAT (7X,4H MSW,7X,5H MASW,10X,4H RSW,7X,4H MSE,7X,5H MASE,10X,
     1 4H RSE)
      WRITE(NOUT,21) MSW,MASW,RSW,MSE,MASE,RSE
C
C     SECTION TO OBTAIN A FINE GRID
C
      VA=(2.0*VA)/10.0
  240 XX=SMAX
      AAA=MNW
      XXX=XX+((AAA-1.0)*VA)
      YY=SMAY
      BBB=MANW
      YYY=YY+((BBB-1.0)*VA)
      MF=1
      MAF=1
C
C     SECTION TO CALCULATE THE CORRELATION OF Z AND THE DISTANCES
C     FROM XXXYYY AND LOCATE IT IN THE MATRIX H
C
  241 DO 242 I=1,N
  242 D(I)=SQRT(((X(I)-XXX)**2)+((Y(I)-YYY)**2))
C
C     RECIPROCAL TRANSFORMATION     KOT
C
      IF (KOT) 249,249,248
  248 DO 981 I=1,N
  981 D(I)=1.0/(D(I)+1.0)
```

```
C
C     PROGRAMME BY E.CASETTI AND R. K. SEMPLE
C     DEPARTMENT OF GEOGRAPHY UNIVERSITY OF TORONTO OCTOBER  1968
C
C     PROGRAMME FOR STUDYING THE RELATIONSHIPS OF SPATIALLY DISTRIBUTED
C     VARIABLES WITH DISTANCES FROM POINTS
      DIMENSION X(500),Y(500),Z(500),R(15,15),NAME(500,3),FMT(6),A(500)
      DIMENSION H(10,10),B(500),G(8,8),C(500),D(500)
      DIMENSION ZZ(500),BAR(500),SLOPE(500)
C
C     SECTION TO READ AND PRINT DATA
      NIN=5
      NOUT=6
      READ (NIN,990) KAT,KOT,KET
  990 FORMAT (3I5)
      READ (NIN,1) N,FMT
    1 FORMAT (I3,6A6)
      WRITE (NOUT,4)
    4 FORMAT (/20X,4HDATA,/)
      READ (NIN,FMT)(((NAME(I,J),J=1,3),X(I),Y(I),Z(I)),I=1,N)
    5 WRITE (NOUT,FMT)(((NAME(I,J),J=1,3),X(I),Y(I),Z(I)),I=1,N)
C
C     SECTION TO CALCULATE THE AVERAGE AND STANDARD DEVIATION OF Z
C
      RA=N
      ZA=0.0
      DO 50 I=1,N
   50 ZA=ZA+Z(I)
      ZA=ZA/RA
      ZVA=0.0
      DO 55 I=1,N
      Z(I)=Z(I)-ZA
   55 ZVA=ZVA+(Z(I)**2)
      ZVA=ZVA/RA
      ZVA=SQRT(ZVA)
C     SECTION TO OBTAIN A GRID
   56 SMAX=AMIN1(X,1,1,N,1)
      BIGX=AMAX1(X,1,1,N,1)
      SMAY=AMIN1(Y,1,1,N,1)
      BIGY=AMAX1(Y,1,1,N,1)
      DX=BIGX-SMAX
      DY=BIGY-SMAY
      IF(DX-DY) 51,51,53
   51 VA=DY/15.0
      GO TO 54
   53 VA=DX/15.0
   54 XX=SMAX
      YY=SMAY
      M=1
      MA=1
C
C     CALCULATE THE CORRELATION OF Z AND THE DISTANCES FROM XXYY AND
C     LOCATE IT IN THE MATRIX R
C
   59 DO 60 I=1,N
   60 A(I)=SQRT(((X(I)-XX)**2)+((Y(I)-YY)**2))
C
C     RECIPROCAL TRANSFORMATION       KOT
C
      IF(KOT) 69,69,68
   68 DO 985 I=1,N
  985 A(I)=1.0/(A(I)+1.0)
   69 CONTINUE
C
C     LOG TRANSFORMATION              KET
C
      IF (KET) 64,64,991
  991 DO 62 I=1,N
   62 A(I)=ALOG(A(I))
   64 CONTINUE
      AA=0.0
      DO 65 I=1,N
   65 AA=AA+A(I)
      AA=AA/RA
      AVA=0.0
      DO 70 I=1,N
      A(I)=A(I)-AA
   70 AVA=AVA+(A(I)**2)
      AVA=AVA/RA
      AVA=SQRT(AVA)
      DO 75 I=1,N
   75 R(M,MA)=R(M,MA)+(A(I)*Z(I))
      R(M,MA)=R(M,MA)/(AVA*ZVA*RA)
      IF (MA-15) 80,85,85
   80 MA=MA+1
      YY=YY+VA
      GO TO 59
   85 IF (M-15) 90,95,95
   90 M=M+1
      MA=1
      XX=XX+VA
      YY=SMAY
      GO TO 59
```

```
   20 FORMAT(7X,5H MFSW,7X,5HMAFSW,10X,4H HSW,7X,5H MFSE,7X,6H MAFSE,
     1  10X,4H HSE)
      WRITE(NOUT,21) MFSW,MAFSW,HSW,MFSE,MAFSE,HSE
   21 FORMAT(7X,I3,7X,I3,7X,F8.4,7X,I3,7X,I3,7X,F8.4)
  505 CONTINUE
C
C     SECTION TO OBTAIN A VERY FINE GRID
C
  509 GA=(2.0*HA)/8.0
      XX=SMAX
      YY=SMAY
      XXX=XX+((AAA-1.0)*VA)
      CCC=MFNW
      XXXX=XXX+((CCC-1.0)*HA)
      YYY=YY+((BBB-1.0)*VA)
      DDD=MAFNW
      YYYY=YYY+((DDD-1.0)*HA)
      MG=1
      MAG=1
C
C     SECTION TO CALCULATE THE CORRELATION OF Z AND THE DISTANCES
C     FROM XXXXYYYY AND LOCATE IT IN THE MATRIX G
C
  541 DO 542 I=1,N
  542 C(I)=SQRT(((X(I)-XXXX)**2)+((Y(I)-YYYY)**2))
C
C     RECIPROCAL TRANSFORMATION        KOT
C
      IF(KOT) 549,549,548
  548 DO 982 I=1,N
  982 C(I)=1.0/(C(I)+1.0)
  549 CONTINUE
C
C     LOG TRANSFORMATION     KET
C
      IF(KET) 547,547,993
  993 DO 546 I=1,N
  546 C(I)=ALOG(C(I))
  547 CONTINUE
      CC=0.0
      DO 543 I=1,N
  543 CC=CC+C(I)
      CC=CC/DA
      CVA=0.0
      DO 545 I=1,N
      C(I)=C(I)-CC
  545 CVA=CVA+(C(I)**2)
      CVA=CVA/DA
      CVA=SQRT(CVA)
      DO 550 I=1,N
  550 G(MG,MAG)=G(MG,MAG)+(C(I)*Z(I))
      G(MG,MAG)=G(MG,MAG)/(CVA*ZVA*DA)
      IF(MAG-8) 555,560,560
  555 MAG=MAG+1
      YYYY=YYYY+GA
      GO TO 541
  560 IF(MG-8) 565,570,570
  565 MG=MG+1
      MAG=1
      XXXX=XXXX+GA
      YYYY=YYY+((DDD-1.0)*HA)
      GO TO 541
  570 CONTINUE
      WRITE (NOUT,22)
   22 FORMAT (56X,35H VERY FINE CORRELATION COEFFICIENTS)
      WRITE(NOUT,23) ((G(MG,MAG),MG=1,8), MAG=1,8)
   23 FORMAT(25X,8F8.4)
C
C     SECTION TO LOCATE AND IDENTIFY THE LARGEST CORRELATION
C     COEFFICIENT IN THE VERY FINE GRID
C
      GRIG=G(1,1)
      MGCOL=1
      MAGROW=1
      MG=1
      MAG=1
  600 IF(ABS(G(MG,MAG))-ABS(GRIG)) 602,602,601
  601 GRIG=G(MG,MAG)
      MGCOL=MG
      MAGROW=MAG
  602 IF(MAG-8) 603,605,605
  603 MAG=MAG+1
      GO TO 600
  605 IF(MG-8) 606,610,610
  606 MG=MG+1
  607 MAG=1
      GO TO 600
  610 CONTINUE
      WRITE(NOUT,24)
   24 FORMAT(10X,7H COLUMN,10X,4H ROW,10X,6H VALUE)
      WRITE(NOUT,25) MGCOL,MAGROW,GRIG
   25 FORMAT (14X,I3,11X,I3,8X,F8.4)
C
C     SECTION TO CALCULATE ACTUAL MAP CO-ORDINATES OF HIGHEST
```

```
  248 CONTINUE
C
C     LOG TRANSFORMATION          KET
C
      IF(KET) 247,247,999
  999 DO 246 I=1,N
  246 B(I)=ALOG(B(I))
  247 CONTINUE
      BB=0.0
      DO 243 I=1,N
  243 BB=BB+B(I)
      BB=BB/DA
      BVA=0.0
      DO 245 I=1,N
      B(I)=B(I)-BB
  245 BVA=BVA+(B(I)**2)
      BVA=BVA/DA
      BVA=SQRT(BVA)
      DO 250 I=1,N
  250 H(MF,MAF)=H(MF,MAF)+(B(I)*Z(I))
      H(MF,MAF)=H(MF,MAF)/(BVA*ZVA*DA)
      IF(MAF-10)255,260,260
  255 MAF=MAF+1
      YYY=YYY+HA
      GO TO 241
  260 IF(MF-10) 265,270,270
  265 MF=MF+1
      MAF=1
      XXX=XXX+HA
      YYY=YYY+((BMF-1.0)*VA)
      GO TO 241
  270 CONTINUE
      WRITE(NOUT,14)
   14 FORMAT (56X,30H FINE CORRELATION COEFFICIENTS)
      WRITE (NOUT,15) ((H(MF,MAF), MF=1,10),MAF=1,10)
   15 FORMAT (10X,10F8.4)
C
C     SECTION TO LOCATE AND IDENTIFY THE LARGEST CORRELATION
C     COEFFICIENT IN THE FINE GRID
      HBIG=H(1,1)
      MFCOL=1
      MAFROW=1
      MF=1
      MAF=1
  300 IF(ABS(H(MF,MAF))-ABS(HBIG)) 302,302,301
  301 HBIG=H(MF,MAF)
      MFCOL=MF
      MAFROW=MAF
  302 IF(MAF-10) 303,305,305
  303 MAF=MAF+1
      GO TO 300
  305 IF(MF-10) 306,310,310
  306 MF=MF+1
  307 MAF=1
      GO TO 300
  310 CONTINUE
      WRITE (NOUT,16)
   16 FORMAT(10X,7H COLUMN,10X,4H ROW,10X,6H VALUE)
      WRITE (NOUT,17) MFCOL,MAFROW, HBIG
   17 FORMAT(14X,I3,11X,I3,8X,F8.4)
C     SECTION TO LOCATE AND IDENTIFY THE CORRELATION COEFFICIENTS
C     DIAGONAL TO HBIG
      MFCEL=MFCOL
      MAFREW=MAFROW
      IF(MFCEL-1)360,360,370
  360 MFCEL=2
  370 CONTINUE
      IF(MAFREW-1)380,380,390
  380 MAFREW=2
  390 CONTINUE
      IF(MFCEL-10) 410,400,400
  400 MFCEL=9
  410 CONTINUE
      IF(MAFREW-10) 430,420,420
  420 MAFREW=9
  430 CONTINUE
      HNW=H(MFCEL-1,MAFREW-1)
      MFNW=MFCEL-1
      MAFNW=MAFREW-1
      HNE=H(MFCEL+1,MAFREW-1)
      MFNE=MFCEL+1
      MAFNE=MAFREW-1
      HSW=H(MFCEL-1,MAFREW+1)
      MFSW=MFCEL-1
      MAFSW=MAFREW+1
      HSE=H(MFCEL+1,MAFREW+1)
      MFSE=MFCEL+1
      MAFSE=MAFREW+1
  440 CONTINUE
      WRITE (NOUT,18)
   18 FORMAT(7X,5H MFNW,7X,6H MAFNW,10X,4H HNW,7X,4HMFNE,7X,6H MAFNE,
     1  10X,4H HNE)
      WRITE(NOUT,21) MFNW,MAFNW,HNW,MFNE,MAFNE,HNE
      WRITE(NOUT,20)
```

```
C     CORRELATION COEFFICIENT
C
      PP=MGCOL
      PQ=MXGROW
      XX=SMAX
      YY=SMAY
      XXX=XX+((AAA-1.0)*VA)
      YYY=YY+((BBB-1.0)*VA)
      XXXX=XXX+((CCC-1.0)*HA)
      YYYY=YYY+((DDD-1.0)*HA)
      PX=XXXX+((PP-1.0)*GA)
      PY=YYYY+((PQ-1.0)*GA)
      WRITE(NOUT,30)
   30 FORMAT(20X,32H MAP CO-ORDS OF HIGH CORRELATION)
      WRITE (NOUT,31)
   31 FORMAT ((7X,14H X CO-ORDINATE,5X,14H Y CO-ORDINATE)
      WRITE (NOUT,32) PX,PY
   32 FORMAT ((10X,2F10.2)
C
C     SECTION TO CALCULATE THE RESIDUALS
C
      RA=N
      DA=0.0
  641 DO 642 I=1,N
  642 D(I)=SQRT(((X(I)-PX)**2)+((Y(I)-PY)**2))
C
C     RECIPROCAL TRANSFORMATIONS   KOT
C
      IF (KOT) 647,647,646
  646 DO 750 I=1,N
  750 D(I)=1.0/(D(I)+1.0)
  647 CONTINUE
C
C     LOG TRANSFORMATION        KET
C
      IF (KET) 701,701,702
  702 DO 703 I=1,N
  703 D(I)=ALOG(D(I))
  701 CONTINUE
      DO 643 I=1,N
  643 DA=DA+D(I)
      DA=DA/RA
      BARA=0.0
      SLOP=0.0
      DO 644 I=1,N
      BAR(I)=(D(I)-DA)*Z(I)
      BARA=BARA+BAR(I)
```

```
      SLOPE(I)=(D(I)-DA)*(D(I)-DA)
  644 SLOP=SLOP+SLOPE(I)
      SLOP=BARA/SLOP
      DO 645 I=1,N
  645 ZZ(I)=Z(I)-SLOP*(D(I)-DA)
      IF (KAT) 649,649,996
  996 WRITE (NOUT,995)
  994 FORMAT (40X,10H RESIDUALS)
      WRITE (NOUT,996) (ZZ(I), I=1,N)
  995 FORMAT ((10X,8F10.0)
  649 CONTINUE
C
C     SECTION TO CALCULATE EXPLAINED VARIANCE
C
      ZAN=0.0
      DO 650 I=1,N
  650 ZAN=ZAN+ZZ(I)
      ZAN=ZAN/RA
      ZVAN=0.0
  651 DO 655 I=1,N
      ZZ(I)=ZZ(I)-ZAN
  655 ZVAN=ZVAN+(ZZ(I)**2)
      ZVAN=ZVAN/RA
  660 EXVAR=GRIG*GRIG
  661 CONTINUE
      WRITE (NOUT,26)
   26 FORMAT (1H ,5X,19H EXPLAINED VARIANCE)
      WRITE (NOUT,27) EXVAR
   27 FORMAT (8X,F10.5)
  662 CONTINUE
  670 IF(EXVAR-0.05) 700,700,680
  680 ZVA=SQRT(ZVAN)
      DO 681 I=1,N
  681 Z(I)=ZZ(I)
      GO TO 56
  700 CONTINUE
      STOP
      END
$DATA
    1    1    0
(2(1X,3A6,F8.0,F9.0,12X,F6.0)
 RICHMOND HILL              320       260        -1403     87
 GEORGETOWN                 285       288        -1375     71
 PORT CREDIT                310       295        -1342     44
 MARKHAM                    332       262        -1217      2
 PRESCOTT                   588       172        -1188     -6
 STONEY CREEK               303       325        -1201    -17
```

AURORA	328	258	-1176	-18
STREETSVILLE	300	290	-1104	-62
CHIPPAWA	348	343	-1189	-67
MILTON	290	300	-1094	-71
GRIMSBY	312	328	-1053	-105
SUNDRIDGE	527	108	-975	-126
WOODBRIDGE	318	260	-1047	-161
ACTON	280	289	-902	-187
ALLISTON	292	236	-900	-203
FONT HILL	328	338	-964	-204
AYLMER	211	366	-805	-256
BANCROFT	428	248	-814	-269
STOUFFVILLE	335	253	-795	-271
STURGEON FALLS	287	22	-729	-276
GODERICH	163	273	-740	-277
GANANOQUE	545	213	-798	-278
ORANGEVILLE	275	258	-725	-296
ESSEX	080	390	-824	-296
DELHI	245	361	-799	-301
DUNNVILLE	309	355	-724	-304
BRADFORD	313	266	-809	-308
PETROLIA	130	354	-716	-326
ELMIRA	243	288	-690	-336
FERGUS	254	278	-705	-339
CLINTON	175	285	-672	-348
AMHERSTBURG	059	427	-692	-351
PICTON	400	155	-677	-355
TECUMSEH	075	455	-685	-355
ST. MARY.S	202	319	-678	-356
NIAGARA	344	326	-776	-363
BEAMSVILLE	319	329	-655	-367
CARDINAL	597	165	-778	-371
CRYSTAL BEACH	347	360	-734	-371
STRATHROY	167	348	-605	-391
HUNTSVILLE	335	122	-642	-406
HESPELER	260	304	-576	-419
EXETER	138	311	-616	-420
ALMONTE	540	128	-587	-422
NAPANEE	482	223	-560	-425
GRAVENHURST	326	167	-607	-426
CARLETON PLACE	543	135	-548	-441
PORT PERRY	350	240	-548	-446
LISTOWEL	215	275	-599	-448
BRACEBRIDGE	330	150	-587	-450
KINGSVILLE	085	435	-650	-451
UXBRIDGE	342	240	-556	-453
WATERDOWN	290	314	-607	-461
CALEDONIA	285	338	-584	-462
HAGERSVILLE	277	358	-548	-474
HANOVER	211	234	-560	-488
KEMPTVILLE	579	145	-717	-496
PORT DOVER	266	356	-468	-499
STIRLING	451	219	-509	-504
LAKEFIELD	401	208	-458	-515
WALKERTON	203	237	-388	-518
BRIGHTON	439	244	-476	-531
CAMPBELLFORD	434	219	-437	-547
CHESTERVILLE	606	135	-470	-561
MITCHELL	199	301	-460	-566
WINGHAM	192	259	-433	-576
MEAFORD	241	151	-399	-578
BRIDGEPORT	240	275	-455	-586
MARMORA	441	203	-478	-588
MORRISBURG	610	158	-486	-588
IROQUOIS	601	150	-478	-591
FRANKFORD	448	230	-410	-593
TWEED	467	207	-396	-599
PALMERSTON	223	266	-410	-599
HARROW	072	434	-410	-603
TILBURY	108	413	-380	-604
PENETANGUISHENE	286	176	-345	-607
WATERFORD	260	353	-337	-619
SUTTON	327	229	-456	-624
BELLE RIVER	88	410	-318	-632
NEW HAMBURG	232	310	-341	-638
ROCKLAND	598	93	-299	-645
MADOC	455	200	-355	-658
PORT ELGIN	186	207	-399	-664
MATTAWA	371	027	-303	-674
ALEXANDRIA	644	104	-277	-676
WINCHESTER	598	137	-343	-676
DESERONTO	486	228	-269	-684
DRESDEN	127	382	-278	-690
BLENHEIM	139	407	-318	-698
WIARTON	205	178	-253	-706
WHEATLEY	106	427	-307	-712
KINCARDINE	169	231	-266	-713
DURHAM	225	233	-225	-739
FOREST	140	340	-272	-741
BEAVERTON	341	208	-260	-747
RIDGETOWN	147	396	-208	-757
SOUTHAMPTON	189	201	-281	-758
STAYNER	230	251	-154	-759
SEAFORTH	184	292	-199	-761
NORWICH	239	348	-148	-772
PORT STANLEY	195	377	-268	-776
MOUNT FOREST	249	218	-193	-780
COLBORNE	479	243	-206	-781
LITTLE CURRENT	155	057	-240	-791
ELORA	276	210	-189	-792
WEST LORNE	163	385	-238	-792
VANKLEEK HILL	640	097	-141	-807
SHELBURNE	267	242	-251	-825
HARRISTON	281	259	-193	-826
EGANVILLE	477	100	-112	-846
FENLON FALLS	370	198	-136	-847
HAVELOCK	427	207	-110	-856
MARKDALE	236	220	-130	-876
WATFORD	153	350	-75	-879
CHESLEY	215	220	-132	-882
MILVERTON	218	291	-109	-883
BARRY.S BAY	439	105	-49	-884
ARTHUR	245	265	-106	-887
BOBCAYGEON	382	199	-47	-897
CASSELMAN	618	117	-28	-951

MICHIGAN INTER-UNIVERSITY

COMMUNITY OF MATHEMATICAL GEOGRAPHERS

DISCUSSION PAPER

Number 12

THE PHILOSOPHY OF MAPS

W. Bunge, R. Guyot, A. Karlin, R. Martin,
W. Pattison, W. Tobler, S. Toulmin, and W. Warntz

Wayne State University

June 1968

Dr. John D. Nystuen, Editor
Department of Geography
University of Michigan
Ann Arbor, Michigan

Discussion Paper Series

1. Arthur Getis, "Temporal Land Use Pattern Analysis with the Use of Nearest Neighbor and Quadrat Methods," July, 1963.

2. Marc Anderson, "A Working Bibliography of Mathematical Geography," September, 1963.

3. William Bunge, "Patterns of Location," February, 1964.

4. Michael F. Dacey, "Imperfections in the Uniform Plane," June, 1964.

5. Robert S. Yuill, "A Simulation Study of Barrier Effects in Spatial Diffusion Problems," April, 1965.

6. William Warntz, "A Note on Surfaces and Paths and Applications to Geographical Problems," May, 1965.

7. Stig Nordbeck, "The Law of Allometric Growth," June, 1965.

8. W. R. Tobler, "Numerical Map Generalization," and "Notes on the Analysis of Geographical Distributions," January, 1966.

9. Peter R. Gould, "On Mental Maps," September, 1966.

10. John D. Nystuen, "Effects of Boundary Shape and the Concept of Local Convexity."

    Julian Perkal, "On the Length of Empirical Curves."

    Julian Perkal, "An Attempt at Objective Generalization." December, 1966.

11. E. Casetti and R. K. Semple, "A Method for the Stepwise Separation of Spatial Trends," April, 1968.

TABLE OF CONTENTS

THE PHILOSOPHY OF MAPS

INTRODUCTION

The tone of the papers presented is extremely disjoint. The only commonality is the subject - philosophical questions arising from maps - and in some cases the authors did not see their comments as especially philosophical. The editor had to underscore the philosophical unity that impressed him and led to the selection of material. The editor did not feel that a review article would do since much of the material is original, and perhaps more deeply, who is to say that the various points of view might not all be fruitful and should be pursued simultaneously? Commalities cannot be forced by the editor dishonestly, that is, before he sees them.

This is a collection of papers written by seven geographers and a philosopher in which from the study of maps philosophical questions arise. Previous philosophical questions of geographers were those of the theory of knowledge and the philosophers were read from a great distance. Kurt Schaefer's long conversations with Gustav Bergmann, which questioned the epistomological exceptionalism of the science of geography, was the initial and the best representation of this dialogue.

But the tables have turned for now the geographers are raising questions to the philosophers. We no longer feel ourselves to be humble students before sages but slightly annoyed critics of philosophy's neglect of not just geography but all the visible, literal fields of knowledge including graphics and geometry. Now short of pugnacity, but past persistance, we return to philosophy in a new wave of interest with questions that we suspect may provide a challenge for philosophy today. We have changed the subject, too, from ground where we were the uninitiate--philosophy of science (and how we used to pour over Cohen and Nagel) to ground that is clearly our home territory--The Map. The older geographers, those that were horrified at our initial furious attack on maps as inferior to mathematical functionals, had substantial position on their side. But they were and are so religious about their commitment to the map--complete with religious persecutions for those that did not genuflect before the fundamentalist map thumpers--that they practically compelled our revolt. We were provoked. Why did Hartshorne's excellent universal methodology ignore maps? Why did the cartographers ignore methodology? All those Leroy Pens and Zipatone and never a philosophical question. What could this mean? Certainly nothing flattering.

The first paper is a short one by Waldo Tobler tying the geographer's traditional concern with "projections" into the main stream of "transformations." William Pattison is next. Pattison brought his notion of the literalness of geographic concepts to the Dedham Conference some years back which William Warntz and I also attended. This started both Warntz and myself to thinking and Warntz came up with the Venn diagram application at the conference itself, which he has developed over the years into the third paper. The fourth contribution is the editor's thinking about spatial prediction. Stephen Toulmin's excerpted article follows. Toulmin is the only philosopher who is quoted. Few philosophers have initiated an interest in geography, Kant and a handful, and no one other than Toulmin has shown an interest in maps. After Toulmin, comes Roland Martin, the interesting geographic mystic and then still another report by the editor. The final two papers are by Andrew Karlin and Richard Guyot, students. All bibliographic references are collected at the end. This is the complete collection of philosophical material on maps that has come to my attention.

(W. Bunge)

X X X X X X X X X X X X X X X

What, if anything, is ultimately invarient in science? The impact of Waldo Tobler's paper is two-fold. To help geographers used to thinking in our ancient terms of "map projections" to broaden their scope, and to imply a question, "What portions of geography are not transformations?" Such a question leaves a classically trained geographer shaken. What a wierd world for geographers raised in a "factual" geography suffocated in parameters such as the highest point in the State of Wisconsin. What remains invariant that is not trivial? Perhaps philosophers will find this question dull after their earlier experiences with physics but if even geography is transforming itself into transformations, is this the ultimate fate of all scientific knowledge? Everything a mapping? The breadth of mapping might then expand all the philosophical discussion to follow, such as spatial prediction of mapable temporal phenomena.

Transformations

Waldo Tobler

A map projection can be considered a transformation applied to spatial point coordinates. The emphasis in the present work is on this class of geometrical transformations. It would be misleading to imply, by omission, that there are no additional types of transformations of interest to geographers. Only two elementary examples are presented here, but these suffice to introduce briefly some additional transformations.

As a first example consider the entire cartographic process. Geographical information is supplied to the cartographer and he then transforms this into a geographical map. Clearly this can be considered a signal processing operation, much like radio, and can be represented diagrammatically as

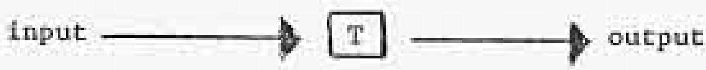
The basic assumption is that the system is of high fidelity; that is, the map user expects that he can use the geographical map as an adequate representation of the input description of the environment. Alternately stated, the inverse of the cartographic transformation is map reading. In practice a map is never a complete representation of the environment. The electrical engineer would inquire about the transfer function characteristics of the signal processing system (the cartographer). Many of the detailed steps in the cartographic process can also be considered within this general framework. Map projection conversions are a particular case. Another step occurs in map generalization. This is worth examination in some detail.

Let G denote an n by n matrix of topographical elevations taken at equally spaced geographical intervals. This matrix can be contoured using bivariate interpolation to produce a topographical map. Let S be an n by n smoothing matrix with entries as follows:

$$S = \begin{matrix} 3/4 & 1/4 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 & & 0 & 0 & 0 \\ 0 & 1/4 & 1/2 & 1/4 & 0 & 0 & & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/2 & 1/4 & 0 & & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 & & 0 & 0 & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & & 1/2 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1/4 & 3/4 \end{matrix}$$

A generalization of the topographical map can be obtained by contouring the result of G pre- and post-multiplied by S, i.e., $G^* = S\,G\,S$. Diagrammatically:

which is equivalent to



The points of interest are that the inverse $(G = S^{-1}\ G^*\ S^{-1})$ exists, and that the transformation can be put in series $(G^{**} = S^2GS^2$, etc.). In other words, generalized maps can be further generalized, and can be ungeneralized.

Hagerstrand's theoretical model of the geographical spread of phenomena provides a second example. Here the transformation consists of transistion probabilities which map one state matrix, to use systems terminology, onto another state matrix. The transformation in this instance is stochastic and generally does not have an inverse. The point is that this example, like that preceeding it, does not disturb the geometrical relationships, as do map projections.

X X X X X XXX X X X X X X X

If reality is not real, what is? Geographers have a sturdy tradition of climbing mountains and exploring distant places. William Pattison somehow strikes that mood. By George, when we geographers say something we really mean it! The emotional impression is the exact opposite of Tobler's lead article. What could be realer than geography?

The Literalness of Spatial Thought

William Pattison

It is usual for practically everyone to organize his experiences spatially, by thinking of events according to their "whereness," or position. By structuring our experiences in this fashion we create for ourselves what may be called a spatial frame of reference. In contemplating the framework, one notices first the remarkable fact that all of its component concepts, starting with position, are to be taken literally. According to strictly spatial rules, a position is a location in space, not a way of designating a job assignment. Similarly, left and right are truly physical orientations, not political tendencies; and up and down are directions in which one can physically move or point, rather than figures of speech for use in describing pretended movement on "the ladder of success," for example. Likewise, distances are to be understood in terms of such units as inches and miles; which is to say that the concept of social distance, for example--valuable though it may be in appropriate frames of reference--has no validity in the context now under discussion.

The literalism of spatial thinking is difficult for many people, especially those in academic life, to acknowledge as legitimate, habituated as they are to using spatial terms almost exclusively in a figurative or metaphorical manner. Academic discourse, perhaps most notably in the humanities and the social sciences, relies heavily upon references to such motions as rising and falling, arriving and departing, emerging and penetrating, without "really meaning" what is said.

The discoursers speak of approaching a brink, of crossing a boundary, of fixing an aim, of reserving an area, or of locating an interface, while they remain entirely in a world of metaphorical usage. They often require almost forcible reminding that there is a prototype world of sense perception in which all of these expressions have a literal signification--and to which the figurative usages themselves must be ultimately referred, for clarification of meanings.

Among essayists and other representatives of belles lettres, acceptance of spatial thinking is not uncommon. As keen an appreciation of its distinctiveness as could be wished for has appeared in the writings of the British author Freya Stark:

> Geologists may see their years in terms of time; musicians in some echo that dies in the air; the poet strains language beyond the bounds of telling;...but to us, who delight in maps, the idea of life inclines to be spatial--we see it moving point to point, like a road if we are inclined to attribute its shaping to men, like a river, if we have more feeling for the unexpectedness of nature.

To be sure, even here a strong suggestion of metaphorical intent is apparent, but happily the author is unambiguous in her allusion to maps, for the map is the expression par excellance of the kind of literal thought that we are seeking to understand.

X X X X X X X X X X X X X X X X X

If even point-set has literal meaning, how abstract are abstractions? Warntz, as usual, is dramatic. As Warntz is more than aware, if distances are measured in cost terms or some other unit, the earth is far from spherical, and therefore the variance of locations is considerable. The antipodal point of England for Mackinder in the proper geodesic coordinate system was not New Zealand, hardly anywhere near "half way around the world" but the heartland of Eurasia. Nordbeck on his visit at Wayne mentioned some work on locational variance that he and Hagerstrand were working on. The globe is as lumpy as a potato. It is not anywhere close to being round nor did Magellan go around it, even conceptually.

William Pattison's literalness bears fruit. What is considered more abstract by logicians-mathematicians or more fundamental to their thinking than point-set? But Warntz makes his own case.

Some Elementary and Literal Notions About Geographical Regionalization and Extended Venn Diagrams

William Warntz

Part I - Geographical Regionalization

All that we have to do, apparently, is to be willing to admit the feasibility of recognizing places on the earth's surface as though they could be elements in sets capable of assignment to variously defined classes. We wish to consider how this is related to regionalization as the geographer recognizes it.

Having done this, we are then in a position to employ graphical methods (cartographical for geographers) to show the results of our regionalizations as we perform the various operations possible such as union, intersection, and so on upon the sets involved. Maps showing regional classification can thus be regarded as logic diagrams. Mapping of sets is a general mathematical concept. Geographical mapping is merely a special case of this.

Consider the sets of places on the earth's surface as _necessarily_ having spatial properties (location--absolute and relative) and _capable_ of having non-spatial properties (climatic, geomorphologic, economic, political, linguistic, _ad infinitum_). How do we transform the former into the latter in intelligent, systematic, meaningful and useful ways?

First let us note (following Gardner, 1958) that a logic diagram may be "a two-dimensional geometric figure with spatial relations that are isomorphic with the structure of a logical statement."

Gardner adds that,

> "These spatial relations are _usually_ (italics mine) of a topologic character, which is not surprising in view of the fact that logic relations are primitive relations underlying all deductive reasoning and topological properties are, in a sense, the most fundamental properties of spatial structures. Logic diagrams stand in the same relation to logical algebras as the graphs of curves (maps of areas) stand in relation to their algebraic formulas; they are simply other ways of symbolizing the same basic structure."

What happens if we say that sets consist of places on the earth's surface as elements and that all geometric properties and locational relationships, however transformed, are to be recognized in our manipulations? We wish to go explicitly from diagrams for non-spatial sets to geographical maps for spatial sets. The inverse of this, i.e., maps to logic diagrams, was recognized very early by the Harvard-educated (Class of 1859) logician and philosopher, Charles Sanders Peirce (1839-1914). At a time when specialization was taking command, Peirce remained universal. In an empirical age and place, he was a rationalist and a theorist. He stressed the role of theory building and how data are to be defined by where they fit into a theoretical structure. Peirce has been called the unexpected, exotic culmination of a physical geography tradition in the United States. I suggest that he be regarded as the inevitable culmination of a physical geography tradition which he not only maintained but also extended beyond his fellows.

At a time when the Venn graphical method of logic diagrams was being perfected for non-spatial sets, Peirce was developing graphical methods for analyzing in detail the structure of all deductive reasoning; breaking structures into their elements and giving each of the elements its simplest, most iconic geometrical representation possible. He considered that in this way the mind could "see" the logical structure in "a fashion analogous to seeing a geographical area when you look at a map."

He wrote that maps put before us pictures of thought. He inferred that maps could be experimented upon in a manner similar to the way a scientist experiments with a structure in nature. By altering mappings in various ways we could discover new properties of the structure not previously suspected. A map may be a device for invention and discovery of new truths as well as an instrument for proving, preserving, and recording old ones.

Alas! Peirce, the philosopher, was not influential among succeeding generations of geographers, or philosophers for that matter, and we find ourselves, at present, having to argue deductively from mathematical-logical mapping to geographical mapping rather than inductively the other way around. Much time has been wasted. The Schaefer-Bergmann (1953) cooperation has marked the rebirth.

Dividing the earth's surface, or parts of it, into meaningful regions of various kinds is a valid and useful endeavor. Many geographers have turned their attention to this problem and as Bunge (1962) has shown, geographers independently have rediscovered the entire logic of classification. This, as Bunge remarks, has been "no mean intellectual feat." Bunge has provided a table of vocabulary equivalences for the terminology in regionalization and in general classification.

It is true that every square inch on the earth's surface differs in its non-spatial properties in some way from every other square inch. No two places, however small, are exactly alike. To know about the phenomena in such minute detail is virtually impossible. Moreover, such unorganized detail would have limited use, and, as in science generally, a method of grouping similar elements, in this case places, is, for many purposes, essential. The classification of the earth's surface lies at the heart of regional geography. The number and characteristics of the geographical classes and regions, so defined, obviously, and of necessity, depend upon the nature and the purpose of the classificatory scheme employed, and thus may be expected to differ from purpose to purpose. In this respect the essence of regional classification is like the problem of classification in general based on the principle of dichotomy in all academic disciplines. There is, however, one factor always present in regional classifications and that is earth location. It is this factor and the concepts related to it, such as concentrated, dispersed, clustered, evenly distributed, contiguous, etc., that are essentially geographic. It can be shown that the terminology and method developed in science generally for classification are relatable precisely by vocabulary equivalences to those developed independently in geography through the years when that factor that is peculiarly geographical is added, namely location of the earth's surface. This view as noted above has been stated explicitly by Bunge and has been reviewed by Greeg who also acknowledged the interests of Gilmour, Cline, Simpson, Hettner, Hartshorne, et al. Golledge, Amadeo, and Haggett have commented meaningfully upon the problem.

Whereas regionalization is akin to classification, it is a mistake to think that anyone versed in classification procedures could thereby construct useful geographical regions. In addition to handling the important geographical factor of location successfully, the regionalizer must make appropriate decisions as to what significant non-spatial features or differentiating characteristics are to be included as criteria for the classification.

An understanding of non-spatial processes is imperative to establish superior classifications for regionalization purposes, but the study of non-spatial processes for their own sake "rather than for the ultimate classification (uniform regions) would appear to be outside the province of geography." Others are better equipped to handle these non-spatial processes than geographers.

All regionalizations like all classifications are arbitrary. No perfectly rational classification of phenomena exists independent of the use to which the classes are to be put. Similarly, there can be no perfect regionalization independent of purpose and a valuable regional division for one purpose may be quite inadequate for another. Geographers now generally agree that a region is an intellectual construct and that the concept of a "region as a concrete unit object" is indefensible. Bunge puts it well when he states that in order to produce "areal classifications of identical sorts no matter what differentiating characteristics are considered, it is necessary that there exist a perfect areal correlation among all phenomena. This condition is not met on the earth's surface. As an alternative it is possible, but absurd, to insist on some one arbitrary areal classification as sacred and immutable."

A technical problem of any good classification (regionalization) is to establish the class limits (regional boundaries) in such a way that the differences among classes (regions) and the differences within classes (regions) are discernible and can be manipulated to desired ends.

From just these few considerations alone from among the number that Bunge has presented we find it easy to accept the idea that it is efficient and indeed necessary to consider "uniform" geographical regionalization as one kind of classification problem.

It seems possible now to go beyond this important beginning to additional concepts that may prove useful in our attempts to understand the essence of regionalization. In particular, it might be of benefit to make explicit the implicit link of geographical regionalization to Boolean algebra and Venn diagrams.

Boolean algebra is the algebra of logic, an abstract mathematical structure appearing in three different forms. It is the algebra of sets and to this we should specifically turn attention in our research.

To the mathematician or philosopher sets are considered to be just collections of objects. The objects which are the elements of the set may be material objects or purely conceptual intangible "objects" (or ideas). In fact, the basic theory does not need to specify. It is enough that a set is a collection of elements. The patterns of relations among sets is the concern. The most important idea involved is that any element in the "universe" is either a member of a given set or it is not. Now, this idea does not seem at first to hold much interest, but it turns out to be fruitful material for the human mind which generates an astonishing array of concepts and techniques in the presence of this basic idea.

Thus, we see that set manipulation, like classification is a dichotomous procedure. And since all information can be coded and transmitted by employing a binary system, and since electrical circuits can be described in these terms, and since electronic computers operate on this basis, the possibilities for elaborate and complex data manipulation systems are great.

Scientists employing various set theory concepts in their various disciplines are interested in the patterns of relations among sets, but they also are inescapably concerned with the actual composition of sets, that is, with the classification of the "objects" of investigation into operational categories based on useful properties for the problem at hand. Thus, classification problem occurs at some stage in the development of every science and the recognition of classes is, in fact, crucial to its completeness.

Logic is not concerned with specific examples or unique individuals because generalization about individuality is per se a contradiction. Logic deals with members of a class. Scientists in their various disciplines may extend logical reasoning to individuals as members of a class, however.

As noted at the beginning of this paper, classification in geography is intimately linked with the concept of regions. Places classified according to their properties, when location is always included as one of the properties produce geographical regions varying according to the kinds of criteria and the limits employed including location.

To supplement the symbolic statements of the patterns of relations among sets when various operations are performed on them, graphical procedures are employed. The special diagrams used are called Venn diagrams after John Venn, a British mathematician of the nineteenth century who refined an earlier procedure of Euler. It is part of our purpose here to extend the use of such diagrams to the mapping of geographical regions by making use of properties already inherent in Venn diagrams but as yet unutilized. We preface the following demonstration by noting again that in conventional set theory the graphical portrayal recognized the topological orderings among sets and portrays these. We shall attempt to utilize not only topological properties but various other geometric properties as well regardless of the statement by C. I. Lewis (1918) that for diagrammatic purposes only, the elements of the algebra of sets can be applied "to spatial entities such as continuous or discontinuous segments of a line, or to continuous or discontinuous regions in a plane" (italics mine). "The applications to regions in a plane gives the more workable diagrams, for obvious reasons. And since it is only (italics mine) for diagrammatic purposes that the application of the algebra of sets to spatial entities has any importance, we shall confine our attention to regions in a plane."

Not wishing to belabor a point, we nevertheless explicitly reject the notion that the spatial entities must be regarded, as above, only as analogies. We intend to apply spatial properties literally to real spatial distributions on the earth's surface when this may be done appropriately.

The conventional interpretation of the algebra of sets applies to classes taken "in extension." That is to say, a given letter symbol, say A, signifies not a class-concept, but rather the aggregate of all the objects, i.e., the entire membership, denoted by the class-concept. To say that A = B means not that the concept of class A may be regarded as a synonym for the concept of the class B, but that the classes A and B must consist of exactly the same members. They, therefore, are regarded as having the same extension.

Conventionally, it is also noted that classes may be divided into logical but not physical parts. We shall consider extension in a physical sense as well. Both number and areal extent are our objectives and while actual physical division on the earth's surface cannot be accomplished, we proceed as though this division can be portrayed graphically by maps. In other words, our consideration of objects in extenso includes not only number but space.

We turn now to several purely existential assumptions:

(1) The existence of a universe class;

(2) The existence of a null class;

(3) The existence of more than one element. (Though this is not essential, it is usually assumed).

By a universe class, or simply universe, we mean a complete collection of all those, and only those, elements which belong within the realm of discourse, in a formal way. The universe is the "set of all sets" in which we are interested. Figure 1 gives us very simply the conventional Venn diagram of the universe of discourse.

Figure 1

The rectangular form is a mere convention. Other shapes, regular or irregular, could have been used. The important thing is that an intended universe is bounded and all of the points or unit areas within the bounded area of the plane may be regarded as elements in the universe. Not only does the shape of this universe have no meaning apart from convenience but also neither does the size of the diagram nor do distances, directions, and specific geometric locations within it have specific meanings. Neither does the location on the paper carry logical significance.

When various elements are arranged into various sets, nearness of elements to each other in the topological sense are important and "neighborhoods" must remain invariant though the geometrical locations of the various neighborhoods is of no consequence. To repeat, the diagrams need only be regarded as having topological properties. These are what are important in the conventional Venn diagrams of the algebra of sets and we may do whatever we find convenient to do with the other geometrical properties of our diagrams to preserve the required topological properties in the patterns of relations among sets.

Now let us turn our attention again to geography. Here the universe is the whole surface of the earth. The earth's surface is the set of all sets and the places on it are the elements of the universe and its sets. Note that a place may be a point or an area on the earth's surface as necessary. For example, a given latitude-longitude designation would specify a point as place. On the other hand, we can regard the area of New Jersey as the place where the inhabitants of New Jersey live. For continuous distributions over the earth's surface, geographical places may be all or certain points and the elements in sets are infinite or finite in number as required. For other distributions geographical places may be all lines or selected lines, being thus infinite or finite. If areas (whether all or some) are the elements on the earth's surface, the sets, of necessity, are finite in number. Place, therefore, is to be defined operationally as required for the use to which the concept is to be put.

Essex County, Massachusetts, is a place in one context, the area of the United States in another, and the whole earth in yet another context. Whether places are points and thus there may be an infinity of them on the whole earth's surface or whether the whole earth's surface constitutes the one and only place when the appropriate non-spatial properties we wish our places to have are defined, is purely an operational consideration. The only property that all places have in common is the locative.

If areas are places then a region contains a finite number of elements with each element containing in itself an infinite number of point members. Additional considerations of denumberable and non-denumerable infinities lies beyond the scope of this introductory presentation.

Now we want to show a diagram of the earth's surface as our universe. Moreover, regardless of the other non-spatial characteristics which the elements subsequently may be recognized as possessing, we insist at the outset that the locations be represented "correctly." We start, thus, with the irreducible geographical requirement, location in its full sense.

Figure 2 is to be regarded as a mapping of the earth's surface in both the mathematical and the cartographical sense. On this map:

Figure 2

there can be, and we assume there is, a one to one agreement between the points on the earth's surface and the points within the ellipse. Immediately, we run into problems in that the earth's surface, being that of a very nearly perfect sphere, is at once finite but unbounded. Our mapped earth diagram represents this finiteness and it is bounded. This is a consequence of the particular projection we have chosen to use to transform the spherical surface to the plane.

On the map in Figure 2 the kind of projection used can not be readily recognized because there are no outward and visible signs such as the conventional latitude and longitude grid or some well-known and equally recognized distributions.

The significant thing is that the full range of geometrical properties such as size (to scale), distance, shape, direction, and so on are now ordered by mathematical functions and we may later begin further consideration of properties of phenomena associated with places.

Let us assume that the earth's surface is portrayed in Figure 2 according to the Aitoff projection which shows the entire earth's surface.

Moreover, the Aitoff projection maps the entire earth according to a projection without singular points. We might be more comfortable with the idea that the area of the plane bounded by the ellipse in Figure 2 is really a representation of the earth's surface though as yet you have only the author's word for it, let us add a reassuring latitude-longitude grid. This, too, is a classification procedure involving defined sets. We continue our elaboration now with the mapped points of the entire earth's surface and a representation of the conventional coordinate system for defining and fixing locations. See Figure 3. We are now able to define places that concern us in terms of the geographic irreducibility--location.

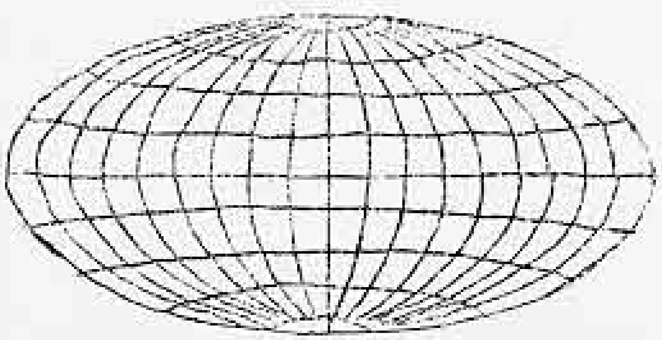


Figure 3

Let us now turn our attention to the important, existential assumption, the existence of a null class. This means simply the class in which no member has a certain given property. The null class contains all of the non-existing elements in the universe of discourse. It is the only class which contains members with incompatible properties. It is, in fact, made up only of such members. Graphically, the null class cannot be portrayed, nor can we map it as a geographical reality, as will be indicated subsequently.

Although it may be difficult to think of a universe of discourse of only one element, but the existence of more than one element is not essential because the algebra would be just as acceptable. The algebra would still be pertinent, however trivial, if we operated on null classes alone.

In addition to the foregoing existential assumptions, Hilton notes certain operational assumptions, namely:

(1) The existence of a complement for every term;

(2) The existence of a sum for any two, or more terms; and

(3) The existence of a product for any two, or more, terms.

When we say that we assume the existence of a complement for every term, we mean that for every class of a given type, there is also a class of elements not of that type. Together, these two classes constitute the universe. The second class may be thought of also as the negation of the first class. This is, of course, exactly the same thing as the principle of dichotomy discussed above.

In Venn diagrams for conventional logic, the relation of a set and its complement are usually shown as in Figure 4. L is the specified property. All elements not in the L set are shown as L', that is, not L.

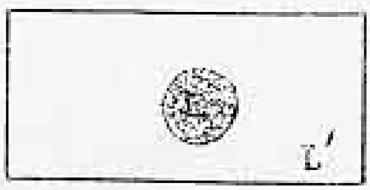


Figure 4

From an empirical earth-surface point of view, let L be the class of all elements (places) on the earth's surface that are regarded as dry land, i.e., solid ground, and L' be its complement, i.e., the places that are regarded as not dry land. And, let us insist upon a graphical portrayal in which the locations on the earth's surface of the elements in set L be regarded as inviolate and that these locations be ordered according to the map projection illustrated in Figure 3. The result of this is certainly familiar to all. It is the map of land and water surfaces shown in Figure 5. It is "accurate," of course, only to the extent that the scale permits.

Of course, such a map as in Figure 5 represents to most people more than "a mere classification" of the earth's surface. But, it is just that--a classification. So incomparably useful has this classification been throughout our history and so great have been the expenditures of time and other resources to secure all possible detail about elements so classified, that maps showing less than this particular classification of the earth's surface are regarded by most people--geographers included--as not being really maps at all.

Such a classification of the earth's surface is an example of an arbitrary classification that is so useful for so very many and such varied considerations that it is not ordinarily thought of in a classifacatory sense but rather regarded as a purely "natural" thing. Land and water as such are, we must admit, examples of non-spatial categories. The locations of these places have thus defined the areas shown, however.

Now when we introduce the full range of geometrical properties to Venn diagram methods of graphical portrayal of sets in which location is made explicit we find that <u>one type</u> may have many regional examples. Thus, the lack of compactness, i.e., the variance, in the distribution of dry land as a category has occasioned the development of different names for different parts of the fragmented land mass. We have given names to certain large pieces and have regarded them as "continents" and have similarly given names to smaller pieces and have regarded them as "islands," either as belonging to some continent or not as the case may be. Oceans, too, may be so regarded. We will investigate the "continent" and "oceans" questions in a bit more detail below. Here, however, let us consider additional aspects of non-spatial properties of places and classification in these terms.

After our first great classification of the places on the earth's surface into that which is <u>terra firma</u> and that which is not <u>terra firma</u>, we see that the number of "geographical regions" representing each of these classes seems to have

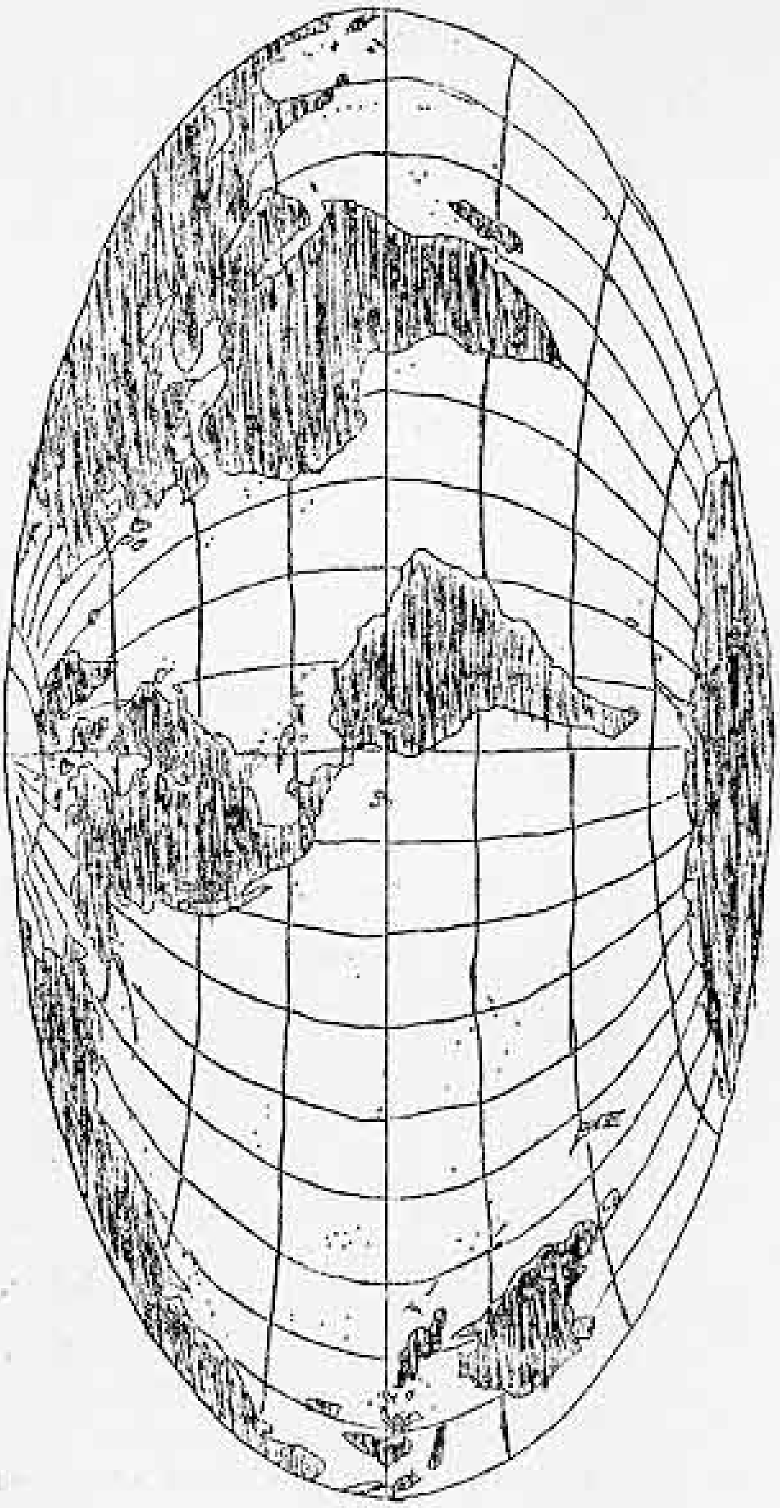


Figure 5

been more or less determined by sheer locational considerations.

Although the mapped distributions based on the land-water classification have proved to be of inordinate value through the ages one can readily think of a "base map" showing neither of these features. The above distinction, of course, becomes less and less significant as the atmosphere adjacent to the earth becomes the preferred space for travel. And, certainly, very long range guided missles for military purposes render the land-water "boundaries" more nearly inconsequential. Horrible thought!

As an example of another kind of classification of the earth's surface, let us consider certain temperature phenomena. Take, for example, as a class, the elements (places) having the property such that the average temperature of the coldest month is above 32 degrees Fahrenheit. The typical Venn diagram portrayal of this in a non-locational frame would be similar to that of figure 4 with whatever letter symbols are chosen to stand for this particular class and its complement. Again let L indicate dry land and H the property associated with elements for which the temperatures are as defined above. Note again that L and H are both non-spatial properties. These circumstances are shown in Figure 6 below.

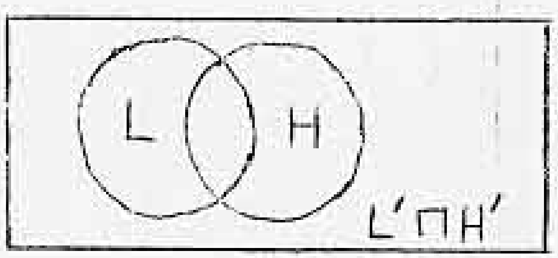


Figure 6

This figure now may be interpreted, for example, as follows:

If L and H are two sets we may derive a new set by considering the aggregate of all elements which are members of set L or set H or both. This new set is called the union of L and H or simply the sum of those terms. In conventional symbols we write $L \cup H$. Portrayed graphically (Figure 7) for elements (places) independent of their geographical location, the following emerges. The shaded area denotes the union.

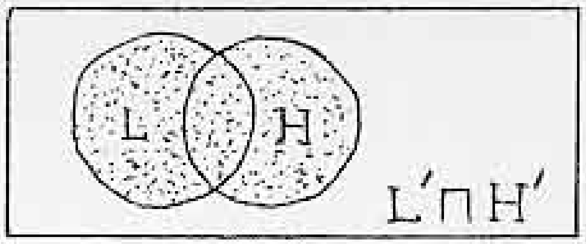


Figure 7

Of course many other concepts and operations exist such as inclusion, subset and proper subset, equality, disjunction, intersection, inverse functions, partitions, steps, miniterms, dualities, circuits, switchings, transformations and so on in an "astonishing array." Virtually every general concept or operation now clearly defined within set theory has its important special case within geography. Geography is in no way exceptional if we regard it as the science of location and spatial relations, i.e., from descriptive through classificatory through theoretical-predictive levels.

We give one more simple example. Figure 8, in its shaded area, shows the

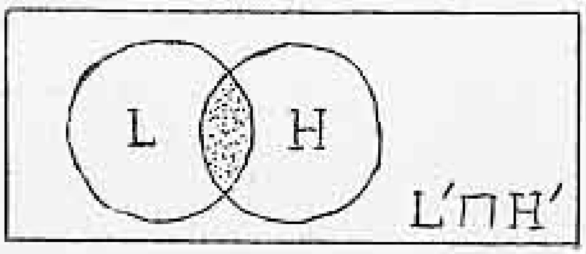


Figure 8

intersection of sets L and H, the product of those terms. In symbols we may write $L \cap H$ and thus specify the set containing elements (places, independent of actual earth location) which are members of both L and H.

Let us now preserve geographical locations of places and portray the above intersection, again according the Aitoff equal-area transformation from the earth sphere to the map plane.

The result is given in Figure 9. The great lesson here is that a single compact intersection non-geographically may map geographically into many "regions" on the earth's surface.

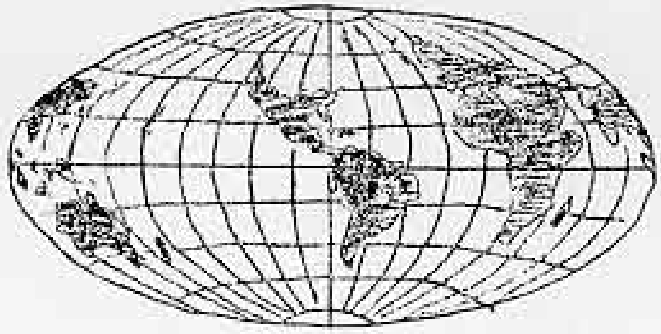


Figure 9

Figure 10 may be more comforting because it shows the same information and has the remaining land areas indicated.

Obviously we could increase the number of sets we wish to consider and arrive at an intersection of sets such that only one point (or even none) on the earth's surface satisfies the non-spatial conditions imposed. In spatial terms such definition is possible and frequently highly desirable. The intersection determined by a specific latitude value and a specific longitude value is positionally uniquely determined, of course. For non-spatial properties with the location of the elements preserved, we may adjust our classifications to get the most meaningful results.

Figure 11 portrays the geographical mapping of that set of non-spatial properties whose elements are sometimes referred to as constituting the "Temperate Marine Type" of climate. These elements are at least a proper subset of those shown in figures 9 and 10. They also meet additional requirements concerning amounts and seasonality of precipitation and maximums imposed on the warmest month average temperatures.

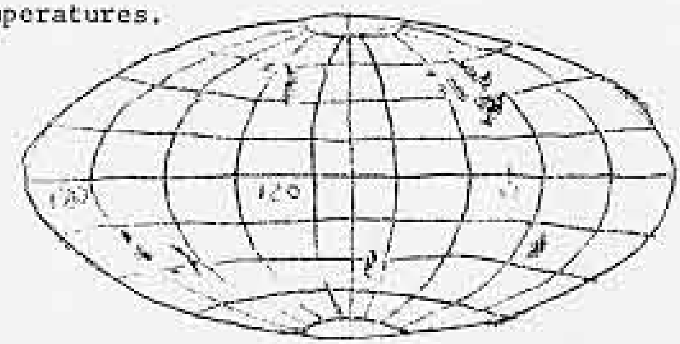


Figure 11

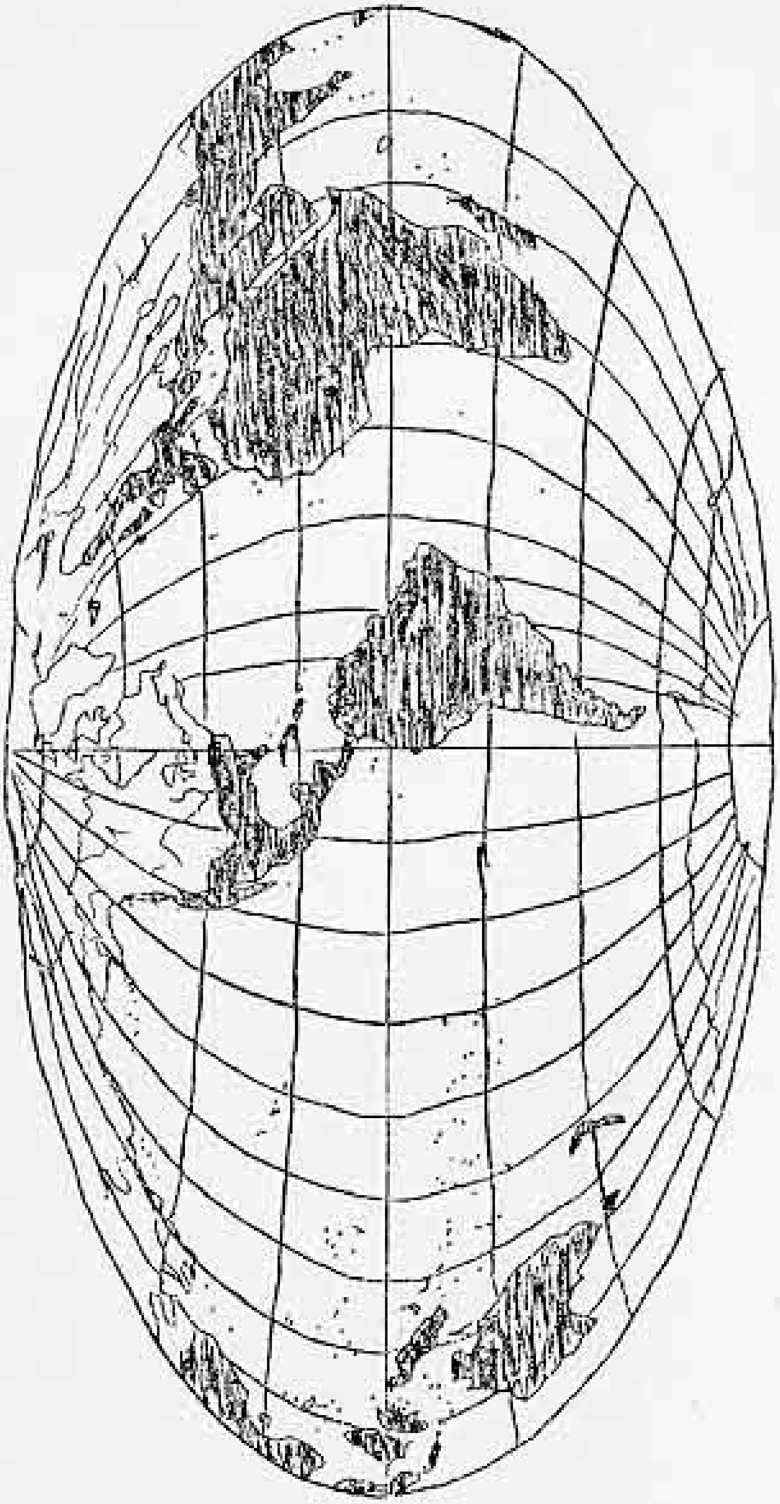


Figure 10

We will not follow this line of reasoning any farther here except to point out again that non-spatial property compactness may lead to geographical fragmentation. What is desirable non-spatially may be undesirable or awkward geographically.

For the degrees of freedom obtaining, we wish to maximize the ratio of variance among sets to the combined variance within sets. If we consider climatic types, income types, etc. for places (non-locationally) the conventional techniques of analysis of variance, of course, are pertinent and have meaning. We can arrive at a "Temperate Marine Type" in such a fashion. The geographical mapping that results, however, as noted, may be satisfactory. How should we group the fragmented "types of areas" (using type in the non-spatial sense) to give geographically useful and significant regions?).

Part II - Extended Venn Diagrams

Let us now consider variance in geographical terms alone. That is to say, we are concerned only with location of elements (places) and, at first approximation, we are willing to forego any consideration of non-spatial properties that places might be considered to have. In simple geographical terms averages are positions, deviations are distance (physical, time, cost, etc.) and variance is spatial. Degrees of freedom and probabilities retain their conventional roles.

If the regions into which we wish to divide the earth's surface are to be meaningful and convenient spatially for, say, planning or administrative purposes, spatial compactness may be desirable. Generally speaking the most efficient regionalization achievable for a given set of places in terms of both their non-spatial and also their geographical variances can be obtained only as a deliberate compromise of these two kinds of variance. To the degree that for the criteria use, the differences among places within regions are insignificant compared to the differences among regions we may regard our regions as individually homogeneous and collectively heterogeneous and our regionalization as a technical success for the criteria used. Again we note that location is the sine qua non of geography. If this factor's variance ratios alone are to be regarded, then the "ideal" regionalization based on location and the metric of distance results in a mosaic of hexagonally-shaped regions on the plane. Another tesselation is required for a closed spherical surface. Such a distribution alone simultaneously maximizes the distance among the centers of regions and minimizes the distances of all points within a region from that region's center. If other criteria of places than just their locations are included in the classification, other spatial patterns emerge depending upon the relative "weights" assigned to the various criteria. Whether we have chosen the appropriate other criteria on which to base a particular regionalization, i.e., whether we have regions of practical value for their intended purposes is, as noted, of crucial importance. We shall not venture here to discuss the various rationales and strategies for effecting the required compromises between the two types of variance and interpreting the results. Rather, we chose to conclude this paper with additional comments about spatial variance alone.

No two places coexist locationally by definition. Places are unique in the limit as to location, but this introduces no new idea or consternation into classification theory. Individuals are unique but grouping is at once necessary and efficient. Minimum distance variation between places occurs for adjacent, i.e., contiguous places. Maximum distance variation for places on the earth's surface obtains, of course, when the places are antipodal. Thus, in locational terms alone, any grouping procedure putting places into sets might regard contiguous places as most likely candidates for mutual inclusion in the same set and antipodal places as least likely candidates for the same set.

If spatial sets are regions then the Temperate Marine type of climate has many geographical regions of the same non-spatial type. Perhaps we would wish to include England and France within the same region. But, this region perhaps should not include its non-spatially identical New Zealand. New Zealand is virtually antipodal spatially to England. It perhaps should be included in another geographical region but of the same non-spatial type.

Can we think of some pertinent problem such that pure spatial variance of places alone is the significant phenomenon independent of their non-spatial properties? Perhaps! But, on a perfectly spherical earth with places differentiated by location but not differentiated in terms of non-spatial properties, every place is an equally efficient average (central) location because the total variance is a constant regardless about which place it is measured.

It seems as though spatial variance of places on a regular closed geometrical surface like that of the assumed spherical earth becomes significant only when we select out certain places (arbitrarily, or in terms of certain non-spatial properties, hypothetically, theoretically, or empirically) such as the places occupied by persons, the places deemed to be within a certain class of vegetation forms, the places regarded as exhibiting certain climatic conditions, and so on.

The most nearly pure spatial example involving the least contrived and perhaps most generally significant non-spatial classification (at least in this and the most recent epoch in human history) is the one based on the land-not land (i.e. water) distinction noted earlier. We return to the "number of continents" problem. This is clearly not an idle question when viewed in terms of spatial variance and the number of sources yielding the most efficient description.

At work through the ages, though never fully consciously announced, has been the understanding by mankind that the specific recognition of the separate continents includes the considerations of the centers and boundaries of continents such that the present classification of the earth with seven continents might approximate closely the "ideal" ratio of the distances among all points within these continents and distances among the continents. This is, of course, subject to appropriate examination and verification. We should test to see whether or not this is so. And, too, should Europe and Asia so-called be regarded separately or as Eurasia? Recognition of Antarctica as a separate continent (or is it a large island?) seems easy enough, but what about North America and South America? Here the isthmus connecting them (artificially pierced by the Panama Canal) renders them contiguous. But, this may not be enough to cause us to regard North America and South America so-called as one continent as the actual spatial distributional shapes of these two land masses perhaps causes the average distances between the points of the two continents to be too great for the narrow land link to overcome. On the other hand, actually non-contiguous land masses such as "adjacent" enough islands may be included with a given continent as for example Japan as part of Asia.

In connection with all of the above it is interesting to note that the meaning of the word "continent" as an adjective is "held together, contained, or restricted."

"How many continents are there?" is thus seen to be not just a rhetorical question. It is valid and geographers classifying the earth's surface should attempt to answer it and support their answers with results obtained by careful methodology based on the analysis of the variance of locations within as compared to the analysis of the variance of locations among continents based on the various assumptions as to what are continents. Appropriate statistical tests would provide pertinent measures of significance based on varying degrees of freedom. The fragmented spatial distribution of our non-spatial class L has made it relatively easy for mankind acting collectively to reach a certain tentative concensus through means not actually explicitly agreed upon for the class L and the associated number of and names for continents.

We are currently designing a small research effort to attack directly this problem by way of an "iterative-weighted functions of locational values" computer program. Research into the literature of the history of continent recognition and naming is also contemplated.

Did Europeans divide Eurasia differently when they knew that Australia existed than they did when they did not? Why? How much larger would Greenland have to be to be recognized as a continent assuming its approximate present location? Or, how much more isolated would Greenland have to be to be recognized as a continent given its present size? It may be that there is no position on the earth that a "Greenland" might assume that would occasion continenthood given its present size. Clearly, continenthood is a function of size, shape, and position as the concepts of the analysis of spatial variance reveal. At optimum location for continenthood how much larger might a "Greenland" have to be?

Suppose all of the dry land of the earth to be gathered together in one maximally compact, and thus circular, continent. (Before continental drift?). Would there be an advantage to recognizing sub-division into continents based on locational factors alone? Clearly not! The idea of continents in a physical sense (apart from political, cultural, economic, etc. considerations) takes on meaning in the real world only in the presence of fragmented land masses and/or the not altogether convex shapes of these land masses. In fact it is the obvious matching of concavities with convexities that makes the continental drift hypothesis such a pleasing notion.

The basic ideas of the theory of convex sets are natuarally and easily appreciated when examined in a geographical context, rather than in a non-spatial one. In geography the meaning is direct, obvious, and literal but nonetheless powerful. This is true in general of a large part of set theory as developed by Georg Cantor (1845-1918) who provides the foundation of the theory of point-sets, of real functions, and of topology. In this respect geography is truly elegant. In fact it may be the least exceptional science. What about the complement of the class L, that is L' and its regions of location? We may regard this at the outset as being not only the negation of L but as constituting in its entirety another class, say W, or the water surface of the earth. This, too, can be mapped. In fact we have done it in figure 5 for the very reason that W and L gave no members in common and together they constitute our universe of discourse, all the places on the surface of earth.

One final comment. In general in statistical analysis of variances, squared values for deviations are customarily used because of their obvious significance in terms of certain model distributions such as the Gaussian normal. Then, too, squared values have lovely algebraic properties--no negatives, for example. In the geography of continenthood, all deviations are themselves positive and squaring to avoid negative values is not necessary although it may still be desirable for other reasons. The power of the distance to be involved is, of course, a matter of considerable importance. We can define a center as a place where the appropriate extreme value for a spatial moment occurs and we define any spatial moment about any point (following Neft, 1966) as:

$$M'_n = \int r^n DdA,$$

where r is distance, n is an exponent and DdA is the density of elements over the infinitesimal bit of area. The assumption of different values for n occasions different centers. What is most reasonable for given circumstances is a difficult problem and is often more clearly dramatized by geographical problems than non-spatial ones.

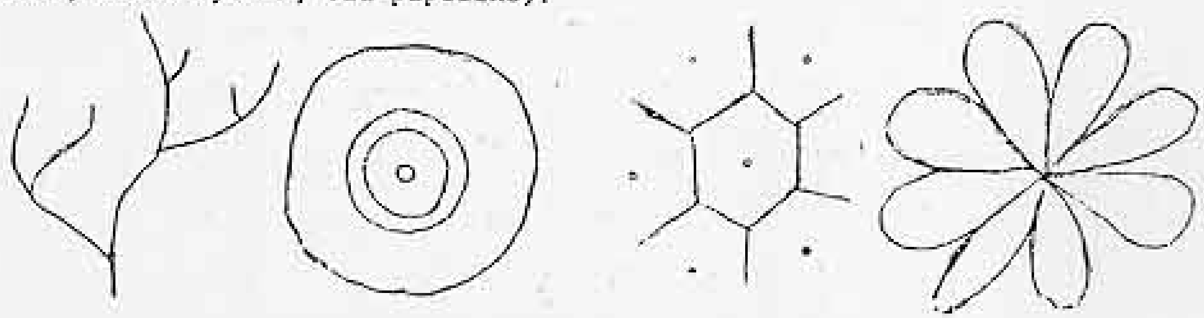
Some one should write a book on regionalization. It might well have the following title:

> "The Algebra of Sets, Geographical Regions, and the Use of Some Obvious But Hitherto Ignored Properties of Venn Diagrams--Or The Logic and Method of Geographical Regionalization as an Extension of Boolean Algebra Examining the Relationship to General Classification Theory, Symbolic Logic, the Algebra of Sets, Computers, Electronic Circuitry, and Logic Diagrams with Special Consideration of Automated Logic Machines and the Representation of Systems of Geographical Regions Based on a Rational Interpretation of Non-Spatial Differences and Geographical Distances, that is to say the Theory and Application of the Geographical Analysis of Variance."

We sometimes make our greatest gains in geography when it is our basic concepts that are simple and natural and our techniques that are sophisticated.

X X X X X X X X X X X X X X X X

What is implied if the concept of pure spatial prediction has validity? Geography, like other mathematized sciences before, is searching for the correct coordinate system and point of origin. Tobler is our Copernicus. The Geographic Projection, the one we still seek, the one much more important than the infinite projections mastered, is the Uniform Plain, which is the geographic equivalent of "other things equal" assumptions in other sciences. Once the space is properly projected, the patterns (our primatives are the dimensions) both probalistic and extremum (with the function to be minimized some concept of nearness) should emerge and be more testable. Somehow the patterns and the coordinate system should be related functionally. This is a Grand Design of Theoretical Geography. In none of this work is it necessary to refer to time. Movements can often be eliminated as well: Space and perhaps movements, but never time. This work led the editor to the work on pure spatial prediction. Time is so damnably invisible, but space we can see. Kurt Schaefer's great sentence, "Patterns are morphological laws." implies what philosophically? A visible law? The law itself seeable? The law of a line and an area as near to each other as possible (with certain other minor restraints) is a dendrite (river, sewer system, oil pipeline)?



The following comments were excerpted from Theoretical Geography (1966).

Spatial Prediction

William Bunge

Abstraction rules supreme these days. With the overthrow of Kant, the rise of non-Euclidean geometry and the scientific success of the new physics, anything that is non-abstract is definitely out of fashion. As students are intoduced to the notions of dimensions and hyperspaces, they are simultaneously admonished not to take them seriously but immediately gain the sophistication of considering dimensions as a little boy's version of variables. Real terrain surfaces are demeaned as organically inferior to mathematical ones--pedagogical crutches for the weak minded.... In mathematics, it is considered the most flagrant gauchery to use a diagram. "Graphics" is thought to be an inflated title for "mechanical drawing." In fact, all the intrinsically visible subjects;

geography, graphics, and geometry, are suspected of being really grade school subjects, fit only for brains that are still undergoing biological maturation and whose harmfully misleading approach will have to be undone later. No one is arguing for a return to Greek Naiveté. Certainly invisible abstractions can be real and are necessary even in geography, but why is this in contradiction to the thought that perhaps concrete visability might also be real? A truth is no truer for being obtuse or sophisticated. In an age of mental Schaum tortes can not nature's fresh peaches also be worth savoring? Whitehead, you have gone too far!

Abstraction is so much the rage that even geographers have difficulty imagining a pure spatial prediction. Asking students for examples is always a disappointment. Their examples typically include the eventual draining of the Great Lakes or the creation of vast cities in Southern California. Always time is present as a variable even if space is present in their examples. But time need not be present at all to make a prediction. Consider the profile, or side map, of a wave pattern. (Figure 1).

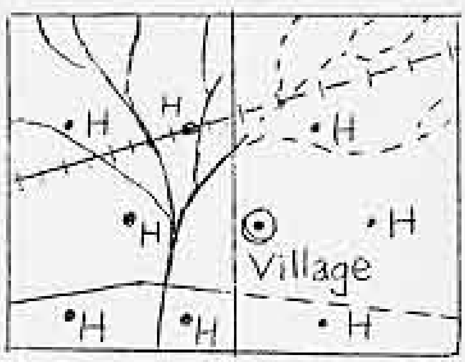
Figure 1

Wave Pattern

This could be the march of temperature over the seasons, a time prediction. It could also be a profile of the Ridge and Valley Region of Appalachia. Standing on the crest of the last wave to the right, what can be predicted? --that summer will be followed by fall and so forth in time or that the peak of the ridge will be followed by the side of the valley and so forth in space. Consider an idealized topographic sheet. (Figure 2). What will the adjacent sheet be like?

Predicted Map:

GIVEN



PREDICTED

Figure 2

The railroad and highway might be continued as straight lines and the river pattern idealized. The farther in space away from the known location the less confident the predictions, just as confidence is lost in the Weather Bureau the farther in the future the prediction is made. Some might object that such a prediction is merely an extrapolation. But most clearly with the spatial prediction of the village, an object that did not appear on the known sheet, this is not true. With improvement of our skills it is conceivable that Koppen's Hypothetical Continent be extended to a Hypothetical Globe. No extrapolation is involved. Geographers predict a village because geographic theory demands it. If geographic theoriticians provide geographic experimenters, including explorers, with the most likely sites of the missing Mayan cities, compute their k's, their rank-size rule, estimate the number of hamlets, compute their average spacing and so forth are they not predicting? To place our finger on the map of Yucatan and say, "There," is no less impressive a prediction than an astronomer pointing in the Heavens to a missing planet or Babe Ruth at the right field wall.

The most compelling case for pure spatial prediction is the seeming isomorphism of concepts with the concepts of temporal prediction. The temporal scientist's past and future corresponds to geography's behind and ahead; their event, our place; their moment, our location. It is the visability of spatial prediction which makes geography so instrinsically visible. The case should not be overstated. For example, it is most likely that the solution to the mapping of the uniform plain will not be a visible solution. Limits exist to literalness just as to abstraction...

X X X X X X X X X X X X X X X X

In what sense is the map a theory? One effect of reading Toulmin is to put mathematics into a proper perspective. Toulmin (1953) writes elsewhere in his excerpted chapter, "...the mathematician remains the servant of the man who knows when and how the results of his computations can be applied. Jeans and Eddington were both primarily mathematicians, and in their popularizations of physics gave prominence to the mathematical side of the subject, but the results were in certain respects misleading: the physics is not in the formula, as they suggested and as we are often inclined to suppose, any more than being able to find your way about is part of a map. The problem of <u>applying</u> the theoretical calculus remains in physics the central problem, for a science is nothing if its laws are never used to explain or predict anything." Science is the queen of mathematics, but the sciences seem so weak in preparing their students in philosophy that it all occurs in the mathematics departments where young minds are pumped full of the Christian Science view of the world that afflicts mathematicians. More power to them if such a Ptolemaic view of the intellectual universe spurs on starving mathematicians, but the sciences should not allow their students to be so mapped by intellectual default. Philosophy of mathematics is what the brighter students must somehow overcome to arrive at philosophy of science. As Toulmin comments, "The physics is not in the formulae..." A question plagued me for years, "Though Newton could predict the motions of the moon, how did the moon know where to go? Can the moon integrate?"

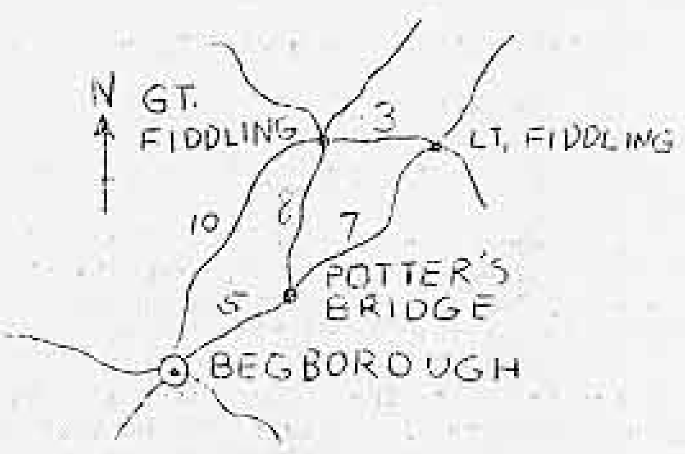
As to possible extensions of Toulmin, he goes a rye when he mis -answers his superb question "...what exactly corresponds in cartography to laws of nature in physics?" He assumes it must be map projections which he justly criticizes but it is patterns, geographic laws, like Christaller's or Thünen's, or the host of new ones, that corresponds to laws of nature in physics. He also makes, the minor-to-the-discussion error, of assuming that the task of map projections is to "preserve certain chosen features" of the surface of the Earth, when, of course, the major task is to create new features that Nature missed, such as Mercator's straight loxodromes accomplish. We need a terrible "crooked" projection, the Uniform Plain maker, that will make the simple "distortions" of physics appear as child's play.

Though for purposes of his discourse there is no error in comparing the fundamental map with complete theory in physics, geography does not view a complete map as its complete theory. The map is merely the experimental data controlled through the laboratory instrument of the projection; the spatially manipulated locations of the earth's surface. Geography as the science of location, in its most sophisticated labor, seeks to predict locations. The most fundamental theory is hueristically defined as placing objects of different dimensions as near to each other as possible, with some feeling for probabilistic convergences with the extremum statement. The following comments were excerpted from Toulmin's book, <u>The Philosophy of Science</u>, 1953.

Theories and Maps

Stephen Toulmin

...Consider, for instance, the imaginary motoring map opposite, showing the town of Begborough and its environs.



(Reproduced with Publisher's permission.)

We can ask about this section of map a question similar to Mach's question: namely, what relation it bears to the set of geographical statements that can be read off it, such as "Potter's Bridge is 5 m. NE of Begborough on the road to Little Fiddling", and "Great Fiddling is 3 m. due West of Little Fiddling."

How are we to answer this question? Certainly the map cannot be said to be deduced from the set of geographical statements nor, in a logic-book sense of the phrase as opposed to a Sherlock-Holmesian one, are the statements deduced from the map. For in a deductive inference, such as "Fish are vertebrates, mullet are fish, so mullet are vertebrates", the same terms appear both in the premises and in the conclusion; whereas here the 'conclusions' read off may be statements, but the 'premise' is a map and contains no 'terms' at all. Only where premises and conclusion are comparable in the way that "Fish are vertebrates" and "Mullet are vertebrates" are comparable, is there room for a deductive connexion, so the relation between the map and the geographical statements must be of a different, non-deductive kind. At the same time, the map need not be said, in Mach's sense, to 'contain' anything which cannot be expressed as a geographical statement of the kind included in our set: everything which one could read off from the map of this sort. Though the map and the geographical statements are not deductively related, one need not conclude that the map goes beyond the surveyor's readings; since it does not present us with additional information of a novel kind, but represents the same information as the statements in a different manner. This example shows that, when we are presented with two logically incomparable forms of expression, the question whether or no one form of expression contains more than the other is quite independent of the question whether or no the one can be deduced from the other. In fact, unless the expressions are of logically similar kinds, there can be no question of such deduction...

The aggregate of discrete observations is transformed into a simple and connected picture, much as the collection of readings in a surveyor's notebook is transformed into a clear and orderly map.

The consequences of this analogy are worth noticing. For if someone asks, "Doesn't the map tells us that Potter's Bridge is 5 m. NE of Begborough, and a whole lot of similar things?", we can only answer "Yes and No." Certainly, if you know how, you can read off from the map a great range of geographical information; but the map on the one hand, and the geographical statements on the other, tell us things in very different ways. A man might own Ordance Survey maps of the whole country, and yet, for lack of a training in map-reading, be quite unable to tell us anything of a geographical kind; likewise, a man might have memorized all the currently accepted laws of nature and even know a vast amount about the calculative side of mathematical physics, and yet not be equipped to explain or predict any of the phenomena observed in the laboratory...

In the traditional logical account of the sciences, one encounters certain difficulties when explaining how it is that experiments are used to establish theories. In the first place, physicists seem to be satisfied with far fewer observations than logicians would expect them to make: one finds in practice none of that relentless accumulation of confirming instances which one would expect from reading books on logic. This divergence is partly to be accounted for by the logicians' confusion between laws and generalization--one would hesitate to assert, say, that all ravens were black if one had seen only half a dozen of the species, whereas to establish the form of a regularity in physics only a few careful observations are needed--but this is not the whole story. There is also a second, related difficulty to be overcome: that of explaining how subsequent applications of a theory are related to the observations by which the theory was originally established.

To take the two difficulties together: it is worth noticing that they arise for theories as much as, and no more than, for maps. Not all the applications to which a theory is put need have been specifically made in the course of the experimental investigation by which it was established. But nor need all the things which can be read off from a map have been specifically put in. A child might wonder how it was possible ever to produce a map at all, since to tread every inch even of a small area, and to measure all the distances and directions that one can read off from a map, would take an unlimited length of time. This, of course, is the marvel of cartography: the fact that, from a limited number of highly precise and well-chosen measurements and observations, one can produce a map from which can be read off an unlimited number of geographical facts of almost as great a precision. But it is not a marvel calling for a general explanation, for only in some regions can the techniques be implicitly relied on. In irregular country it is always possible to be misled, and the number of observations which have to be made per square mile will be much greater in some areas than others--just how many are needed being something the practising cartographer must be able to judge.

Correspondingly, it is a fact that many physical systems have been found whose behavior can be similarly 'mapped.' Having made a limited number of highly accurate observations on these systems, one is in a position to formulate a theory with the help of which one can draw, in appropriate circumstances, an unlimited number of inferences of comparable accuracy. Thus it is always possible that the next time Boyle's Law is applied, the particular combination of pressure and volume concerned will be being observed for the first time. But again, though this fact is in its way a marvel, it is not one requiring a general explanation, any more than is the possibility of mapping. For here, too, how far the behavior of a given system consists of phenomena which can be mapped in a simple way, and just how many observations will need to be made before we can be confident that our theory is a trust-worthy one, are things which will vary very much from system to system and which it is part of a physicist's training to learn to judge...

The imaginary road map of the region between Begborough and the Fiddlings which we discussed a few pages back, need not be the only map of the region.

There will also be some more elaborate physical maps drawn to a larger scale and showing a great deal more detail. In such maps as these, roads will perhaps be drawn to scale, not represented by lines of purely conventional widths, while towns and villages will be marked, not as mere dots and blobs of standard sizes, but as having definite shapes and made up of individual streets and blocks of houses.

Now a number of things should be noticed about the relation between the road map and a physical map of the same region. In the first place, many things can be mapped on the physical map which there is no way of putting into the road map: this is a consequence of the ways in which the two maps are produced, and of the comparative poverty of the system of signs used on the road map. On the other hand, given the physical map, one could produce a satisfactory road map: all that appears on the road map has its counterpart on the more elaborate map, even though in a different form. But this does not mean that the road map is not, of its kind, an unexceptionable map of the region. Providing that it is not thought of as having irrelevant pretensions, there is nothing wrong with it: indeed, for some applications one will be able to discover the things one wants to know, e.g., distances by car, more easily from the road map than from the physical one. Finally, it is worth noticing what happens if we mix up the systems of signs used on two different kinds of map. There are some motoring maps in which one finds town-outlines and other features sketched in on top of the simple road pattern: but since only distances along roads can be given a satisfactory interpretation on such maps, the result is usually confusing, and the simply blob for a town is more consistent with the general scheme of the map.

The relation between geometrical optics and the wave-theory is not unlike that between a road map and a detailed physical map. Thus the fact that one can explain on the wave-theory, not only all the phenomena that can be accounted for on the geometrical theory, but also why the geometrical account holds and fails to hold where it does, is like the fact that one can construct a road map from a physical map; but again it is not a sign that the geometrical theory need be superseded for all purposes. Road maps did not go out of use when detailed physical maps were produced. It shows only that, as one can produce a road map from a physical one but not vice versa, so one could produce a ray-diagram from the wave-theory picture of an optical system, but not vice versa. The conceptual equipment of the geometrical theory, like the system of signs on a road map, is too poor for one to do with it all that can be done with the wave-theory. Indeed, the notion of a light-ray is an artificial one in very much the way that the conventional-width road is, and has to be abandoned in the wave-theory because the accuracy with which one wants to answer questions about optical phenomena is too great for the conventional picture to be retained. No more can one, from a simple motoring map, answer questions about the distance from the northern verge of one road to the middle of another--these are things that a map of that type does not pretend to show. Again, since there is no room within geometrical optics for representing the phenomena of diffraction, a physicist would hardly think it worth while to give any indication on a ray-diagram of the shapes of any diffraction-fringes he observed: they would be just as out of place there as town shapes are on a bare motoring map.

If we look at the relation between different theories from this angle, we can notice some points of importance about the notion of a 'fundamental' or 'basic' theory. One finds that, at a given stage in the history of physics, there is commonly one theory, at any rate in a particular field, which is regarded as the basic theory: this theory is thought of as capable of accommodating all the phenomena to be observed in that field. Now two questions need to be asked. Since it will never be the case that all the phenomena have in fact been explained, all that need be claimed is that the basic theory can in principle explain them all: the first question is, what are we to understand by this claim? Secondly, when physicists talk about explaining everything, what are the criteria by which they would judge that everything had in fact been explained?

It is helpful to compare the basic theory with the fundamental map on which the Ordance Survey might record all the things which it is their ambition to record. This would, of course, be a map drawn on the very largest scale, but it would not be the only true map of the country: rather it would be the one which most fully and precisely represented the region mapped, and the one from which by appropriate selection and simplification all others could be produced. For many purposes it will be too elaborate to be of practical use, but for some purposes none else will do, and the lover of cartography for its own sake must have a special place for it in his heart.

The value of the comparison lies in this: it suggests that the standards of what constitutes a complete theory in physics may change. For we could say that the fundamental map was complete only if it showed all the things which in that region it was the cartographer's ambition to record. Now it is always possible for cartographers to develop fresh ambitions: the criteria of the completeness of a map are, accordingly, at the mercy of history. So are they with the theories of physics. One is at first inclined to suppose that the physical sciences have a definite goal, the same for Aristotle, Newton, Laplace, Maxwell, and Einstein, but a closer look at the history of the subject will show the mistakenness of this idea. Rather there is at any given stage a standard of what sorts of things require explaining: This is something with which scientists grow familiar in the course of their training, but which is hardly ever stated. The standard accepted at any time determines the horizon of physicists' ambitions at that time, the goal which for them would have been reached if 'everything'--i.e., everything thought of as requiring explanation--had been found a place in the theories of physics...

It is, then, still in cases where our interest is in how one might 'get somewhere,' i.e., produce or counteract some spotlighted development, that we talk about causes--though the destination need not be one that we care about either way. From this we can see why the term 'cause' is at home in the diagnostic and applied sciences, such as medicine and engineering, rather than in the physical sciences. For the theories of the physical sciences differ from those of the diagnostic and applied sciences much as maps differ from itineraries. If the term 'cause' is absent from the physical sciences, so also a map of South Lancashire does not specifically tell us how to get to Liverpool. To a man making a map, all routes are as good as each other. The users of the map will not all be going the same way, so a satisfactory map is route-neutral: it represents the region mapped in a way which is indifferent as between starting-points, destinations and the like. An itinerary, however, is specifically concerned with particular routes, starting-points and destinations, and the form it takes is correspondingly unlike that of a map. Often enough, of course, a map be used to work out the itinerary for a particular journey, and from one map an indefinite number of routes may be read off, as occasion requires. But, from its form, there is nothing about a map to show that it is to be used for this, rather than any other of a wide range of purposes...

This analogy shows us something about the relation between the fundamental and applied sciences, and about such phrases as 'applied physics.' For in many fields of science practical skills preceded theoretical understanding, and even provided the first data for systematic study. Sundials were in use for centuries before their operation was properly understood, and there are still plenty of familiar processes, in cooking for instance, about whose physico-chemical nature we have only the sketchiest of ideas. There is therefore only a part of engineering which can be called 'applied physics', even though this part may be continually growing and may in some divisions, such as atomic energy, be all but exhaustive. This state of affairs also has its natural counterpart in cartography. For a long time, travellers relied on itineraries rather than on maps; Greek seamen and Roman legionaries as often as not followed set routes for which itineraries had been written out; there must still be today a few more remote parts of the world which are totally unmapped, but around which a guide could take one; and even in our own well-mapped country we all know some short cuts and refinements that are shown on no map. So though the preparation of itineraries may in fact often be applied cartography, it need not be. Itineraries preceded maps. The development of cartography has given us a way of understanding the relations between different routes, and at the same time a source of new itineraries whose possibility had not previously been recognized. And there may be some parts of the world so remote, so mountainous, that one could hardly hope to work out itineraries for them except by first mapping them from the air...

In cartography, too, there is a good deal which has to be contributed by us before there can be a map at all, and this contribution is again of an unmysterious kind. Cartographers and surveyors have to choose a base-line, orientation, scale, method of projection and system of signs, before they can even begin to map an area. They may make these choices in a variety of ways, and so produce maps of different types. But the fact that they make a choice of some kind does not imply in any way that they falsify their results. For the alternative to a map of which the method of projection, scale and so on were chosen in this way, is not a truer map--a map undistorted by abstraction: the only alternative is no map at all. To draw an analogy between a cartographer's method of projection and the ichthyologist's fish-net would accordingly be misleading. There is no question of falsification here. Quite the reverse: it is only after all these decisions have been taken and a map has been produced, that the question can even be raised, how far the product of the cartographer's work is true to the facts, for only then will there by anything which can be true to <u>or</u> falsify them...

The existence of the Absolute Zero can be compared with the existence of the boundary in a map of the World drawn to a stereographic or orthographic projection. On these projections, the surface of the Earth does not cover the whole of any sheet of paper you use, as a Mercator's map is capable of doing, but fills only two circles. If there is blank space round the circles, that is not because the cartographer has chosen to cut off the map half-way up Greenland, say, but because, the nature of the projection being what it is, no point on the Earth can be mapped outside the circles. One can, of course, decide to make the circles as large as one chooses; but, however large one decides to have them, there will still be a boundary, whereas a map drawn to Mercator's projection is capable of going on indefinitely.

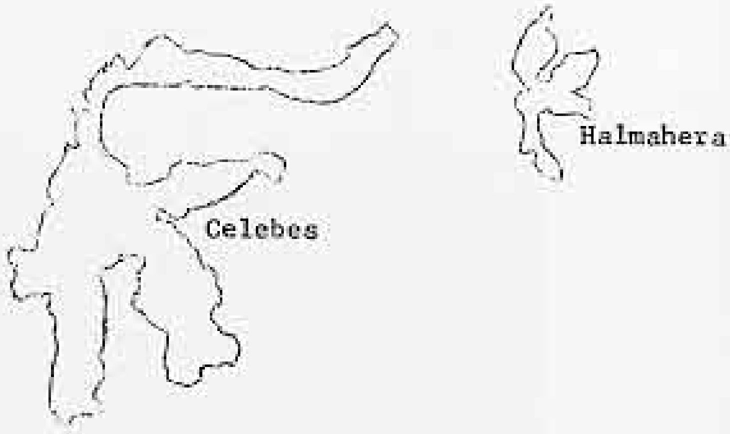
If we prefer, it is open to us to stop using a map of one kind and start using one of the other kind; and to abolish the boundary in this way shows nothing about the area we are mapping. The presence or absence of such a boundary tells us nothing about the surface of the Earth. The same is true in physics. One can, if one chooses, change over from the ordinary ideal gas scale to a logarithmic scale, which extends without limit in both directions; and to make this change implies nothing about actual thermal phenomena. In neither case does one, by changing the method of representation, burke any facts about the World. . .

What, then, of the question, "Do electrons exist?" How is this to be understood? A more revealing analogy than dodos or Ruritania is to be found in the question, "Do contours exist?" A child who had read that the equator was'an imaginary line drawn round the center of the earth'might be struck by the contours, parallels of latitude and the rest, which appear on maps along with the towns, mountains and rivers, and ask of them whether <u>they</u> existed. How should we reply? If he asked his question in the bare words, "Do contours exist?", one could hardly answer him immediately: clearly the only answer one can give to this question is "Yes and No." They 'exist' all right, but do they <u>exist</u>? It all depends on your manner of speaking. So he might be persuaded to restate his question, asking now, "Is there really a line on the ground whose height is constant?"; and again the answer would have to be "Yes and No", for there is (so to say) a 'line', but then again not what you might call a <u>line</u>. . .

And so the cross-pruposes would continue until it was made clear that the real question was: "Is there anything to show for contours--anything visible on the terrain, like the white lines on a tennis court? Or are they only cartographical devices, having no geographical counterparts?" Only then would the question be posed in anything like an unambiguous manner. The sense of 'exists' in which a child might naturally ask whether contours existed is accordingly one in which 'exists' is opposed not to 'does not exist any more' or to 'is non-existent', but to 'is only a (cartographical) fiction'. . .

X X X X X X X X X X X X X X X X

What, if anything, besides language can we read? Martin's contribution is maddening. It is original and sweeping, a man who sees "letters arranged into sentences on the face of the map." What of Wegner whose continents literally fit like a jog saw puzzle and who died trying to prove a "mere" jig saw puzzle idea himself? At the Brighton meeting of MICMOG where Martin presented his material again in the spring of 1967, we all were messmerized by his concepts and confessed the little secret mysteries we had snatched from Mother Map over the years. For instance, since a child I have personally been struck by the enormous similarity between Celebes and Halmahera, both in the same part of the world, both with impossibly similar shapes and orientations. The similarity is so striking as to call for a small map.



Martin refers to his work as a "new concept" in geography. He looks at the map fresh, like a child might year after year and he sees things. "The Quebec shore of the Ottawa River expresses the opposite E. Ontario Shore, as can be expected of any river. The corresponding Africa and America shores, though across oceans do the same."

Philosophically, must all languages be written by men? Do scientists literally try and "read" nature? What does a geographer mean when he says he is "reading" a map? Is a map a language? Is the landscape? The following comments were excerpted from Martin's thesis.

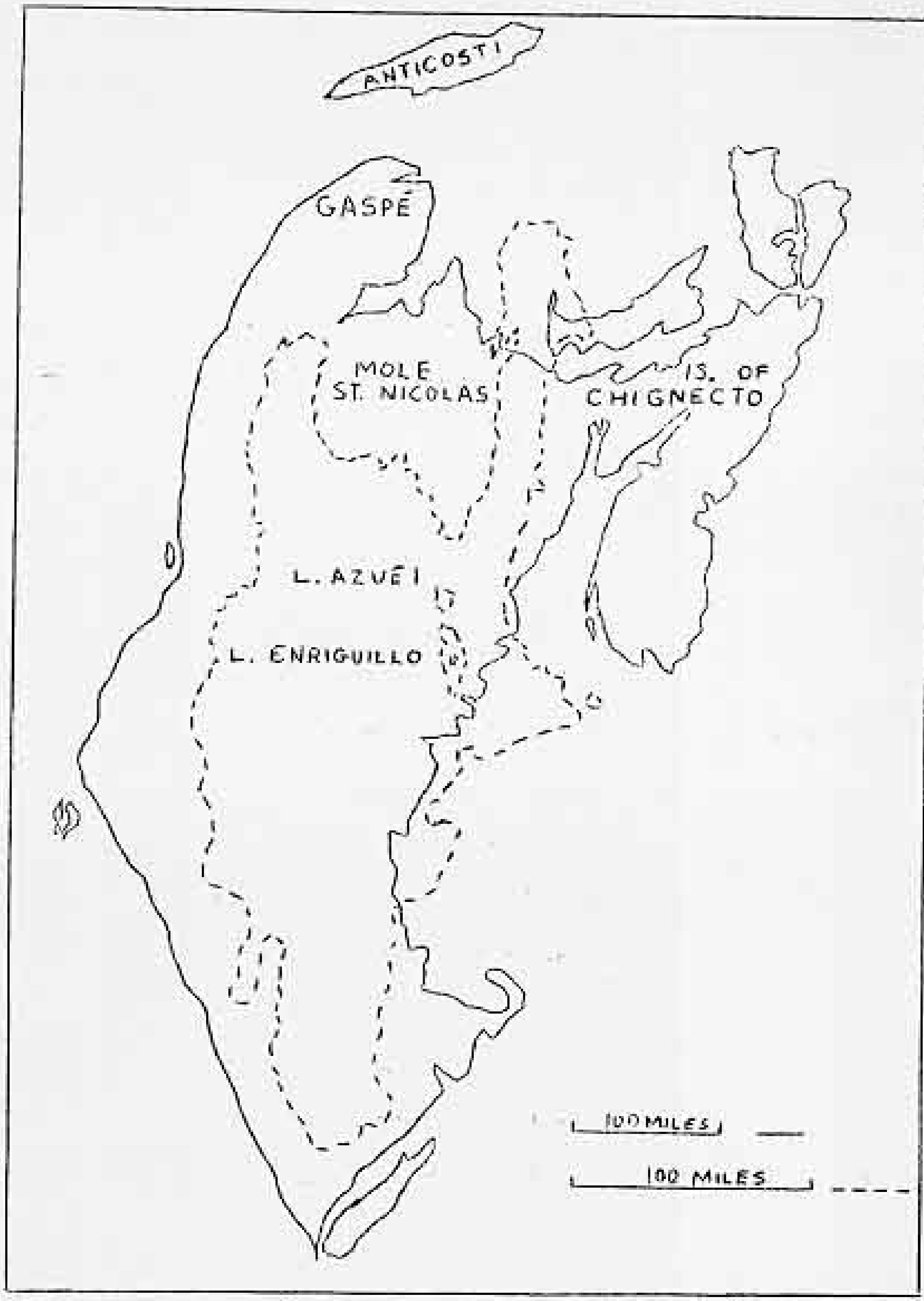
The Earth as a Living Body

Roland Martin

The first difficulty confronting us was in finding a starting point in the physical actuality of the world represented on the map. Ask any geographer if he knows of any two lands in the world having the same features of same relative proportions. Not only will he deny such knowledge but he is likely to go on to deny the very possibility of the existence of such two similar lands, or at best he will only state that he had never given the matter any thought or consideration, let alone the benefit of a tentative investigation.

If the relationships are uncomplicated, why did not someone perceive them merely by close observation? First of all, in order to find something one has to look for it. In this case, what one would be looking for is not the features of a particular land entity--such as geographers normally look into--but a relationship. The idea of a relationship is entirely abstract and exists only as a concept of the mind until it rests between two objects that enter in relation.

In addition, it still takes more than an abstraction implanted in the mind to prime the process. There are several other impediments to be overcome before gaining access to this first conjugated observation. The Rockies and the Appalachians are oriented primarily from north to south, while the Notre Dame Mountains and Ridge of Nova Scotia are oriented generally to the northeast, and the Cordillera Central and La Hotte--La Selle Mountains of Haiti are oriented from west to east. With such divergent orientations as are the case between Haiti and North America, the similarity between the two land units hardly jumps to the eye. The observer seeking to investigate the relationship between the two units would have to confront a map of the island with a map of the continent and wheel them freely. He might then orient the two western peninsulas of Haiti in line with Alaska and the northern end of the Appalachians so that the mountain ranges fall parallel to one another.



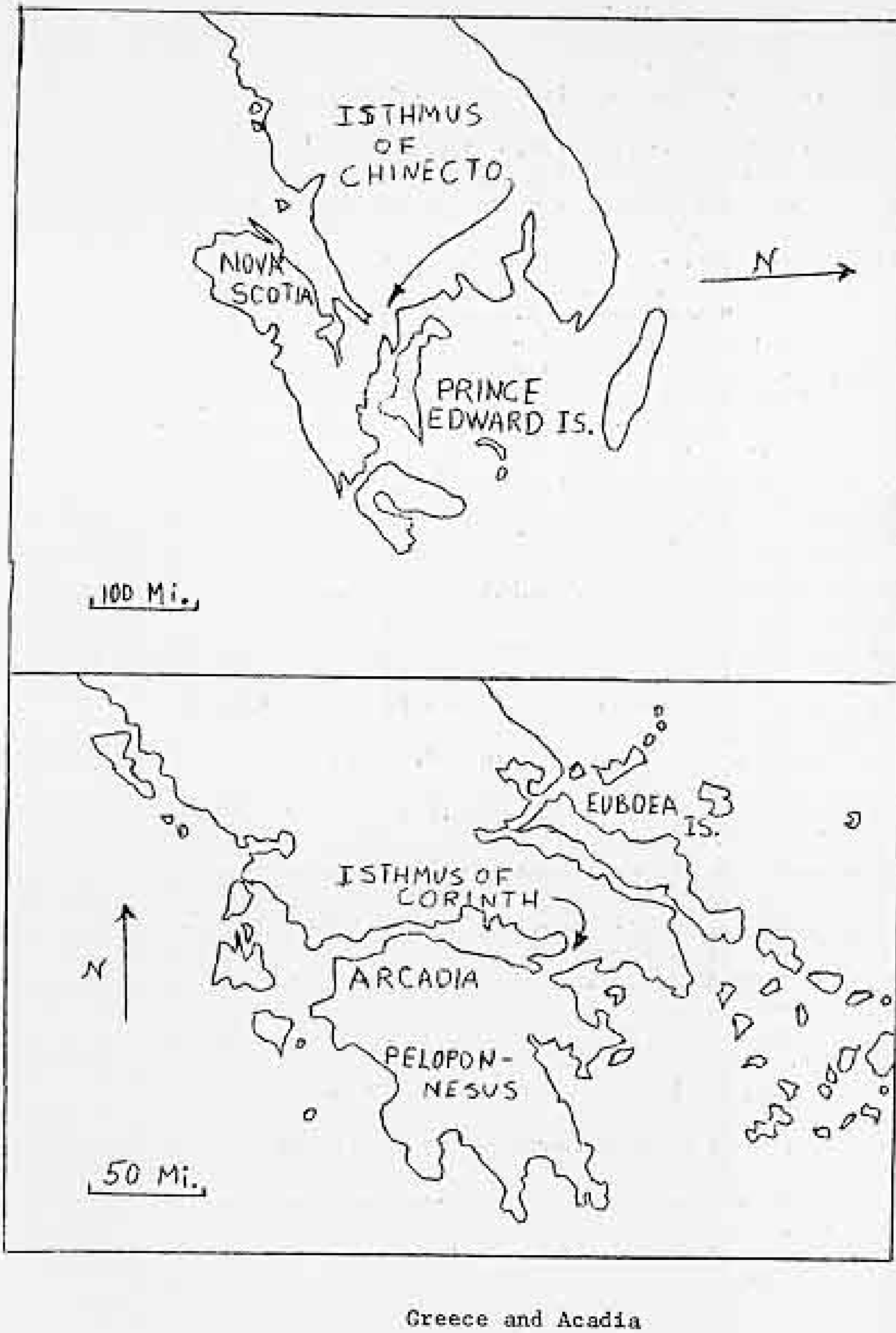
Haiti and Acadia

What is the task of the investigating geographer? To learn to separate the systematic aspect of each land from the physiognomic. The physiognomy throws a thousand veils over that self-same ever-recurrent system. The distinction is essential, fundamental. How will the geographer deal with it?

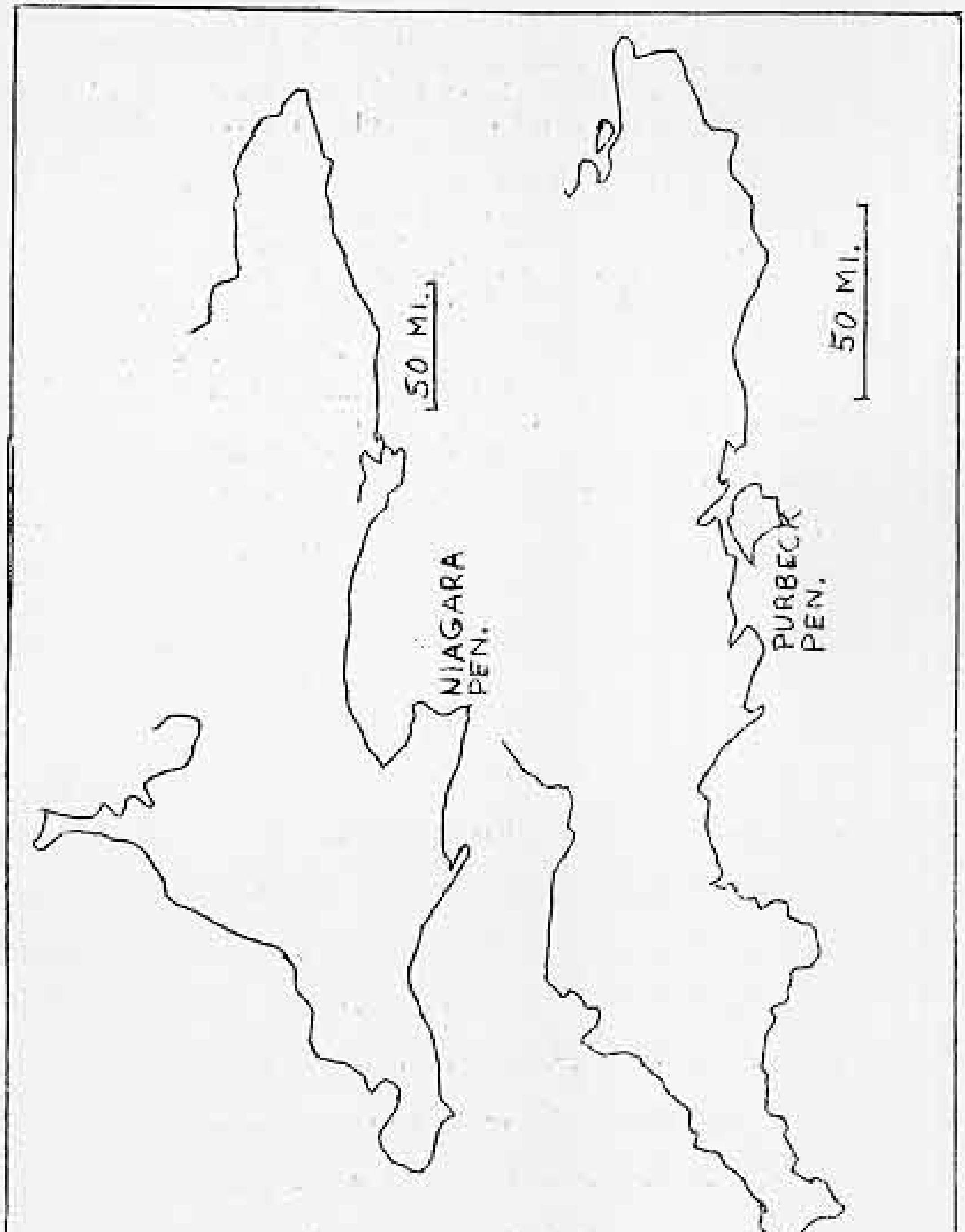
Our own way offered no short-cuts. It consisted of spending thousands of hours spread through consecutive days and weeks and months and years amounting to two decades in the perusing of maps: these of varying scales, symbolisms, contour lines; hydrographic or orographic; physical, political, economic or historical, etc. Once one becomes thoroughly familiar with the individual features of the world, and their numerous images are stored in the mind, the close observation of any area on a map brings back to the consciousness related images stored in the memory. Should I mention Brittany, will the reader be tempted quickly to think of Alaska or Turkey, to name only two others, will he be tempted to go a step further and think of the county of Retz, across the Loire, appended to and belonging to Brittany? Will he think of the Alaska Panhandle as another appendage? Will he think of the Sandyak of Alexandrette returned to Turkey two decades ago? These countless items of reference constitute a fund of experience indispensable to build on further and more intricate experiences. The investigator needs to be involved with the more obvious experience that areas such as Alaska, Brittany and Turkey afford before he can move on to less obvious ones such as Wisconsin or the German coast of the North Sea. Between this last area and Alaska, the structure is not the same, the physiognomy is not the same, but the profile of both coastlines is the same.

Thus the process is successively compounded from the more obvious to the less obvious, and thence to the introduction of concepts. Does the Alaskan type of morphological occurrence constitute a natural appanage of the west? Along the same line, but as an opposite, do such peninsulas as Kamchatka or Gallipoli constitute a natural appanage of the east?

The process of identification of parts is followed by one of distribution



Greece and Acadia



South Channel Coast and South Coast of Great Lakes and organization. A cluster of features two or three or four times repeated is a clue to what to expect (or what yardstick to bring) in an additional area. Thus are identified what we have called land isolates, or self-contained areas presenting a recognizable complex of large-scale features.

X X X X X X X X X X X X X X X X

How do we prove truth? In an age of philosophical abstraction a rather strong materialistic rebound is not unexpected. Besides, as graphic a subject as geography, with its field boots, landscapes and feel of the earth's surface, this earthy subject, could not be held in Whitehead's hypnotic trance for long.

While the discussion of maps as related to proof is somewhat buried in the middle of the effort, it was the triggering notion of the article and opens the possibility of proving certain geometries by direct geometric methods again. Yes, we geographers know what happened to hidden theorems in the straight edge and compass. We also know just how impossibly immature the state of the mathematical arts is relative to some of our basic problems.

One great advantage of formal mathematical proofs not mentioned is the possibility of those mysterious simplifiers popping out such as i's and e's, the strange parameters. But maps might have their own parameters, strangely reoccurring patterns.

The material is part of a larger effort in preparation.

Truth

William Bunge

The human mind, like the other organs of the body, evolved for reasons of straight Darwinian biological survival. There is no reason to believe that the brain contains any more adornment than the human stomach... It exists no more for its own sake, for self-gratification, than does the liver. The separation of mind and matter is a false separation. No such dichotomy exists. Only matter went into the creation of the brain. The mind need not consist of magical nonsensical stuff for religion to be meaningful. Whispy, imaginary vapor in the cranium is too crude a foundation for religion. The mind is obviously not in perfect balance with its environment or it would never have produced the radioactive poisons that so threaten it with extinction, but then neither is the rest of the human organism so well adjusted. The feet too seem half way from hands to hoofs, from prehuman to human. Still, "thoughts" with possibly some efficient low level random noise, and certainly all "systems of thought" are either a direct response to current survival or past, that is, thoughts are as much evolutionary features as fingernails.

What system of thoughts exist in the human mind? It is not necessary that the brain divide its survival work as neatly and with as much even-handed balance as a college catalogue. A mixture of computers of varying practical-survival importance might be the most efficient system with considerable cross communication between the organ's subparts.

The table shows examples of systems of thought. What is the history of these forms? Many appear in prehuman evolutionary stages. Chess and go, the greatest games, came into being considerably after man came down out of the tree. The first maps, drawn in the dust of some cave, have been vastly improved. Certain truth systems, such as science and mathematics, came into conscious division of labor only most recently. It is to be expected that other aspects of the human mind will be discovered and new truth systems emerge into consciousness.

On the surface, the truth systems seem totally unrelated with the exception of the already stressed fact of their Darwinian common origin, but other commonalities emerge. Each truth system has a "pure" form. Many practitioners of the special circuits in the brain feel themselves to be totally motivated by "impractical" impulses. "Art for art's sake," "Mathematics for its own sake," and so forth. But the ultimate survival purpose of the function being performed does not have to be clear to the practitioner. A farmer might take pride in his ability to plow a straight center furrow and might even enter a plowing contest "for its own sake," but we can plainly see this activity is related to growing food. The farmer hardly thinks to himself every second he is farming, "I'm growing food so the species can survive." Such a constant thought signal would not be survival efficient. Idealism, no matter how strong the subjective pull of it, can be explained as materially efficient, but materialism cannot be explained, with any conviction, on the basis of idealism. It is true our minds might be

Thought System	Science	Humor	Religion	Games	Justice	Art	Senses
"Pure" Form	theory	contra- diction	faith	play	laws	rules	sensations
Proofs	experiment	jokes	worshipping	contests	trials	works of art	feeling
Truth death life	disproven proven	obscene humorous	blasphemous spiritual	losing winning	illegal legal	ugly beautiful	painful pleasurable

Thought System	Logic	Instinct	Space	Emotion	Wisdom
"Pure" Form	relation- ships	reactions	patterns	feelings	judgments
Proofs	"mathematical" proofs	living	maps	getting to know someone or something	experience
Truth death life	false true	a don't a do	lost oriented	hated loved	foolish wise

sitting in a warm saline solution in the year 8,000,000 and our sense nerves being fed artificial perceptual stimuli from an enormous computer, but this is too tortured. Rod Serling does not have the believability of Charles Darwin.

The "for its own sake" feeling has been extremely seductive and has led to all sorts of nonsense including the essentially unflattering assertion that much mental work is "useless." But this is so much in contradiction to the facts, such as the tremendous utility of "useless" mathematics, that a mysterious explanation had to be offered that the conscious pursuit of useless curiosity was useful. How metaphysical! Yet this "rationale" has much currency even among scientists...

Some calculations have been made as to the amount of digital computer apparatus that would be necessary to recapitulate the mental processes of the human mind. The calculation indicates an enormous computer requiring great energies and producing serious quantities of heat. But why not a series of computers specially designed for special survival problems; a mix of computers pragmatically evolving with some abilities to intercommunicate. Such a strategy would explain several mysteries, including the tiny size of the brain relative to its ability to compute.

Scientists and mathematicians often achieve success by direct reliance on a sense of symmetry or even more deep feeling of general beauty. How could this be? Is this magic? But assuming that the mind can afford no more ornamentation than the rest of the organs, perhaps the aesthetic sense is survival prone, a direct analogue computer cutting through the "normal" logic of "formal proof."

Mach writes:

> ...In every symmetrical system every deformation that tends to destroy the symmetry is complemented by an equal and opposite deformation that tends to restore it.... One condition, therefore, though not an absolutely sufficient one, that a maximum or minimum of work corresponds to the form of equilibrium is thus supplied by symmetry. Regularity is successive symmetry. There is no reason, therefore, to be astonished that the forms of equilibrium are often symmetrical and regular.

Toth writes:

> Besides...classical theory, regular figures may be approached in another way, starting from the observation that extremum postulates often involve regularity. Classic theory starts with a more or less arbitrary definition of regularity. Here, in turn, regular arrangements are generated from unarranged, chaotic sets by the ordering effect of an economy principle, in the widest sense of the word. This theory may be called the genetics of regular figures.

When Maxwell "mysteriously" balanced the partial differential equations he was being efficient. Tobler and I and countless others have had similar experiences in dealing with mathematical solutions and I have had a sense of symmetry and similitude especially in discovering why rivers and other lines are dendritic. It was originally on purely aesthetic grounds that I favored Christaller's "fixed k" assumption over Losch's. This removes the mystery. The human mind developed a set of circuitry that short cuts the logic of extremum and simply "instinctively" selected that which was "beautiful." But the deeper question is not the question as to why aesthetics helps scientists and mathematicians but is artistic expression a direct path to truth? If symmetries and perhaps much deeper instincts toward beauty help "formal" logicians, what do they do to artists themselves used directly in their hands....

Much special computing capacity is devoted to spatial problems. The seeking of game, the hiding from enemies, the searching for food, the problem of being lost, all are deadly biological survival situations. A cab driver can find a close approximation by "instinct" of the least time path down a polar geodisic coordinate system that is constantly changing shape. He goes extremely "straight" in an extremely complicated crooked world without even the benefit of high school calculus....

The fundamental reason for proof is that the human mind-survival computer needs help in directly perceiving the truth when making certain survival judgments. Proof helps reduce human error and therefore helps survival. Error is the key; the total error in judging the truth so that the organism can properly respond to its environment and not make so many deadly mistakes that the species perishes. Notice perfection is not the goal. All organisms make mistakes, even individual deadly ones. Squirrels fall out of trees to their deaths, birds fly into mountains, fish drown in air and so forth. If a "perfect" solution to any one problem were computed, the brain would have to be too large in the circuitry to afford the "correct" answer. "Perfect" answers are too expensive. This is the Error of Mathematicians in spite of their modern acceptance of probability theory. It makes no survival sense to have a perfect computer if other errors are going to nullify the "logical" perfection as we know from the theory of errors. Consider the errors. One, the errors of perception of the real world, data input. This error is usually much greater than the form in which the data is presented suggests. For instance, maps of rainfall or anything else are almost totally believed but a great deal is not known about the data: The individual error in each rain gauge; the great distance between rain gauges: since no one has ever done a detailed study of the falling rain, say every meter over 100,000 square kilometers, who knows if the shape of the approximating function so blithely drawn by the TV weatherman bears any resemblance to the actual surface whatsoever? The behavior of the second derivatives of such a map look suspiciously like the contours of the earth's terrain and the actual rainfall might be more closely approximated by a step function for all we know. Two, the errors of approximation in order to fit the system of logic to be used. In practice, we always abstract from the data, already at best dimly perceived, in order to fit our system of proof. In applied mathematics even continuous functions of continuously conceived real life variables are only approximately fit. Not only do the least square errors enter, but the selection of the function to be fit can give great error, for instance, the least squares fit of the downtown skyline to a sine function! Once the imperfectly recorded data is squeezed into some form of proof, we encounter the third error, the imperfection of the logic. The practical problems of survival that mankind faces prove intractable to the proof systems developed. Certain cheating emerges. For instance, in the proof system of mathematics, the mathematicians insist on stupendous errors in cramming the data into the existing logical machinery. If a problem from the real world of survival arises as to the location of a highway between city A and city B, all known factors involved affecting the highway location, pollution, travel time, cost to builder and so forth can be placed in some common utile measure and this continuous distribution mapped. If the surface is not topologically simple, and it will not be, analysis will not be possible to apply to trace out the total least cost line. So, blithely, the problem is "simplified" into a series of squares with some average utile cost assigned to each square. Now the mathematics might prove manageable and assuming not too big a matrix and a muscular computer, the "exact" solution might be forthcoming. But look what the mathematician has done! In order to secure a "perfect" mathematical solution in the algebras, he has forced an increase in the error of the second type. He has forced the data into a discontinuous algebra. But if the error of the second type is increased to the point where it is larger than the "logical" errors, what is the survival-gain? The "practical" solution is in greater error though the mathematical part is perfect. The <u>fourth error</u> is that mankind has no complete theory. Even if perception were perfect, data fitting perfect, the proof mechanism perfect, the theory would be incomplete. With the highway problem perhaps future switching to electric batteries in automobiles will eliminate the pollution consideration and thus this future effect must be partially discounted from the present location problem and on and on. Since the death of the illusion that Newton had determined the Universe to be a Great Clock, science is increasingly pessimistic about its ability to have a universal theory of Truth. Science's work will never be done. How much Truth can the human mind be expected to understand? It has been a great human error to assume that the function of the brain was to discover The Ultimate Truth. Blessed are the humble, baby. The function of the brain is to enable the human species to survive, a modest goal, just like the function of the nervous system of the ant is to enable that creature to survive. We therefore must stop looking for the Ultimate Truth in theory. The Mathematician's Error is his refusal to accept that the best the human mind can do is to distribute the error in such a way as to arrive at the nearest thing to Truth that its limited capacity will ever allow so that we can continue to exist as a species. All the mathematician's insistence on "perfection" has done is up the errors among the other three sources of error, thus enfeebling the Truth of the applied product.

Having faced a Total Theory of Errors, we can sensibly raise the question of what is proof. The mind has found Darwinian virtue in what we called "formal proof," mathematics-logic. Crows can count up to a certain number of hunters entering and leaving the woods and avoid a shooting up to a small number. We do not have to recapitulate the evolutionary history of this advantage. The proof of the pudding lies in the mind's ability to think in this direction just as the best proof of the utility of the human eyelash lies in its existence. But the mind needs help. The help is called "proof." For example,

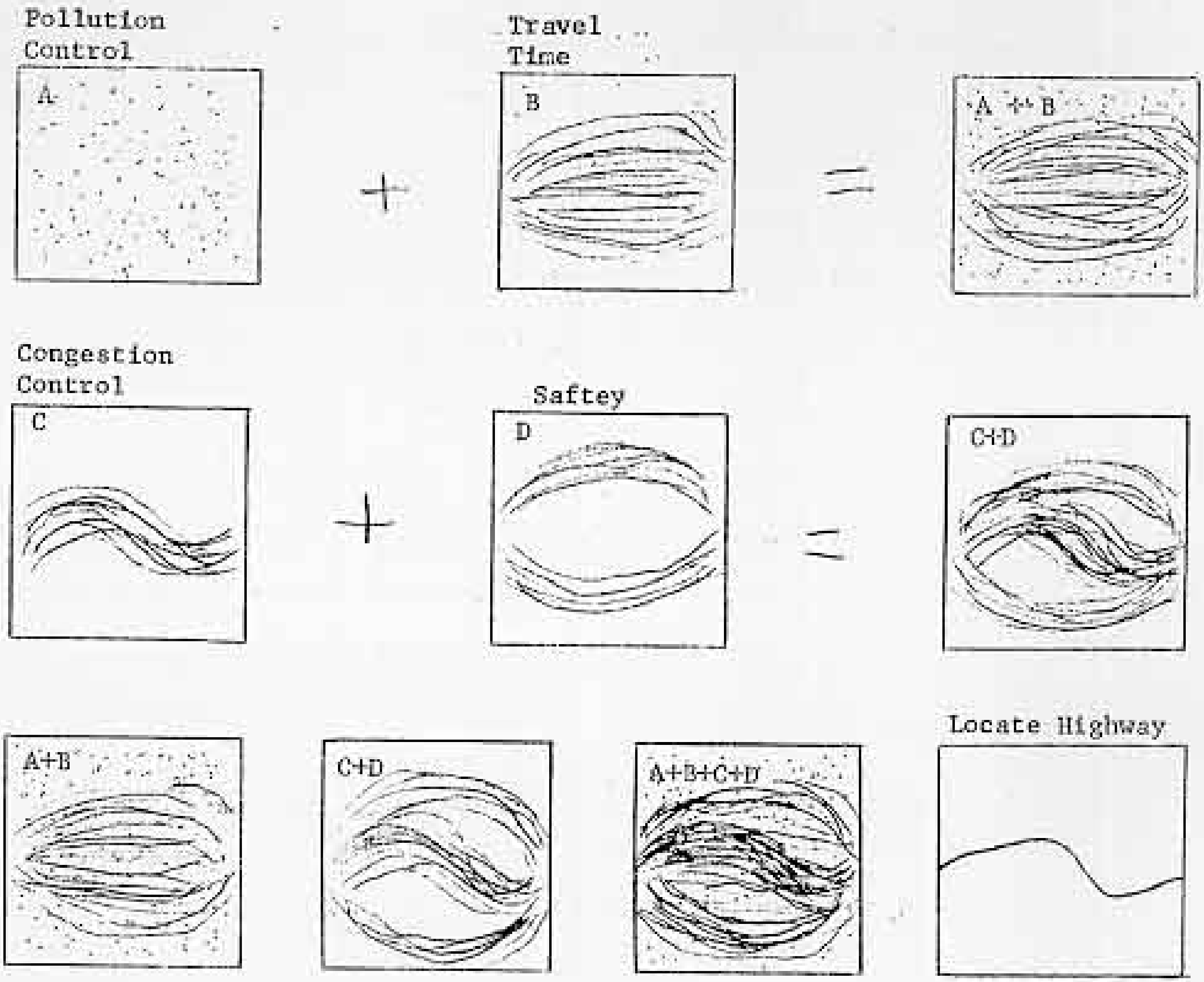
Prove: $8a + 16b - 3c = -10$
Given: $16a + 32b = 6c - 20$

1. $16a + 32b = 6c - 20$
2. $8a + 16b = 3c - 10$
3. $8a + 16b - 3c = -10$ Q.E.D.

You might object that the proof was trivial and unnecessary. But you might argue that all proofs are trivial and unnecessary, as R.A. Fisher did. Proofs are a waste of time since mathematical truth can be directly perceived. But most logicians feel better with the proofs, the step by step visual aid so that the human computer, which did not develop this memory in many of us, can be aided. Not all "formal" proofs are so full of certitude. In the problem of finding the best location of a highway, the problem might be even too much for the best modern computer and use of the most modern algorithm in finding the single best solution. Then an approximately good solution might be found. The approximate solution might contain an estimate of the error, the approximation of the approximation as in probability theory, and it might not. Simple simulation techniques might be used to establish a tolerable error. Scientists are rather well aware that all is fair in love and science. Any dirty method will do.

The proof lies on the sheet of paper but the truth of the proof lies in the human mind. The proof does not give birth to itself. Placing Step One on the paper and leaving in a warm saline solution does not give rise by birth to the other steps. The "proof" is a visual aid and the judgment of the Truth is always in the human mind. All proof systems, even "abstract" mathematics, are perceptual. Mathematicians should examine what they really do. Put little black marks on white pieces of paper.

But logic-mathematics is not the only Darwinian brain circuitry that exist. Men have different kinds of thoughts, different brain processes for arriving at survival-truth. Toulmin points out that the map does not tell the viewer anything that the viewer does not compute in his mind. To "prove" that some town is so many miles from another one can submit a map and measure the distance on the map. Assuming the map is free of error (preposterous) one can exactly (ridiculous) prove the distance. But to come closer in illustration to mathematical proof, return to the problem of locating the highway between two towns. A map of every factor that seems to enter can be made as Alexander has done. Each factor can be weighed by its relative merit as Alexander has not done but Roberts has. The weighing can be done perceptually so that a factor three times more important is three times blacker. One map can be placed on another and the two combined. If the darkness of the imposed maps begins to make the map solid black in appearance in any part, the value can be lightened on the combined maps. After all are combined, the best looking, darkest appearing route can be traced by a pencil. The proof is complete. Man starting with the same approximate value systems, the given set of original maps, will draw approximately the same ultimate highway system.



Q.E.D.

But think of the error! But proof has nothing whatsoever to do with error elimination, only error reduction! Just listen to the outcry of the mathematician. To be "correct" why not find a numerical utile value for each point on the map (really, each point!), that is, an approximate utile value for approximately each value on the map and then find the line integral (intractable!), well, all right an algebraic approximation, and assuming the computer will not smoke too badly, you will be given an "exact" and "provable" best single solution, provided your dirty old theory was true. But how can the error of the map proof be estimated? Well, it obviously is not an infinite error, in fact, it seems to be pretty close to the Truth, besides we do not know of a better way to estimate highway location, and on top of that a better way may never be a "formal" say, it might simply be an improved visual way. Look what is going for the maps as systems of proofs--all the evolutionary apparatus of space computing. Error might be estimable by psychological testing, but remember error in the proof might be desirable: why have a proof that is better than data input or incomplete theory. Maybe the human mind also evolved its own practical theory of error. It can be argued that replication is difficult. What Total Problem in real life is not, but approximate replication should be achieved. It can be argued that novices would not do as well as the practiced. This is different with mathematical proofs? It can be argued that none of the steps are necessary. Well good, if you happen to be an R. A. Fisher of maps, but most of us find perceptual errors reduced enough to make the map steps worthwhile. It can be argued that we have no sound theory at all for drawing in the line of the highway from the darkened area of the map. It looks pretty good to my evolved eye-mind and that is good enough for most highway locations. After all, how do we locate highways in real life?

Artists also use proofs. The beauty-truth is perceived in the eye and mind of the beholder and if we all were R. A. Fisher type artists we would not have to have the proof of the "work of art" to see the artistic truth. We could all be direct artists and in part we all are. What do artists say about their work? They say that they start with a certain situation, a set of "givens" and the worl more or less forces itself if they, the artist, have the courage to be honest. The novel writes itself, the play itself, the painting draws itself, step by ste and it just has to come out the way it does. There are more random artistic works, more error in the proof and perhaps less in the viewer, for errors must be balanced and cannot this interpretation be given to the debate against the perfection of classical realistic art? The truth may be perfect on the canvass but this merely shifts the error to the perceiver, or the original data, or the total theory of truth, which is incomplete in artists as well as scientists.

How does one know, when does the mind tell us, that the proof is complete?

Again, coming back to evolutionary essentials greatly simplifies the problem. Proof is a step by step transformation that cuts down perceptual error. It is complete when the problem, given by life, is transformed to the point where the person or group puzzled by the problem knows what to do. All thought systems--science, humor, religion, games, justice, art, senses, logic, instincts, space, emotion and wisdom tell us that war in this age is disproven, obscene, blasphemous, losing, illegal, ugly, painful, false, a don't,lost, hated and foolish. All systems of truth signal Death.

But not all truths require proof. It is foolish to prove what is known to be true. Now immediately people are going to come running out of the woods displaying every paradox in the book. Yes, intuition can be terribly wrong. But more impressively, it can be terribly right. How many decisions are made by "formal" proof? What is the percentage of error from intuitive failures? And even what about Goedel and truths that lie outside proof even in formal logic? It is wasteful and ultimately deadly dangerous, to spend time proving the obvious.

X X X X X X X X X X X X X X X

What is a "natural" language? The last two articles are the works of students. Young minds, like young atheletes, are the best ones. Karlin is going to peal off the map patterns and read the language of the maps like a latter day Rosetta Stone. This paper, the most formal, somehow pulls much together.

Shapes as a Group

Andrew Karlin

Geographers have always studied shapes, but we have rarely worked directly with them in a rigorous way. There are probably several reasons for this neglect, but perhaps the two most important are, first, our tendency to regard a region's shape as the spatial limits to some phenomena under study, and, second, a lack of systematic and rigorous methods. Currently, however, we are discovering the importance of shape in its own right. Two good examples of this are Bunge's *Theoretical Geography* and Alexander and Manheim's *The Use of Diagrams in Highway*

Route Location.

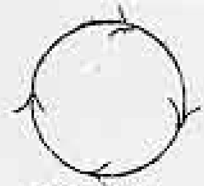
If, however, we find it useful to manipulate shapes, or what is the same thing here, figures, how are we to do it? But this is really two questions. First, if we manipulate shapes are our answers meaningful? For example, if we impose a circle over a triangle does the result have any significance? Moreover, is any curlicue a figure in the same sense as a circle, with its tidy geometry and simple equation? Second, if we attribute some meaning to the sum, what are the mechanics of the addition? In practice, of course, these questions of meaning and technique are bound together.

Bunge has suggested that we give formal answers to these questions through group theory. A group, according to Keyser, is a special type of system. That is, a group is a class, or collection of things, with some definite rule, or way, in accordance with which any member of the collection can be combined with either itself or any other member. More precisely:

> Let S denote a system consisting of a class C (whose members we will denote by a, b, c and so on) and of a rule of combination (which rule we shall denote by the symbol o, so that by writing, for example, a o b, we shall mean the result of combining b with a). The system S is called a group if and only if it satisfies the following four conditions:
>
> (a) If a and b are members of C, then aob is a member of C: that is, aob = c, where c is some member of C.
>
> (b) If a, b, c, are members of C, then (aob)oc = ao(boc)... that is, the rule of combination is associative.
>
> (c) The class C contains a member i (called the identical member or element) such that...aoi = ioa = a...
>
> (d) If a be a member of C, there is a member a' (called the reciprocal of a) such that aoa' = a'oa = i...
>
> Other definitions of the term "group" have been proposed and sometimes used. The definitions are not all of them equivalent but they all agree that to be a group a system must satisfy condition (a).

Let us assume that any figure or shape has an equation, though perhaps unknown. This resolves the semantic problem of figure and shape raised earlier. Further, let us assume that we can assign vectors to figures, an assumption first suggested by Nystuen. For convenience' sake, a "counter-clockwise" figure is positive and a "clockwise" figure negative, as the figures below illustrate.

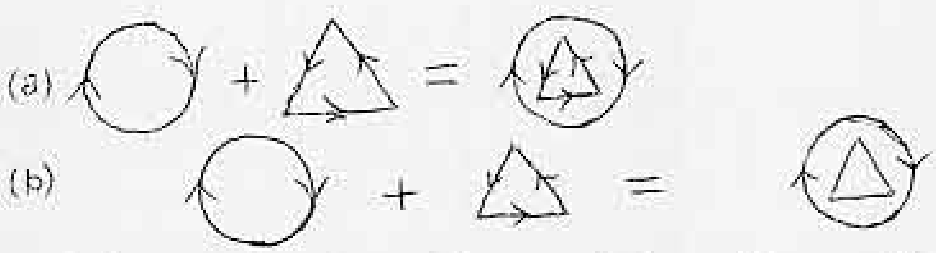
positive



negative

Our rule of combination, then, is addition. To add two figures, we impose one upon the other. For example, put each pattern on a separate sheet of paper, and then lay one sheet upon the other. Place the two sheets on a light table. The pattern showing through is the sum, and may be traced directly. Adding a positive to a negative figure which is otherwise identical results in a blank or an undirected figure, the counterpart of zero in the number system. (Although we normally think of two equal but opposite vectors as canceling each other and would expect the result to be no figure, it is perhaps useful to let an undirected figure act as a place-holder equal to blank paper.)

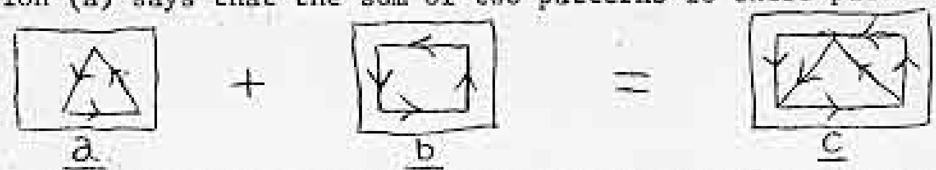
Two questions remain, however. First, what is the algebra? Alexander and Manheim look for the number of times shaded areas coincide, measuring coincidence by the intensity of the shading. They use the simple additive algebra we are accustomed to, $\underline{a} + \underline{a} = 2\underline{a}$. But a Boolean algebra is also possible: $\underline{a} + \underline{a} = \underline{a}$. That is, when we lay a pattern over an identical pattern the result is one pattern, not two. The lines are not twice as black as before. This first question of algebras suggests the second. When we talk about figures we may mean either the outline alone or the area within the outline. Their different additions are shown below.



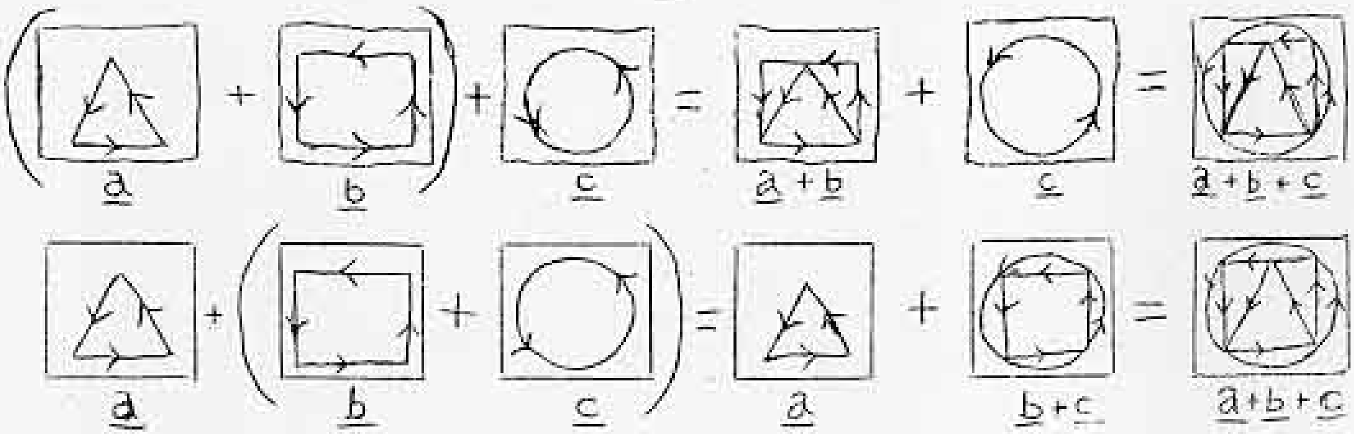
In case (a) we add just the outlines, and the sum is a positive triangle "inside" a negative circle. In case (b), the addition of areas, the sum is a negative circle with a triangular "hole" or zero-area within itself. The intersection of a positive and a negative figure is an undirected zero vector; that is, an undirected vector whose length is zero.

Returning to Keyser's criteria with all this in mind, let us take the simpleest case--adding outlines by Boolean algebra. The other cases are all analogous to this.

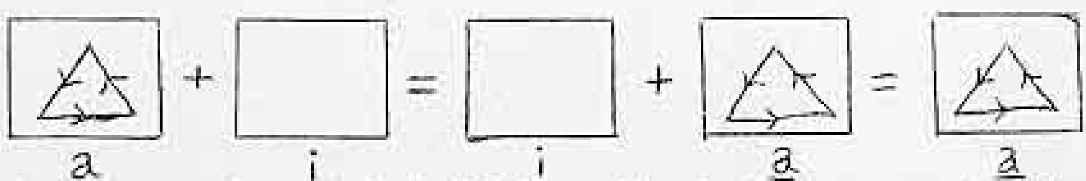
Condition (a) says that the sum of two patterns is third pattern.



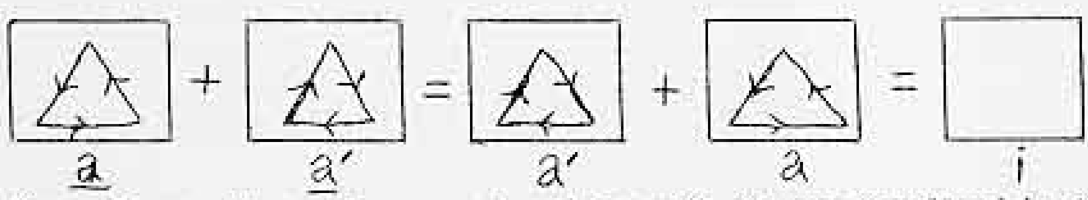
(B) says that overlaying a with b and then overlaying the result with c yields the same pattern as overlaying b first with c, then a.



(C) is fulfilled by using a blank paper as the identical member.



The reciprocal required in (d) is an identical figure except with a different direction.



There is an alternative way of working with figures developed by John Pfaltz and Azriel Rosenfeld for computer applications. It is conceptually simple, although not conveniently workable with a pencil and paper. "Any region can be regarded as a union of maximal neighborhoods of its points, and can be specified by the centers and radii of these neighborhoods." This set of centers of maximal neighborhoods often forms a centrally located stick figure and so is often called a "skeleton." There are algorithms both for converting regions (or figures) to skeletons and for regenerating regions from skeletons. Pfaltz and Rosenfeld describe the set-theoretic operations for determining the union, the difference, and the intersection of two regions. The identical member here is a skeleton of centers with zero radius. Thus, these skeleton figures are also a Group.

The immediate point of this paper has been that figures are a group. But more importantly, what I have hoped to show is this: we geographers can handle shapes in an exact, simple, and direct way. When we talk about patterns on the map we mean just that, but we have almost always either spoken of shapes in a vague and abstract way, or, attempting to be more precise, converted shapes to abstruse mathematical functions. But there is no point, no gain, in moving away from our data before it is collected.

X X X X X X X X X X X X X X

The last paper is filled with the traditional cry of pain over the basic disarray that prevades the logic of geography. But to the editor it raised an interesting question. Is the four color problem "really" geographic?

Two Theorems for Geography

Richard Guyot

Geography has forsaken its core for its adjectivally applied fields: physical, economic, urban, cultural, perceptual, ad. nauseum. The literal meaning of geography is: earth + graphe (description). According to Cassirer, the philosopher, the core of description is some generic concept (1923). The description and study of form is morphology, therefore the morphology of things on the earth is geography. Science is a way of knowing and the science of geography is the study of patterns as morphological laws (Schaefer, 1953, p. 226-49).

Geometry is nominally the "measure" of, or a calculus for figures "on the earth". There are actually many kinds of geometries, some are metric, others, such as topology, are relational. If patterns can be conceptualized by metrics or relations then there is a need to accumulate useful morphology theorems under a geographic geometry.

Geographic geometry can be discerned in the stacking of airplanes over airports and pistons in the Ford Rouge Plant. Few geographers do field work let alone see the pure spatial implications in sub-assembly and assembly lines. The dendrite can be discerned in sewers, trees, commuters (Warntz and Bunge, Geography, The Innocent Science, 196?) and in crystal growth in metals (Bell Telephone Laboratories' ad in Scientific American, Sept., 1967, p. 33).
Time and motion studies could be called applied geography. For some reason geographers rarely consider anything shown at a scale larger than 1:24,000. They seldom consider anything with an R.F. greater than one or anything that occurs indoors. Since geography can't logically be a jack-of-all-trades it cannot be a study of everything on the surface of the earth. If geography is the study of processes on the surface of the earth then it must contain practically all knowledge. If it is interactions then it should logically be all interactions. It should not be the study of everything no one else wants to study (e.g., climates, superficial sociology, and left-over landforms from geology). Thus geography can fall in less academic disfavor by proceeding logically with the use of patterns and other spatial constructs for description. The solution to geography's lack of academic legitimacy is to develop a core field of theoretical or general geography. This geography must be made up of patterns of elements.

In reality this has been done through terms as Mackinder's Heartland, Colby's Centrifugal and Centripetal (which have been renamed gravity models), and time honored site and situation to cite a few examples not used by "mathematical geographers". The new generation of mathematical geographers neither recognizes its predecessors' accomplishments nor passes much beyond neanderthal mathematics as compared to the level used by freshman engineering students. A quantitative flash-in-the-pan resulted from Sputnik in 1957 when all the humanities lost out to psuedo-science. The mathematical geographers have failed because they discovered well-known statistical techniques (to the social sciences) and borrowed pieces of theory without building generic concepts.

It is wrong to require a "new" science to provide all the answers instantaneously. However, it is also wrong to allow people to feel they must mathematize course titles without providing them the elements of a logical system. Mathematics is but one logic system. One can't "prove anything with statistics". Statistics and mathematics are abstract tools. Any fault or gain lies in the intellectual work done in applying abstractions.

Such criticism of quantitative geography has already been made by many. "The changing of the key concept is in itself more likely to be a recombination of older knowledge rather than something completely new" (Bunge, "Simplicity," 1968). The "old geography" will work provided there is at least a framework. Geographic knowledge must be set down in the form of definitions and axioms from which geographic theorems can be built.

As a start, what follows is an attempt to define mapping and map features. In the spirit of the Micmog Discussion Papers two working theorems are presented without formally determining proper axioms. Only through further work can the definitions used be worked up into good axioms.

Definitions

A dimension is normally defined in dimension theory as "... $\leq$ n if an arbitrarily small piece of the space surrounding each point may be delimited by subsets of dimension $\leq$ n-1." The dimension 0 is bounded by the empty set $\{-1\}$ (Hurewicz and Wallman, 1941, p. 10-24). This definition is not rigorous enough for mathematicians but will provide a guide for an intuitive geographic definition. In geography, the definition of an object (herein called a figure) of n dimensions is that it is bounded by a figure of dimension n-1. For the sake of visual conceptualization only, a figure may be considered an object. For example, a line is bounded by points. Areas are bounded by lines. Volumes are bounded by surfaces (Warntz and Bunge, Geography the Innocent Science, The dimension $n < 0$ will be defined later. All dimensions are integers.

Each n-figure may be considered as an uniform region according to set theory or Bunge (1966, p. 14-26). For example if an agricultural area ($n = 2$) is divided by a line (n-1) it is implied that a difference exists on either side of this line. This line is a common boundary between two different types of agriculture such as wheat and corn. Therefore each common boundary of dimension n-1 separates distinguishable figures (regions). If two figures are not distinguishable then their separation is not distinguishable. That is, a partition that bounds nothing does not exist. Conversely if two figures are not distinguished by a partition then they are uniform throughout and exist as one figure only. This concept has already been established in Hudson's Unit Area method. It only has to be expanded to n dimensions.

The Meeting Theorem

This theorem only states the upper limit to the number of dimensions that figures might meet in. Meeting is defined as the contact in common in the form of a mutually inclusive figure of dimension n_c (called a meeting figure) between figures of all the same dimension n. That is, a figure of dimension n_c exists such that every part of it will be simultaneously common to m figures, all of the same n.

Where: n_c is the dimension of any meeting figure.
m is the number of figures, all of the same n, that are meeting.

n_{c_m} is the maximum dimension of a meeting figure for m, n dimensional figures. This maximum dimension is the highest absolute value of n_c. Absolute value is the highest integer regardless of sign.

Everywhere in the meeting figure is common to all m of the n dimensional figures.

For example: two lines ($n = 1$) meet in a point ($n_{c_2} = 0$), two surfaces ($n = 2$) meet in a line ($n_{c_2} = 1$), and two volumes ($n = 3$) meet in a surface ($n_{c_2} = 2$). Therefore, when $m = 2$ the maximum dimensioned meeting figure between <u>two</u> n-figures is n-1. This can be assumed since the highest dimension that is common to both n-figures is their n-1 boundary. Further, <u>one</u> figure of dimension n meets itself in itself ($n_{c_1} = n$).

Parallel lines, surfaces and other similar cases in higher dimensions obviously don't meet. However two lines lying together, having no "width" (n-1), must meet in their length. If they are both the same length then they can be classified as the same identical line. If a line is laid along the middle of a longer line the meeting theorem delimits three different segments: the left end of the long line, separated by a point (n-1 boundary) from the end segment is the short line segment (which is identical for both lines), and after another $n-1 = 0$ point the right end of the long line. This is simply Hudson's unit "area" method in one dimension.

Since two figures of n dimensions meet in $n_{c_2} = n-1$, then <u>this</u> n-1 meeting figure will meet the n-1 boundary of <u>a third</u> figure of the same n. Such that:

$(n-1)'$ meets $(n-1)''$ in $(n-1) -1 = n-2$. Thus $n_{c_3} = n-2$.

Since:

$$n_{c_1} = n$$
$$n_{c_2} = n-1$$
$$n_{c_3} = n-2.$$

By induction:

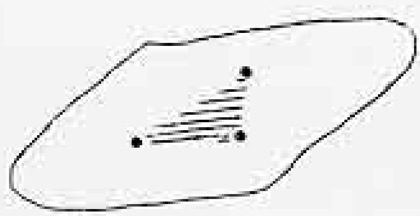
$$n_{c_m} = n - a, \quad \text{where } a = m-1.$$
$$n_{c_m} = n - (m-1)$$
$$n_{c_m} = n - m + 1.$$

It is useful to construct a table of the values of $n-m+1$ for ordinary values of m and n:

Where: n is the dimension of the figures.
m is the number of figures of the same n that are meeting.

m \ n	1	2	3	4	5	6 .... m
0	0	-1	-2	-3	-4	-5
1	1	0	-1	-2	-3	-4
2	2	1	0	-1	-2	-3
3	3	2	1	0	-1	-2
4	4	3	2	1	0	-1
.						
.						
.						
n	n	n-1	n-2	n-3	n-4	n-5 ...... n-m+1

After n_{c_m} approaches zero the dimension of meeting becomes negative. A negative dimensioned figure is defined as an imaginary figure, much like the imaginary number i. For simple combinations it is possible to show the need for imaginary figures. Given three points, $n_{c_3} = -2$. Three points determine the shaded surface (Fig. 1a) which will "connect" or could "meet" these three points. Given three lines, $n_{c_3} = -1$. This implies (fig. 1b) that there "exists"



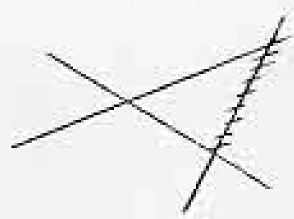


Figure 1

at least one imaginary meeting figure of -1 dimension which will be common to ("connect") the three lines (one of the three possible ones is shaded).

The imaginary dimensions of n_{c_m} provide the highest absolute valued dimension which could connect any given number of figures of the same dimension n. For example four points could be co-planar and would only need a -2 surface to connect them. At the worst, the fourth point could be not co-planar and a -3 dimensional imaginary figure would be common to the four points. This agrees with the theorem which states $n_{c_4} = -3$ is the highest dimensioned imaginary figure needed for <u>any</u> four points. The same can be done intuitively for lines and surfaces.

The range for real meeting figures ($n_c > 0$) is: $n \geq n_c \geq 0$. The range for imaginary meeting figures is: $0 \leq n_c \leq n_{c_m}$ (assuming the absolute value of the imaginary dimensions is taken). For example, eighteen volumes might meet in a line (the axis of a cake cut in eighteen pieces). However, no eighteen volumes can meet in a surface (n-1). They can meet in a point or "less" (imaginary figures). Thus any number of figures can meet in the <u>boundary of a boundary</u> or n_{c_m} for $m \geq 3$. Only one n-figure can meet in itself $n_{c_1} = n$.

The dimension of meeting for figures of different n is the n_{c_m} for the lowest common n. Thus if a line, a surface, and a volume all meet they will meet according to n_{c_m} for three lines.

A mathematical group must be both associative and commutative with respect to operations. Meeting figures are commutative but not associative. For example:

Where: o indicates the operator of "meeting".

$$n' \circ n'' \Rightarrow (n-1)', \text{ and } n''' \circ n'''' \Rightarrow (n-1).$$

Combining these two n-1 figures associatively:

$$(n-1)' \circ (n-1)'' \Rightarrow (n-1) - 1 = n-2.$$

But, $n_{c_4} = n-m+1 = n-3.$

Further, each time one sub-set of m, n-figures is combined with another subset, a degree of freedom is lost. n_{c_m} holds for any m regardless of configuration or constraints. As can be seen above, only with certain configurations can four volumes meet in n-2 (a line). For example:

$$n^1 \circ n^2 \Rightarrow (n-1)' \searrow$$
$$(n-1)' \circ (n-1)'' \Rightarrow (n-2)^1 \circ (n-2)^2 \Rightarrow (n-2)-1 = n-3$$
$$n^3 \circ n^4 \Rightarrow (n-1)'' \nearrow$$

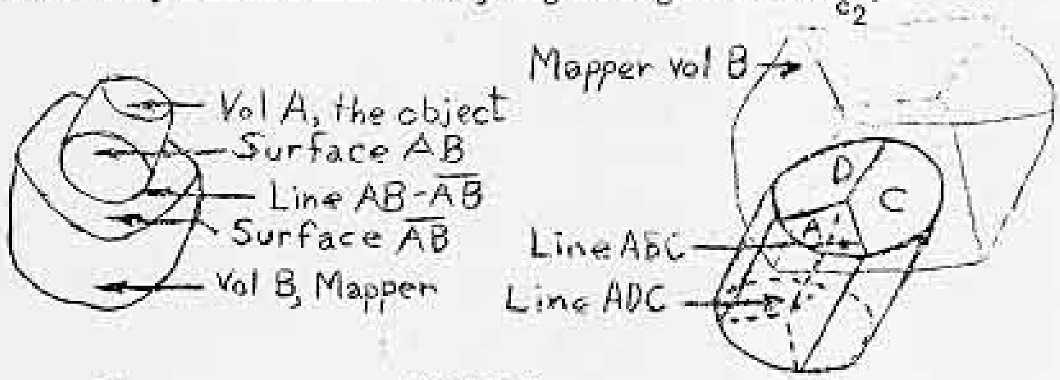
(lowest common dimension)

$$n^5 \circ n^6 \Rightarrow (n-1)''' \longrightarrow (n-1)'''$$

But, $n_{c_6} = n-m+1 = n-5$. Therefore <u>any</u> six volumes may be connected by an imaginary figure of dimension no larger than $n-5 = -2$. However only a certain six volumes can meet in n-3 = 0.

The Mapping Theorem

Mapping is the recording of distinguishable objects such as occurs in the "unit area method". A map is defined as all or part of the n-1 figures which bound any n-figure that has been designated as a mapper. Any n-figure can be a mapper. A mapper of dimension n will map or record in the form of meeting figures all n-figures with which it meets. Therefore, again in the identity case, a map of a map is a map, since the meeting figure for n_{c_1} is n. A map is the common figure or interface between figures of the same n that meet. For example two volumes (n=3) A and B meet in a common surface (n_{c_2} = 2). If volume B is the mapper its n-1 surface will be the map on which m-1 other three-dimensional "objects" will touch. The map will show the surface of B composed of two regions: the common surface AB and the surface "not AB" (Fig. 2a). The regions AB and $\overline{AB}$ (read not AB) being surfaces have their own meeting figure, (n-1-1) = n-2, line AB-$\overline{AB}$. This line is common to both surfaces and also common to both volumes. This illustrates the distinction of the meeting figure n_{c_m} as the maximum dimensioned figure which is mutually inclusive. Two volumes could meet in a line or a point but they can not meet in any figure higher than n_{c_2}.



a Figure 2 b

If three volumes A, B, and C, with B the mapper meet, then the n_{c_3} meeting figure will be line ABC (Fig. 5b). Since the surface of mapper B not touching A or C is feature-less, it may be disregarded. Thus a map feature is defined as the meeting figure which defines (recognizes) any change in the mixture making up a common mapper surface. Where surface feature AB changes to surface feature CB the feature-bounding line ABC exists. Figure 2b can be expanded to include four volumes, three "objects" plus the mapper B. The mapper will be excluded from the description but will be implied in calculating m for n_{c_m}. The three volume-objects A, C, and D all meet in line ADC. Since $n_{c_4} = 0$, three volumes can meet the mapper, the fourth volume, in a point. These three volumes form three different surface features or regions. Therefore these three regions meet in a maximum dimension of $n_c = 0$. Any three or two surfaces may meet in a line but these regions are bound by the constraint that they are the boundary of four volumes meeting at a point. The n-1 features must follow the constraints placed on them by the set of m, n-figures.

Implications of the Mapping Theorem

Most usual maps are two dimensional maps of the interface between the air volume (mapper) and the earth volume. Corn and wheat are actually three dimensional crops. A map of water, corn and wheat with air the mapper is shown in Figure 3.

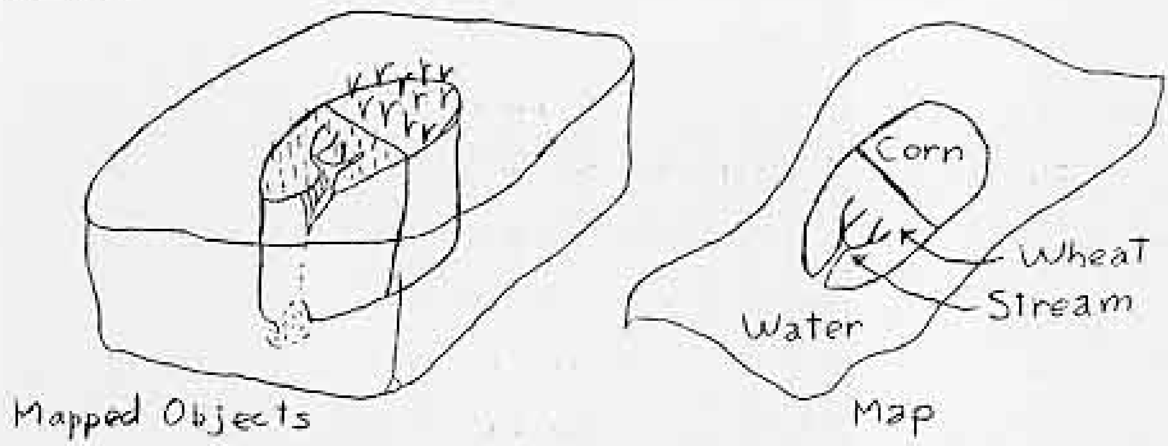


Figure 3

A contour "map" is simply the air volume mapping its interfaces with interval elevations volumes. From sea level to, say, 10 ft. is one volume. Where this volume and the water and air volumes meet is a line called the shoreline (Warntz and Bunge, Geography, The Innocent Science, 1967). The next line is where the 0-10 ft. volume meets the 10-20 ft. volume, and so forth. The resulting two dimensional surface is a bit lumpy but it is a surface and can be flattened by transforming it (projecting on) to a flat surface.

A geologic cross section (n=2) and a planimetric map (n=2) meet in a line (n_{c_2} = 1) (Fig. 4).

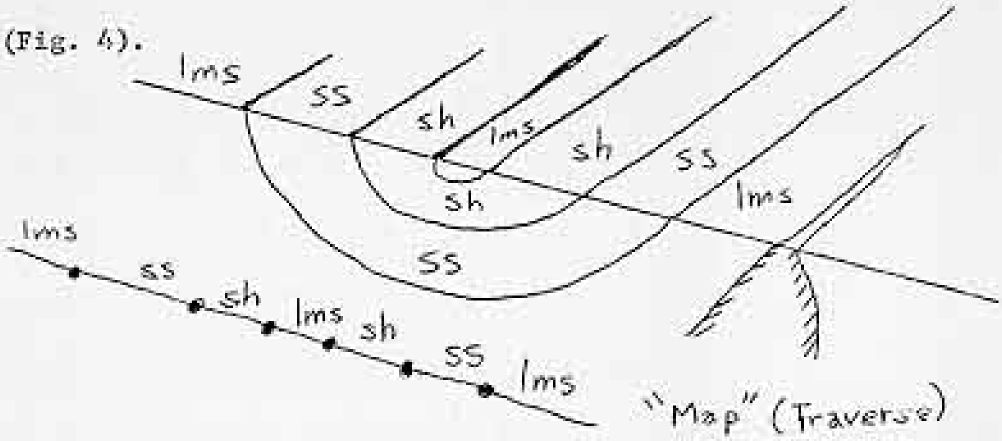


Figure 4

Imaginary figures (n<0) are maps. Given a group of points a planimetric map can be made. If the points have data associated with them a contour map (n=2) or even a three dimensional model can be made. The mapper volume for an aerial photo is the cone of reflected light from the objects that "meet" it. Oil and coal seams can be mapped using limestone for the mapping volume. Mariners sometimes still use an "armed lead" (sounding lead with a tallow filled depression in the base) to record the sea bottom. The lead is the mapper, the tallow is the meeting figure and it picks up sand, mud, or nothing if the bottom is bare rock.

Assume Figure 4 is a block diagram and an intrusive dike (which is a volume) touches the air volume in a line. The surrounding mantle rock intervenes so that the maximum meeting figure n_{c_3} = 2 is not realized. This line begins at a point and ends at a point but does not partition the earth's surface. However it is, in fact, a boundary of a boundary (n-2). The mapper only maps n-1 interfaces. Thus lines on maps may be from volumes or from surfaces intersecting. The stream in figure 3 changes from lines to an estuary. The stream is still everywhere an earth (wheat), air, and water interface.

If lines and areas can be explained as meeting figures, then patterns are meeting figures. Patterns in two dimensions will have counterparts in n dimensions. Therefore all objects may be mapped in some way. Table 1 shows that two 3 dimensional figures and three four dimensional figures both meet in two dimensions. In fact, sixteen 17 dimensional figures also may meet in two dimensions. This implies that the same 2 dimensional map may represent an infinite number of objects of the appropriate dimensions. The same map patterns and morphological laws might run through a wide number of interacting objects.

Coloration and the Four Color Problem

The four color problem suffers in part because of a lack of definition for coloring. The following definition is offered in the context of distinguishable objects.

Although a line distinguishes two regions, both regions cannot be colored the same color because it would imply they are identical (Cayley, 1879). Different colors would improve the distinguishing characters of lines. The same color could be used again if regions could be particularly well distinguished. A line bounds regions, a point bounds lines, thus a point bounds a boundary. A boundary of a boundary should provide the necessary distinguishing power for using the same color.

If four regions, corn, oats wheat, and rye meet at a point (Fig. 5a) then the corn-oats boundary changes to a corn-rye boundary at a n-2 bounding point.

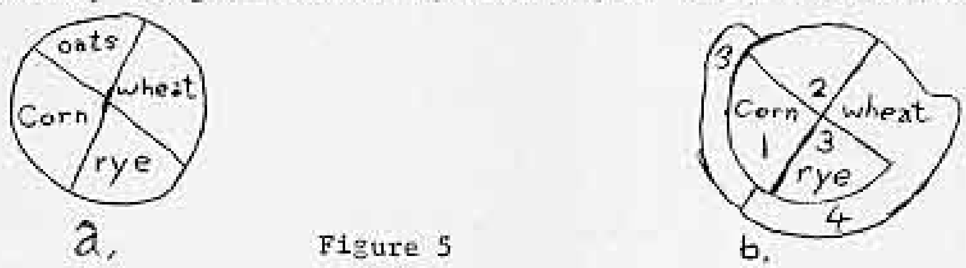


Figure 5

If the n-1 boundary of the corn does not share any n-1 boundary with the wheat then both may use the same color. It can be seen that the oats and rye intervene when the meeting figure is n-2 (the center point). Thus only three colors are required provided the m+1th region does not share a n-1 boundary with whichever color is available.

It is possible, as in Figure 5b, that a 4th region may share some of the n-1 boundary at some place removed from the n-2 meeting figure. This contingency may be handled by allowing four colors to be used. Since the last region does join up due to its configuration (Fig. 5b) the set of m figures is meeting associatively. Each time the m+1th region, due to configuration, meets associatively with a "distant" region one degree of freedom is lost. If no foresight is used, an infinite number of colors could be used.

On a line map, made from intersecting surfaces, the number of colors needed may only be two. Such a map could be made from a string with alternate black and white yarn segments. If the number of segments is even then the ends of the string may be tied together. If the number of segments is odd then a black will be tied to a black unless one black is changed to a third color. This is because meeting cannot be done associatively and maintain a maximum number of choices.

Four colors may be necessary on a finite surface. Since the surface of a cylinder joins itself four colors are also needed at most. The typical solution on a sphere is to cut a hole in the sphere or expand a point. However, this hole or point must be colored since a solid sphere (or hollow ball, like a geoid) is the mapper. A torus requires six colors since the inside of the hole permits an additional joining confrontation. Needing more than 6 on a torus, 5 on a sphere, 4 on a finite surface, or 3 on a line is due merely to careless choosing of colors. Two or more starting points for the coloring will require additional colors as does starting at one point and returning to that point a multiple number of times. Each return to the starting point may require an additional color.

Conclusion

Arguments of primacy of labor seem fruitless. Who is the queen to whom, mathematics to science, science to philosophy? It is like asking which part of the watch is more vital? Which parts of a watch which if they were removed would still allow the watch to function? Geographers seem to clearly need the skills of philosophers. Hopefully we have raised interesting philosophical questions for why else would they bother with us? The commerce should be mutually helpful.

The simplest review is a listing of the questions raised by the authors. Tobler asks, "What, if anything, is ultimately invarient in science?" Pattison's question is "If reality is not real, what is?" Warntz asks, "If even point-set has literal meaning, how abstract are abstractions?" Bunge's first question is, "What is implied if the concept of pure spatial prediction has validity?" Toulmin partially answers his own question, "In what sense is the map a theory?" Martin wants to know, "What, if anything, besides language can we read?" Bunge's second question reads, "How do we prove a truth?" Karlin implies the question, "What is a 'natural' language?" And Guyot asks, "Is the four color problem really geographic?" It is not likely that geographers will do as well solving philosophical problems as philosophers, helpful as the answers might prove to geography.

Bibliography

Alexander, Christopher, "The Coordination of the Urban Rule System," Center for Planning and Development Research, University of California, Berkley, July, 1966.

Alexander, Christopher, "The Most Stable Decomposition of a System Into Subsystem," submitted to Information and Control, 1963 (b).

Alexander, Christopher, Notes on the Synthesis of Form, Harvard University Press, 1963 (a), Appendix 2.

Alexander, Christopher, and Manheim, Marvin, The Design of Highway Interchanges: An Example of a General Method for Analysing Engineering Design Problems, Research Report R62-1, Cambridge, Mass., Civil Engineering Systems Laboratory, M.I.T., 1962.

Alexander, Christopher, and Manheim, Marvin, Hidecs 2: A Computer Program for the Hierarchial Decomposition of a Set Which has an Associated Linear Graph, Research Report R62-2, Cambridge, Mass., Civil Engineering System, Laboratory, M.I.T., 1962.

Alexander, Christopher, and Manheim, Marvin, The Use of Diagrams in Highway Route Location: An Experiment, Research Report R62-3, Cambridge, Mass., Department of Civil Engineering Systems Laboratory, 1962.

Alexander, Christopher, and Poyner, Barry, "The Atoms of Environmental Structure," Center for Planning and Development Research, University of California, Berkley, July, 1966.

Arnheim, Rudolph, Art and Visual Perception, Berkeley: University of California Press, 1954.

Bell Telephone Laboratories, Advertisement in Scientific American, Sept., 1967.

Boole, George, An Investigation of the Laws of Thought, London, 1854.

Bunge, William, "Simplicity" Unpublished article, 1968.

Bunge, William, Theoretical Geography, Lund, Sweden, 1966.

Cassirer, Ernst, Substance and Function, Chicago, Open Court, 1923.

Cayley, "On the Colouring of Maps" Proceedings of the Royal Geographical Society, New Monthly series, Vol. 1, 1879.

Gardner, Martin, Logic, Machines, and Diagrams, New York, 1958.

Gottschaldt, Kurt, "Gestalt Factors and Repetition," in D. Ellis, A Sourcebook of Gestalt Psychology, London: K. Paul, Trench, Trubner and Company, 1938, p. 109-135.

Hilton, Alice Mary, Logic, Computing Machines, and Automation, Cleveland, 1963.

Hurewicz, Withold and Henry Wallman, Dimension Theory, Princeton, Princeton Univ. Press, 1941.

Keyser, Cassius T., "Group Concept," The World of Mathematics, Vol. 1, (ed. James Newman), Simon & Schuster, 1956, p. 1538.

Lewis, Clarence Irving, A Survey of Symbolic Logic, Berkeley, California, 1918.

Mach, Ernest, The Science of Mechanics, 1942, p. 490.

Neft, David S., Statistical Analysis for Areal Distributions, Philadelphia, 1966.

Pierce, Charles Sanders, Collected Papers, 8 Vols., Cambridge, Mass., 1931-1958.

Ritter, Paul, Planning for Man and Motor, New York: MacMillan, 1964.

Ritter, Paul, "Social Patterns and Housing Layouts," Thesis prepared for the University of Nottingham, England, 1957.

Roberts, Paul O., "Using New Methods in Highway Location," Photogrammetric Engineering, June, 1957, p. 563-569.

Shaefer, Fred K., "Exceptionalism in Geography: A Methodological Examination," Annals, Association of American Geographers, 1953, p. 226-49.

Stark, Freya, The Journey's End, 1964, New York: Harcourt, Brace, and World, p. 97.

FejesToth, L., Regular Figures, New York: 1964, p. X.

Venn, John, Symbolic Logic, 2nd ed., London, 1894.

Warntz, William and William Bunge, Geography, the Innocent Science, forthcoming.

Webber, Melvin, "Order in Diversity: Community Without Propinquity," in London Wing, Jr. (ed.) Cities and Space, Published for Resources for the Future, Inc. (Baltimore: The Johns Hopkins Press, 1963), p. 23-54.