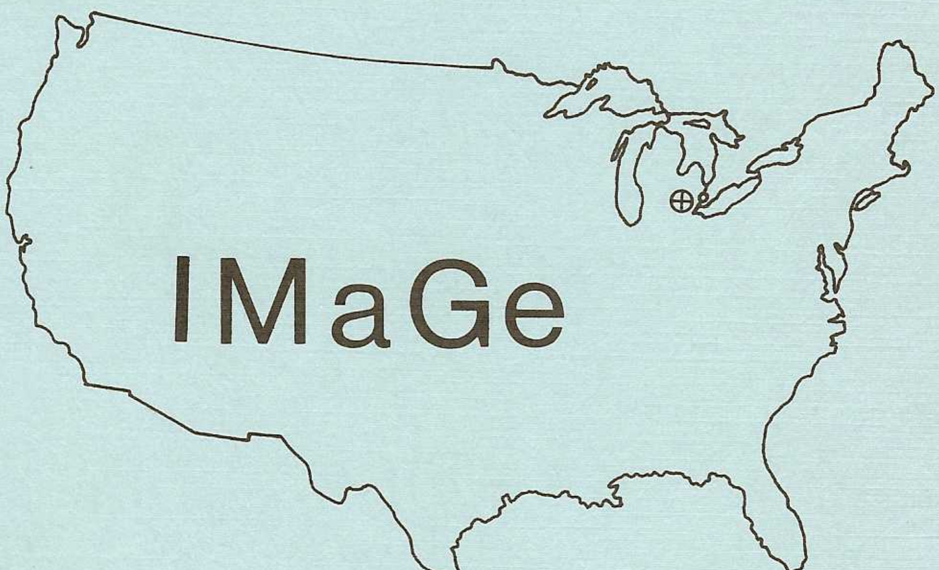
Institute of Mathematical Geography

MONOGRAPH SERIES

ESSAYS ON MATHEMATICAL GEOGRAPHY

by: Sandra Lach Arlinghaus, Ph.D.

Monograph #3



"IMaGe-in-nation"

? IMaGe

An MDS Publication

ESSAYS ON MATHEMATICAL GEOGRAPHY

BY

SANDRA L. ARLINGHAUS (Ph.D.)

Director
Institute of Mathematical Geography

INSTITUTE OF MATHEMATICAL GEOGRAPHY

MONOGRAPH #3

A MICHIGAN DOCUMENT SERVICES PUBLICATION

ANN ARBOR, MICHIGAN

1986

ESSAYS ON MATHEMATICAL GEOGRAPHY

Library of Congress Identification Number: DCLC86188450-B
Library of Congress Card Number: 86188450
Library of Congress Call Number: G70.23.H75 1986.

ISBN: 1-877751-06-5

ACKNOWLEDGMENT

IMaGe reviewer W. Arlinghaus read the entire monograph manuscript; IMaGe reviewer J. Nystuen read, individually, seven of the ten essays. H. S. M. Coxeter, Professor of Mathematics at the University of Toronto, read the first essay. To them I am deeply indebted for constructive commentary which has improved significantly the content of this monograph. Others over the years (mentioned in "References"), and members of the Colloquium in Mathematical Geography, have contributed interesting insights, periodically. Their combined effort has been critical in bringing about this monograph; omissions, errors in fact or in interpretation, or other blunders are, of course, mine alone.

Sandra L. Arlinghaus

Ann Arbor, Michigan

1986

CONTENTS

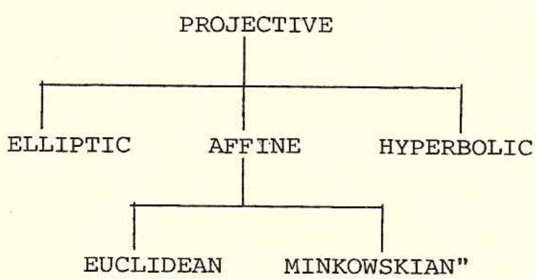
THE WELL-TEMPERED MAP PROJECTION*

"For there is a music wherever there is a harmony, order or proportion; and thus far we may maintain the music of the spheres; for those well ordered motions, and regular paces, though they give no sound unto the ear, yet to the understanding they strike a note most full of harmony."

Sir Thomas Browne, 1605-1682.

INTRODUCTION

Non-Euclidean geometries have served, most notably in application, to characterize abstract relationships underlying theoretical structure in physics; Hermann Minkowski's specialization of affine geometry resulted in the "space-time" geometry used in special relativity theory [Einstein, 1961, pp. 57, 122]. The geometer H. S. M. Coxeter identifies a geometric " 'genealogy' [in which] each geometry (save the first) is derived from its parent by some kind of specialization.



[Coxeter, 1965, p. 19]. Use of the Minkowskian metric is also of current concern to cognitive mappers who see advantages in employing it to represent more realistically the shape of perceived spaces in mental maps [Golledge and Hubert, 1982; Muller, 1984]. This paper, however, works within the broadest non-Euclidean geometry in Coxeter's hierarchy: projective geometry.

Generally, the transformational approach to using mathematics to explore geographic and cartographic relationships, as in the work of Waldo Tobler [Tobler, 1961, 1962, 1963], follows the direction taken by Felix Klein using the notion of a mathematical transformation to uncover global structure. Specifically, the Polish cartographer-geodesist, Franciszek Biernacki, understood the suitability of employing the transformational projective approach to analyzing various classes of map projections as he commented that "The problem of projective representation is the concern of a mathematical discipline--projective geometry." Unfortunately, however, he stopped short of execution [Biernacki, 1965, p. 297]. This article does align, through a sequence of theorems, perspective map projections with fundamental projective geometric concepts. The process of forming this theory presents opportunities for others to employ the full spectrum of theorems from projective geometry, which dates (as a systematic discipline) from the seventeenth century.

All non-Euclidean geometries treat the notion of infinity in a manner different from that of the Euclidean approach. Thus, while rigorous axiomatic development for these geometries is available in the mathematical literature, it is difficult for us to visualize these theorems in our Euclidean-trained minds; therefore, it will be useful, where possible, to interpret non-Euclidean relations in terms of abstract models that are easier to visualize. This essay will introduce a mathematically structured set of concepts from projective geometry and will apply them to perspective map projections, where the notion of infinity is the same as the projective geometric view of infinity. It will prove those projective geometry theorems that either are not easily available or that emphasize the differing view of infinity. The reader will be referred to appropriate literature for the proofs of others. The main result, the

Harmonic Map Projection Theorem (Theorem 7), will show how to use one of the fundamental transformations of projective geometry, that of harmonic conjugacy, to obtain one perspective map projection from another, totally within the plane of projection and without reverting to the sphere. It will thereby prove that the entire set of perspective projections may be derived in the projective plane from the subset of projections with centers of projection contained within the sphere of projection; it will reduce an unbounded set of possibilities to a bounded set. At a broader level, this alignment will suggest the advantages to be gained from applying this highly symmetric geometry, that does not distinguish the ordinary from the infinite, to real-world situations that exhibit some sort of symmetry in underlying relations and that embrace the concept of infinity as part of an attainable system of fundamental values.

NON-EUCLIDEAN CONCEPTS

Inversive geometry

Suppose that a circle of radius r and center O has been drawn in the Euclidean plane (Figure 1). Draw a ray OP emanating from O, and designate as P' the point on OP such that $|OP| \times |OP'| = r^2$. The points P and P' are called inverses with respect to this circle of inversion; for, if r=1 and $|OP| = x$, then $|OP'| = 1/x$ so that appropriate distances are reciprocals or multiplicative "inverses." The mathematical transformation indicated by this procedure has the effect of sending, simultaneously, points inside the circle to points outside the circle, points outside the circle to points inside the circle, and points on the circle to themselves. Thus the disc inside the circle is turned inside-out to cover the entire unbounded set outside the circle, while, at the same time, the disc is filled, with no overlap, by the entire collection of points outside the circle. The circle itself remains fixed under this transformation of inversion. Only the inverse of the center of the circle is unclear.

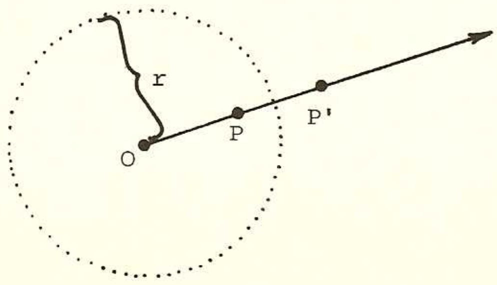


FIGURE 1

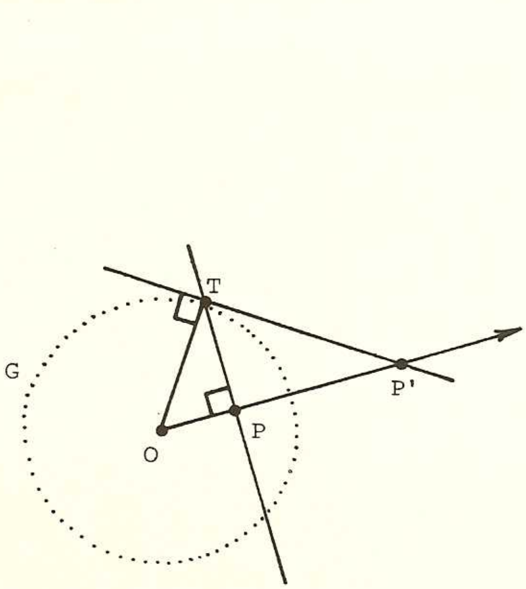


FIGURE 2

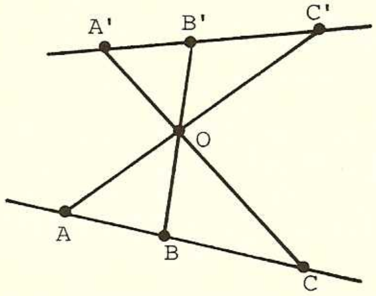


FIGURE 3

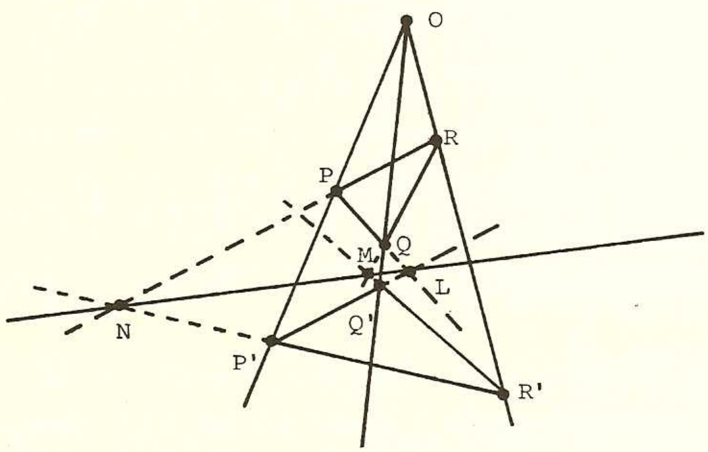


FIGURE 4

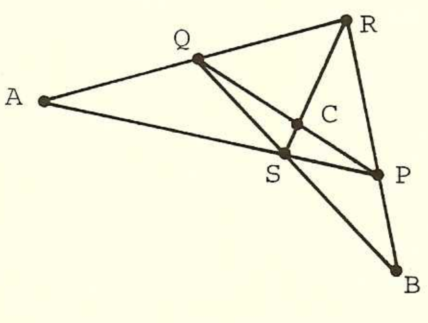


FIGURE 5

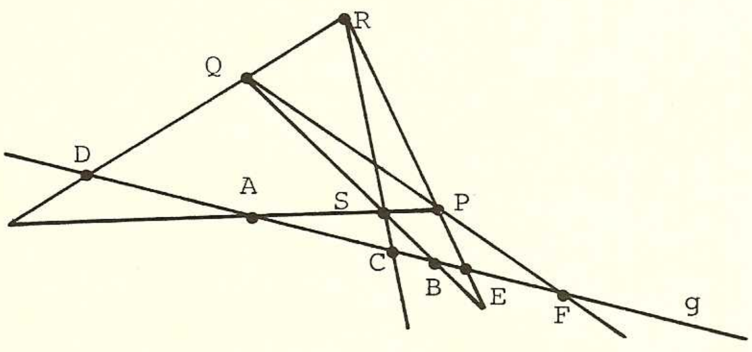


FIGURE 6

<u>Definition 1</u> [Coxeter, 1961, pp. 78, 84].

In a fixed circle of radius r and center O, in the Euclidean plane, suppose that P is an arbitrary point, $P \neq O$. Then the point P' on the ray OP is the inverse of P with respect to the circle if and only if $|OP| \times |OP'| = r^2$. If the Euclidean plane is extended by joining to it a point O', at infinity, to be used as the inverse of O, then this extended plane is called the inversive plane.

Thus any line passing through O contains also the point O'; this line inverts to itself although the only points fixed by inversion are the ones on the circle of inversion. (As a three-dimensional analogue, the intersections of meridians at antipodal "poles" on a sphere corresponds to the notion of rays emanating from O containing O'.) Any line not passing through O has ends at O' and so inverts to a circle passing through O. Abstractly the plane may be completely inverted; alternately, this transformation is called reflection in a circle. For a point P close to a mirror on the fixed circle reflects to P' which appears to be not far back in the mirror, while a point P closer to O reflects to a more distant position in the mirror [Lyndon, 1985]. The following construction will permit easy determination of inverses, given a point P.

<u>Construction 1</u> [Coxeter, 1961, p. 78].

To construct the inverse of a point P within a circle G, first construct the perpendicular to OP at P (Figure 2). Let T denote an intersection point of this perpendicular with G. Construct a tangent to G at T. The intersection point P' of this tangent with OP is the inverse of P with respect to G (of radius r) since $\Delta OPT \simeq \Delta OTP'$. For then $|OT|/|OP| = |OP'|/|OT|$, and since $|OT| = r$, $|OP| \times |OP'| = r^2$ as required by Definition 1. This construction reverses step for step if P is outside G.

The material above describes the basis of a geometry that exists in a space formed by extending the Euclidean plane by one extra point at infinity. It seems natural to ask next what happens if the Euclidean plane is extended by more than one point at infinity.

Projective geometry

The well-known example of parallel railroad tracks converging at the visual horizon, that is often used as a starting point in expanding the spatial awareness of elementary students, makes a convenient beginning for visualization in the projective plane. For, the possibility that parallel lines "meet" at infinity suggests the following extension of the Euclidean plane. Suppose that all lines parallel to a given line m meet at $\infty(m)$; suppose that all lines parallel to n (not parallel to m) meet at $\infty(n)\neq\infty(m)$. The line joining $\infty(m)$ to $\infty(n)$ will be composed of an infinite number of other points that may be viewed as intersection points of other families of parallel lines. These points at infinity will be referred to as "ideal" points, and the line consisting of all ideal points as the "ideal" line. When points and the line at infinity are not distinguished from Euclidean points and lines, the geometry of this Euclidean plane extended by the ideal line is the two-dimensional geometry of the projective plane.

It is a geometry that possesses a remarkable degree of symmetry; it may be studied from a coordinatized (analytic) approach or from a coordinate-free (synthetic) approach. Since selection of coordinates tends to introduce bias toward one coordinate scheme or another, and since such bias may be reflected in application, it is preferable to function in the synthetic, transformational, approach which reveals more clearly the power of the fundamental abstract ideas [Mac Lane, 1982]. The hallmark of synthetic projective proofs is mathematical elegance; proofs of theorems may appear "simple" once they are discovered, because they are clear. They are not, however, easy to find.

The extent of symmetry present in the projective plane may be measured abstractly by constructing a lexicographic model by which to interpret the following "metaprinciple" which is a principle about the entire set of theorems that can be proven in the projective plane.

Meta-principle--The Principle of Duality [Coxeter, 1974, p. 4].

"In the projective plane, in which ideal points and the ideal line are indistinguishable from other points and lines, every true statement about points and lines may be replaced by a corresponding true statement about lines and points."

Thus the entire geometry possesses a bilateral symmetry folded along an axis that provides for interchanging "point" concepts with "line" concepts. For example, since the statement "two points join to determine a line" is true, it follows by duality that "two lines meet to determine a point" is also true in the projective plane, but is certainly not true for parallel lines in the Euclidean plane. All that is needed to translate a true theorem to its projective dual is the appropriate dictionary, an abridged version of which is given below.

Term	Dual term
Point	Line
Join	Meet
Collinear	Concurrent
Quadrangle	Quadrilateral
Center	Axis

For a more complete discussion see Coxeter, Projective Geometry; only terms used in the subsequent development have been included in this partial listing. An implication of this principle is that if a particular geographic map may legitimately be characterized in the projective plane, its dual may immediately be derived from it.

<u>Definition 2</u> (dual words noted in parentheses) [Coxeter, 1974, pp. 10-12].

A perspectivity is a transformation in the projective plane that relates one set of collinear (concurrent) points (lines) to another through a center (axis) O (o) that joins (meets) each of the points (lines).

For example, in Figure 3, the set of points A, B, C is sent by perspectivity through O to the set of points A', B', C'. This is denoted $ABC \overset{O}{\bar{\bar{\wedge}}} A'B'C'$.

The following theorem, Desargues's two triangle theorem is frequently taken as an axiom of projective geometry, since proofs of it in two-dimensions generally rely on three-dimensional arguments. This theorem is due to Girard Desargues, a seventeenth century French architect, who together with the physicist Johannes Kepler, seized the significance of the point at infinity thereby moving this non-Euclidean geometry away from its descriptive origins in the fine arts toward a more rigorous abstract developemnt [Coxeter, 1974, pp. 2-3].

<u>Theorem 1</u> (Desargues's Two Triangle Theorem) [Coxeter, 1974, pp. 18-19].

If two triangles PQR and P'Q'R' are such that PP', QQ', RR' are concurrent at O, then the points of intersection, L, M, N, given as PQ · P'Q' = L, QR · Q'R' = M, and RP · R'P' = N (where "·" denotes intersection of lines) are collinear (Figure 4).

This remarkable theorem, useful in proving many other theorems, is stated here without proof. In addition to theorems about triangles, there are many theorems about quadrangles that rest on Desargues's theorem. Theorem 2 below is typical of this sort of theorem.

<u>Definition 3</u> (Dual words in parentheses). [Coxeter, 1974, p. 7].

A quadrangle (quadrilateral) in the projective plane is composed of four points (lines) P,Q,R,S (p,q,r,s) and their joins (meets) PQ, RS, PS, SQ, QR, RP (p·q, r·s, p·s, s·q, q·r, r·p) (Figure 5). The remaining meets (joins), A,B,C (a,b,c) are called the diagonal points (lines) of the quadrangle (quadrilateral). (This is not true in the Euclidean plane for quadrangles with a pair of parallel sides.)

Definition 4 [Coxeter, 1974, p. 20].

If a line g cuts across a quadrangle, not through one of the vertices P,Q, R, or S, it will generally contain six intersection points, one from each side of the quadrangle. This set of intersection points is called a quadrangular set; in Figure 6 the points A,B,C,D,E,F form a quadrangular set denoted (AD)(BE)(CF) where notational pairings reflect opposition in sides of the quadrangle.

Construction 2 [Coxeter, 1974, p. 21].

Any five collinear points, A,B,C,D,E on a line g may be viewed as belonging to a quadrangular set from which a sixth, F, may be constructed (Figure 7). To do so, draw a triangle QRS so that QS passes through B, RS passes through C, and RQ passes through D (the choice of these lines is arbitrary as long as they are not concurrent). Construct P = AS · ER, so that F = g · PQ.

Theorem 2 [Coxeter, 1974, pp. 21-22].

"Each point of a quadrangular set is uniquely determined by the remaining points."

The formal proof of this theorem rests on Desargues's theorem and its converse. Experimentation with the construction shown in Figure 7 displays the plausibility of this theorem.

Quadrangular sets generally contain six points; however, they may contain as few as four points if the line g passes through two of the diagonal points. This particular specialization leads to the following definition.

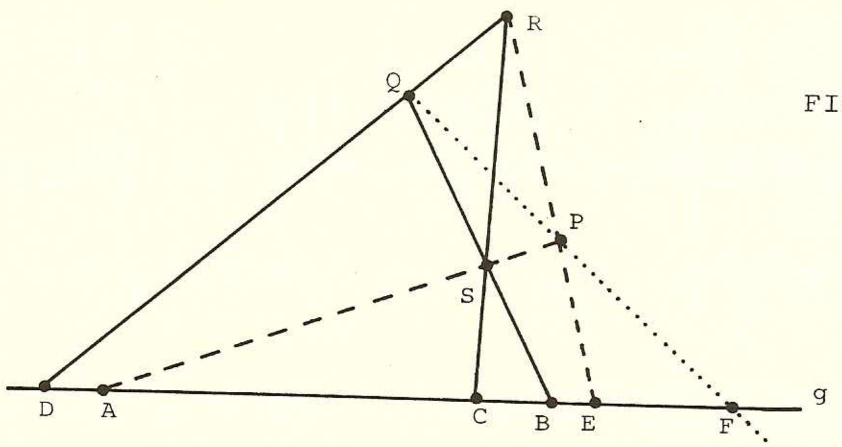


FIGURE 7

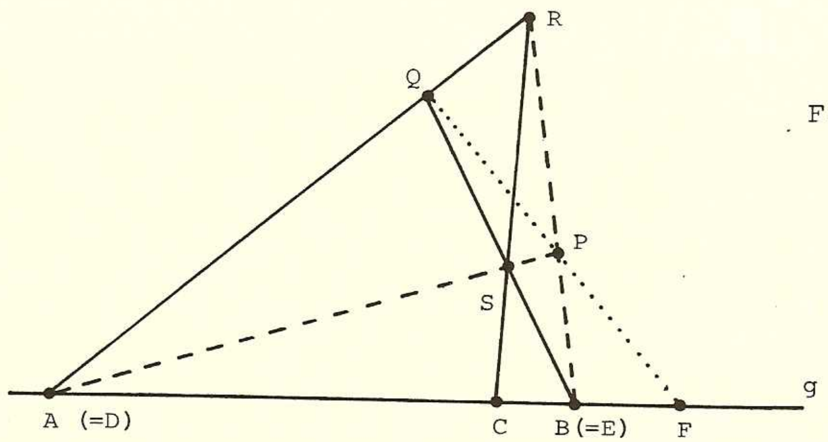


FIGURE 8

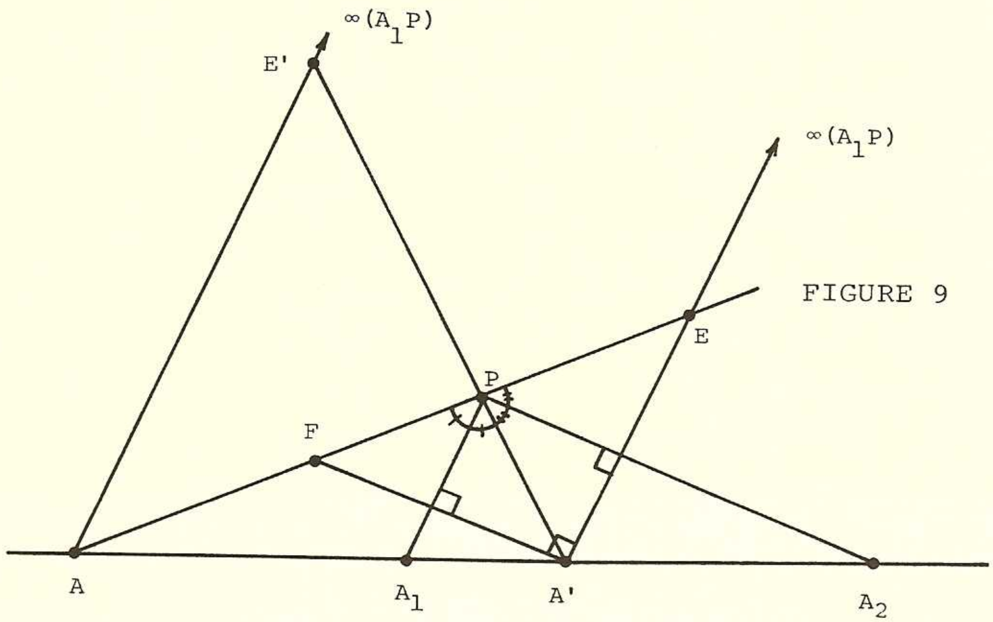


FIGURE 9

Definition 5 [Coxeter, 1974, p. 22].

"A harmonic set of four collinear points is the special case of a quadrangular set when g joins two diagonal points of the quadrangle." (Figure 8). This is denoted (AA)(BB)(CF) and F (or C) is called the harmonic conjugate of C (or F) with respect to A and B.

Construction 3 (due to Philippe de La Hire, 1640-1718) [Coxeter, 1960, p. 18].

Any three collinear points A,B, and C on a line g may be viewed as belonging to a harmonic set from which the fourth may be constructed (Figure 8). To do so, draw a triangle QRS so that QS passes through B, RS passes through C and RQ passes through A. Again, construct P = AS · BR so that F = AB · PQ. Theorem 2 ensures uniqueness of the construction.

Harmonic conjugates are critical to the development of projective geometry because they guarantee that there are at least four points on every line. It will be useful to have properties about harmonic conjugates; first, it will be noted that harmonic conjugates are invariant under perspectivity, and second, it will be proved that the notion of harmonic conjugacy can be expressed in a metric form. Further it will be of interest to note that La Hire not only is credited with Construction 3, but also that he developed perspective projections from centers of projection in a variety of positions [Steers, 1962].

Theorem 3 [Coxeter, 1974, pp. 28-29].

If $ABCF \overset{O}{\overline{\overline{\wedge}}} A'B'C'F'$ and if F is the harmonic conjugate of C with respect to A and B, then F' is the harmonic conjugate of C' with respect to A' and B'.

Theorem 4 (metric representation for the harmonic relation)

Given four collinear points A, A', A_1, A_2. They form a harmonic set, $(AA)(A'A')(A_1A_2)$ if and only if $(AA_1)/(A_1A') = (AA_2)/(A'A_2)$. [Coxeter, 1961, pp. 242, 88-89, exercises).

Proof:

A) Suppose $(AA_1)/(A_1A') = (AA_2)/(A'A_2)$ (Figure 9).

Let P be any point not on AA' so that the internal and external bisectors of <APA' meet AA' in A_1 and A_2. Locate E and F on AP so that $A'E||A_1P$ and $A'F||A_2P$. Since interior and exterior angle bisectors are perpendicular, $<A_1PA_2 = 90°$ as are the remaining angles of the parallelogram containing this angle. Therefore (FPA') and (A'PE) are isosceles triangles with vertex angles at P, so that FP = A'P = PE.

Since $\Delta AA_1P \simeq \Delta AA'E$ it follows that $(AA_1)/(A_1A') = (AP)/(PE)$ and since $\Delta AA'F \simeq \Delta AA_2P$ it follows that $(AA_2)/(A_2A') = (AP)/(FP)$. Thus, since PE = FP, $(AA_1)/(A_1A') = (AA_2)/(A_2A')$ as in Figure 9.

It remains to show that A, A', A_1, A_2 form an harmonic set. To do so, we show that they form a quadrangular set on a line passing through two diagonal points of a quadrangle. To this end, label as E' the point on A'P that produces $AE'||A_1P$.

Then reflection in A_2P transforms A_2A into A_2E' since A_2 lies on the axis of the reflection and since $AE' \perp A_2P$ with A and E' equidistant from A_2P on opposite sides. The point A' lies on AA_2 and the point A' is transformed to the point E by reflection through A_2P. Thus E lies on A_2E'.

Therefore A_2 is the harmonic conjugate of A_1 with respect to A and A' as these form a quadrangular set, derived from the quadrangle with vertices E', P, E, $\infty(A_1P)$, on a line passing through diagonal points A and A'.

B) Suppose A_1 and A_2 are harmonic conjugates with respect to A and A'.

Use the projective construction of Construction 3 to find the harmonic conjugate of A_1 with respect to A and A' (shown below the line in Figure 10).

Let P be any point on the circle of diameter A_1A_2 so that $<A_1PA_2 = 90°$. Label, respectively, E (E') as the intersection point of the line through A', (A)

parallel to A_1P, with AP (A'P). Label F as the intersection point of the line through A', parallel to A_2P, with AP. The parallelogram with two vertices at A' and P must be a rectangle since $\angle A_1PA_2 = 90°$ (Figure 10). Now the configuration above the line in Figure 10 is identical to that in Part A of this proof. Since A_2 is the harmonic conjugate of A_1 with respect to A and A', the points E', E, A_2 are collinear by the uniqueness of the harmonic construction given in Theorem 2.

Thus this harmonic set, determined projectively, is the same as the set determined affinely (with the "infinite" quadrangle) in which the ratio in which A_1 and A_2 separate A and A' is such that $(AA_1)/(A_1A') = (AA_2)/(A'A_2)$. Q.E.D.

Corollary 1 [Coxeter, 1961, p. 89].

The circle with diameter A_1A_2 from Part B of Theorem 4 inverts A to A' (the "circle of Apollonius").

Proof: (Figure 10)

Since $\angle A_1PA_2 = 90°$ as in Theorem 4, P lies on a circle of diameter A_1A_2; call its center O and its radius r. To show that A and A' are inverses, we need, by Definition 1, to show that $|AO| \times |A'O| = r^2$. Since $(A_1A)/(A_1A') = (AA_2)/(A'A_2)$ by Theorem 4, and since $(A_1A)/(A_1A') = (|AO| - r)/(r - |A'O|)$ and $(AA_2)/(A'A_2)$ $= (|AO| + r)/(|A'O| + r)$ it follows from cross-multiplying, that $|AO|\cdot|A'O|$ $+ r|AO| - r|A'O| - r^2 = r|AO| - r|A'O| + r^2 - |A'O|\cdot|AO|$ or $|AO|\cdot|A'O| = r^2$ as desired. Q.E.D.

One example of an harmonic set comes from music theory; the positioning of musically harmonic triads may be expressed in terms of length of string plucked [Coxeter, 1974, p. 23, exercise 6]. If a stretched string is fixed at point O, and if it plays C at a distance 15 units away from O, then it plays E when stopped

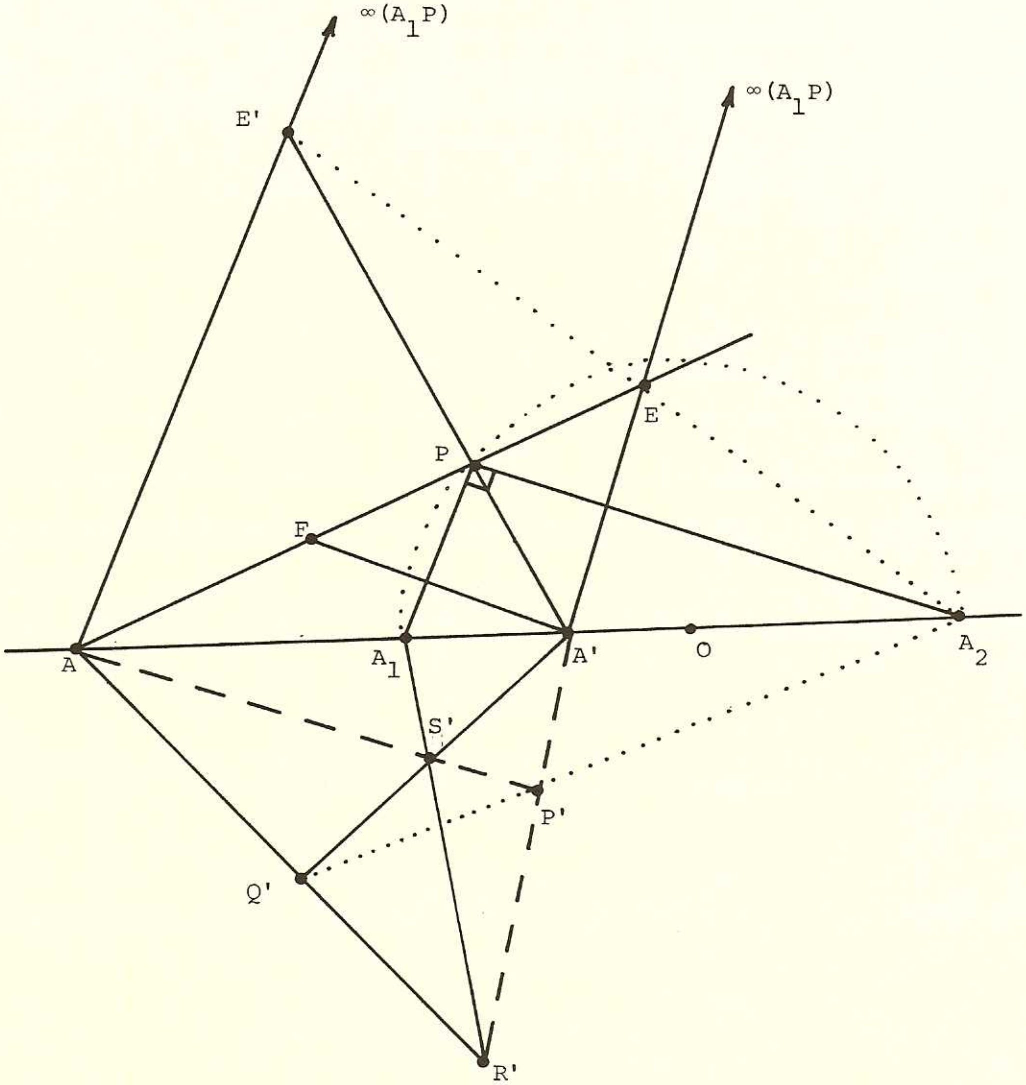


FIGURE 10

three units back from C and it plays G when stopped 5 five units back from C. The three notes G, E, C form, musically, a major triad, and projectively, G and C are harmonic conjugates with respect to O and E. For, by Theorem 4, 2/3=(GE)/(EC) =(GO)/(CO)=10/15, demonstrating metrical, as well as musical, harmony. Another example will appear below, and what will emerge is that the infinite set of perspective projections, with center of projection within the sphere of projection, has as harmonic conjugates perspective projections, with center outside that sphere, forming a closed system of map projections in the projective plane.

AN APPLICATION TO MAPS

Perspective map projections use a center C of perspectivity to send points on the sphere to points in the plane. The position of C within the sphere produces various maps in the plane; if C is located at the center of the sphere the corresponding map projection is gnomonic, if C is on the sphere the corresponding map projection is stereographic, and if C is at infinity the corresponding map projection is orthographic. The center C could also be located at any of an infinite number of other points; without loss of abstract generality, the following material shows how to determine coordinates for a projected point $P'(\phi',\lambda')$ in the plane from a point $P(\phi,\lambda)$ on the sphere projected from an arbitrary position for C along tha ray ON emanating from the center of the sphere O and passing through a pole, N. (Rotation matrices make corresponding alignments for positions of C tilted from the vertical.)

Definition 6.

Suppose C lies on ON inside the sphere of projection. The parallel centered on C will be called the bounding parallel on the sphere; it serves as an upper bound for the set of points that can be projected from C into the tangent plane. Its latitude will be denoted θ.

Suppose C lies on ON outside the sphere. The parallel at which the sphere is tangent to the ruled surface (cylinder or cone) formed by rays emanating from C, used to project points on the sphere into the tangent plane, will be called the parallel of contact. Its latitude on the sphere will be denoted $\overline{\theta}$ and its projection in the plane will serve as an outer possibility for the boundary of a map in the tangent plane projected from C.

Lemma 1

The notions of bounding parallel and parallel of contact coincide at N. That is, if C = N then $\theta = \overline{\theta}$.

Proof:

By the first part of Definition 6, as C approaches N from within the sphere, the corresponding sequence of bounding parallels is a set of parallels, of decreasing radius, centered on N. Passing to the one-seded limit shows that when C = N, $\theta = \pi/2$.

By the second part of Definition 6, as C approaches N from outside the sphere, the corresponding sequence of parallels of contact is a set of parallels, of decreasing radius, centered on N. Passing to the one-sided limit shows that when C = N, $\overline{\theta} = \pi/2$.

The two one-sided limits exist and are equal; thus, the notions of bounding parallel and parallel of contact extend to the point N, where they are identical. Q.E.D.

Thus we split the set of available positions along the ray at N into intervals $[O, N)$, $\{N\}$, $(N,\infty]$ (where a bracket indicates that the endpoint is included and a parenthesis indicates that it is omitted). Then, arguments from spherical trigonometry may be employed to prove, using the ideas of Lemma 1, formulas and their arithmetic (rather than their geometric) inverses for projecting the sphere to the plane (and back) [Deetz and Adams, 1931; Snyder, 1981].

Theorem 5

A point $P(\phi,\lambda)$, with latitude ϕ and longitude λ, on a sphere of radius r is projected through C to a point $P'(\phi',\lambda')$ in the tangent plane such that

a) when $C \varepsilon [O,N)$,

i) $\lambda' = \lambda$,

ii) the latitude θ of the parallel on the sphere whose image bounds the projection in the tangent plane is $\theta = \sin^{-1}(|CO|/r)$,

iii) ϕ' is determined at the intersection of a meridian of longitude λ' with a parallel centered on S in the tangent plane of radius $|P'S|$ $= (r\cos\phi(1+\sin\theta))/(\sin\theta \mp \sin\phi)$. The minus sign applies if P is in the southern hemisphere, and the plus sign if it is in the northern hemisphere;

b) when $C \varepsilon (N,\infty]$,

i) $\lambda' = \lambda$,

ii) the latitude $\bar{\theta}$ of the parallel of contact of the cone with the sphere, whose image bounds the map in the tangent plane, is $\bar{\theta} = \sin^{-1}(r/|CO|)$,

iii) ϕ' is determined at the intersection of a meridian of longitude λ' with a parallel centered on S in the tangent plane of radius $|P'S|$ $= (r\cos\phi(1 + \sin\bar{\theta}))/(1 \mp \sin\phi\sin\bar{\theta})$. The minus sign applies if P is in the southern hemisphere, and the plus sign if it is in the northern hemisphere;

c) if $C = N$,

i) $\lambda' = \lambda$,

ii) ϕ' is determined at the intersection of a meridian of longitude λ' in the tangent plane with a projected parallel centered on S of radius $|P'S| = (2r\cos\phi)/(1 \mp \sin\phi)$. The minus sign applies if P is in the southern hemisphere, and the plus sign if it is in the northern hemisphere.

Corollary 2

For any set of arbitrarily chosen centers of projection $\{C_0, C_1, \ldots, C_\infty\}$ on ON, and for a point P fixed on a sphere, the images of P in the tangent plane, $\{P_0', P_1', \ldots, P_\infty'\}$ are collinear and the line containing them passes through S. In fact, $C_0C_1\ldots C_\infty \overset{P}{\overline{\overline{\wedge}}} P_0'P_1'\ldots P_\infty'$ (Figure 11).

Corollary 3

With respect to a single center of perspectivity C, $\theta = \overline{\theta}$ if and only if C = N.

Proof:

a) If C = N then $\theta = \overline{\theta}$; this follows from Theorem 5.c.

b) If $\theta = \overline{\theta}$ then C = N.

If $\theta = \overline{\theta}$ it follows that $\sin\theta = \sin\overline{\theta}$. But $\sin\theta = |CO|/r$ by Theorem 5a and $\sin\overline{\theta} = r/|CO|$ by Theorem 5b. So $|CO|/r = r/|CO|$, and since both values are positive, $r = |CO|$. The only value C for which $r = |CO|$ is C = N. Q.E.D.

Thus if $C \neq N$, the only way in which it is possible to obtain $\theta = \overline{\theta}$ is to use a second center of perspectivity. And, for the bounding parallel to be the same as the parallel of contact, one center, C_1, must lie in [O, N) and the other, C_2, in (N, ∞]. The theorem below will link the classes [O, N) and (N, ∞] by characterizing those positions for C_1 and C_2 where $\theta = \overline{\theta}$ in terms of inversive geometry.

Theorem 6 (Linkage Theorem)

Suppose that $C_1 \in$ [O, N) and $C_2 \in$ (N, ∞], and that the bounding parallel associated with projection from C_1 is the same as the parallel of contact associated with projection from C_2. Then

a) the points C_1 and C_2 are inverses with respect to any great circle G through N.

b) the plane containing the great circle G, together with the point at ∞ along ON, is the inversive plane.

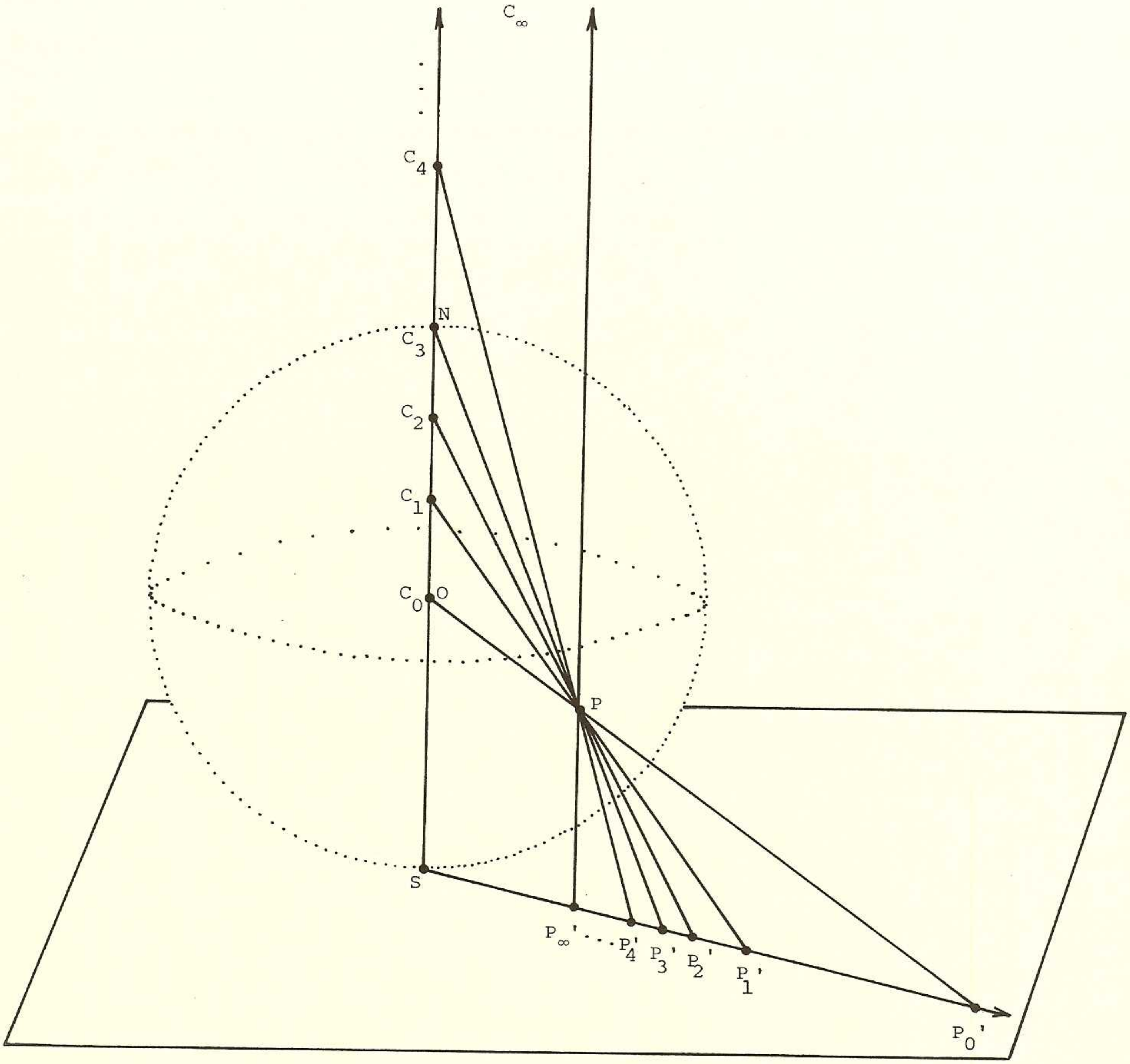


FIGURE 11

Proof:

a) To show that C_1 and C_2 are inverses with respect to G requires (according to Definition 1) that $|OC_1| \times |OC_2| = r^2$. From Theorem 5a, $|OC_1| = r\sin\theta$ and from Theorem 5b, $|OC_2| = r/\sin\bar{\theta}$. Since $\bar{\theta} = \theta$ from the hypotheses and Theorem 5c, it follows that $|OC_1| \times |OC_2| = r\sin\theta \times r/\sin\theta = r^2$ as required.

b) By Construction 1, if P = O a perpendicular to ON intersecting G at T leads to a line tangent to G at T that is parallel to ON. This tangent intersects ON at O', the inverse of O, which is the point ∞. Q.E.D.

Corollary 4 (Uniqueness)

The linkage between the intervals [O, N) and (N, ∞] given in Theorem 6 is the only one which may be made through the transformation of inversion.

Proof:

If C_1 and C_2 are such that associated bounding and contact parallels are not identical, then $\theta \neq \bar{\theta}$, and linkage in the style of Theorem 6 is not possible, by Corollary 3.

Corollary 5

The linkage in Theorem 6 provides a 1-1 correspondence of the set [O, ∞] onto itself in which [O, N) corresponds to (N, ∞]; (N, ∞] corresponds to [O, N); and, {N} corresponds to itself--it is invariant under the correspondence.

Proof:

This follows from Theorem 6, Corollary 4, and Construction 1.

Corollary 6

The centers of gnomonic and orthographic projection are inverses with respect to G, a great circle on the sphere of projection.

Proof:

This follows directly from Theorem 2, since O and ∞ are inverses.

Since the behavior of centers of perspectivity both within and between classes on either side of N has been determined, as well as at N, the characterization of the centers of perspective projections in terms of inversive geometry is complete. Beyond this, it seems natural to ask, what sorts of geometric relationships link points in the tangent plane that have been projected from inverse centers of projection.

Theorem 7 (Harmonic Map Projection Theorem)

Suppose a point P on the sphere is projected from inverse centers of projection, C and C', to points P_C and $P_{C'}$ in the plane tangent to the sphere at the South Pole, S. Suppose the stereographic image of P, projected from the North Pole N, is denoted P_N. Then $(SS)(P_NP_N)(P_CP_{C'})$; P_C and $P_{C'}$ are harmonic conjugates with respect to P_N and S in the tangent plane (Figure 12).

Proof:

Let C be a point in [O, N) and C' its inverse in (N, ∞]. Let P be an arbitrary point on the sphere chosen so that projection from all of C, N, C' is well-defined.

Now the points C and C' are harmonic conjugates with respect to N and S. For $|CN| = r - |CO|$, $|C'N| = |C'O| - r$, $|CS| = |CO| + r$, $|C'S| = r + |C'O|$ so that $(|CN|)/(|C'N|) = (|CS|)/(|C'S|)$ from which, it follows, by Theorem 4 and Corollary 1, that C and C' are harmonic conjugates with respect to N and S.

By Definition 2 and Corollary 2, we have $SCNC' \overset{P}{\overline{\overline{\wedge}}} SP_CP_NP_{C'}$. Thus, by Theorem 3, which shows that harmonic sets are invariant under perspectivity, it follows that P_C and $P_{C'}$ are harmonic conjugates with respect to P_N and S. Q.E.D.

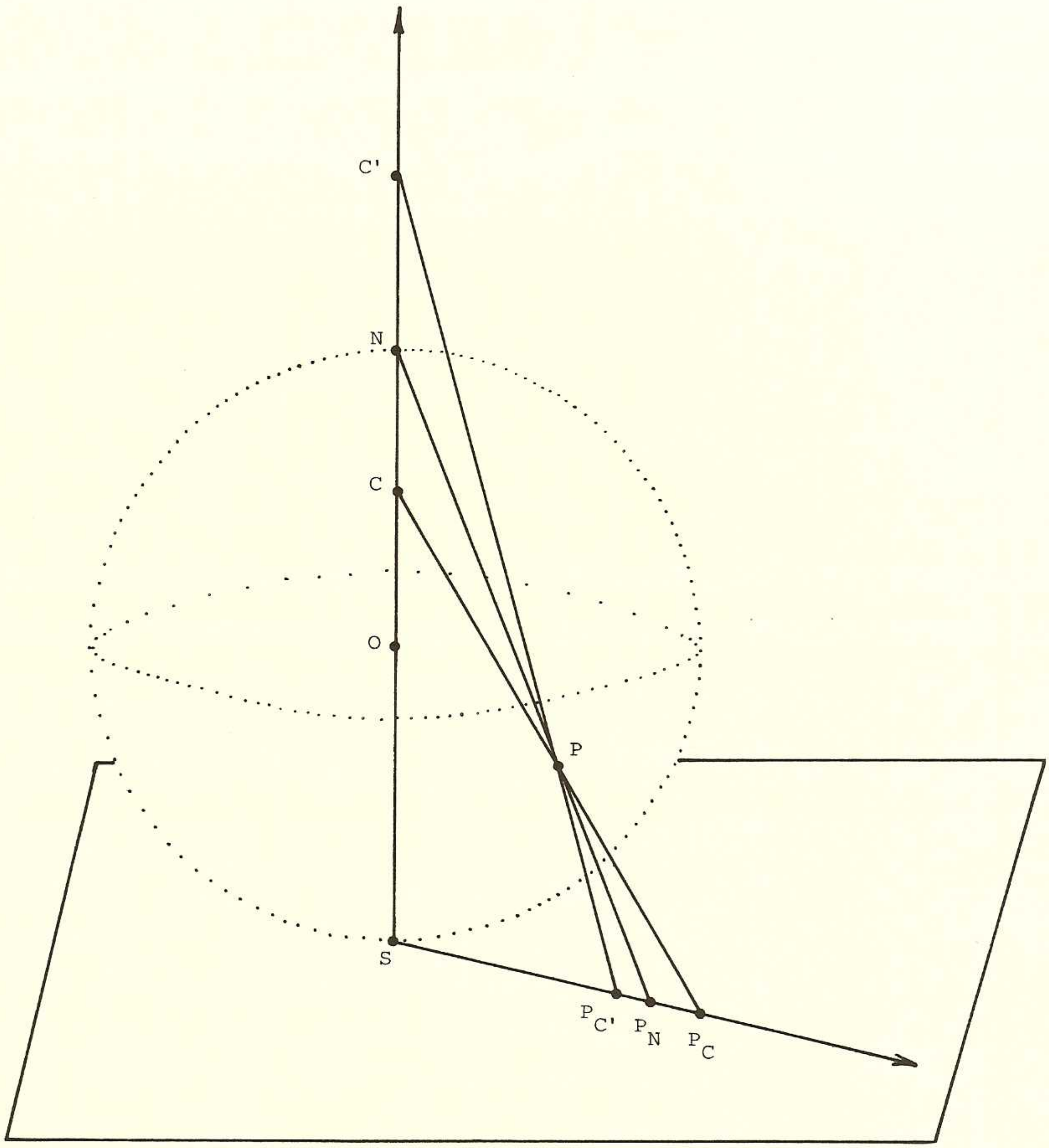


FIGURE 12

CASE STUDY

Suppose a point P, located at 30°S. Latitude on a unit sphere, is to be projected into a plane tangent at the South Pole, S. Using the conventional formulas based in spherical trigonometry, displayed in Theorem 5a, 5b, 5c, it is a simple matter to calculate the distance in the plane from the projected image of P to S. From Theorem 5a (with $\phi = -30°$), when the center of projection C is at the center of the sphere, O, the image under this gnomonic projection, P_O, is such that $|SP_O| = \sqrt{3}$. From Theorem 5b, the orthographic image of P, P_∞, is such that $|SP_\infty| = \sqrt{3}/2$. From Theorem 5c, the stereographic image of P, P_N, is such that $|SP_N| = 2/\sqrt{3}$. This yields positions for points on SP_O determined in Figure 13 via "perspective projections."

The Harmonic Map Projection Theorem (Theorem 7) promises that, given S, P_N, and P_∞, P_O may be constructed in the tangent plane using the Harmonic Construction (Construction 3), as long as O and ∞ are inverse points with respect to the sphere (Definition 1). Positions for P_∞, P_N, and P_O may be determined in the projective plane, along SP_O, without recourse to Theorem 5, as follows. In $\Delta(OSP_O)$, $\angle SOP_O = 60°$ since $\phi = -30°$, so that $|OS| = 1$ and $|SP_O| = \sqrt{3}$. Further, P bisects $|OP_O|$ since $|OP|$ is a radius of the unit sphere; thus, $|SP_\infty| = 1/2 \times |SP_O| = \sqrt{3}/2$. Finally, Theorem 4 shows that $|P_\infty P_N|/|P_N P_O| = |P_\infty S|/|P_O S|$. Cross-multiplying, and substituting $|P_\infty P_N| = |P_N S| - |P_\infty S|$ and $|P_N P_O| = |P_O S| - |P_N S|$ where appropriate, yields $|P_O S| \cdot |P_N S| - |P_O S| \cdot |P_\infty S| = |P_\infty S| \cdot |P_O S| - |P_\infty S| \cdot |P_N S|$ or, $2|P_\infty S| \cdot |P_O S| = |P_N S|(|P_O S| + |P_\infty S|)$ so that $2/|P_N S| = (|P_O S| + |P_\infty S|)/(|P_\infty S| \cdot |P_O S|) = (1/|P_\infty S|) + (1/|P_O S|)$. Thus, $2/|P_N S| = (2/\sqrt{3}) + (1/\sqrt{3}) = \sqrt{3}$, and $|P_N S| = 2/\sqrt{3}$. The positions for S, P_O, P_N, P_∞ are the same as those determined using Theorem 5; the positions for these points on SP_O are determined in Figure 13, in the tangent plane, using the notion of harmonic conjugacy. Equality in the metric representation of harmonic conjugacy, given in Theorem 4, will follow naturally. For, $|P_\infty P_N|/|P_N P_O|$ $= (2/\sqrt{3} - \sqrt{3}/2)/(\sqrt{3} - 2/\sqrt{3}) = 1/2$ and $|P_\infty S|/|P_O S| = (\sqrt{3}/2)/\sqrt{3} = 1/2$. Thus, Construction 3 permits the construction, in the projective plane, of a gnomonic projection as the harmonic conjugate of an orthographic projection with respect to stereographic projection and the fixed point of the South pole (Figure 13).

CONCLUSION

Through the case study we have seen that the gnomonic, stereographic, and orthographic are more than merely a set of useful perspective projections; for, together with the fixed point S, they exhibit the same natural mathematical harmony as the major musical triad. An infinite number of other harmonic triads of perspective map projections may be built on any ray emanating from S, and, by the Harmonic Map Projection Theorem (Theorem 7), these harmonic conjugates in the plane will have centers of projection that are inverses relative to the sphere. Building from these triads, one can imagine using theorems about harmonic conjugates, from projective geometry, to orchestrate the non-Euclidean "Look of Maps" [Robinson, 1952], lending a geographic interpretation to the harmonious notion of the "music of the spheres."

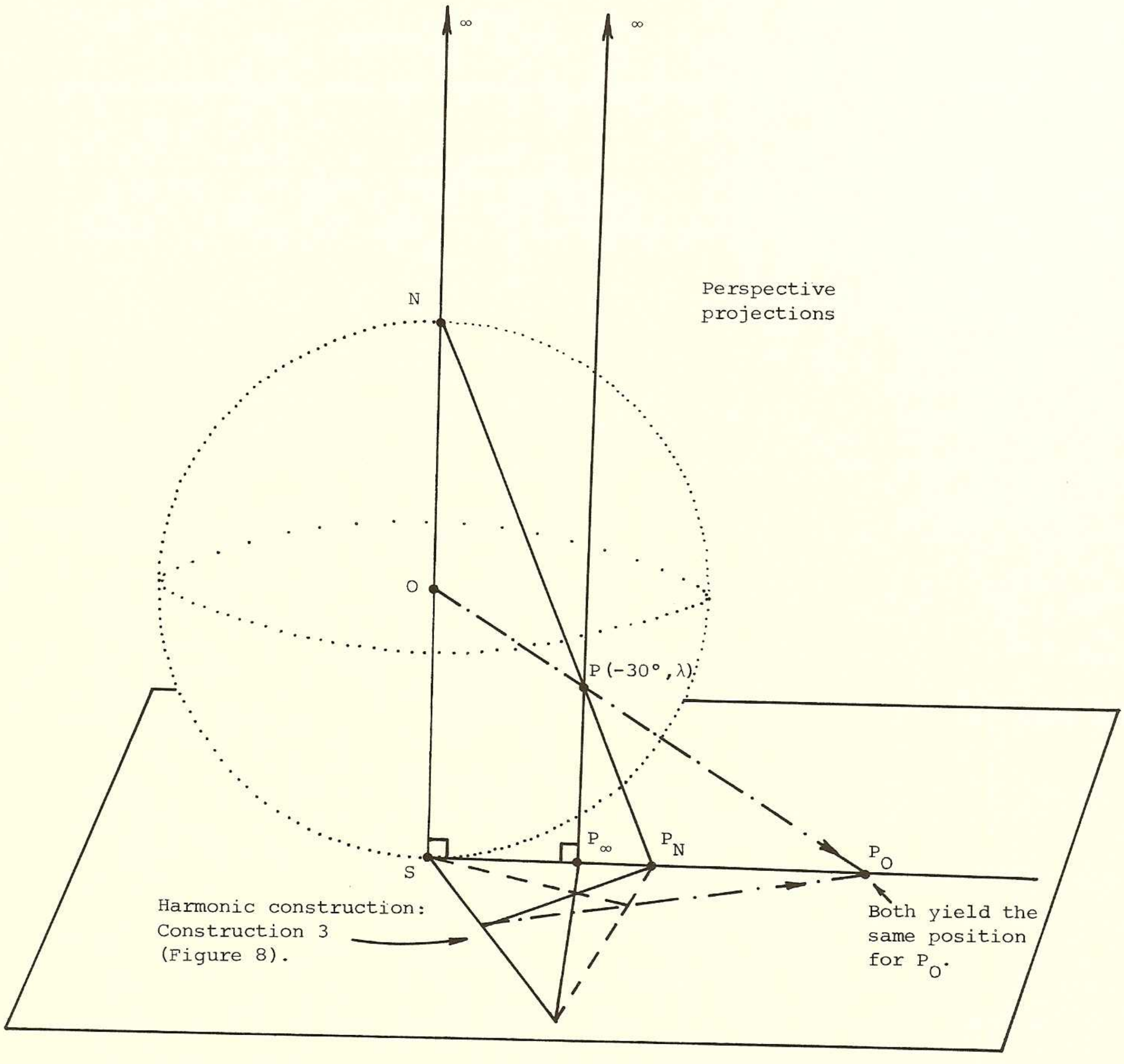

∞
∞
N
Perspective
projections
O
P(-30°,λ)
P_∞
P_N
P_O
S
Harmonic construction:
Construction 3
(Figure 8).
Both yield the
same position
for P_O.

FIGURE 13

LITERATURE CITED

* The author wishes to thank Professor H. S. M. Coxeter, Department of Mathematics, University of Toronto, for his constructive comments leading to significant improvement of the "Case Study" section. His generosity of time and effort are greatly appreciated.

1. Biernacki, F. Theory of Representation of Surfaces for Surveyors and Cartographers. Warsaw, Poland: U. S. Department of Commerce and the National Science Foundation, Scientific Publications Foreign Cooperation Center of the Central Institute for Scientific Technical and Economic Information, 1965.

2. Coxeter, H. S. M. The Real Projective Plane. Second ed. Cambridge: The University Press, 1960.

3. Coxeter, H. S. M. Introduction to Geometry. New York: John Wiley, 1961.

4. Coxeter, H. S. M. Non-Euclidean Geometry. Fifth ed. Toronto: University of Toronto Press, 1965.

5. Coxeter, H. S. M. Projective Geometry. Second ed. Toronto: University of Toronto Press, 1974.

6. Deetz, C. H. and Adams, O. S. Elements of Map Projection. Washington, D. C.: U. S. Department of Commerce, Coast and Geodetic Survey Special Publication no. 68, third ed., 1931.

7. Einstein, A. Relativity: The Special and the General Theory. Translation by R. W. Lawson, New York: Bonanza Books, 1961.

8. Golledge, R. G. and Hubert, L. J. "Some comments on non-Euclidean mental maps." Environment and Planning A 14, (1982), 107-118.

9. Lyndon, R. personal communication via attendance at lectures on "Groups and Geometry," The University of Michigan, 1985.

10. Mac Lane, S. "Proof, truth, and confusion: the 1982 Ryerson lecture." Chicago: The University of Chicago, 1982.

11. Muller, J.-C. "Canada's elastic space: a portrayal of route and cost distances." Canadian Geographer, 28, (1984), 46-62.

12. Robinson, A. H. The Look of Maps: An Examination of Cartographic Design. Madison, Wisconsin: The University of Wisconsin Press, 1952.

13. Steers, J. A. An Introduction to the Study of Map Projections. London: London University Press, 1962.

14. Snyder, J. P. "The perspective map projection of the earth." The American Cartographer, 8, (1981), 149-160.

15. Tobler, W. R. "World map on a Moebius strip." Surveying and Mapping 21, (1961), 486.

16. Tobler, W. R. "A classification of map projections." Annals, Association of American Geographers, (1962, #2), 167-175.

17. Tobler, W. R. "Geographic area and map projection." Geographical Review, 53, (1963), 59-78.

ANTIPODAL GRAPHS *

"The earth, that is sufficient,
I do not want the constellations any nearer."

Walt Whitman, "Song of the Open Road," 1860.

Geographical puzzles, from strolls in Königsberg parks to tours of cities "Around the World," have served, historically, to help stimulate graph theoretic ideas [Harary, 1969]. In today's technological setting, which carries the physical realm of geography far beyond earthly confines, one wonders what sorts of similar interaction, between mathematics and geography, might emerge.

For example, satellite navigation systems, such as Transit (U. S. Navy system) and the Global Positioning System, NavStar (Department of Defense), employ satellite locations that provide highly symmetric tesselations of their orbital spheres [Laurila, 1976, 1983; Wenninger, 1979]. Evenness in satellite spacing, suggested by engineering demands, may produce, as an upper bound for the number of satellites visible to an earth-based observer, half the number of the total in the entire configuration. Thus, when the satellites appear in antipodal pairs, so that as one satellite descends beyond the horizon another ascends, a flat geographical map that conforms to the earth-based observer's view of the satellite sphere, and that represents simultaneously the entire satellite configuration, would appear with antipodal identified (abstractly glued together). Klein's model of the elliptic plane, an object of non-Euclidean geometry in which "parallel" lines meet, formed by identifying antipodal points on a sphere, would suit these maps well [Mac Lane, 1982]. Geographical maps that represent discrete phenomena, such as satellite constellations, can be compressed abstractly as graphs in some surface (once adjacency relations have been specified). The natural surface in which to embed such satellite graphs is the elliptic plane, for like the elliptic plane, the eye of the satellite-camera knows no parallel lines [Arlinghaus, 1985].

Definition 1

Suppose a graph G can be embedded in a closed ball with its vertices on the boundary of the ball so that each vertex of G has an antipodal point which is also a vertex of G. The antipodal graph of G, denoted ⅁, has vertices obtained by identifying corresponding antipodal vertices of G, and edges determined by adjacencies present in G.

The graph composed of vertices and edges of a cube satisfies Definition 1; however, the graph composed of vertices and edges of a tetrahedron does not. For, as Figure 1a shows, each of a, b, c, and d on the cube has, as antipodal points, a', b', c', d', which are also on the cube. In Figure 1b, the four vertices, a, b, c, d of the tetrahedron, have a', b', c', d' as antipodal points; however, none of these lies on the tetrahedron.

Lemma 1:

Any graph that satisfies Definition 1 has 2n vertices, n a positive integer. The proof is a direct consequence of Definition 1.

Corollary 1:

The converse of Lemma 1 does not hold.

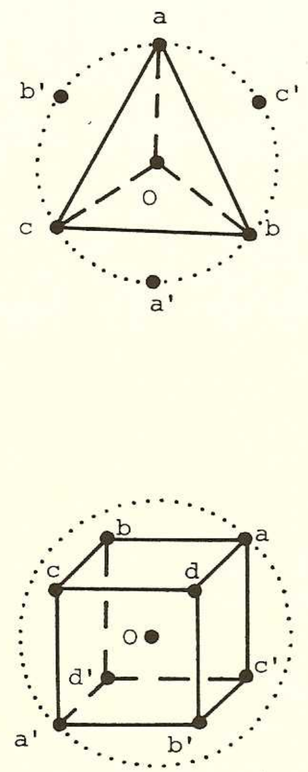
The tetrahedron has 2 x 2 vertices; however, it does not satisfy Definition 1.

Lemma 2:

There exists an unambiguous labelling of any graph G that satisfies Definition 1.

Proof:

By Lemma 1, G has 2n vertices for some positive integer, n. Let $X = \{x_1, \ldots, x_n\}$ be a set of distinct numerals; label n vertices of G in a single



Tetrahedron inscribed in sphere

Cube inscribed in sphere

FIGURE 1

hemisphere of the ball using a single label from X for each. Then

a) label the vertex antipodal to x_i as x_i', $i = 1,\ldots,n$;

or b) label the vertex antipodal to x_i as $x_1 x_2 \ldots x_{i-1} x_{i+1} \ldots x_n$.

Corollary 1:

There exists an unambiguous labelling of $\mathcal{D}$.

Proof:

To obtain $\mathcal{D}$ from G identify antipodal points so that either

a) x_i and x_i';

or b) x_i and $x_1 x_2 \ldots x_{i-1} x_{i+1} \ldots x_n$

represent the same point in $\mathcal{D}$.

Definition 2: (Coxeter, 1950].

The Schläfli symbol $\{p,q\}$, $(p,\ q > 2)$ represents a regular solid with a typical face of p edges and a typical vertex with q vertices adjacent to it. The solid $\{q,p\}$ is said to be the dual of $\{p,q\}$. Solids for which $\{p,q\} = \{q,p\}$ are said to be self-dual.

Proposition 1

For a graph G, representing a solid with Schläfli symbol $\{p,q\}$, if

a) $n = p \cdot 2^{\alpha}$

or b) $n = q \cdot 2^{\alpha}$

where $\alpha \geq 0$, then

a) p

or b) q+1

distinct labels will provide an unambiguous labelling for G.

Proof:

a) If $n = p \cdot 2^{\alpha}$, $\alpha \geq 0$, then $X = \{x_1,\ldots,x_p\}$.

Label the p vertices in a single face of G using a single label from X for each. Label the vertices antipodal to these as in Lemma 2b. This provides labels for 2p vertices, leaving $2n-2p = p \cdot 2^{\alpha+1} - 2p = 2p(2^{\alpha} - 1)$ to be labelled.

To label another set of p vertices, use p pairs of labels from among the $\binom{p}{2} = \frac{p!}{2!(p-2)!}$ possible pairs. It follows that $\frac{p!}{2!(p-2)!} > p$, since $p > 2$ (Definition 2) and so this labelling is possible. Use Lemma 2b to label the p points antipodal to these with strings composed of (p-2) labels. This procedure may be continued, using strings composed of r labels, and of (p-r) labels for the antipodal points, so long as $\binom{p}{r} > p$. There are enough levels for this procedure to continue, as long as $2^\alpha \leq [p/2]$ (greatest integer in p/2).

b) If $n = q \cdot 2^\alpha$, $\alpha \geq 0$, then $X = \{x_1, \ldots, x_{q+1}\}$.

Label the q vertices of G that are adjacent to a vertex with q of the (q+1) single labels, and label the vertex itself with the remaining single label. Label the vertices antipodal to these as in Lemma 2b. This provides labels for 2(q+1) vertices, leaving $2n-2(q+1) = 2(q(2^\alpha - 1) - 1)$ to be labelled. To label another set of (q+1) vertices, use (q+1) pairs of labels from among the $\binom{q+1}{2}$ possible pairs. It follows that $\binom{q+1}{2} = \frac{(q+1)!}{2!(q-1)!} > (q+1)$ since $q > 2$ (Definition 2), and so this labelling is possible. Use Lemma 2b to label the (q+1) points antipodal to these with strings of labels composed of ((q+1) -2) = (q-1) elements. This procedure may be continued, using strings composed of r labels, and of ((q+1) - r) labels for the antipodal points, so long as $\binom{q+1}{r} > q+1$. There are enough levels for this procedure to continue as long as $2^\alpha \leq [(q+1)/2]$ (greatest integer in (q+1)/2).

Corollary 2:

Under the hypotheses of Proposition 1, there exists an unambiguous labelling of $\mathfrak{D}$.

Proof:

To obtain $\mathfrak{D}$ from G identify antipodal points so that both the labels $x_i \ldots x_j$ and $x_1 \ldots x_{i-1} \ldots x_{j+1} \ldots x_n$ represent the same point in $\mathfrak{D}$.

The set of regular solids which can be inscribed in a sphere is the set of Platonic solids, comprised of the tetrahedron, the cube, the octahedron, the dodecahedron, and the icosahedron [Coxeter, 1961]. Because of their high degree of symmetry, these might serve as models for satellite placement on the sphere with radius, for example, that of the geostationary orbit.

The Proposition and the Lemmas above will permit consistent labelling of the vertices of these solids. When the solids are projected into the plane, they are represented as a graph, in what is known as a Schlegel diagram [Coxeter, 1961].

<u>Construction</u> (Coxeter, 1961].

Within the set of graphs, G, of the Platonic solids, only those of the non self-dual solids have antipodal graphs. The antipodal graphs, $\mathcal{D}$, of the cube, octahedron, and icosahedron are, respectively, the complete graphs on four, three, and six vertices. The antipodal graph, $\mathcal{D}$, of the dodecahedron is the so-called 'Petersen' graph. These are displayed below, using labelling schemes from Lemma 2b and from Proposition 1, as indicated in Figure 2.

When adjacency between vertices is viewed as communication between satellites (presumably including bouncing back to earth), the less susceptible is the entire constellation of satellites (with antipodal points identified) to fragmentation, the more secure is the transmission of messages. The following material from graph theory, when applied to the set of Platonic solids, classifies the antipodal graphs of the Platonic solids, according to extent of mathematical fragmentation.

<u>Definition 3</u> [Harary, 1969]

The degree (or valency) of a point v in a graph G is the number of edges of G incident with v.

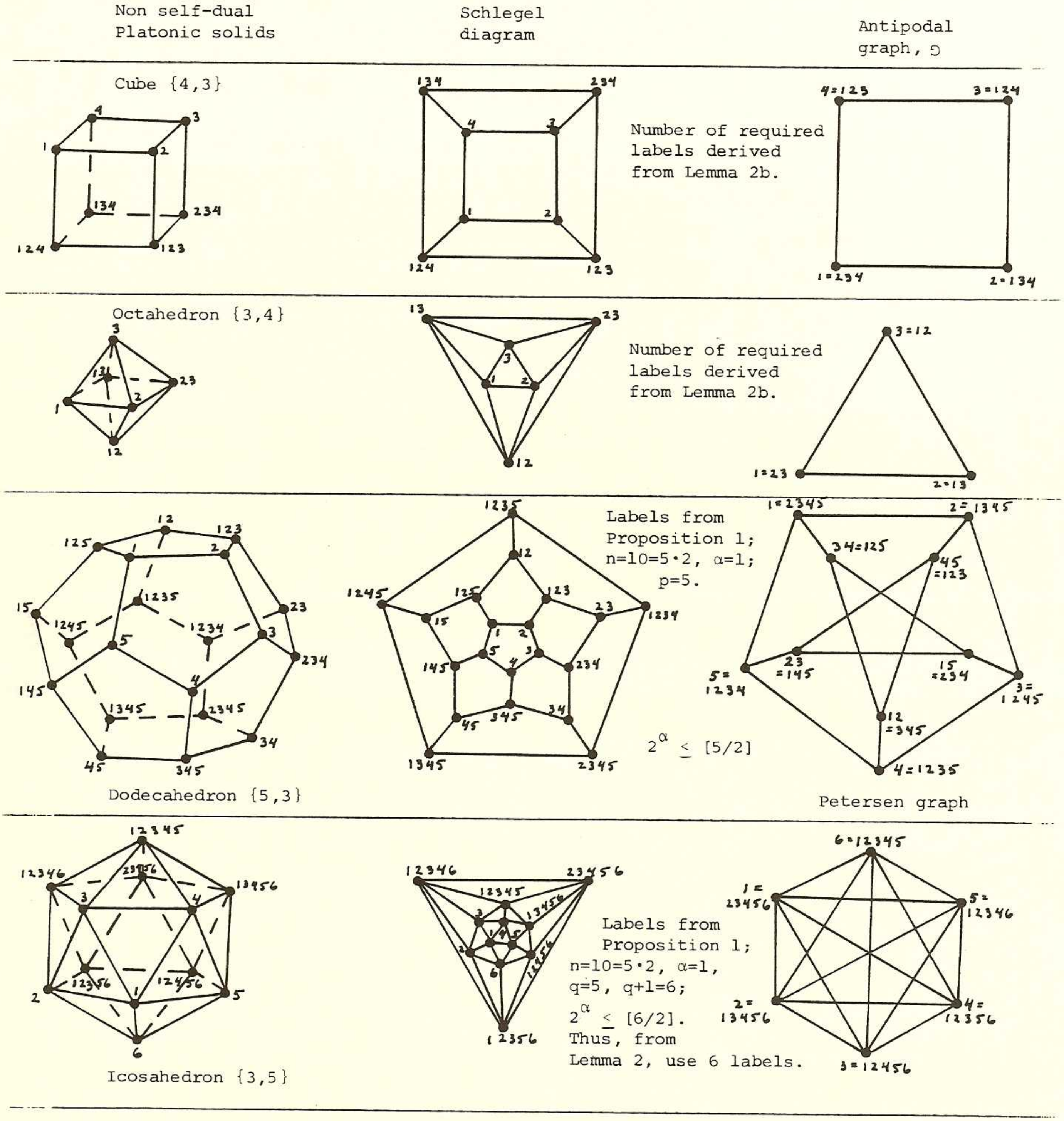

Non self-dual
Platonic solids
Schlegel
diagram
Antipodal
graph, ɔ
Cube {4,3}
Number of required
labels derived
from Lemma 2b.
4=123
3=124
1=234
2=134
Octahedron {3,4}
Number of required
labels derived
from Lemma 2b.
3=12
1=23
2=13
Labels from
Proposition 1;
n=10=5·2, α=1;
p=5.
2^α ≤ [5/2]
Dodecahedron {5,3}
Petersen graph
Labels from
Proposition 1;
n=10=5·2, α=1,
q=5, q+1=6;
2^α ≤ [6/2].
Thus, from
Lemma 2, use 6 labels.
Icosahedron {3,5}
6=12345
1=23456
5=12346
2=13456
4=12356
3=12456

FIGURE 2

<u>Definition 4</u> [Harary, 1969].

A graph G in which all points have degree n, is said to be regular of degree n.

<u>Definition 5</u> [Harary, 1969].

A bridge, x, in a graph, G, is an edge of G whose removal forces an increase in the number of connected components of G.

<u>Definition 6</u> [Harary, 1969].

An n-factor of a graph G is a subgraph of G, spanning the vertices of G, which is regular of degree n.

<u>Definition 7</u> [Harary, 1969].

A graph G is n-factorable if it can be expressed as a sum of n-factors.

Thus, in Figure 3, the antipodal graph of the cube is 1-factorable, the antipodal graph of the octahedron is 2-factorable, the antipodal graph of the icosahedron is 1-factorable, and the antipodal graph of the dodecahedron is not factorable. The Petersen graph may be decomposed as a sum of a 1-factor and a 2-factor; the 2-factor, however, is irreducible--it may not be further decomposed as a sum of two 1-factors.

<u>Theorem</u> [Harary, 1969].

Every bridgeless graph, which is regular of degree three, can be decomposed as a sum of a 1-factor and a 2-factor.

Petersen used the graph, labelled with his name in Figure 3, to show that the stronger result, 'Every bridgeless graph, which is regular of degree three can be decomposed as a sum of three 1-factors' is false [Petersen, 1891; Harary, 1969]. This result may hold for some graphs which satisfy the hypotheses, but not for all of them; in particular, it does not hold for the graph which is the antipodal graph of the dodecahedron. Thus the decompositions shown in Figure 3 are all irreducible.

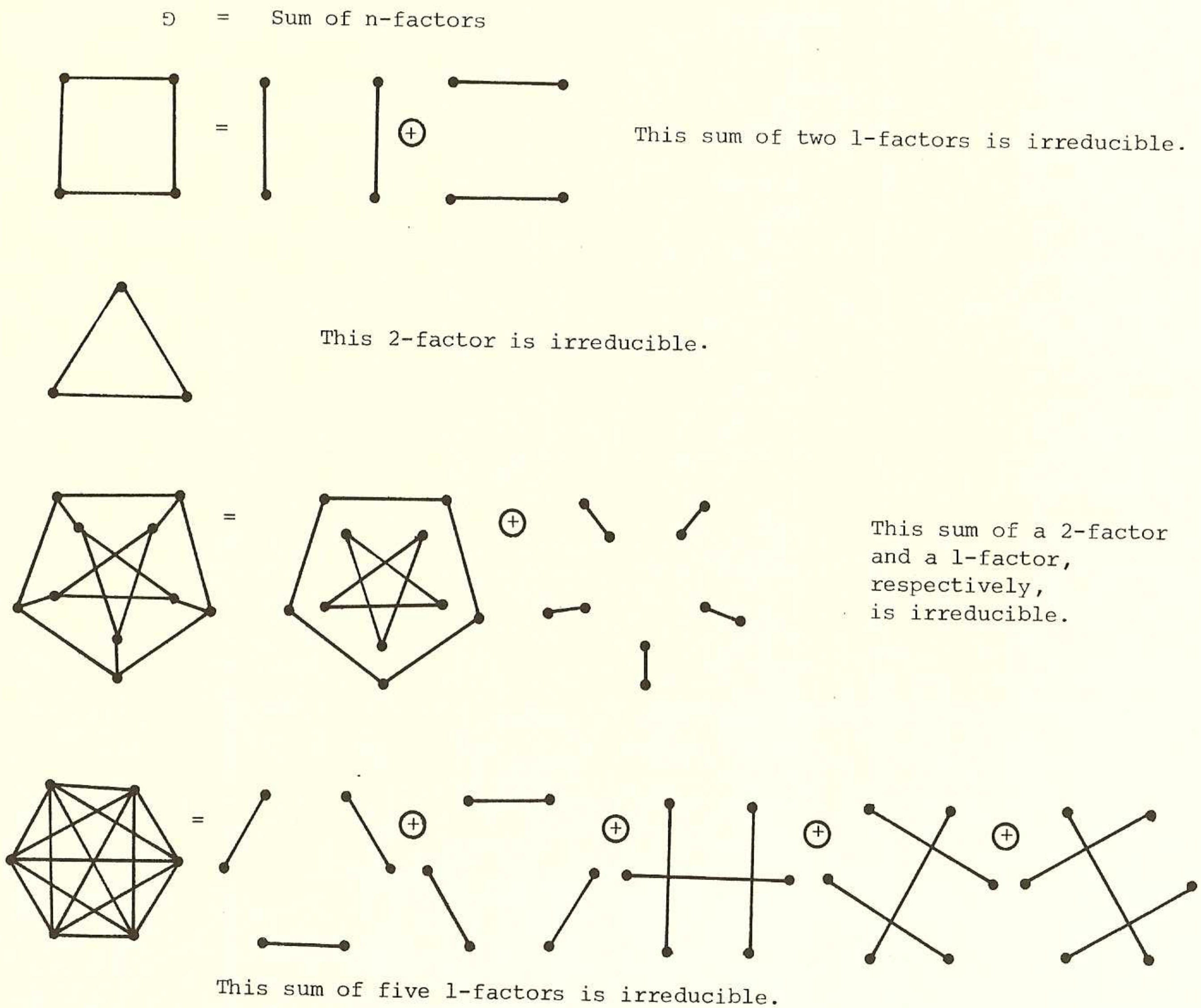

G = Sum of n-factors
=
This sum of two 1-factors is irreducible.
This 2-factor is irreducible.
=
This sum of a 2-factor and a 1-factor, respectively, is irreducible.
=
This sum of five 1-factors is irreducible.

FIGURE 3

An interpretation of this Theorem suggests that satellite communications systems whose antipodal graphs have bridges are most vulnerable to mathematical fragmentation and to physical disarmament; the bridge is mathematically weak, since removal of it necessarily increases the number of connected components in the underlying graph (Definition 5). Those graphs which are 1-factorable are next most vulnerable (the antipodal graphs of the cube and of the icosahedron are 1-factorable). Those antipodal graphs which are 2-factorable are next most susceptible to mathematical fragmentation (such as the antipodal graph of the octahedron). Those which are neither 1-factorable, nor 2-factorable, but which can be decomposed as a sum of 1-factors and 2-factors (such as the Petersen graph) are not as vulnerable, as disarmament is required at two different levels of connectivity. Any graph which resists decomposition is least vulnerable. The same sort of ordering of fragmentation would hold for graphs that were n-factorable and (n+1)-factorable.

Thus, within the set of Platonic solids, the dodecahedron has the antipodal graph which is mathematically the most resistant to decomposition. This suggests that a dodecahedral constellation of satellites is relatively difficult to disarm.

LITERATURE CITED

* Based on a paper given by the author January 13, 1982, Cincinnati, OH at the 88th Annual Meeting of the American Mathematical Society. Abstract #792-92-404, Abstracts of the American Mathematical Society, Vol. 3, No. 1, 1982. The notation for the graph antipodal to G was suggested by William C. Arlinghaus, 1982. John Nystuen discussed satellite navigation in informal sessions of the Colloquium in Mathematical Geography, 1983.

1. Arlinghaus, S. "Terrae antipodum," 1985.
2. Coxeter, H. S. M. Introduction to Geometry. New York: Wiley, 1961.
3. Coxeter, H. S. M. "Self-dual configurations and regular graphs," Bulletin of the American Mathematical Society, 56, (1950), 413-55.
4. Harary, F. Graph Theory. Reading, Mass.: Addison-Wesley, 1969.
5. Laurila, S. Electronic Surveying and Navigation. New York: Wiley, 1976.
6. Laurila, S. Electronic Surveying in Practice. New York: Wiley, 1983.
7. Mac Lane, S. "Proof, truth, and confusion: the 1982 Ryerson lecture," University of Chicago., 1982.
8. Wenninger, M. J. Spherical Models. Cambridge: Cambridge University Press, 1979.

Background material on satellite navigation, not cited here, was introduced to the Colloquium in Mathematical Geography by Richard P. Nesper, Ph.D. Candidate in the Program in Geography, The University of Michigan, 1983.

MEASURING THE VERTICAL CITY*

"Die Strahlen der Sonne vertreiben die Nacht"
W. A. Mozart, The Magic Flute

INTRODUCTION

Forces that shape the physical form of a central business district are largely economic; high demand for accessible land that is relatively short in supply creates high costs per unit of land. Vertical construction is a natural response to this situation; repetition of vertical construction produces skylines. Although skylines are often unplanned in global form, large cities can usually be identified from skyline silhouettes [Gottmann, 1967].

The arrangement and the form of buildings in a central business district can create positive feelings within the surrounding urban population [Fuller, 1975; Forgey, 1978]. Abstractly, this notion is an extension of the principles of Japanese landscape gardening [Feldt, 1974], to the urban scale, in which the harmonious placement of buildings is designed to produce a unit of urban space that creates a positive attitude in most who enter it [Conway, 1977].

Thus, the skyline might emerge in response to a set of aesthetic as well as to a set of economic forces. The procedure described below uses matrices of 0's and 1's to measure the form of existing skylines and to evaluate the impact of new construction on a skyline. The possibility of using these matrices rests fundamentally on the necessity of the elevator in skyscraper construction. It precludes holes, or gaps, in buildings and in corresponding matrix structure which would prevent accurate outlining of a skyline profile.

SKYLINE MATRICES

Buildings set back from the streets permit light to enter, offer open spaces in which to design gardens, rest areas, or sculpture, and generally produce a more relaxing urban environment than do their counterparts with forty stories directly abutting the sidewalks [Lynch, 1960]. Measures of the sort of terracing required in building set-backs to permit a fixed amount of sunlight to enter from a given compass direction would, of course, vary with latitude. Further, local empirical evidence might suggest that steep rises can be tolerated better farther from the street than closer to the street. Thus, what would emerge is a terraced, stair-case hull with the width of a step and the height of the rise dependent on various local conditions that respond favorably to light and to openness requirements. Buildings built within this hull (which would vary in shape according to position within the central business district relative to the sun, prevailing winds, and so forth) would work together to create the desired positive unit of space, while those which pierce the hull would not. The material below shows how to measure violations of a simple hull formed on abstract bases; the same procedure, modified in obvious ways, would apply to specific terraced hulls formed on empirical bases (Figure 1).

Given a parcel, P, of central business district land of arbitrary size; obtain maps of P that show both position and height of the buildings. Partition the map into cells which reflect light and openness requirements; assume that P is rectangular and that there is a series of uniform-width horizontal strips, with width reflecting how much set-back is required to accommodate another story of height, which covers P. Assume also that there is a corresponding set of vertical strips of uniform-width (not necessarily the same width as that of the horizontal set) which covers P. Thus, in Figure 2, the strips on the edge are best left empty; hence each contains a 0. Move back a bit--a single story does not

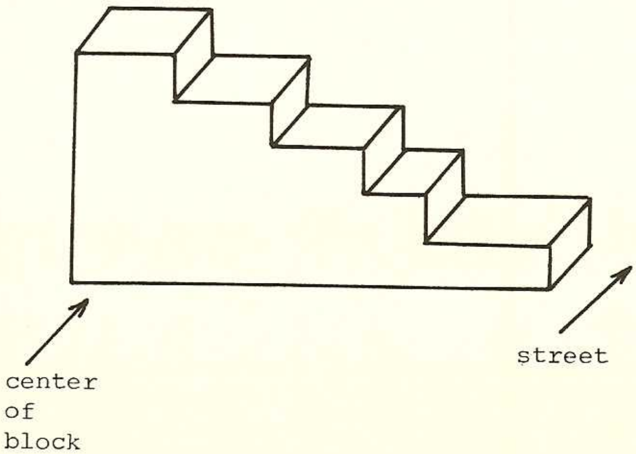


FIGURE 1

street

street (left) / street (right)

0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	1	1	1	0
0	1	2	2	2	2	2	2	2	2	1	0
0	1	2	3	3	3	3	3	3	2	1	0
0	1	2	3	4	4	4	4	3	2	1	0
0	1	2	3	4	5	5	4	3	2	1	0
0	1	2	3	4	5	5	4	3	2	1	0
0	1	2	3	4	4	4	4	3	2	1	0
0	1	2	3	3	3	3	3	3	2	1	0
0	1	2	2	2	2	2	2	2	2	1	0
0	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0	0	0	0	0

street

FIGURE 2

intrude too much; these cells have a 1. Move back more and add another story; these are coded 2. Continue until no more cells are left. In the example of Figure 2, the center of P supports five stories. Visualizing this as a three-dimensional object produces a terraced mountain with vertical profiles that conform to the direction of the boundaries in the cellular partition, similar to the situation in Figure 1. This represents a terraced hull on P, formed on abstract bases; it is topologically equivalent to a hull that might arise from empirical set-back conditions, such as that shown in Figure 3. The idea of terracing is the same in both Figures 2 and 3; however, the actual shape of the vertical profile is not the same, even though the structure of the vertical profiles, both with five levels, is the same. In an attempt to create uniformity, matrices of 0's and 1's will be used to represent vertical profiles of the hull.

Definition 1

Suppose that the tallest permissible building in parcel P has n stories. An n x n profile matrix is a matrix of 0's and 1's which represents a vertical profile of a terraced hull over P. An entry of 1 in the (i,j) matrix position means that there is part of a building within the corresponding position in the terraced hull; an entry of 0 in that position indicates the absence of a building in the corresponding position in the terraced hull.

Figure 4 shows profile matrices for selected profiles of the terraced hulls represented in Figures 2 and 3.

Definition 2

Suppose that the tallest building in parcel P has k stories. A kxn (with n as in Definition 1) skyline matrix is a matrix of 0's and 1's which represents a vertical profile of the actual skyline over P. (P is partitioned laterally as it was for the terraced hull.) An entry of 1 in the (i,j) matrix position

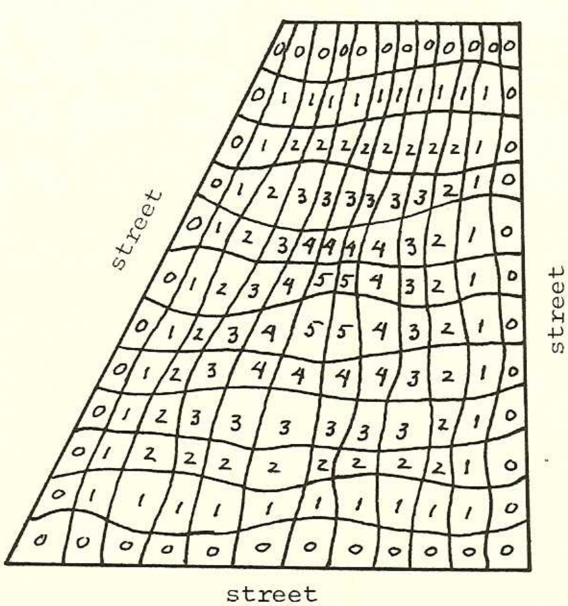


FIGURE 3

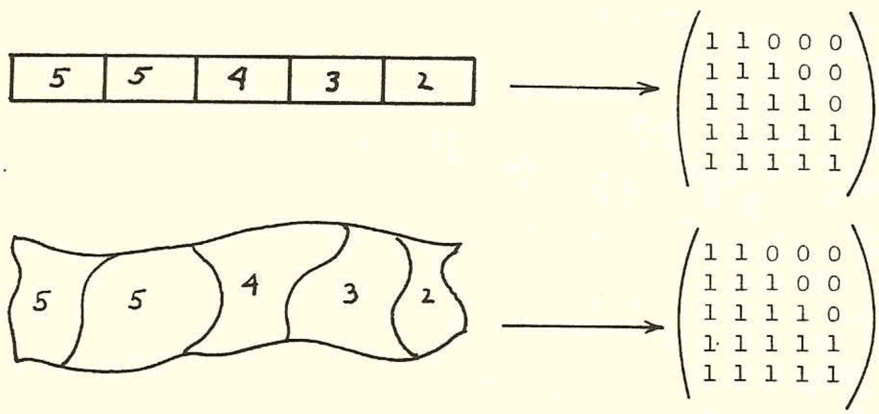


FIGURE 4

means that there is part of a building within the corresponding position in the skyline; an entry of 0 indicates the absence of a building.

Figure 5 illustrates Definition 2; here, the resulting profile matrix of the actual skyline has dimension (8x5) reflecting the presence of an eight-story building.

Definition 3

Profile matrices of dimension (nxn) of the terraced hull may be made conformable with (kxn) skyline matrices as follows:

i) if $k > n$, add k-n rows of zeroes to the top of the (nxn) profile matrix;

ii) if $k < n$, add n-k rows of zeroes to the top of the (kxn) skyline matrix.

The added rows record empty space and do not alter the faithful representation of building structure in either case.

Theorem 1

Given a (kxn) profile matrix of either the terraced hull, or of the actual skyline. If there is an entry of 1 in the (i,j) matrix position ($i \neq k$), then all positions (h,j), with $i < h \leq k$, also contain a 1.

Proof:

The theorem is a direct consequence of the fact that skyscrapers have no "missing" floors. Q.E.D.

Any (kxn) profile matrix may be converted, through a sequence of elementary transformations (adding integral multiples of one row to other rows), to a matrix in which the only 1's are in rooftop positions. Generally, begin with the

$$\begin{array}{|c|c|c|c|c|}\hline 5 & 8 & 4 & 3 & 2 \\ \hline \end{array} \longrightarrow \begin{pmatrix} 0&0&1&0&0\\ 0&0&1&0&0\\ 0&1&1&0&0\\ 1&1&1&0&0\\ 1&1&1&0&0\\ 1&1&1&1&0\\ 1&1&1&1&1\\ 1&1&1&1&1 \end{pmatrix}$$

Partition same as in Figure 4

FIGURE 5

$$\begin{pmatrix} 1&1&0\\ 1&1&1\\ 1&1&1 \end{pmatrix} \xrightarrow{-R_1+R_2} \begin{pmatrix} 1&1&0\\ 0&0&1\\ 1&1&1 \end{pmatrix} \xrightarrow{-R_1+R_3} \begin{pmatrix} 1&1&0\\ 0&0&1\\ 0&0&1 \end{pmatrix} \xrightarrow{-R_2+R_3} \begin{pmatrix} 1&1&0\\ 0&0&1\\ 0&0&0 \end{pmatrix}$$

FIGURE 6 a

$$\begin{pmatrix} 1&1&0\\ 0&1&1\\ 1&1&1 \end{pmatrix} \xrightarrow{-R_1+R_2} \begin{pmatrix} 1&1&0\\ 1&0&1\\ 1&1&1 \end{pmatrix} \xrightarrow{-R_1+R_3} \begin{pmatrix} 1&1&0\\ 1&0&1\\ 0&0&1 \end{pmatrix} \xrightarrow{-R_2+R_3} \begin{pmatrix} 1&1&0\\ 1&0&1\\ -1&0&0 \end{pmatrix} \xrightarrow{R_3+R_2} \begin{pmatrix} 1&1&0\\ 0&0&1\\ -1&0&0 \end{pmatrix} \xrightarrow{-R_3} \begin{pmatrix} 1&1&0\\ 0&0&1\\ 1&0&0 \end{pmatrix}$$

FIGURE 6b

first row, R_1, in which all entries represent rooftops. Add (-1) x R_1 to each of the (k-1) remaining rows. This annihilates all entries below the 1's in the first row, without introducing any new elements into these (k-1) rows (Theorem 1). Use the second row, R_2, to annihilate entries of 1 in any of the (k-2) rows below R_2. Repeat this procedure until R_k is reached. By Theorem 1, only rooftop entries of 1 will remain in the matrix. For example, Figure 6a shows a 3x3 profile matrix. Adding $(-1)xR_1$ to the second row annihilates the stories directly below roof-top level in the two buildings on the left, and leaves the lower rooftop level in the third column unchanged. Adding $(-1)xR_2$ to the third row removes the bottom two stories in the first two columns and leaves the third unchanged. Finally, adding $(-1)xR_2$ to R_3 annihilates the entry below roof-top level in the third column and leaves all else unchanged. The final matrix in this sequence outlines the roof-tops; were not Theorem 1 true, entries of 1 in positions other than roof-top positions might have been introduced, as Figure 6b shows.

Definition 4

The rooftop matrix is a row-reduced form of the profile matrix; the rank of this matrix, representing the number of linearly independent rows, is given by the number of distinct rows containing at least one entry of 1. In the landscape, the rank represents the number of rises in the skyline, along the given profile.

Theorem 2

Given a (kxn) rooftop matrix representing actual skyline and a conformable (kxn) rooftop matrix representing the corresponding profile of the terraced hull. Suppose both matrices have the same rank. Add the (row, column) coordinates of all the matrix entries of 1 to obtain (row sum, column sum) in each matrix; denote these sums by (R,C) in the rooftop matrix of the terraced hull, and by (R',C') in the rooftop matrix of the skyline. Then, when C=C', (i) if $R' < R$, the skyline pierces the terraced hull; (ii) if $R' > R$, the skyline fails to fill the terraced hull.

Proof: (Figure 7 shows an example)

i) Suppose R' < R, and that (h,j) is an element of the skyline rooftop matrix and (i,j) is an element of the terraced hull rooftop matrix. For R' to be less than R requires that h < i for at least one pair of entries. Thus the matrix position (h,j) in the skyline lies above the matrix position (i,j), and the skyline pierces the terraced hull.

ii) Suppose R' > R and that (h,j) is an element of the skyline rooftop matrix and (i,j) is an element of the terraced hull rooftop matrix. For R' to be greater than R requires that h > i for at least one pair of entries. Thus the matrix position (h,j) lies below the matrix position (i,j), and the skyline fails to fill the terraced hull.

ROOFTOP MATRICES

FIGURE 7

```
1 1 1 0 0 0 0
0 0 0 1 1 0 0
0 0 0 0 0 0 0
0 0 0 0 0 1 1
0 0 0 0 0 0 0
```

actual

```
1 1 1 0 0 0 0
0 0 0 1 0 0 0
0 0 0 0 1 0 0
0 0 0 0 0 1 0
0 0 0 0 0 0 1
```

terraced hull

The matrix positions corresponding to rooftop positions contribute to the row sum R and to the column sum C.

Here, these matrix positions are: (1,1), (1,2), (1,3), (2,4), (2,5), (4,6), (4,7).

Thus, R'=1+1+1+2+2+4+4=15
C'=1+2+3+4+5+6+7=28

Here these matrix positions are: (1,1), (1,2), (1,3), (2,4), (3,5), (4,6), (5,7).

Thus, R=1+1+1+2+3+4+5=17
C=1+2+3+4+5+6+7=28.

Thus, R' < R, and in fact, the actual skyline does pierce the terraced hull as Theorem 2 predicts.

Some skylines will both pierce, and fail to fill, a given terraced hull. In those cases, comparison of the extent of the violation shows which is dominant. The extent to which parts of skylines fit, or fail to fit, the terraced hull, has obvious implications for planning both the location and the height of new buildings within the skyline.

A subset of all rooftop skyline profile matrices over P, taken from any one vantage point, produces a numerical shadow of the silhouette from that vantage point. Taking the complete set of rooftop skyline profile matrices over P produces, simultaneously, a numerical composite of skyline silhouettes from spatially opposed vantage points. This would present a numerical model of the whole skyline, in a single city, against which parts might be tested.

To classify American skylines across the set of urban areas, smaller cities which possess a skyline silhouette similar to a segment of that of a larger city might benefit from planning guidelines for the appropriate part of the large city. An abstract approach to describing such similarity might involve using the notion of self-similarity, used to overcome scale differences in fractal geometry [Mandelbrot, 1983].

LITERATURE CITED

* Based on participation in the Metropolitan Communities Seminar (funded by the Mott Foundation, John D. Nystuen and Allan Feldt, directors), The University of Michigan, Ann Arbor, 1974-75.

1. Conway, D. J. (ed.) Human Response to Tall Buildings. Stroudsburg, PA; Dowden, Hutchinson and Ross, Inc., Community Development Series, Vol. 34, 1977.

2. Feldt, A. Personal communication, Metropolitan Communities Seminar, 1974.

3. Forgey, B. "Sculptor Isamu Noguchi's elegant world of space and function," Smithsonian, 9, 1 (April, 1978), 46-55.

4. Fuller, R. B. Synergetics: Explorations in the Geometry of Thinking. New York: Macmillan, 1975.

5. Gottmann, J. "The skyscraper amid the sprawl," in Metropolis on the Move: Geographers Look at Urban Sprawl, Gottmann, J. and Harper, R. A. (eds.), New York: Wiley, 1967.

6. Lynch, K. The Image of the City. Cambridge, Mass.: MIT, 1960.

7. Mandelbrot, B. F. The Fractal Geometry of Nature. San Francisco: W. H. Freeman, 1983.

CONCAVITY AND HUMAN SETTLEMENT PATTERNS

The shape of the northern coastline of Australia has been generalized, by biologist Joseph Birdsell, as two concave-down lobes of land separated by a concave-up lobe of water (Figure 1) [Birdsell, 1950]. Through this geometric observation, and through a study of the hunting and gathering practices of early migrants, he argued that a strong sense of territoriality would force migrants toward the interior, away from points of entry on the coast, and that coastline shape would force concentration of migrant settlement under concave-down northern coastline segments and dispersal of migrant settlement under concave-up northern coastline segments (Figure 1). One of the implications of these geometric observations for genetics was that greater genetic diversity would occur in settlements located under concave-down northern coastline segments than would occur in settlements located under concave-up northern coastline segments [Birdsell,1950; Kolars, 1975].

In an urban setting, when interurban arterials play the role of the northern Australian coastline, bridges across expressways force concentration of traffic flow from surface routes under concave-down parts of the curve and dispersal under concave-up parts of the curve (Figure 2). Naturally, a region exhibiting dispersal is opposed, across the expressway, by a region of concentration (and vice-versa--Figure 2). With the concentration of surface-route traffic flow would come concentration of commercial activity (locating shopping centers at high traffic-volume intersections) and, ultimately, of general settlement patterns (building high-cost houses, with large lots, far from high-volume intersections).

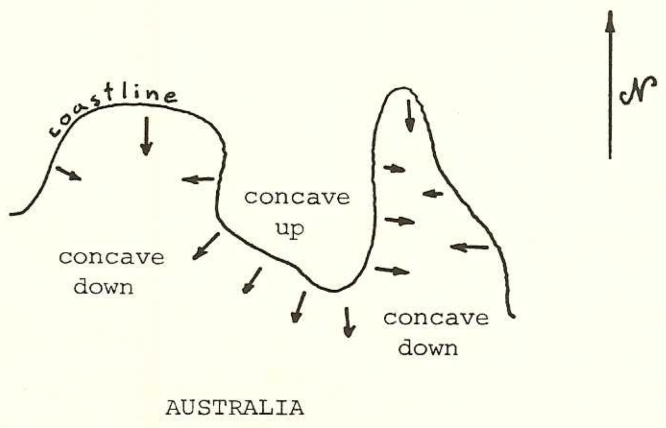


FIGURE 1

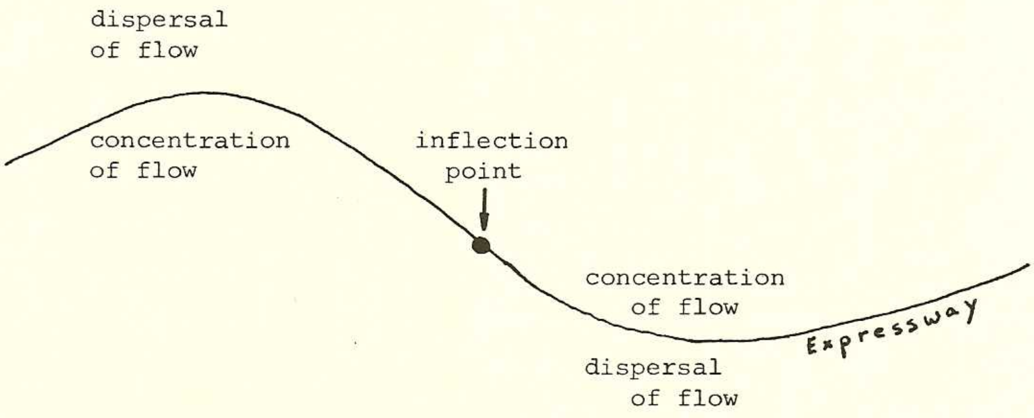


FIGURE 2

Thus, the position of concave-up, and of concave-down, segments of an expressway helps to guide the direction of concentrated settlements within a metropolitan region. A new expressway, planned to speed cars past an already densely-settled urban area, could force an additional concentration of surface-route flows in this area (Figure 3), and would have the potential to affect the traffic pattern of the entire metropolitan region. Simple reshaping of the curvature of the expressway, as in Figure 4, minimizes the shock of concentrated additional flow across expressway bridges into the already densely-settled region.

Further, a study of the history of an "inner city," within a northeastern American city, might profit from an analysis of the shape of interurban arterials (such as rivers, rails, and expressways) and their geometric impact on underlying concentrations of population and land-use during the era in which each artery-type was the dominant mode of interurban transport. Determining, in an historical context, whether or not new arterials forced additional concentration of surface flows within densely-settled areas might offer lessons for planning interurban arterials to link cities emerging in the latter half of the twentieth century.

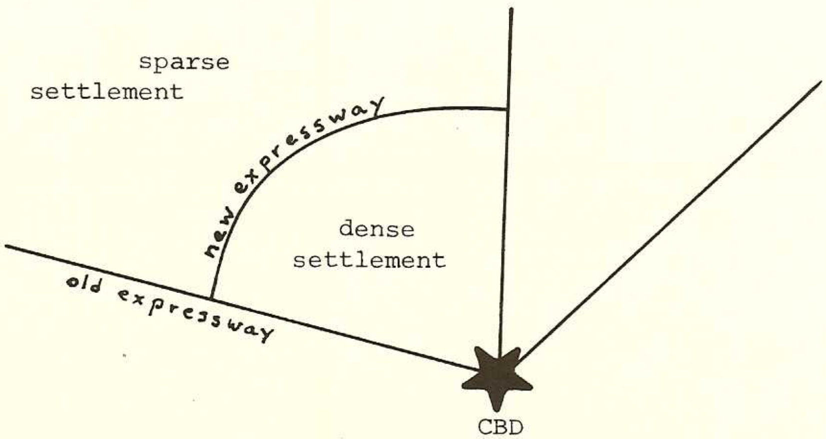


FIGURE 3

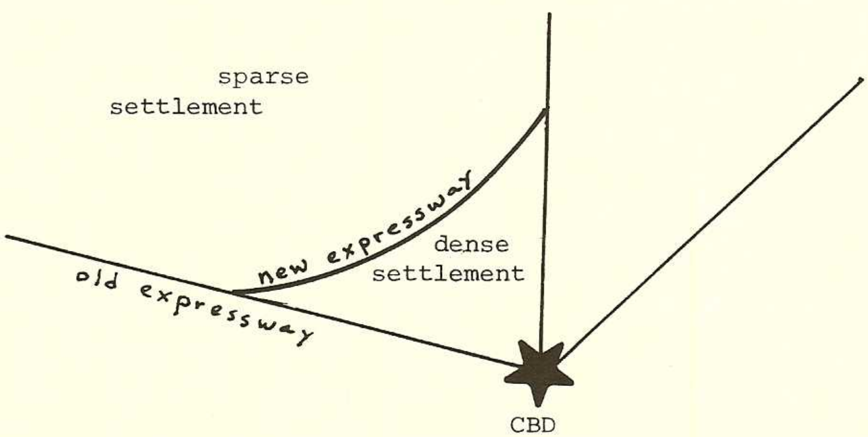


FIGURE 4

LITERATURE CITED

1. Birdsell, J. B. "Some implications of the genetical concept of race in terms of spatial analysis," Symposia on Quantitative Biology, Vol. 15, Origin and Evolution of Man. Long Island, New York: The Biological Laboratory, Cold Springs Harbor, 1950.

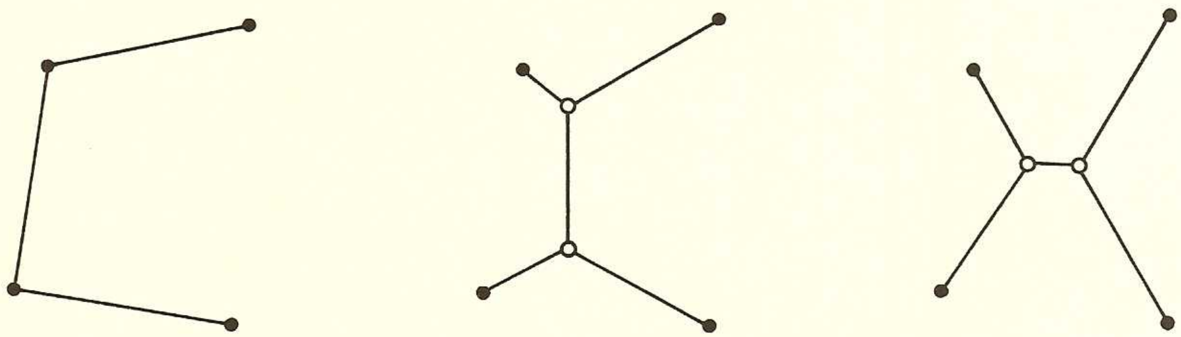
2. Kolars, J. F. Personal communication, September 8, 1975.

Professor Donald R. Deskins, Jr., of The University of Michigan, has suggested references, not cited here, appropriate to developing a ten-year study, (applying these ideas in Ann Arbor) that is in progress.

STEINER TRANSFORMATIONS

STEINER TREES

Networks that minimize total length of linkage joining a finite number of locations are graph-theoretic trees; there are no circuits providing redundancy of network connection. Figure 1 shows three possible "shortest" networks joining four vertices. Each is shortest within prescribed patterns of connection: in Figure 1a, linkage joins the four vertices with no additional vertices; in Figure 1b, linkage joins the four vertices, using two additional vertices, in such a way that the upper and lower vertices are grouped; and, in Figure 1c, linkage joins the four vertices, using two additional vertices, in such a way that the left-hand and right-hand vertices are grouped. Each of these is minimal within prescribed connection constraints, and each is a candidate as the shortest tree. The shortest tree, in any complete set of candidate trees, is called the Steiner tree, after Jakob Steiner, a nineteenth century mathematician [Courant and Robbins, 1958].



a b c

FIGURE 1

Determining the Steiner tree is generally difficult; there are an infinite number of points available, to serve as elements of a tree, within the convex hull of any finite set of points in the Euclidean plane. The set of candidate Steiner trees gets large rapidly as the number of vertices increases; there are 105 candidates in the six vertex problem [Gilbert and Pollak, 1968]. The theory for locating Steiner trees is available in the mathematics literature [Cockayne, 1967, 1969, 1970]; programs for generating Steiner trees are available from Bell Laboratories [Boyce and Seery, 1975].

AN APPLICATION OF STEINER TREES

One use to which Steiner trees might be put is to design routes through urban parklands. Urban neighborhood parks typically have exterior boundaries controlled by surrounding land acquisitions; however, the interior is open to design, and is basically Euclidean [Nystuen, 1983] as movement is possible in all directions from every point not on the boundary.

Basic Assumption

In urban parklands it is desirable to have a minimum of parkland replaced by cement.

Thus Steiner trees, or candidate trees, that are shortest forms within a prescribed connection pattern, would respond to the Basic Assumption as a design tool for route location in parks.

Figure 2 shows how an urban neighborhood park might be designed using Steiner procedures. Suppose there are four points of entry to the open area, produced at street intersections in a grid pattern. People who use the park have open access to it from these points only; entry is blocked at other points along the convex hull by fences or buildings. Routes through this park are required to serve pedestrian traffic and bicycle traffic; pedestrian traffic might include couples going for a walk, mothers with children in strollers or on tricycles,

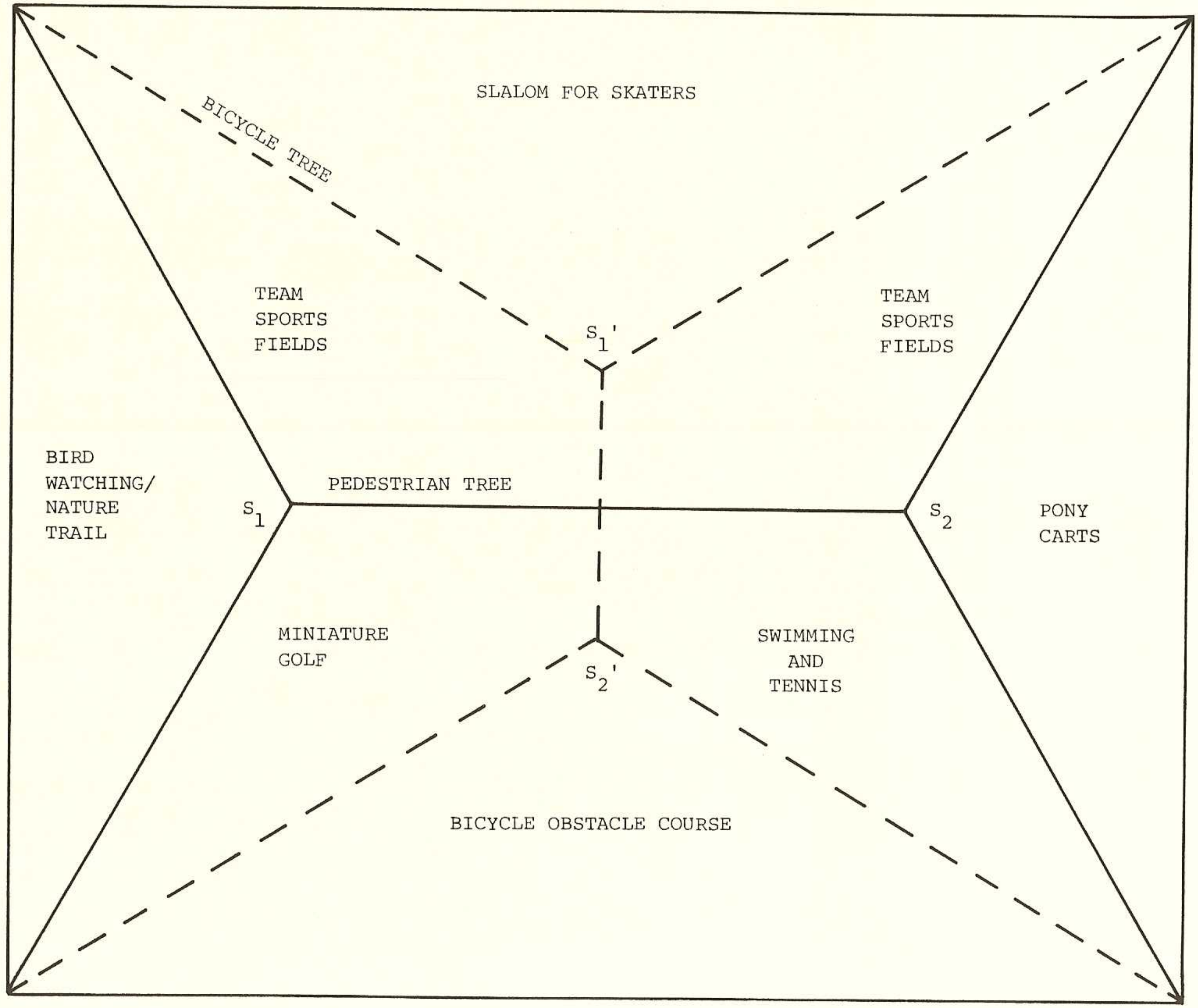


FIGURE 2 Steiner NeighborhoodPark

or people in wheelchairs. Bicycle traffic might include cyclists of all ages, skate-boarders, or roller skaters. Generally, these two basic flow types do not mix well--pedestrians are threatened by bicyclists, and bicyclists are frustrated by slow-moving pedestrian traffic which holds them back. Thus, to separate the flows, we need routes for each flow which seldom intersect and which are not near each other. If the Basic Assumption is also to be met, candidates for Steiner networks will serve well.

The pedestrian-tree and the bicycle-tree shown in Figure 2 are Steiner candidate trees; the bicycle-tree is the longer of the two, as bicyclists usually cover ground faster than do pedestrians. To guide flows to the appropriate tree, signs might suggest which tree to use; however, passive design criteria, such as route-surface texture might serve to sort flows even more effectively onto the appropriate tree [Nystuen, 1983]. In addition, facilities of particular interest to pedestrians, such as groupings of chairs and benches might be placed at the interior (Steiner) points (S_1, S_2) of the pedestrian-tree, while facilities such as bike racks might be placed at S_1' and S_2' in the bicycle tree. The single interior intersection of these two trees might be surrounded by facilities of interest to both groups, such as eating areas, restrooms, first-aid, a clock, or telephones. It might also house bicycle racks and bicycle rentals for individuals wishing to switch, comfortably, from one tree to the other. Activity areas would be organized around the routes: those likely to be of interest to both groups, such as for basketball, tennis, handball, or baseball, would be located in regions bounded by both trees and would have direct access from both trees. Those likely to be of interest predominantly to one set, such as a slalom for skaters, an obstacle course for bicyclists, pony carts for small children, or bird-watching areas for quiet people (away from roller skaters and playgrounds), would be located in regions bounded by the appropriate tree, only (Figure 2).

Neighborhood parks are one sort of park found in urban areas; another, the metropark, is usually located away from heavily built-up areas and contains a large number of low cost-per-unit parcels of land. Often there is an entry fee and entry is almost always by car or van. Cars converge at the entrances and then follow a park-wide main car route across which the objectives of global routing are to minimize path length while distributing that path evenly through the park [Kirkpatrick, 1983]. From the main route, cars branch off into various parking areas, around which park facilities have been organized as sub-parks [Huron-Clinton Metroparks Map, 1983]. Because the metropark might be viewed as being formed from smaller sub-parks, two strategies for positioning park-wide car routes arise naturally. One alternative is to use the Steiner tree linking the distinguished points of entry A_1, A_2, A_3, A_4, A_5 (Figure 3a), and the other is to link Steiner networks of each of these distinct regions, for example, as upper, middle, and lower subparks in triangles $(A_1A_2A_5)$, $(A_2A_3A_4)$, $(A_2A_4A_5)$ (Figure 3b). The first possibility (global) disregards access to this route from local subparks, while the second possibility (local) has linkage redundancy introduced by the presence of circuits (Figure 3c), and therefore responds neither to Kirkpatrick nor to the Basic Assumption. To eliminate this redundancy and obtain a tree that retains many of the characteristics of the individual local networks, yet exhibits the general structure of the global tree, a tool called a Steiner transformation will be introduced.

STEINER TRANSFORMATIONS

The shift in park-route scale, from global to local network, might be represented using the notion of self-similarity from fractal geometry

FIGURE 3a

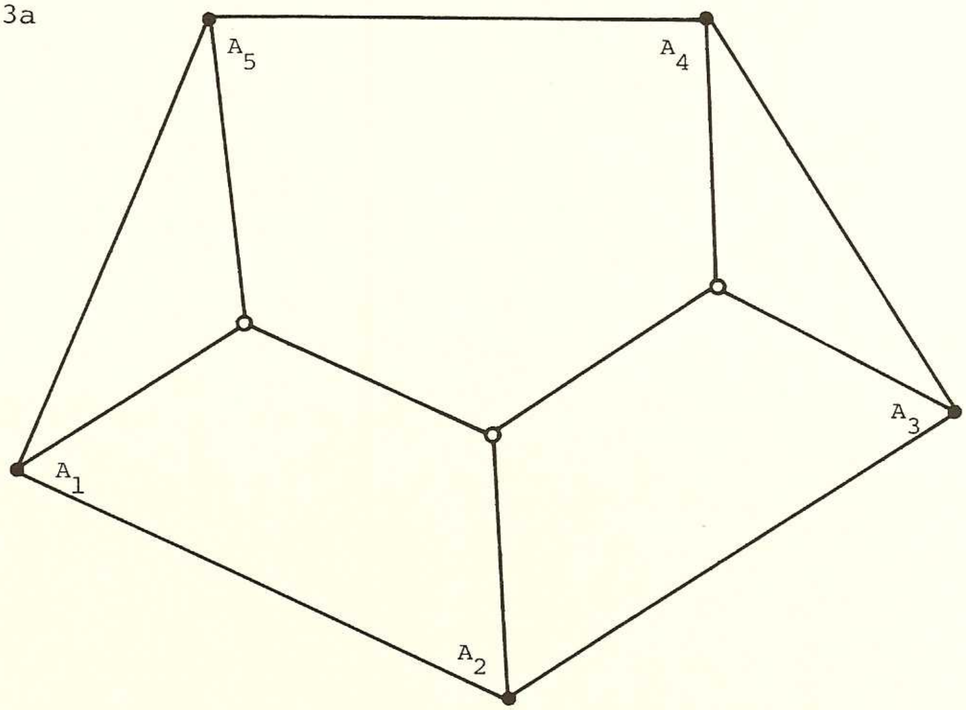
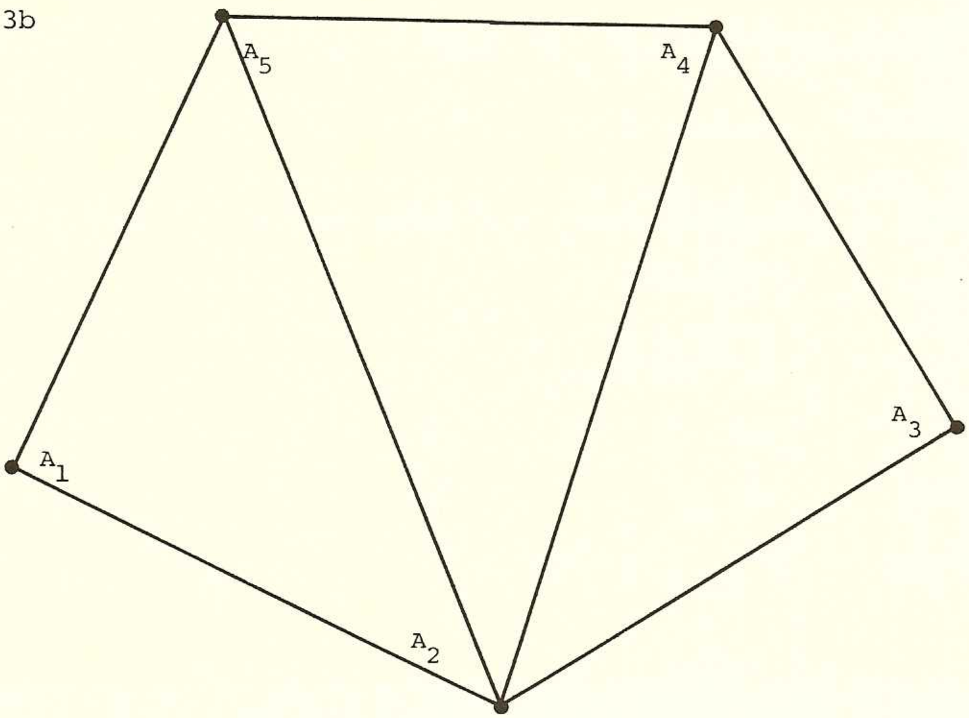


FIGURE 3b



[Mandelbrot, 1983]. In this approach, a geometric shape (a generator) is used to replace the sides of a given shape (initiator). With successive replacements, scale is enlarged or reduced. Like the fractal transformation, the Steiner transformation will be based on successive replacement which will permit it to transcend scale problems; unlike the fractal transformation, it will involve using the Steiner procedure, rather than a specified geometric shape, as that which is replaced.

<u>Steiner transformation</u> [Arlinghaus, 1977].

Given a network of contiguous, closed, polygonal cells. Locate the Steiner network within each cell and discard the initially given structure. Examine the new network; if closed polygonal cells remain, repeat the procedure. Continue until no closed polygonal cells remain. Network edges not included in the boundary of a closed polygonal cell remain invariant under the transformation. The resulting tree is said to be irreducible.

Successive applications of the Steiner transformation to the set of contiguous triangles in Figure 3b results in the sequence of networks, which results in the reduction of closed cellular matter, shown in Figures 3c to 3e. In Figure 3c, $\overline{S_1A_1}$, $\overline{S_1A_2}$, $\overline{S_1A_5}$, are the edges of the Steiner network in cell $(A_1A_2A_5)$ of Figure 3b; $\overline{S_2A_2}$, $\overline{S_2A_5}$, $\overline{S_2A_4}$, are the edges of the Steiner network in cell $(A_2A_5A_4)$ of Figure 3b; $\overline{S_3A_2}$, $\overline{S_3A_4}$, $\overline{S_3A_3}$, are the edges of the Steiner network in cell $(A_2A_3A_5)$ of Figure 3b. With the cell boundaries of Figure 3b discarded, the network, under the Steiner transformation, is as shown in Figure 3c. Then, apply the Steiner transformation to the two quadrangular cells in Figure 3c; cell $(S_1A_2S_2A_5)$ in Figure 3c is replaced by the Steiner network $\overline{A_5S_1'}$, $\overline{S_1'S_1}$, $\overline{S_1'S_2'}$, $\overline{S_2'S_2}$, $\overline{S_2'A_2}$ in Figure 3d, and cell $(S_2A_2S_3A_4)$ in Figure 3c is

FIGURE 3c

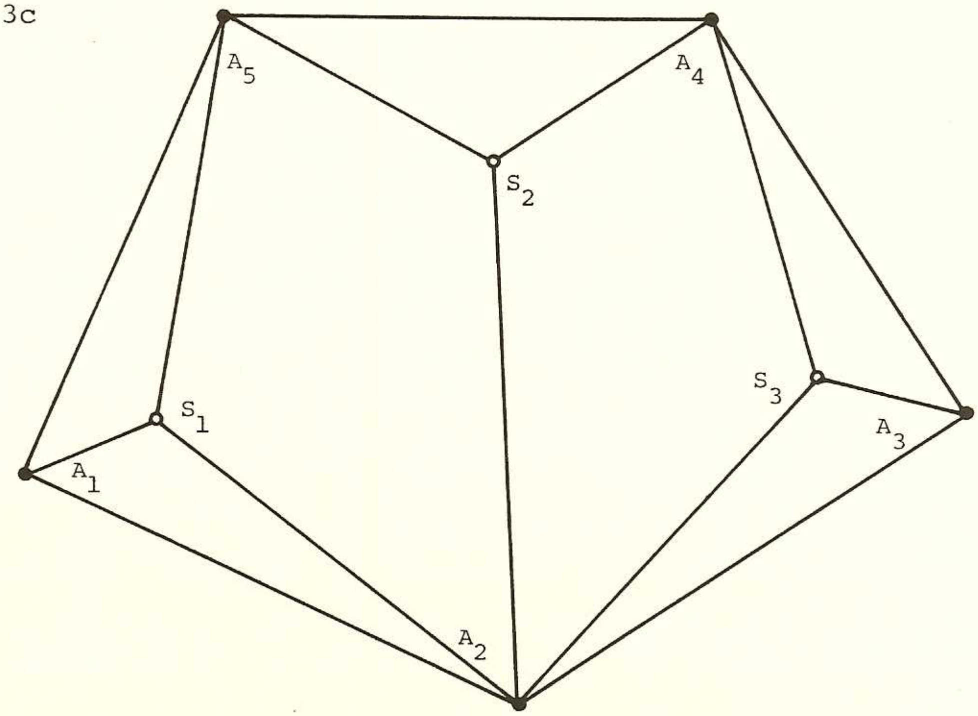


FIGURE 3d

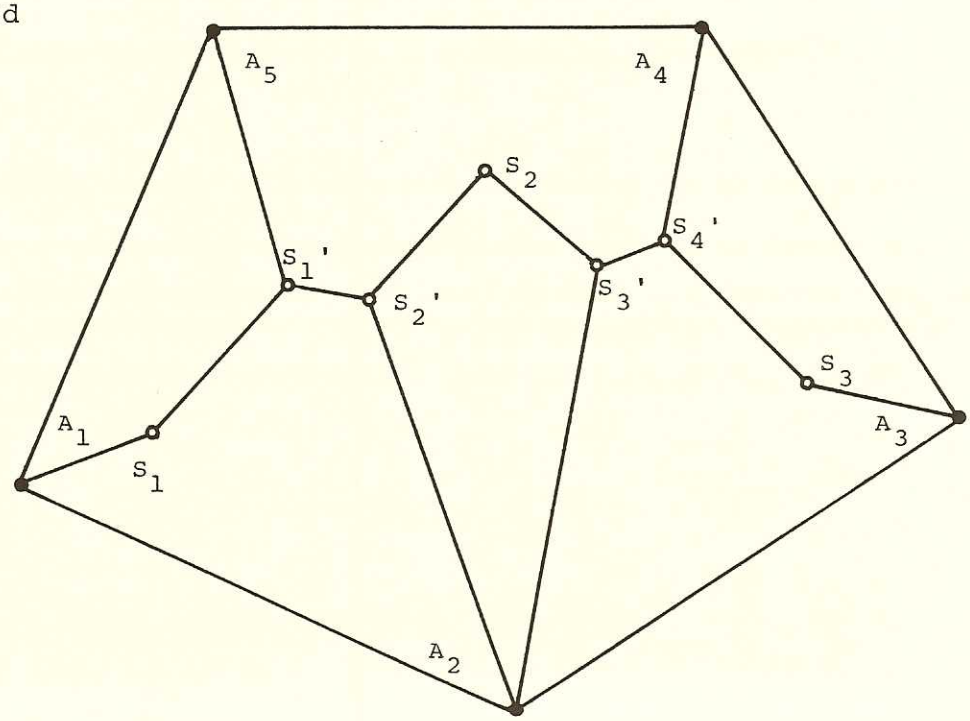
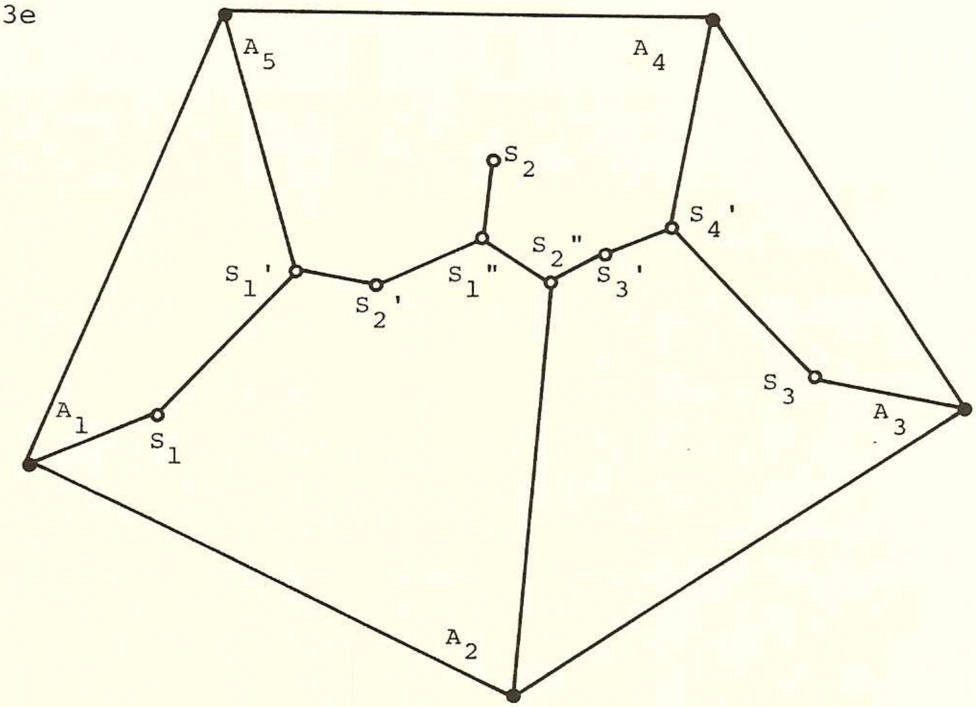


FIGURE 3e



replaced by the Steiner network $\overline{A_2S_3'}$, $\overline{S_2S_3'}$, $\overline{S_3'S_4'}$, $\overline{S_4'S_3}$, $\overline{S_4'A_4}$ in Figure 3d. The edges $\overline{A_1S_1}$, $\overline{S_3A_3}$ in Figure 3c are invariant under the transformation and thus appear once again in Figure 3d. Finally, apply the Steiner transformation to the one closed cell in Figure 3d; all else remains invariant and appears agaon in Figure 3e. Cell $(A_2S_2'S_2S_3')$ in Figure 3d is replaced by the Steiner network $\overline{S_2'S_1''}$, $\overline{A_2S_2''}$, $\overline{S_1''S_2''}$, $\overline{S_2S_1''}$, $\overline{S_3'S_2''}$ in Figure 3e. At this stage, the reduction is complete; no further cells remain.

Further examination of the Steiner transformation reveals conditions under which successive applications of the Steiner transformation generates a finite sequence of reductions. For otherwise, this style of network transformation would create a global network with enough links to choke the entire region from an infinite regeneration of cellular network growth (Figures 4a to 4c). Examination of Figure 4 shows that application of the Steiner transformation is not removing circuits; the vertex A_2 began as a vertex in four circuits in Figure 4a, remains a vertex in four circuits in Figure 4c, and apparently will remain so forever. However, the network structure surrounding A_2 will continue, out of control, through successive applications of the transformation, to fill the region around A_2 (an open question is to calculate a fractal-like dimension for this space-filling process).

Definition 1 [Tutte, 1966, p. 102].

A wheel W_n of order n, $n > 3$, is a graph obtained from an n-gon, P_n, by inserting one new vertex h, the hub, and by joining h to at least two of the vertices of P_n by a finite sequence of edges (A_2 is the hub of a wheel formed in P_4 in Figure 4a).

FIGURE 4a

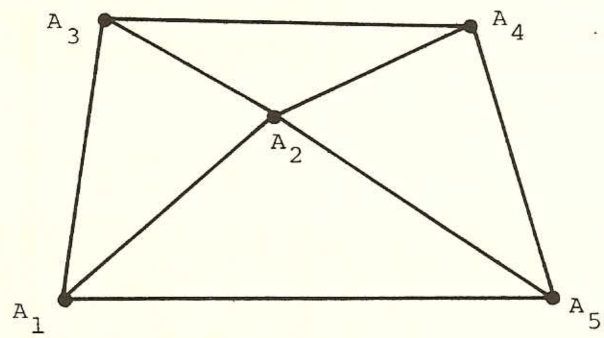


FIGURE 4b

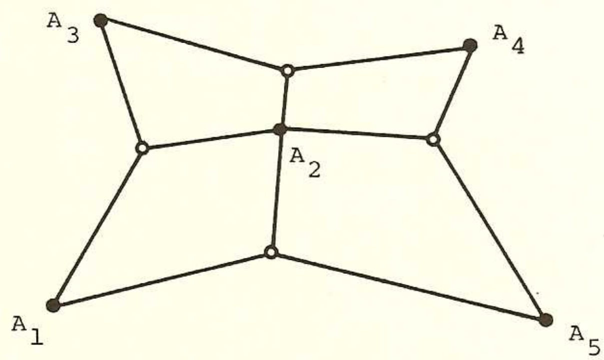
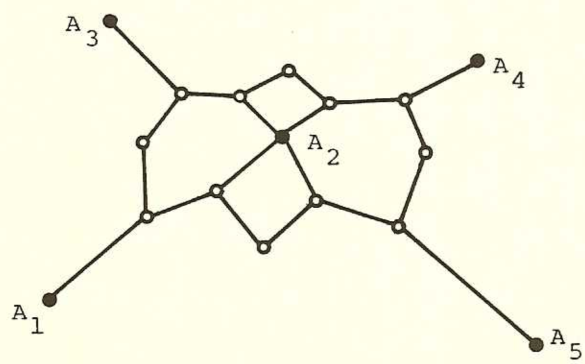


FIGURE 4c



Lemma 1 [Arlinghaus, 1977]

Hubs of wheels are invariant, as hubs of wheels, under any sequence of successive applications of the Steiner transformation, S.

This is clear since no polygons incident with the hub are removed in applying S to a wheel; polygon shape may change through successive application of S, but the degree of the hub is never reduced.

Theorem 1 [Arlinghaus, 1977].

Suppose that there exists a finite set of contiguous triangles $T = \{P_1,\ldots,P_m\}$ with vertex set $V = \{A_1,\ldots,A_n\}$, $n > m$ (as in Figure 3b, $m = 3$, $n = 5$). If T contains a wheel, then a sequence of successive applications of S to T fails to produce an irreducible tree; the sequence fails to terminate.

This is a consequence of Lemma 1.

Corollary 1 [Arlinghaus, 1977].

Suppose that T and V are as in Theorem 1. If T contains a wheel, and if a degenerate Steiner network arises during a sequence of successive applications of S to T, then this sequence may terminate.

Theorem 2 [Arlinghaus, 1977].

Suppose that T and V are as in Theorem 1, and that T does not contain a wheel. The number of steps M, in the sequence of successive applications of S to T, required to reduce T to a tree, is

$$M = (\max(\mathrm{degree}(A_i))) - 1.$$

Indication of proof:

Since T does not contain a wheel, it follows from Theorem 1 that the reduction sequence is finite. Proof of the actual size of M is by induction on the number of cells in T and on the degree of A_i.

Figure 3a shows the global Steiner network joining A_1, A_2, A_3, A_4, A_5; when it is compared to the network in Figure 3e, the same general structure is evident. Thus, a sequence of successive applications of the Steiner transformation helps to resolve scale problems by drawing together local and global network forms, reflecting that

> In nature, parts clearly do fit together into real structures, and the parts are affected by their environment. The problem is largely one of understanding. The mystery that remains lies largely in the nature of structural hierarchy, for the human mind can examine nature on many different scales sequentially, but not simultaneously [C. S. Smith, in A. Loeb, 1976, p. xiv].

AN APPLICATION OF STEINER TRANSFORMATIONS

Steiner MetroPark (Figure 5) emerges from the Steiner transformation, within the boundaries of Figure 3, in much the way that Steiner NeighborhoodPark arose from a set of candidate Steiner trees. The irreducible Steiner tree (Figure 3e), formed from breaking the parkland into upper, middle, and lower, parcels of land (Figure 3b), exhibits the general structure of the Steiner tree (Figure 3a), yet retains much of the Steiner network structure specific to each parcel (Figures 3b to 3e).

Suppose A_1, A_2, A_3, A_4, A_5, are all distinguished as entrances in Figure 5; A_4 and A_5 are entrances from a road while A_1 and A_3 are entrances from bridges across a river which forms one edge of the park. The vertex A_2 represents a boat landing and rental area. Here, the exterior geometrical boundary conforms to environmental guidelines; the river is a bar to wheel formation, so that a park-wide tree with both global and local characteristics can be found as the end of a reduction sequence of applications of the Steiner transformation (by Theorem 2). Each Steiner point involved in the reduction sequence, S_1, S_2, S_3;

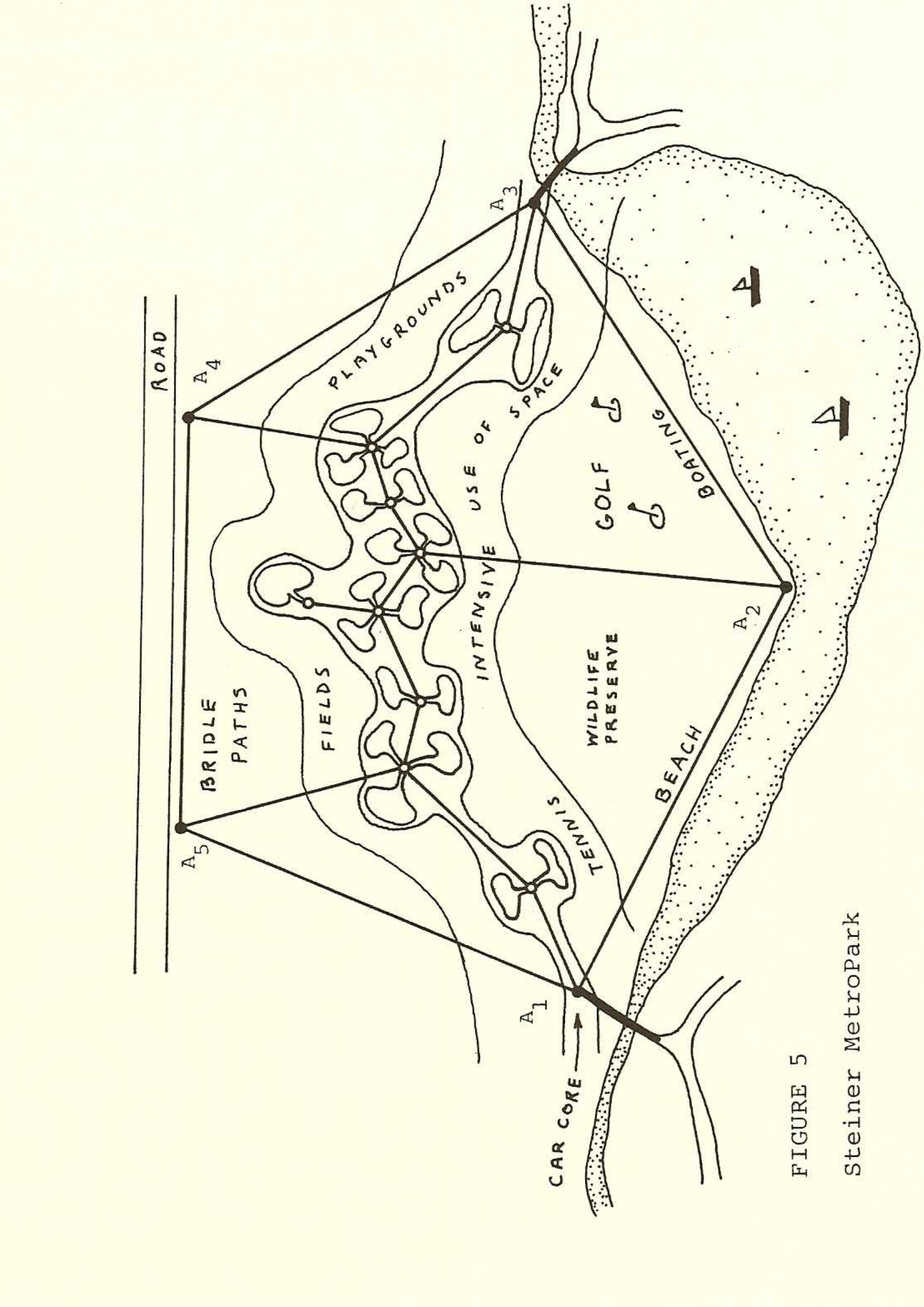


FIGURE 5

Steiner MetroPark

S_1', S_2', S_3', S_4'; S_1'', S_2'' would serve as a base for a parking area around which a variety of activities could be organized. The number of lots incident with a vertex would be the same as the degree of that vertex in the irreducible tree. The clustering of parking areas along a central car-core suggests locating activities which make intensive use of the land (such as those found in a neighborhood park) near these lots, and locating activities that require larger expanses of land near the edges of the region. For example, playgrounds, tennis courts, and team sports fields would reflect activities requiring intensive use of parkland and would be located near the car-core, while golf courses, bridle paths, and wild life preserves would reflect activities requiring less intensive use of space, and would be located near the park's periphery. The land-use pattern induced by this irreducible tree derived from the Steiner transformation is reminiscent of the ideas underlying von Thünen's isolated state; only here, it is the park that is isolated, as metroparks often are, from the underlying urban population which supports them.

LITERATURE CITED

1. Arlinghaus, S. L. "On Geographical Network Location Theory," Ph.D. Dissertation, Department of Geography, The University of Michigan, 1977.
2. Boyce, W. M. And Seery, J. B. An Improved Version of Cockayne and Schiller's Program STEINER for the Minimal Network Problem. Bell Laboratories, Computing Science Technical Report #35, 1975.
3. Cockayne, E. J. "Computation of minimal length full Steiner trees on the vertices of a complex polygon," Math. Comp., 23, 107, (1969), 521-31.
4. Cockayne, E. J. "On the efficiency of algorithm for Steiner minimal trees," SIAM, Jour. of Appl. Math., 18, 1, (1970), 150-59.
5. Cockayne, E. J. "On the Steiner problem," Canad. Math. Bull., 10, (1967), 431-50
6. Courant, R. and Robbins, H. What Is Mathematics? London: Oxford University Press, 1958.
7. Gilbert, E. N. and Pollak, H. O. "Steiner minimal trees," SIAM, Jour. Appl. Math., 16, (1968), 1-29.
8. Huron-Clinton Metroparks and Five County Road Map, 1983.
9. Kirkpatrick, S., Gelatt, C.D. Jr., and Vecchi, M. P., "Optimization by simulated annealing," Science, 220, (1983).
10. Loeb, A. L. Space Structures: Their Harmony and Counterpoint. Reading, Mass.: Addison-Wesley, 1976.
11. Mandelbrot, B. F. The Fractal Geometry of Nature. San Francisco: W. H. Freeman, 1983.
12. Nystuen, J. D. Personal communication, via Colloquium in Mathematical Geography, Ann Arbor, 1983.
13. Tutte, W. T. Connectivity in Graphs. Toronto: Univ. of Toronto Press, 1966.

ANALOGUE CLOCKS*

To see a World in a Grain of Sand
And a Heaven in a Wild Flower
Hold Infinity in the palm of your hand,
And Eternity in an hour.

William Blake, "Auguries of Innocence."

Watches and clocks that keep track of the passing time using a digital display have achieved widespread distribution in recent years. Clock counters in department stores exhibit some traditional clocks and watches with a face containing twelve numerals and hands (analogue clocks), although frequently their stock is dominated by digital products. The distinction between these two types, made by those who market clocks, is that digital provides discrete tracking of the time while analogue produces a continuous display [Ford Motor Company, 1983].

Beyond that, however, the word "analogue" means "something that is similar to something else;" thus the question is, to what else is an analogue clock similar? [Webster, 1965]. One obvious answer is a sundial, the forerunner of the mechanical clock [Cipolla, 1967]. Physical evidence from sundials of both the Northern and the Southern hemispheres suggests two types of dial: the horizontally mounted face frequently found in gardens and the vertically mounted dial often embedded in building walls. In the Northern hemisphere (north of 23.5° North Latitude), horizontal faces require clockwise orientation of numerals to record the time and, vertical dials need a counterclockwise arrangement. In the Southern hemisphere (south of 23.5° South Latitude) the opposite holds; horizontal dials need counterclockwise orientation of numerals and, vertical dials require a clockwise pattern. The reversal in orientation of numerals, which distinguishes a horizontal from a vertical sundial of the same hemisphere, is a result of the switch in position of the background on which the shadow is cast (the reader can verify this experimentally using one pen inclined at a 90° angle to a piece of paper in a horizontal position, and another inclined at a 90° angle to a vertical paper, to mark the passing time). Between the Tropics, the Northern heimsphere approach holds as long as the direct ray of the sun is overhead south of the dial, while the Southern hemisphere orientation applies when the direct ray is north of the dial. The orientation of the numerals on the analogue clocks commonly in use today corresponds to that of a Northern hemisphere garden sundial, or, equivalently, to a Southern hemisphere wall mounted model, although occasional models such as the "O.K. Now Alternative Analog Timepiece," designed by the artist Victor 1111 of Amsterdam, offer a clock face with counterclockwise orientation of the numerals [Ann Arbor News, 1984]. In contrast to analogue clocks, however, the sundial is not portable, and it records hours of varying length depending on the season and on the latitude; thus more powerful analogy is sought.

Clocks that were portable and that measured a standard hour (rather than a varying or "temporary" hour) led to the solution, by about 1750, of the 2000 year old problem of measuring longitude as sea [Brown, 1956]. Suppose, for example, that a ship at location A, with an accurate clock on board, set sail at 11:00 A.M. (on the clock) and sailed west until it reached local noon, determined from a sequence of readings of a sextant, at location B [Forbes, 1974]. At the local noon meridian through B, the clock on board read 11:27. Since this local noon meridian through B will coincide with the actual noon meridian through A when the clock reads 12:00, it follows that in 33 minutes more of rotational time the meridian through B will arrive at the meridian through A. Thus B is 33 minutes of rotational time west of the longitude of A, or B is 8°15' of longitude west of the longitude of A. As early as 1530, Gemma Frisius understood the theory of how to measure longitude at sea using a clock, but he had no clock suited to that task [Brown, 1956].

Christiaan Huygens's application of Galileo's discoveries of physical laws governing the motion of a pendulum led to the pendulum clock by 1656. These clocks were quite accurate on land, but they were difficult to transport and certainly were not precise on a turbulent sea [Brown, 1956; Forbes, 1974]. The use of Robert Hooke's principle that the force exerted by a spring is directly proportional to the spring's stretched length minus its length at rest permitted John Harrison (carpenter and clockmaker) to construct a sequence of clocks with springs, rather than pendulums, as regulatory mechanisms [Brown, 1956; Forbes, 1974]. These were portable, and by 1756 Harrison and his supporters had proven them accurate at sea to within three seconds a day over a period of six weeks [Brown 1956]. Thus, by the middle of the eighteenth century, all the equipment necessary to measure longitude at sea was available. Clocks based on Harrison's construction permitted continuing exploration westward into the New World, and they formed the basis for navigational fixes until the development of radio and atomic clocks in the twentieth centruy [Cipolla, 1967].

In addition, the twelve hour analogue clock face serves directly as a structural replica of the relationship between longitude and time. This is not immediately apparent; lack of clarity arises from the simultaneous partitioning of (i) the clock face into twelve equal central angles each containing 30° of angular measure and each representing one hour of elapsed time, and of (ii) the equatorial diametral plane of the earth into twenty-four equal central angles each containing 15° of angular measure (longitude) and each representing one hour of elapsed time. If one supposes the center of the analogue clock face to be superimposed on the center of the earth, within the equatorial diametral plane, then the partitions do not mesh, and this clock is not a structural model of this relationship. However, if this natural, but unnecessary, supposition is discarded, and a theorem from Euclidean geometry is invoked, then this twelve hour clock face is not only a precise analogue model of the relationship of longitude to time but is the most efficient one as well.

Theorem of Euclid

"In a circle the angle at the center is double the angle at the circumference, when the rays forming the angles meet the circumference in the same two points" [Coxeter, 1961, p.7], so that $\angle POQ = 2(\angle PO'Q)$ in Figure 1.

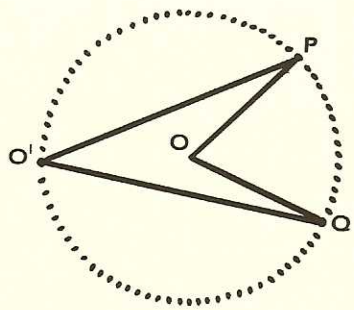


FIGURE 1

Basic Construction

The following basic construction for placing a configuration of clock faces in the earth's equatorial diametral plane will permit alignment of longitude and time (Figure 2).

a) Inscribe two circular clock faces, with centers O_1 and O_2, in the equatorial diametral plane of the earth in such a way that these circles are mutually tangent at the center of the earth, O. The centers O, O_1, and O_2 are collinear.
b) Partition the globe in the standard way with parallels and meridians.
c) These clock faces will be labelled as analogue clock faces modulo 12. The choice of label for the point of tangency for the two clock faces is arbitrary. When the numeral 6 is chosen as the label for this position, the choice of

noon and midnight for the remaining two points of contact is forced, and the configuration will be said to be in standard position. The locations for the remaining numerals are not specified here; either orientation is possible.

Fundamental Theorem

Given a configuration of clock faces in the earth in standard position, let R be an arbitrary point in the equatorial diametral plane used to make points (numerals) on the clock's perimeter correspond with points on the equator (Figures 3 and 4). Let P and Q be two consecutive numerals on the clock's perimeter. Let P' and Q' be the points where the lines RP and RQ intersect the equator. Then P' and Q' have longitudes differing by 15° if and only if R = O (i.e., the spacing between consecutive numerals on an analogue clock face is a precise measure of longitude if and only if R = O).

Proof:

I) Assume R = O (Figure 3).

Choose two arbitrary consecutive numerals, P and Q on the clock face centered on O_1. Linking these to O_1 forms $\angle PO_1Q = 30°$, which is a measure of one hour of time on the clock face. Link P and Q to O. By the Theorem of Euclid, $\angle POQ = 15°$. Extend the sides OP = RP and OQ = RQ to pierce the equator at P' and Q' respectively; thus $\angle P'OQ' = \angle POQ = 15°$, since $\angle P'OQ'$ is central within the equator. Therefore the spacing between P' and Q' is 15° of longitude, providing the desired conversion of clock time to longitude. The same argument applies for this position of R if P and Q are chosen on the clock face centered on O_2 (if P is on one clock face and Q is on the other then one of P and Q must be O since P and Q are consecutive).

FIGURE 2

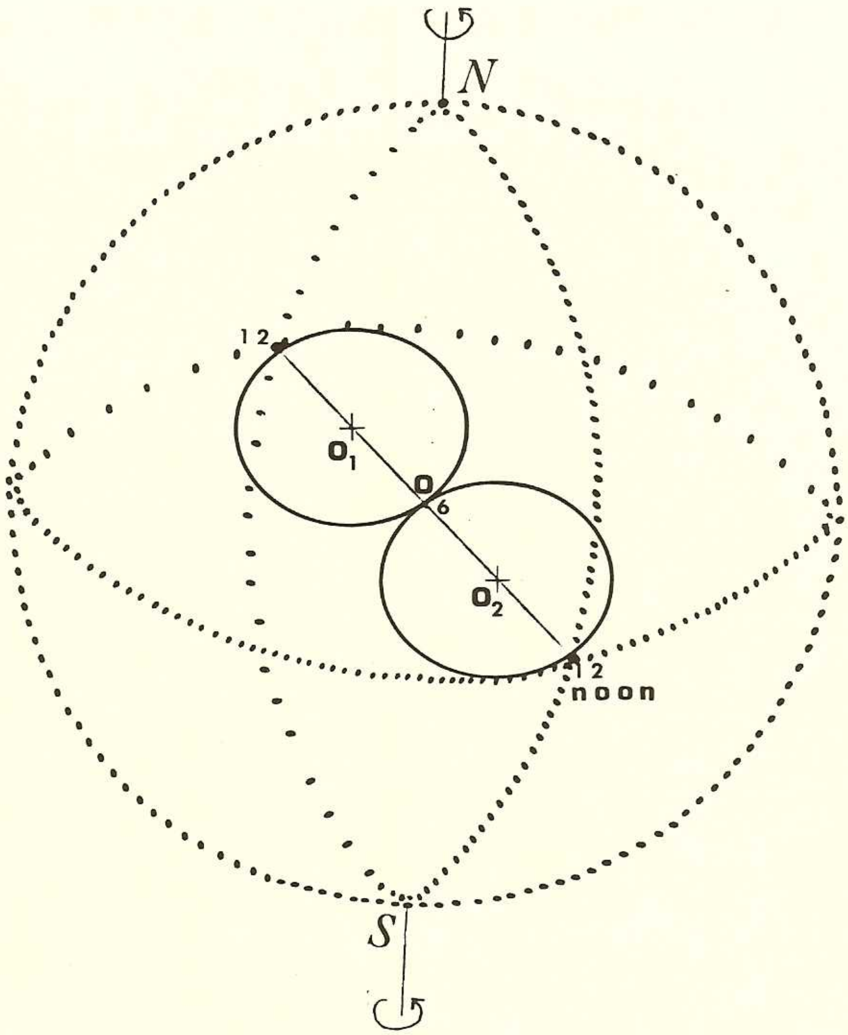
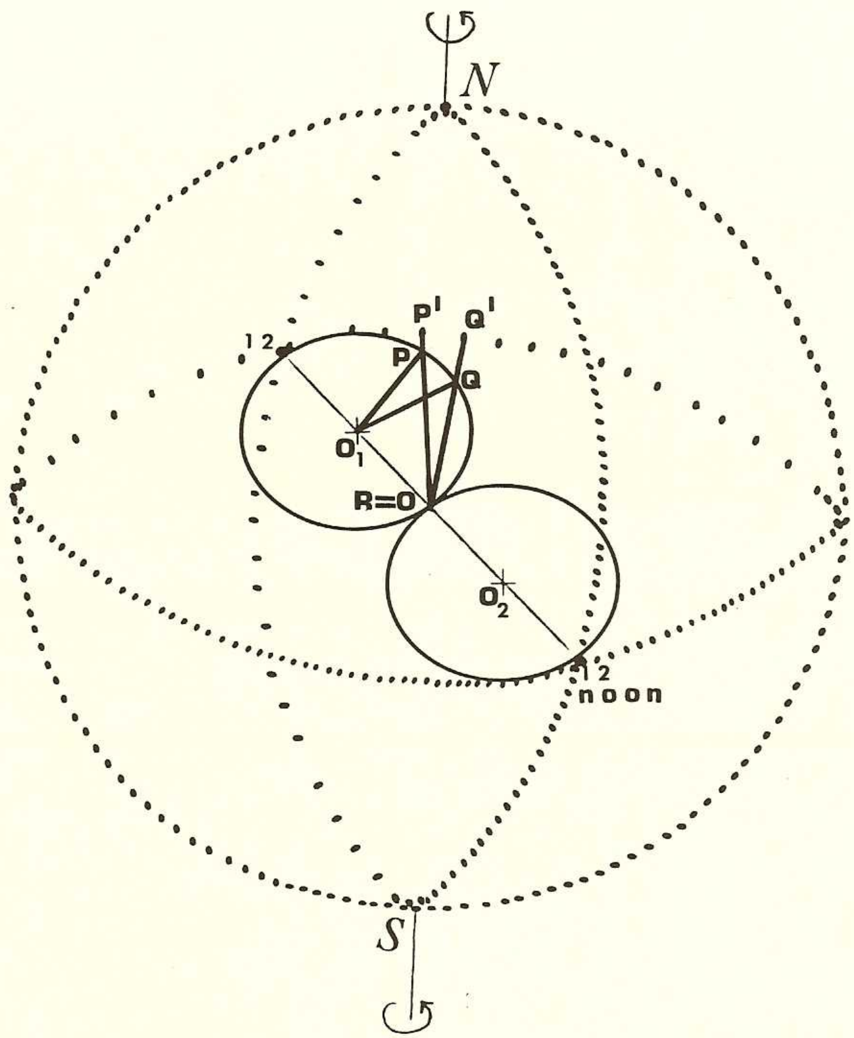


FIGURE 3



II) Assume R $\neq$ O (Figure 4).

a) Suppose R is not on the perimeter of either clock face.

i) Suppose R is outside one clock face.

Join R to the pair of consecutive clock-face numerals, P and Q, that are nearest it. Extend the segments RP and RQ to intersect the equator at P' and Q'. Locate the points $\overline{P}$, $\overline{Q}$ that are antipodal to P and Q, with respect to the clock face. Since the segments $P\overline{P}$ and $Q\overline{Q}$ have the same length and bisect each other, the quadrangle $(PQ\overline{P}\overline{Q})$ is a rectangle. Extend the segments $R\overline{P}$ and $R\overline{Q}$ to intersect the equator at $\overline{P}'$ and $\overline{Q}'$. Because the quadrangle $(PQ\overline{P}\overline{Q})$ is a rectangle, the short arc $\overline{P}'\overline{Q}'$ is properly nested within the short arc P'Q'. Thus the two arcs are not equal, and so at least one of them does not have measure 15°. Therefore this style of position for R does not convert time measured on the clock face to time measured by shifts in longitude resulting from the rotation of the earth on its axis.

ii) Suppose R is within one clock face.

Then R is outside the other clock face and this case reduces to case II.a.i.

b) Suppose R is on the perimeter of a clock face.

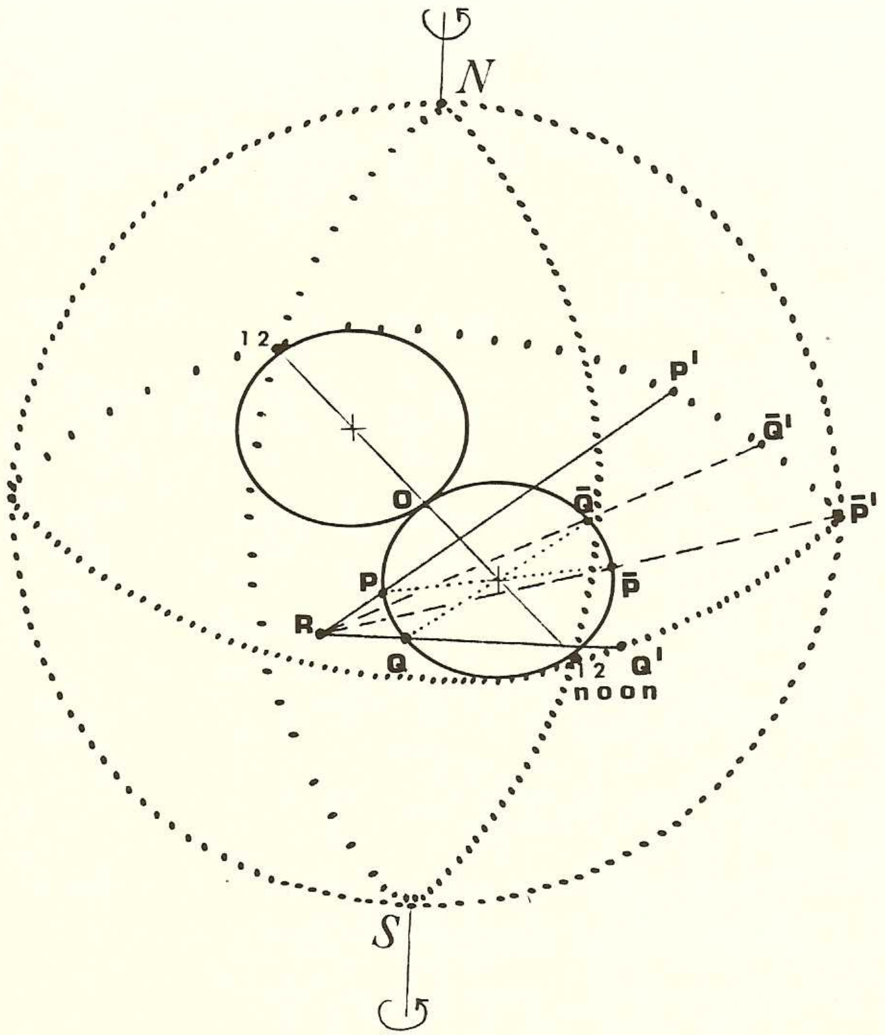
Since R $\neq$ O, R is outside one clock face, and this reduces to case II.a.i.

Q.E.D.

<u>Corollary</u>

Further refinement of the partitions of clock-time and of longitude into minutes and seconds leads to interpretations of the minute and second hands similar to those given by the Fundamental Theorem for the hour hand.

FIGURE 4



The Fundamental Theorem ensures that the configuration in standard position provided by the Basic Construction will permit direct use of the spacing of numerals on a clock face to measure longitude. Next the positions of the remaining numerals for the clock face will be determined from longitude as the earth rotates on its axis.

Ordering Theorem:

There exists a succession of positions measured from O in which clock-face numerals will coincide with meridian positions appropriate to the natural ordering established by the rotation of the earth on its polar axis (Figure 5).

Proof:

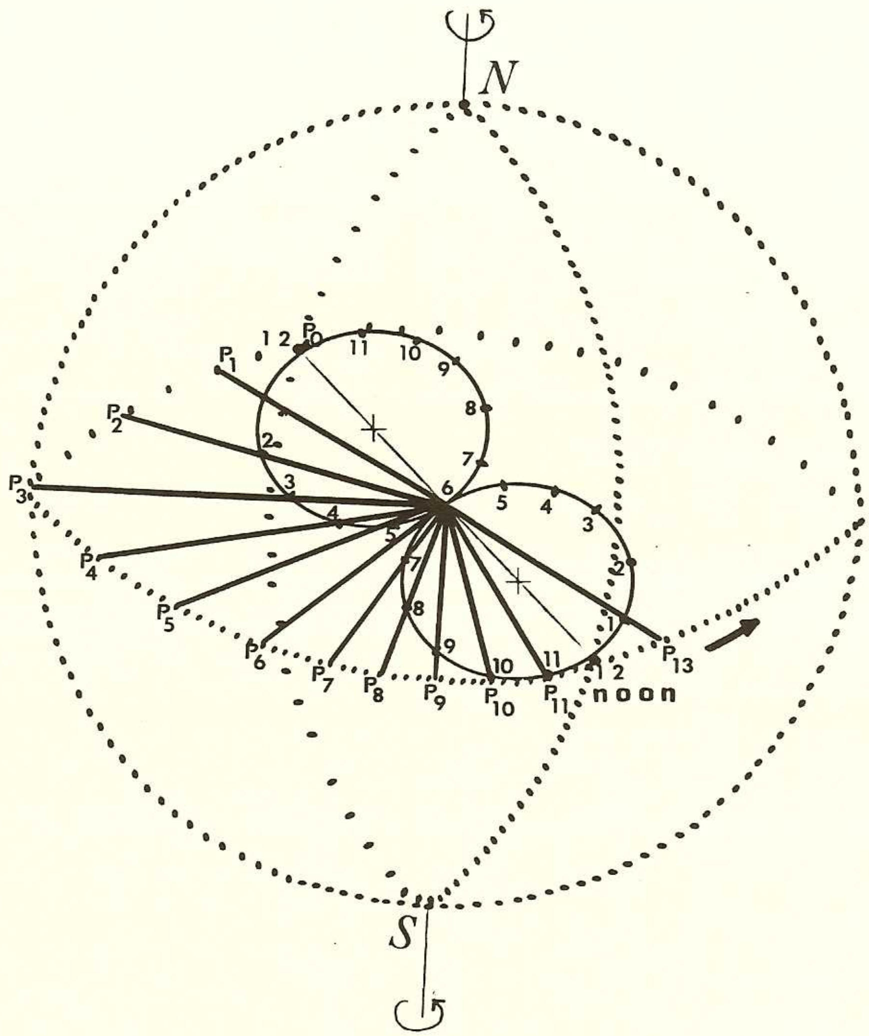
Proof is by construction of a succession of positions of an hour hand centered at O. Initial position: join O to 12 midnight, and call this position P_0 on the earth. After one hour, position P_0 will have rotated 15° to position P_1. Label the corresponding position on the clock face with the numeral 1. By the Fundamental Theorem the spacing between 12 and 1 on the clock face will be such that $\angle P_0 O_1 P_1 = 30°$. After another hour, P_1 will have rotated through 15° to P_2. The corresponding clock position will be labelled 2. Continue this procedure, shifting to the clock face centered at O_2, once 6, the position corresponding to P_6, has been reached. Thus, after another hour, P_6 will have rotated to P_7 and the corresponding position of the clock face centered on O_2 will be labelled 7. After 13 hours the point P_0 will have rotated to P_{13}, through 195° of longitude, and two alternate labelling strategies arise.

I) Label the corresponding clock position 13. Continuing this labelling would produce a scheme requiring 24 different position for numerals on the associated clock face.

FIGURE 5

Western Hemisphere

Eastern Hemisphere



II) The position P_{13} is antipodal in the sphere to the position P_1. Identify (glue together, abstractly) the labels of the associated antipodal clock positions. So, label the clock position corresponding to P_{13} with the numeral 1. Continuing this process requires only 12 different numerals on the clock faces to describe 24 different positions for P_0. To distinguish the time at position P_{13} from that at its antipodal point P_1, two conventions are in use:

a) read the time at P_{13} to be 13:00, as is done by the U. S. military and in continental Western Europe;

b) label the time at P_{13} as 1:00 after the sun's noon position, or 1:00 P.M., as is done in the United States. Positions in the hemisphere preceding arrival at the noon meridian are assigned A.M. suffixes to distinguish them from the times at their antipodal points in the hemisphere succeeding arrival at the noon meridian.

Q.E.D.

From the proofs of the previous theorems, it is clear that either a twelve-hour clock or a 24-hour clock may be used as an analogue model for the relationship between longitude and time that follows the natural ordering created by the rotation of the earth on its axis. Certainly the use of twelve distinct numerals, rather than twenty-four, is more efficient on small clock faces, such as wrist watches, and anywhere reduction of clutter of symbols is significant.

The previous theorem provided means for enumerating the clock face positions not labelled in the Basic Construction. Extracting such a clock face from the sphere produces the following theorem.

Orientation Theorem:

The use of 12 numerals arranged consecutively around the perimeter of a circle serves as an analogue model of (a) the relationship of longitude to time; (b) meridian positions corresponding to the natural ordering established by earth-sun relations. The orientation given to the numerals around the circle will depend on whether the observer is in the Northern or Southern hemisphere (Figure 6).

Proof:

The proof of (a) follows from the Fundamental Theorem. The proof of (b) follows from the Ordering Theorem. Extract a 12-hour clock face centered on O_1 from Figure 5; the orientation of numerals around this circle is a Northern hemisphere view of this clock, and so, for a clock to be an analogue model for the Northern hemisphere, it must be as in Figure 6a. Viewing Figure 5 from below produces the Southern hemisphere analogue clock of Figure 6b. (And, as with sundials, the reader may trace one of the clocks in Figure 6 and then view it from the other side to see, experimentally, the change in orientation resulting from change in hemisphere location.)

Q.E.D.

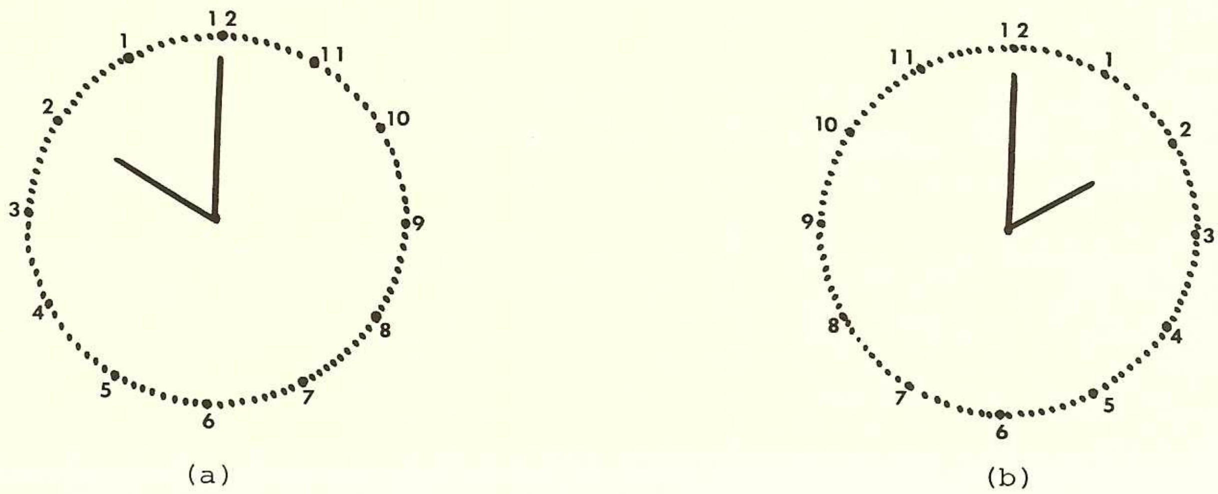


FIGURE 6

Uniqueness Theorem:

The twelve-hour clock face is the clock face with the fewest numerals that can serve as such an analogue model.

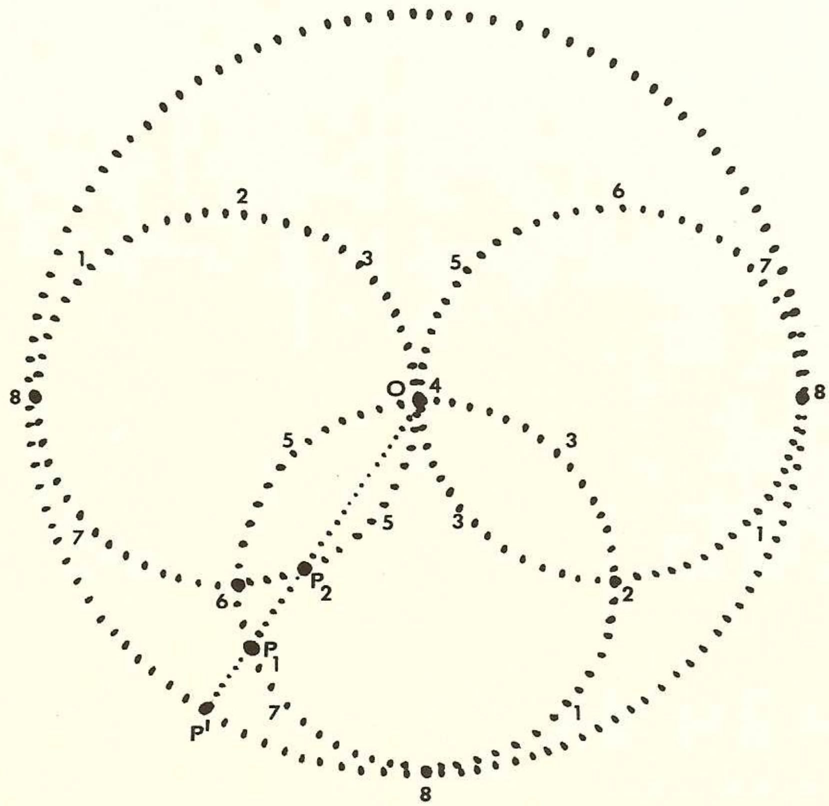
Proof:

This proof follows directly from the fact that no more than two circles in the (diametral) plane can be mutually tangent at a single point. For, suppose the number of numerals used, n, is a divisor of 24 that is less than 12. Then the Basic Construction will be possible, but the Fundamental Theorem will not. For suppose n = 8; then three complete rotations of the hour hand are necessary to account for the 24 hours in a day, and the Basic Construction would require three clock faces, labelled modulo 8, in the equatorial diametral plane. If these three circles intersect at O, in a number of cases (as shown in Figure 7), a single point P' on the equator corresponds to two distinct clock positions, P_1 and P_2, and the Fundamental Theorem cannot hold. To overcome this difficulty the circles must be tangent to each other; but then the center of the sphere, O, from which longitude is measured, cannot be included on all the clock faces simultaneously since three circles cannot be mutually tangent at a single point. Thus the Fundamental Theorem cannot hold. Identical arguments work for n = 6 requiring four clock faces, n = 4 requiring six clock faces, n = 3 with eight clock faces, n = 2 with twelve clock faces, and n = 1 with twenty-four clock faces.

If n is not a divisor of 24 and it is less than 12, then there is no possible representation of the Basic Construction.

Q.E.D.

FIGURE 7



The previous theorems all assume that a day, one complete rotation of the earth on its axis, is partitioned into 24 hours. The same theoretical structure holds in the more general case stated below.

Extension Theorem:

If one complete rotation of the earth on its axis is partitioned into 2n equal units of time, then a circular clock face with n equally spaced numerals arranged consecutively around the perimeter is the clock face with the fewest numerals that can serve as an analogue model of the relationship of longitude to time and of meridian position corresponding to the rotation of the earth on its polar axis.

Abstract open questions that remain could involve the suitability of partitioning the day into an odd number of equally spaced time units, and in particular, into a prime number of time units. Cultural and technological questions that remain could involve determining

I) why the orientation of numerals on present day analogue watches follows that of Northern hemisphere horizontal sundials rather than that of Northern hemisphere vertical sundials;

II) what sets of conditions have led Northern hemisphere inhabitants to wear Southern hemisphere analogue watches.

LITERATURE CITED

* Based on a paper given by the author January 25, 1984, Louisville, KY, at the 90th Annual Meeting of the American Mathematical Society. Abstract #809-51-468, Abstracts of the American Mathematical Society, Vol. 5, No. 1, 1984.

1. Brown, L. A. "The longitude," in Newman, J. R. (ed.) The World of Mathematics, Vol. 2. New York: Simon and Schuster, 1956.

2. Cipolla, C. M. Clocks and Culture: 1300-1700. New York: Walker and Company, 1967.

3. Coxeter, H. S. M. Introduction to Geometry. New York: Wiley, 1961.

4. Forbes, E. G. The Birth of Navigational Science: the solving in the 18th century of the problem of finding longitude at sea. Maritime Monographs and Reports, 10. Greenwich, England: National Maritime Museum, 1974.

5. Ford Motor Company. "Market Research and Ergonometric Studies Summary--Electronic Instrumentation and Features Products," Electrical and Electronic Division, Report, 2/18/83. Communicated by Mr. John Cleary, Product Development Team Manager, Ford Motor Company.

6. "Unusual watch the work of an artistic eccentric," Ann Arbor News, June 25, 1984.

7. Webster, Seventh New Collegiate Dictionary. Springfield, Mass.: G. and C. Merriam Co., 1965.

References 2 and 7 were suggested by Donald F. Lach, Bernadotte Schmitt Professor of History, The University of Chicago.

FAD AND PERMANENCE IN HUMAN SYSTEMS

The processes that lead to the evolution and decay of human systems seem as complex as the society from which these systems emerge. As time progresses the increase in degree of complexity of process reflects not only the level of internal interaction within that system but also the level of interaction of that human system with others of disparate cultural, economic, and political backgrounds.

Within this complexity, the following premise is basic to future work (both theoretical and empirical), and it is the cornerstone upon which the material below will rest.

Basic Premise

Given any degree of complexity of process, the evaluation of the relative permanence of human systems should produce significant insight into the processes of evolution and decay of these systems.

It is the goal of the present material to examine the theoretical potential of this premise, and of some of its implications, through definitional alignment of mathematical with social material.

One structural aspect of human systems might have appearance similar to that of a graph.

Definition 1

A graph G is a mathematical object formed from a finite number p ($p \neq 0$) of nodes (points, vertices) and from a set of q edges (lines) that join pairs of distinct nodes [Harary, 1969].

There are many definitions of the word 'graph', and there are many additional ideas, such as weighted edges, additional edges, and edges that are loops, that may be superimposed on this basic mathematical object. There are extensions of graph-theoretic ideas to simplicial complexes that might apply to a larger component of human systems [Atkin, 1976]. Only the simplest structure, the graph, will be dealt with here, to demonstrate concepts.

Definition 2

A human landform, H, is a structure that has its origins based in human activity and is one that requires the existence of that human activity for its continued presence on the surface of the earth.

A skyscraper, a college campus, a church, the institution of marriage, and inflation are a few examples of human landforms. Scale may vary, but the basic character of the human landform is reliant on human activity, be these structured elements of a region (either dense or diffuse) whose boundaries are determined by diversified human activity such as in a city, or be they elements of a region whose boundaries are determined by specialized human activity such as in a college campus, a political rally, or an accumulation of money. Human landforms represent a set of structures, rooted in human activity, that can be distinguished from one another. And, in that regard, they have formal, structural, properties similar to nodes in a graph.

Definition 3

A social network is formed from an arbitrary number of human landforms, represented structurally as nodes, that are joined by lines of channeled human interaction.

These links joining human landforms may be physical or non-physical. Rail, television, and sewage lines provide natural examples of physical linkage while systems of communication within academic, political, and economic groups suggest some non-physical linkages.

The expression of social networks in terms of human landforms leads to the development of procedure, as cited in the Basic Premise, for the evaluation of the relative permanence of social networks (as one component of human systems). Means for such evaluation will be based on one approach to the nature of the abstract structure of social networks and will employ a modification of recent work in which symbolic logic is used in graph theory to uncover classes of graphs that possess particular types of properties [Blass and Harary, 1979]. The material below exhibits the general idea of material that is dealt with mathematically by Blass and Harary [1979], and by Fagin [1976]; the statements are cast in language so that extension of them to the set of human landforms should appear natural.

Definition 4 [Blass and Harary, 1979]

A graph G is said to be an n-point graph if it has n nodes. In a set of n-point graphs, let P be a property of graphs. Form the ratio F(n), dependent on n, as F(n)=(# of n-point graphs having property P)/(# of n-point graphs).

If $\lim_{n\to\infty} F(n) = 1$, then almost all graphs are said to have P.

If $\lim_{n\to\infty} F(n) = 0$, then almost no graphs are said to have P.

For example, almost no graphs are trees; intuitively, this is not surprising. For, as n gets large, more and more links are introduced joining the n points in all possible ways, and one might imagine that the number of trees produced by such activity is small relative to the total number of possible graphs formed through such linkage procedure. However, proving that the magnitude of the numerator is insignificant relative to that of the denominator (that $\lim_{n\to\infty} F(n)=0$) requires combinatorial analysis resting in expressing the number of trees as a generating function stated in terms of analytic functions of a complex variable

[Harary and Palmer, 1973]. The statement that almost no graphs are trees is typical of a set of theorems that rest on the ideas of Definition 5.

A question of abstract interest greater than the formulation and proofs of specific "almost all" or "almost no" theorems, is to determine conditions for a style of property P that will yield the result that either almost all graphs have P or almost no graphs have P. Precise formulation of style of property is a significant question that has been dealt with more generally from a logical viewpoint by Fagin [1976] and that has been considered specifically with respect to graphs, and then extended to simplicial complexes, by Blass and Harary [1979].

Theorem 1 [Blass and Harary, 1979].

Given a suitable property P of graphs, either almost all graphs have P or almost no graphs have P.

For Fagin, and Blass and Harary, suitability must be expressed in terms of logical language that is first-order definable (Blass and Harary indicate directions for future research in expressing suitability in terms of other logical languages) [Blass and Harary, 1979]. For purposes of application to social problems, the crucial point is to note that it is the determination of suitability of a property P that is fundamental.

Also, note further that Theorem 1 is similar in form to the law of the excluded middle in that it deals with classes of statements that either hold or do not hold. It is not merely a mechanical tool, such as an index, that describes technical characteristics of networks. Since it deals with classes of properties, it appears to be the sort of theorem that could elevate the level of abstraction in application of graph theory, and of other formal structure, to social problems.

Analysis of the 'relative permanence' of a social network will be approached in terms of human landforms and graphs. Within mathematical structures 'permanence' describes the same idea as 'invariance', and invariant mathematical structure is that which remains fixed relative to a given transformation from one mathematical space into another [Renfrew and Cooke, 1979].

Definition 5

Permanence of any quality of a set of human landforms is recognized by identifying elements of that quality which remain invariant through time, relative to some appropriate transformation.

That is, a property of human landforms will be said to be permanent under a given transformation if application of that transformation to the property results in only limited distortion of the landform; the character of the landform is neither destroyed nor altered in such a way that it appears to be a different human landform.

The following empirical observations motivate Definition 6, below. Suppose that the set of all traditions in style of clothing is a set of human landforms covering the period of time from the primitive to the present. As suggested by evidence from art, the idea that at least some clothing is worn by living adults (and even by marble representations of such adults) is a property of this set of human landforms that is permanent--from loincloth and fig leaf, to slacks, skirts, and blue jeans. However, any one particular style of clothing, such as a Nehru jacket or a mini-skirt, viewed as a property of human landforms, might be a fad; only the time interval over which general disappearance occurs varies.

Definition 6

In a set of human landforms, $\{H_i | i \varepsilon I\}$, that exist over time, t, let P be a property of human landforms. Let H_t = (# of H_i at time t) and let $P(H_i)$ = (# of H_i at time t with P).

$$P \text{ is called a fad if } \lim_{t\to\infty} (P(H_t))/(H_t) = 0;$$

$$P \text{ is called permanent if } \lim_{t\to\infty} (P(H_t))/(H_t) = 1.$$

In the case of a fad, almost no H_i have property P as time progresses, reflecting the waning of interest characteristic of a fad; in the case of permanence, almost all H_i have property P as time progresses, reflecting the universal acceptance characteristic of permanence.

Definition 6 may also be applied to music. Auditory stimulation that produces a favorable response is permanent over time within the set of human landforms, although many specific types of music are fads that exist only in some relatively small time interval. Whether the music of Mozart is a fad that exists in a relatively long time interval, is permanent, or occupies a position between permanence and fad is a question that motivates the material that follows.

Certainly Theorem 1 applied to the class of human landforms becomes,

Theorem 2.

Given a suitable property of human landforms, either P is a property of human landforms that is permanent, or it is a fad.

In this form the theorem is apparently not true for a wide variety of qualities of human landforms, as with assessing the enduring quality of the music of Mozart. As in the mathematical case, the problem is to determine suitable properties for which Theorem 2 is true.

This is an open question, whose solution offers to shed light on the following conjectures.

1) The hierarchy of central places and associated sets of hexagonal nets, determined by variable spacing between rival centers (k-values) are (spatially) relatively permanent under the transformation of population growth [Christaller/Baskin, 1966].

The problem is to determine the class of time intervals in which distortion is controlled sufficiently that increase in population does not destroy pattern [Tobler, 1963]. Classical central place theory would then become a central place principle, valid in certain time classes within the larger, spatial, fractal geometry, thereby putting central place theory in a space-time context [Arlinghaus, 1985].

2) A geographical map is a human landform with the property P of potential to communicate.

Application would be to the class of time intervals in which distortion of communication through channels, represented as edges, is controlled sufficiently that understanding of map content and shape dominates other change in attitude resulting from landform changes over time [Tobler, 1963].

LITERATURE CITED

1. Arlinghaus, S. L. "Fractals take a central place," Geografiska Annaler, B 67, (1985), 83-88.

2. Atkin, R. H. Mathematical Structure in Human Affairs. New York: Crane, Russak, 1974.

3. Blass, A. and Harary, F. "Properties of almost all graphs and complexes," Journal of Graph Theory, 3, (1979), 225-240.

4. Christaller, W. Central Places in Southern Germany. (Baskin translation). Englewood Cliffs: Prentice-Hall, 1966.

5. Fagin, R. "Probabilities on finite models." Journal of Symbolic Logic, 41, (1976), 50-58.

6. Harary, F. Graph Theory. Reading, Mass.: Addison-Wesley, 1969.

7. Harary, F. and Palmer, E. Graphical Enumeration. New York: Academic Press, 1973.

8. Renfrew, C. and Cooke, K. L. (eds.) Transformations: Mathematical Approaches to Culture Change. New York: Academic Press, 1979.

9. Tobler, W. R. "Geographic area and map projections," The Geographical Review, 53, (1963), 59-78.

TOPOLOGICAL EXPLORATION IN GEOGRAPHY*

Discovery is the ultimate goal of topological exploration in geography, where point set topology "is concerned with those intrinsic qualitative properties of spatial configurations that are independent of size, shape, and location."[1] Geographic use of topology that superimposes 'size' (in the form of a metric) on topological structure could be viewed as quantitative geography or as applied location theory, that which superimposes 'shape' as cartography, and that which superimposes 'location' as abstract, qualitative, location theory.

The work which follows is general and attempts to exhibit correspondence between topology and geography without the introduction of 'size', 'shape', or 'location', until the final section, when 'location' is introduced to examine, in some detail, the mechanics of geographic uses of point set topology. This essay aligns some basic topological and geographical definitions; future work would attempt to make this correspondence strong enough that theorems from point-set topology might be used to discover new patterns within the associated geographic structure. For, if the prefixes in the pair of words 'topography' and 'geology' are switched, the pair 'geography' and 'topology' emerges--a linguistic duality suggesting that knowledge of topological relations underlying human activity is as critical to understanding groupings of human relationships as is topographic structure to understanding groupings of physical landforms.

GEOGRAPHICAL TOPOLOGICAL SPACES

Topological ideas to be exhibited are among those that are basic to the development of point set topology itself; consistent alignment of fundamental concepts is viewed as a beginning necessary to a systematic approach to characterizing geographic regions with diffuse boundaries.

<u>Definition 1</u> (definition of a topology).[2]

Let X be a non-empty set and let T be any collection of subsets of X. Then T is called a topology for X if

1) $G_\gamma \in T$ for all $\gamma \in \Gamma$, then $\bigcup_{\gamma \in \Gamma} G_\gamma \in T$;

2) $G_i \in T$ for all $i \in I$, I finite, then $\bigcap_{i \in I} G_i \in T$;

3) $X \in T$ and $\emptyset \in T$.

<u>Definition 2</u> (definition of a topological space).[3]

If T is a topology for the set X, then the couple (X,T) is referred to as a topological space.

<u>Definition 3</u> (definition of open sets).[4]

Suppose (X,T) is a topological space. The sets G are called the open sets of the topological space.

The letter M, often used to denote an arbitrary set, is from the German "Menge;" "Menge" means set within this context, although more common meanings of "Menge" such as "crowd" suggest mass or content of the set.[5] This in turn suggests that while the nature fo the content of mathematical sets is unimportant to formal operation, a geographical interpretation might focus, additionally, on the content of sets. Thus, to form geographical sets in a space X one must first determine a differentiating characteristic (or set of differentiating characteristics) by which to distinguish content and on which to base set formation. Identification of such a set of differentiating characteristics will be such that to decompose an individual open set into a union of sets would be to destroy the quality of the differentiating characteristic.

Once a collection of sets can be formed within X, on the basis of geographical differentiating characteristics, Definitions 1 and 3 emphasize the mathematical significance of identifying collections of open sets; again, notational origins reinforce this from a geographical point of view. For an open set of (X,T), often denoted by G, is from the German "Gebiet," which might reasonably be translated "region" rather than "open set."[6] Thus, identification of a collection of open sets might be thought of as selection of a set of regions fundamental to the geographical processes being considered.

One collection of geographical open sets that occurs naturally is composed of an arbitrary collection of uniform regions, and such selection will permit generation of geographical topological spaces. For suppose that in a geographic set X the differentiating characteristic of spatial homogeneity, relative to scale and to natural form, is used to separate homogeneous sets of individuals from one another in forming homogeneous, or uniform, regions.

With open sets and uniform regions in correspondence, a geographical topological space will be formed once it is verified that Definition 1 applies. Let X be the union of all uniform regions and of all other regions contained in that union. Let T be the collection of all uniform regions, that are uniform relative to scale and natural form. Then T is a topology for X if

1) the union of an arbitrary number of elements of T is once again a member of T.

Union of uniform regions represents expansion or aggregation of these regions; nothing new is added, so the aggregate is once again uniform.

2) the intersection of a finite number of elements of T is once again a member of T.

This is clearly true; whether or not "finite" may be replaced by "infinite" is not clear.

3) $X \in T$ by definition of X; $\emptyset \in T$ where $\emptyset$, the empty set, represents the potential uniform region.

To address problems that focus on clustering, the following additional basic topological definition will be useful.

Definition 4 (definition of neighborhood)

A neighborhood N in (X,T) of a point p is a subset of X that contains p and an element G or T, for which $p \in G$. It is important to note that a neighborhood of p is not necessarily an open set of T; however, any open set of T is a neighborhood of each of its elements.[7]

Within the geographical topological space (X,T), outlined above, let neighborhood correspond to functional region. Neighborhoods N of (X,T) are functional regions based on a node p that is contained within a uniform region (open set of T) throughout which distribution of the phenomenon is uniform; this in turn is surrounded by a hinterland in which activity declines.

Internal structure of a geographical topological space

The definitional association of open set to uniform region and of neighborhood to functional region will permit development of a taxonomy as a framework in which to view interaction of these objects. It appears that such interaction unifies (X,T) as a whole composed of an aggregate of basic structural geographical components. This taxonomy will deal with internal structure of a geographical topological space; external structure, via transformations, will be considered later (Appendix A).

To develop classification procedure for characterizing neighborhood interaction, or interaction among geographical units, we proceed as follows. Suppose that N_1 is a neighborhood of p_1 and that $p_1 \in G_1 \subset N_1$, while N_2 is a neighborhood of p_2 and $p_2 \in G_2 \subset N_2$. Only two neighborhoods will be dealt with; this situation is complex mechanically but generalization of it should be fairly clear conceptually.

Definition 5 (Neighborhood domination)

Suppose N_1 and N_2 are two neighborhoods of (X,T), with $p_1 \in G_1 \subset N_1$ and $p_2 \in G_2 \subset N_2$. The neighborhood N_2 will be said to dominate N_1 if $p_1 \in N_2$ and $p_2 \notin N_1$.

The taxonomical framework will be based on forming all possible logical combinations of the statements below, that describe the spatial relations that can occur as two neighborhoods come into contact. Visualization will be made easier if one thinks of N_2 as 'sliding across' N_1, or, as superimposing itself on N_1.

TABLE OF NEIGHBORHOOD INTERACTION

1)	$N_1 \cap N_2 = \emptyset$	5)	$p_1 \in N_2$
∿1)	$N_1 \cap N_2 \neq \emptyset$	∿5)	$p_1 \notin N_2$
2)	$G_1 \cap G_2 = \emptyset$	6)	$p_2 \in N_1$
∿2)	$G_1 \cap G_2 \neq \emptyset$	∿6)	$p_2 \notin N_1$
3)	$G_1 \cap N_2 = \emptyset$	7)	$p_1 \in G_2$
∿3)	$G_1 \cap N_2 \neq \emptyset$	∿7)	$p_1 \notin G_2$
4)	$N_1 \cap G_2 = \emptyset$	8)	$p_2 \in G_1$
∿4)	$N_1 \cap G_2 \neq \emptyset$	∿8)	$p_2 \notin G_1$

The combination of these statements, as given below, will omit, for the sake of clarity, cases that are symmetric with another case; in cases of this sort, the case which includes more of N_1 in N_2 will be considered--that is, the case in which N_2 dominates N_1.

The following outline presents the logical possibilities for contact between N_1 and N_2, under the assumption that N_2 dominates N_1 whenever appropriate. (The stages in the outline correspond to the stages in Figure 1).

NEIGHBORHOOD INTERACTION TAXONOMY

I) (1), or $N_1 \cap N_2 = \emptyset$;

II) (∿1) and (2), or $[N_1 \cap N_2 \neq \emptyset] \wedge [G_1 \cap G_2 = \emptyset]$

A) (3) and (4), or, $[G_1 \cap N_2 = \emptyset] \wedge [G_2 \cap N_1 = \emptyset]$

B) (∼3) and (4), or, $[G_1 \cap N_2 \neq \emptyset] \wedge [G_2 \cap N_1 = \emptyset]$

i) $p_1 \notin N_2$

ii) $p_1 \in N_2$;

C) (∼3) and (∼4), or, $[G_1 \cap N_2 \neq \emptyset] \wedge [G_2 \cap N_1 \neq \emptyset]$

i) $p_1 \notin N_2$

a) $p_2 \notin N_1$

b) $p_2 \in N_1$--not considered; N_2 dominates N_1

ii) $p_1 \in N_2$

a) $p_2 \notin N_1$

b) $p_2 \in N_1$

III) (∼1) and (∼2) or, $[N_1 \cap N_2 \neq \emptyset] \wedge [G_1 \cap G_2 \neq \emptyset]$

(It follows from this assumption that we also have (∼3) and (∼4)).

A) $p_1 \notin N_2$

i) $p_2 \notin N_1$

ii) $p_2 \in N_1$--not considered; N_2 dominates N_1

B) $p_1 \in N_2$

i) $p_1 \notin G_2$

a) $p_2 \notin N_1$

b) $p_2 \in N_1$

1) $p_2 \notin G_1$

2) $p_2 \in G_1$--not considered; N_2 dominates N_1

ii) $p_1 \in G_2$

a) $p_2 \notin N_1$

b) $p_2 \in N_1$

1) $p_2 \notin G_1$

2) $p_2 \in G_1$

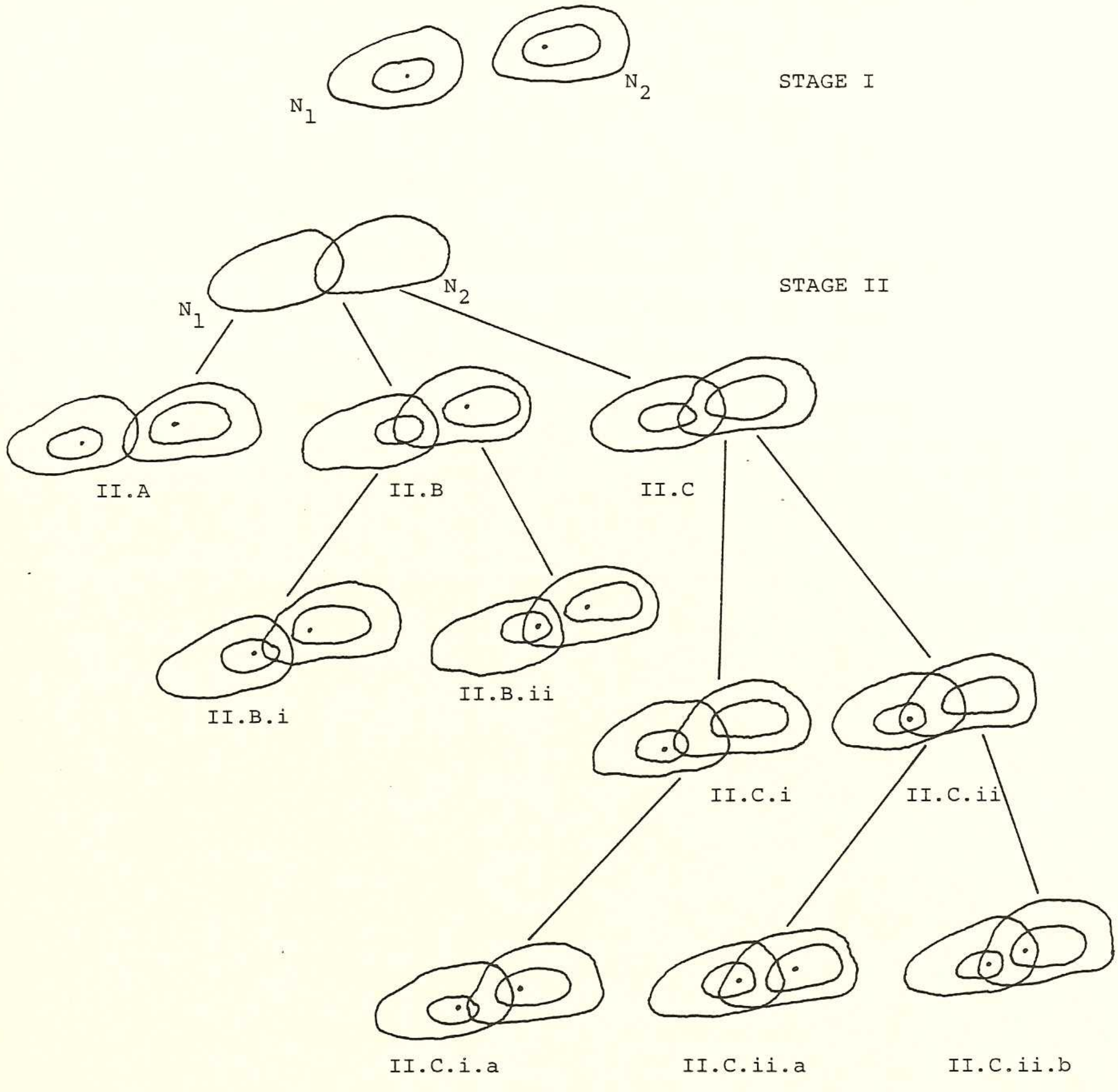

N1
N2
STAGE I
N1
N2
STAGE II
II.A
II.B
II.C
II.B.i
II.B.ii
II.C.i
II.C.ii
II.C.i.a
II.C.ii.a
II.C.ii.b

FIGURE 1

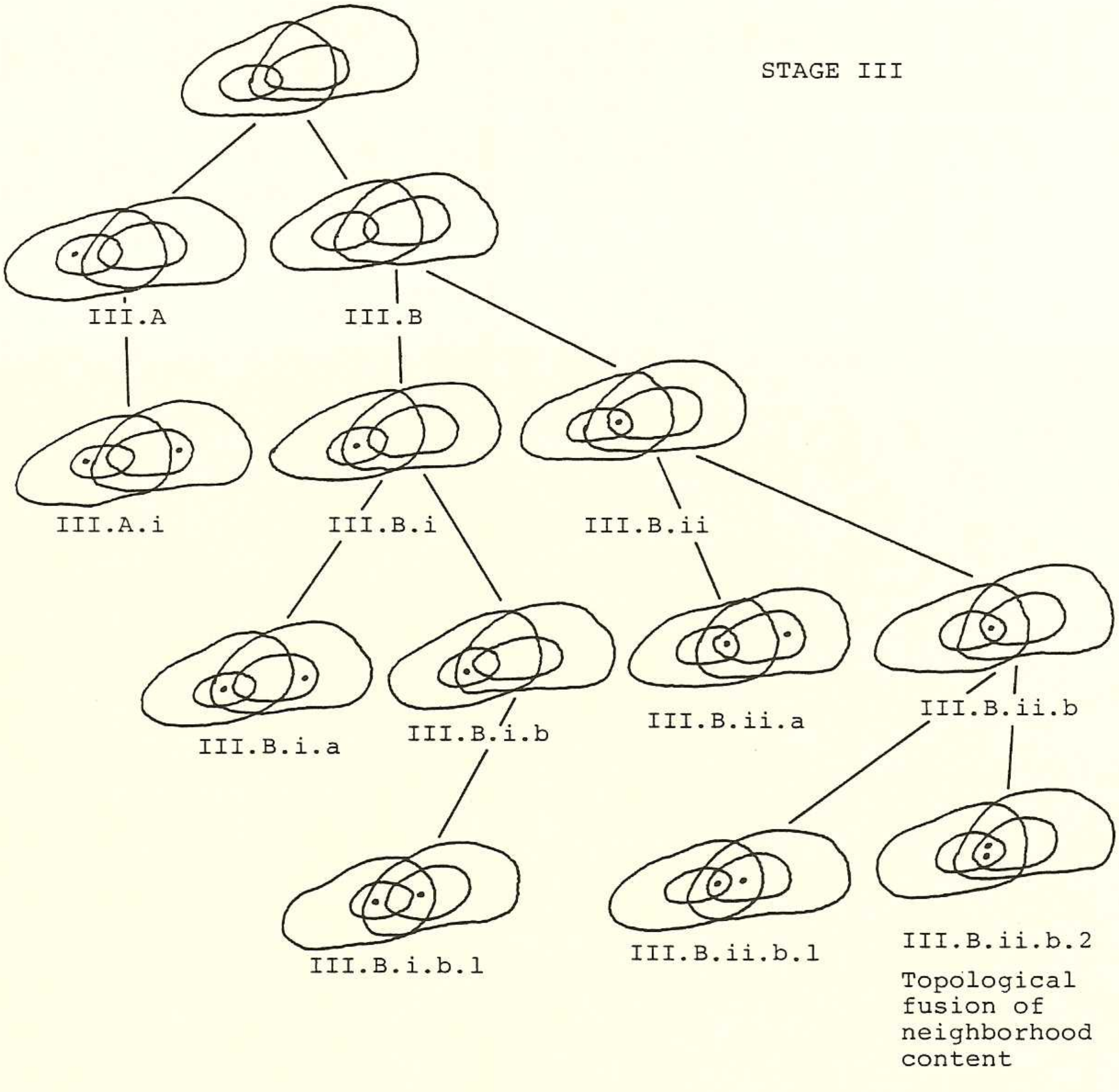


FIGURE 1

(continued)

The hierarchy constructed above is to be used for classifying neighborhood interaction, where the interaction between two neighborhoods is such that hinterlands are joined first (up to stage II.A), then uniform region and hinterland (stage II.B to stage II.C.ii.b), and then both uniform regions (stage III to stage II.B.ii.b.2). The final phase in stage III (III.B.ii.b.2) represents assimilation of the two nodes into a common uniform region resulting in topological fusion of neighborhood content. This taxonomy is viewed as useful for dealing with geographic phenomena that do not rely on a metric, for it is based only on ideas of inclusion.

GENERAL DIRECTIONS FOR EXPLORATION

Geographic avenues to be explored here are of two types: those that appear to require largely geographic work, and those that also require further topological development.

Case 1

The material exhibited below is neither complete nor are the correspondences unique; many of the words represent complex ideas but are given simple interpretation as a beginning. Where there is proof, the structure of the proof is formal but its deals with non-mathematical entities. This use of non-formal proof content within formal proof structure reflects earlier emphasis on neighborhood and open set content within formal topological structure.

Before forming propositions it will be useful to specify, in a general manner, what X and T are. For the rest of this section, we suppose that X consists of the set of all human beings that live, have lived or ever will live, together with all human systems, traditions and institutions (i.e., civilizations). The collection T will consist of 'societies.' Formation of societies will be used as a differentiating characteristic for open set formulation in T. Selection of this particular differentiating characteristic is consistent with the idea that decomposition of an open set should destroy the quality of that characteristic, when viewing a society as "a broad grouping of people having common traditions, institutions and collective activities or interest."[8]

We now have X and T. It remains to show that T is a topology for X, using Definition 1, and to determine what a neighborhood of a point p of (X,T) will be. Then proposition formulation in (X,T) will be approached.

Lemma 1

The collection of all societies, S_γ, $\gamma \varepsilon \Gamma$ forms a topology for the set X of all human systems and traditions.

Proof:

To show this each of the three conditions of Definition 1 must be verified.

1) if $S_\gamma \varepsilon \Gamma$, for all $\gamma \varepsilon \Gamma$, then $\bigcup_{\gamma \varepsilon \Gamma} S_\gamma \varepsilon T$.

For this to be true, we must show that any aggregation of these societies is again a society. An aggregation of "broad groupings of people" is again a "broad grouping" so that part of the definition is satisfied. The quality of "having common traditions , institutions, and collective activities or interest" remains in the the aggregate, although the degree to which such commonness is present may change. In this way $\bigcup_{\gamma \varepsilon \Gamma} S_\gamma$ represents a society and so is an element of T.

2) if $S_i \varepsilon T$ for all $i \varepsilon I$, I finite, then $\bigcap_{i \varepsilon I} S_i \varepsilon T$.

For this to be true, we must show that the intersection of any finite number of societies is itself a society. In contrast to the case for aggregation, the quality of "broad grouping" remains but is changed in degree while the definition of intersection assures us intensification of "commonness." In this way $\bigcap_{i \varepsilon I} S_i$ represents a society and so is an element of T.

3) $X \varepsilon T$ and $\emptyset \varepsilon T$.

We see by the definition of X that X conforms to our definition of "society" and so is a member of T.

The empty set $\emptyset$ is in T if it is viewed as the set of societies which have never occurred (but which could conceivably). Thus two entirely separate societies, S_a and S_b are disjoint, $S_a \cap S_b = \emptyset$, if there has never been any common tradition within the two; however, that potential is still there, so $\emptyset \varepsilon T$.

Q.E.D., Lemma 1.

Lemma 2

A neighborhood N, of a point $p \varepsilon X$, where p represents some human formation, will consist of a culture generated from p where p is adopted relatively uniformly by the society G which formed p.

Proof:

A neighborhood N of a point p within the topological space (X,T) determined in Lemma 1 is a subset of X that contains p and an element $S \varepsilon T$ such that $p \varepsilon S$. Thus if p represents some characteristic human formation, then a neighborhood N of p within (X,T) is that subset of characteristic features that developed in X, based on p. Thus N represents culture based on p, where for "culture," we use "the characteristic features of a particular stage in the advancement of civilization."[9] Such an N contains an element of T that contains p; namely, the society S which generated the formation p and derivative concepts. In this way, $p \varepsilon G$ and $G \subset N$ so that N is a neighborhood of p in (X,T).

Q.E.D., Lemma 2.

Lemma 3

Suppose N_1 and N_2 are two neighborhoods in (X,T). The culture N_1 of society S_2 dominates the culture N_2 of society S_1 if $p_1 \varepsilon N_2$; ie., if interaction between N_1 and N_2 has reached, at least, to stage II.B.ii or to stage II.C.ii or to stage II.B.i of the taxonomy of Figure 1.

Proof:

"Domination" refers to the "supremacy or ascendancy over another or others."[10] Such supremacy of N_2 over N_1 sill necessarily occur if N_2 controls or engulfs that characteristic, p, which is the generating element of N; i.e., if $P_1 \varepsilon N_2$. (In this case one could also say that N_1 is subordinate to N_2). That is, within the taxonomy developed above, domination of N_2 over N_1 occurs in all cases where $p_1 \varepsilon N_2$.

Q.E.D., Lemma 3.

Proposition 1

Suppose N_1 and N_2 are two neighborhoods of (X,T) with $p_1 \varepsilon S_1 \subset N_1$ and $p_2 \varepsilon S_2 \subset N_2$. Then S_2 is said to acculturate S_1 if $S_1 \cap S_2 \neq \emptyset$ and if $p_1 \varepsilon S_2$, i.e., if interaction between N_1 and N_2 has reached stage III.B.ii of the taxonomy of Figure 1.

Proof:

Acculturation refers to "the process of interaction between two societies by which the culture of the society in the subordinate position is drastically modified to conform to the culture of the dominant society."[11] (Migration is not necessarily implied). That is, the societies within N_1 and N_2 must interact, or, within $N_1 \cap N_2$ we must have $S_1 \cap S_2 \neq \emptyset$ and the dominant culture must engulf the heart of the subordinate culture or, in this case, S_2 must contain p_1. Within the taxonomy, acculturation of N_1 by N_2 occurs in all cases where $S_1 \cap S_2 \neq \emptyset$ and $p_1 \varepsilon S_2$.

Q.E.D., Proposition 1.

Clearly Proposition 1 is a refinement of Lemma 3. One could think of these in terms of the introduction of Spanish influence into Latin America. One could begin at stage I of the taxonomy and imagine Indian culture (as N_1), as separate from the Spanish (N_2). Then one could trace gradual interaction leading up to domination (but not acculturation) of N_1 by the Spaniards before 1519. In the Conquest of 1519, the Spaniards capture control of the heart of N_1 and begin a process of acculturation, culminating in the final stage of topological fusion of neighborhood content, represented in the physical landscape by the superimposition of Roman Catholic churches, constructed in the Spanish style of architecture, on sites of previously existing Indian temples.

Case 2

A different approach is suggested below. A topology on a set X might describe regulation of permeability of the set X (viewing X as a barrier) that separates a region into two subregions, each contiguous with one face of the barrier. To consider such an approach, it will be useful to introduce a topological hierarchy of separation axioms, and then to explore directions for geographic application by considering how openings in the barrier are separated from one another.

Definition 6

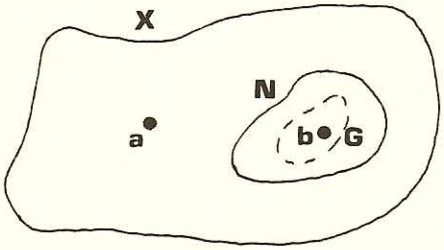
A subset F of X in (X,T) is called closed in T if X-F is open in T ((X-F) denotes all of X not in F).

Definitions 7 to 11 provide a topological hierarchy of separation axioms (The 'T' is from 'Trennungsaxiome'). (Figure 2)

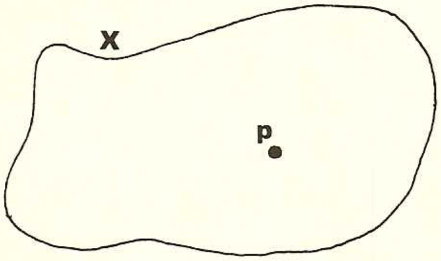
Definition 7 (Kolmogoroff)

A topological space (X,T) is said to be T_0 if and only if, for distinct points there exists a neighborhood containing one point but not the other.

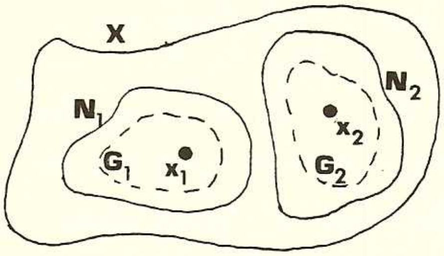
A HIERARCHY OF TOPOLOGICAL SPACES BASED ON SEPARATION



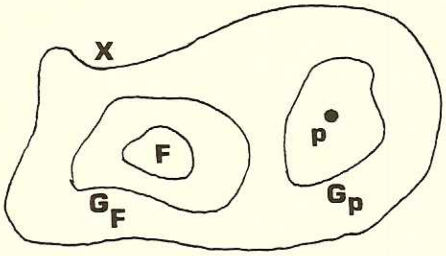
T_0-PROPERTY



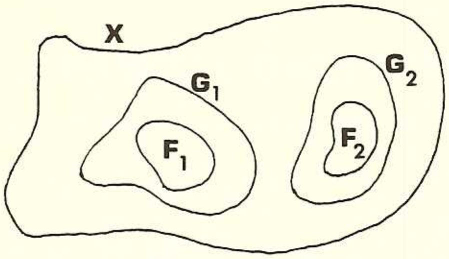
T_1-PROPERTY



T_2-PROPERTY



T_3-PROPERTY



T_4-PROPERTY

FIGURE 2

Definition 8 (Fréchet)

A topological space (X,T) is said to be T_1 if and only if the set $\{p\}$ is closed for every $p \varepsilon X$.

Definition 9 (Hausdorff)

A topological space (X,T) is T_2 if and only if distinct points have disjoint neighborhoods.

Definition 10 (Vietoris)

A topological space (X,T) is T_3 if and only if (a) for each point p and each closed set F there exist disjoint open sets G_p and G_F such that $p \varepsilon G_p$ and $F \subset G_F$; (b) it is T_1 (so points can be separated).

Definition 11 (Tietze)

A topological space (X,T) is T_4 if and only if (a) for each disjoint pair of closed sets F_1 and F_2, there exist disjoint open sets G_1 and G_2 such that $F_1 \subset G_1$ and $F_2 \subset G_2$, and (b) it is T_1.[12]

If (X,T) represents a barrier which is permeable, let the elements of T correspond to holes in the barrier through which flow may pass. If (X,T) is a T_0-space then for any pair of points that attempts to pass the barrier there is an opening that accommodates one member of the pair and excludes the other. Any pair of points partitions the space in such a way that neither element of the partition is an extreme.

If (X,T) is a T_1-space $\{p\}$ is closed and $X - \{p\}$ is open. This provides a partition of X in which both elements are extremes. The set $X-\{p\}$ is an element of the class of 'largest' open subsets of X. Thus all points but p may pass through the barrier via the opening $X-\{p\}$. The substance of the barrier is a mass of points.

If (X,T) is a T_2-space, pairs of points pass through the barrier through any openings G_1, G_2 contained in N_1, N_2. If (X,T) is a T_3-space, then pairs of points and closed sets may pass through the barrier. If (X,T) is a T_4-space, then pairs of closed sets may pass through the barrier (see Figure 2).

Suppose that the topology T on X is based on the differentiating characteristic of property inheritance. If the style of inheritance is that of equal division of property among male heirs, then (X,T) may be viewed as a T_0-space. For as man and wife approach this barrier only the man may penetrate the barrier. In this case, the wife is reflected back to the original space by the barrier. As a man and his son approach this barrier, the son will penetrate it while the father will be absorbed into the barrier as he dies (permitting the son to pass through). The only men who will be reflected by the barrier are those who are no longer in a position to inherit anything because anyone from whom they could is already dead. This topology on X is a regulator of one-way flow; it permits flow only as time progresses--inheritance is not usually a reversible property.[13]

If the style of inheritance is different, then the manner of regulation of flow through the barrier is different. The topology T will be viewed as a T_3 regulator on flow in the following case. Suppose that a man is free to will his property to any of his family or to anyone or anything else. In this case, a man and his family approach the barrier; the man is absorbed into the barrier at point p (dies) while his family penetrates the barrier through the opening created by the open set G_p; any non-members of his family that are to inherit property (institutions, servants, pets) pass through X via G_F, reflecting their dependence on p as well as their separation ($G_p \cap G_F = \emptyset$) from his family.

To view corporate (rather than individual) death, (X,T) could be considered a T_4-space. Here F_1 would play the part of p in the T_3-case.

Analysis of property inheritance in terms of separation of styles of inheritance, resulting in partitioning of land and consequent change in overall pattern of land use, might be of use in studying changing land use over time.

A different realization of separation, and one in which the topology on a barrier X serves as a two-way regulator on flow across the barrier, comes from using as a differentiating characteristic for topology formulation ideas of black-white racial prejudice and discrimination. Here such a topology has two regulatory functions; white acceptance of blacks, and black acceptance of whites.

The view of a white individual walking along the street might be T_0 (Definition 7); as a pair of blacks approach, he is willing to admit one to his territory, but not both. He doesn't wish to be outnumbered.[14]

More refined separation would be available at higher T-levels. Suppose the opposing spaces on either side of X are composed of young black single men on one side, and young white single women on the other side. A bigoted view of white parents' acceptance of black potential sons-in-law, as one aspect of the differentiating characteristic, might exercise T_1 regulatory control (Definition 8) for no individual single black male, p, could pass across the barrier, but groups of single black males, X-{p} of any size could (assuming one woman marries only one man at a time).

A T_2 regulator (Definition 9) of flow across X assures that points are separate and can pass across X through, passible, more than one disjoint pair of openings. A T_2 regulator of black acceptance by whites would represent a situation in which any pair of blacks always passes through the barrier, and, each is accepted individually, as reflected in the separation of the openings through which each passes. Also some groups of blacks, of which each point (individual) is a member, permeate the barrier, while others are rejected (i.e., reflected or absorbed). For example, some groups of black males may cross this acceptance barrier while others, such as large groups of black teen-age boys, may be blocked by the barrier.

A view of black-white relations in terms of separation axioms could be used to map a city according to each regulator on the differentiating characteristic; those areas which showed regulators of the same general T type might be considered to be in greater racial harmony than those areas in which the regulators belonged to different T classes. This might be useful in dealing with current problems in which race relations are crucial, such as in cross-district school bussing.

At a more global scale, analysis of this sort might provide insight into other situations, especially when such T spaces are mapped to other human situations. If the mapping τ is a homeomorphism (Appendix A) and if an early stage of black-white relations is in X, then in the image of X under τ we might speculate that a line of societal development similar to the black-white one could occur. The idea of homeomorphism would permit recognition of invariant structure and extension of internal analysis of structure over time.[15]

KIOSKLAND

Since it is difficult, without extensive exploration, to envision what sorts of problems would arise in topological examination of the internal structure of the spatial development of relatively large segments of human civilizations it will be useful to enlarge the scale of analysis and to consider a local example.

Substantive material for this case study is drawn from direct observation of the University of Michigan campus (1977).

The case study to be presented deals with examining the spatial distribution of M-Kiosks, cylinders of concrete used as message boards, which are scattered around University of Michigan campuses. It will begin by considering the Kiosk distribution from the viewpoint of general observation and mapping, then will show that this descriptive viewpoint and the topological description derived from the neighborhood taxonomy can be made incident. It will conclude by exhibiting the mechanics of (and by suggesting problems associated with) geographic use of point set topology as an attempt to provide a small amount of additional insight into this locational problem.

Analysis of kiosk function

A kiosk is a line-of-sight means of communication within a space in which travel on foot is dominant. It interacts with pedestrians that are sufficiently close to read its messages, has the potential to interact with those who are aware of its location (usually those who can ses it), and generally has no attraction for those who cannot see it. These observations suggest a natural way in which to form a functional region based on a kiosk. The kiosk is the node, the area surrounding the kiosk from which signs can be read is the core of the functional region in which interaction of kiosk and pedestrian is fairly uniform throughout, and the area enclosed by lines-of-sight based on the kiosk forms the hinterland of the functional region. The space containing all kiosk locations, and potential kiosk locations, and their associated functional regions will be referred to as "Kioskland."

Boundaries of Kioskland

Kioskland will be partitioned into the following regions, at least partially matching the partitioning of The University of Michigan into a central Ann Arbor Campus, a satellite North Ann Arbor Campus and separate campuses in Dearborn and Flint, Michigan.

1) Rural Kioskland: the part that is totally removed from the Ann Arbor heart of the University of Mighigan. This would include kiosks and their associated functional regions on the Dearborn and Flint campuses.

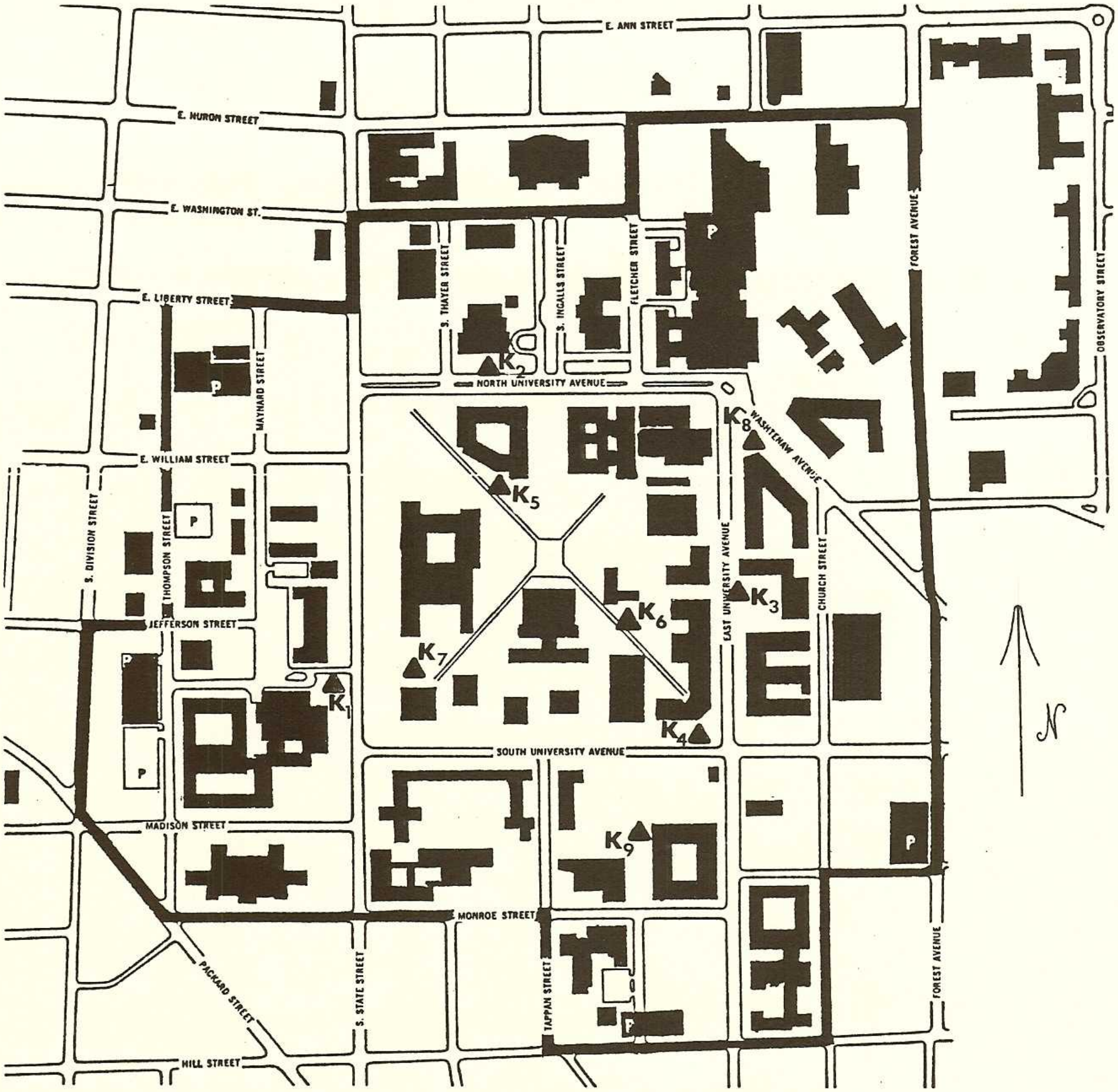
2) Suburban Kioskland; this will consist of those kiosks and their functional regions that are neither rural nor urban (indicated below).

3) Urban Kioskland: the boundaries of this will be determined by the location of parking structures around the Central Campus. For within the hull formed by these structures, pedestrian movement dominates. The edges of this hull will be formed along streets since pedestrian flows move along the street pattern in those parts where there are streets. So to form the urban boundaries, we use Manhattan distance to connect the parking structures (Map 1) and, if two distinct paths have the same length, we choose the one that includes the maximal number of University buildings without including blocks that contain no University buildings.

4) Central Kioskland; the boundaries of this area, in which there is no provision for automobile movement, consist of the boundaries of the "Diag" (or central quadrangle): State Street, North University Avenue, East University Avenue, and South University Avenue (Map 1).

This case study will consider only those kiosks visible from Central Kioskland and their associated functional regions constrained within the edges of Urban Kioskland.

A DISTRIBUTION OF M-KIOSKS



△ K_i, i=1,...,9--kiosks

▬ boundary of urban kioskland

Source: base map, University of Michigan Map, copyright 1967.

0 200 400
feet

Map 1

Location of kiosks within the study area

From direct observation of the study area, the following kiosks were located on a University of Michigan Campus map (Map 1). The kiosks were plotted on the map according to their locations relative to surrounding buildings. Nine kiosks were visible from the Diag. They will be numbered from K_1 to K_9 as shown in Map 1.

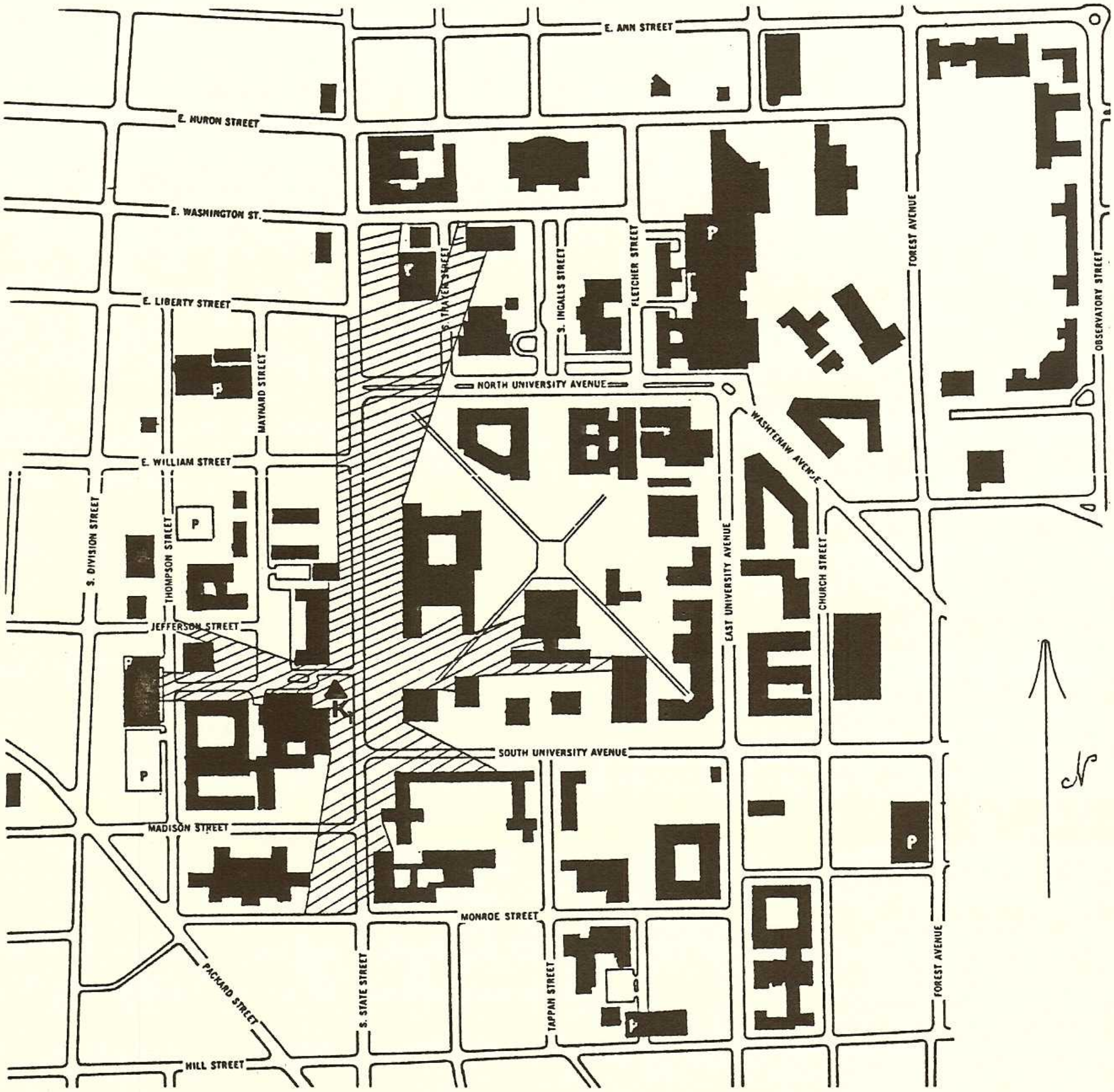
Mapping of associated functional regions

Direct observation of the study area was useful for determining kiosk, or node, location. It was also useful for assessing extent of the core area surrounding the kiosk. The size of sign and lettering on the signs showed a great deal of variation but within forty feet of all kiosks, I could read some part of some sign. So one estimate of core size might be a circle forty feet in radius, centered on the kiosk. Variation in sign size, individual eyesight and a variety of other factors, would warp this circular core boundary.

Direct observation was notespecially useful in determining the line-of-sight edges of the hinterlands of these regions. Changes in vegetation would greatly alter the boundaries of some of these areas. It seemed better to plot line-of-sight regions around each kiosk based on information in Map 1 and use actual vegetation (building) patterns to reduce the size of these regions when appropriate. Maps 2-10 show these line-of-sight neighborhoods for each of K_1 to K_9.

Map 11 shows the composite, derived from overlaying Maps 2-10, with the base map removed. The rough outlines of many of the buildings on the Diag show up as areas not within line-of-sight of any kiosk, as if they were intersections of Kiosk pneumbrae.

LINE OF SIGHT NEIGHBORHOOD OF K_1



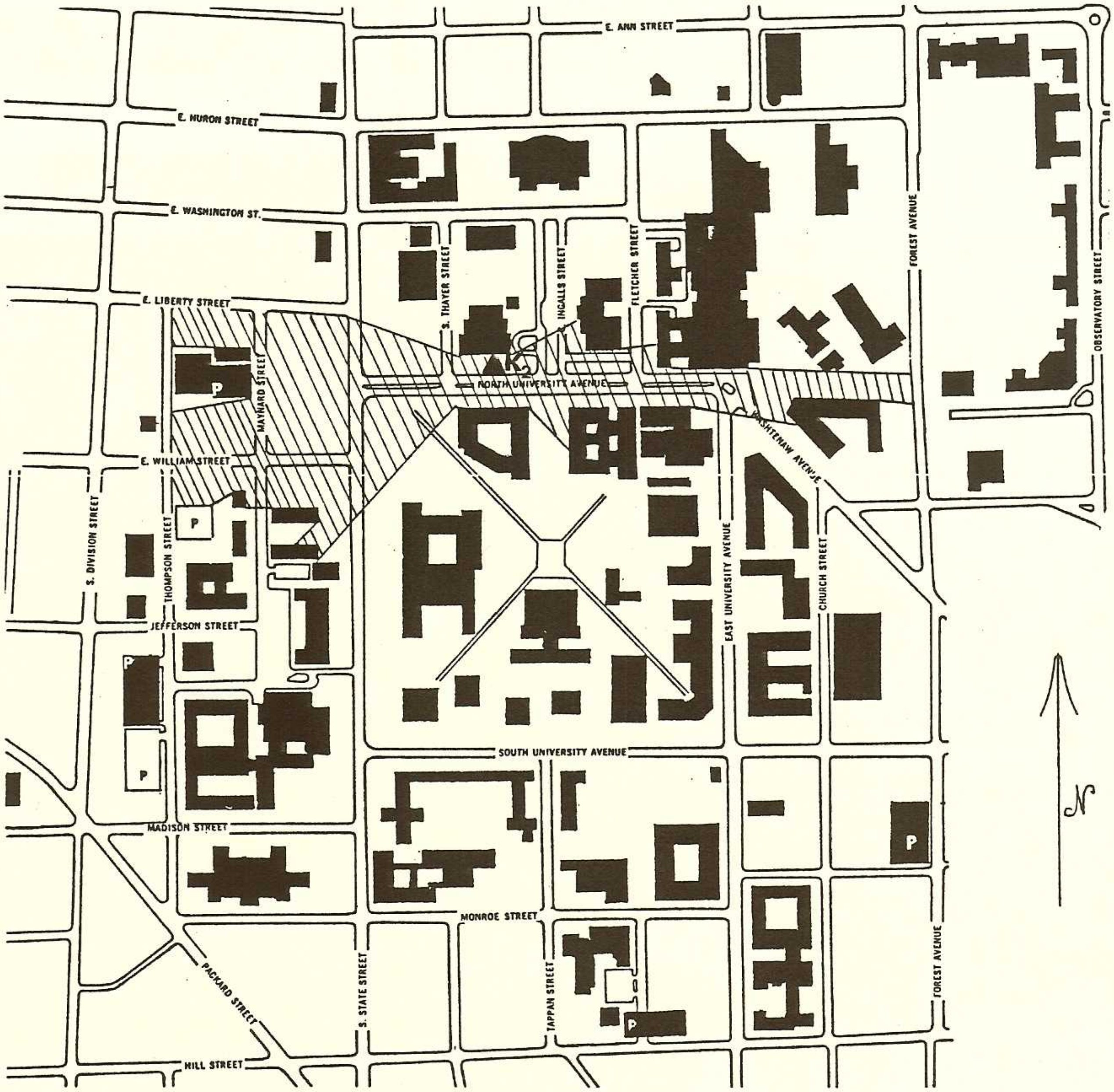
Source: Same

Neighborhood of K_1

0 200 400
feet

Map 2

LINE-OF-SIGHT NEIGHBORHOOD OF K_2



Source: same

Neighborhood of K_2

0 200 400
feet

Map 3

LINE-OF-SIGHT NEIGHBORHOOD OF K_3

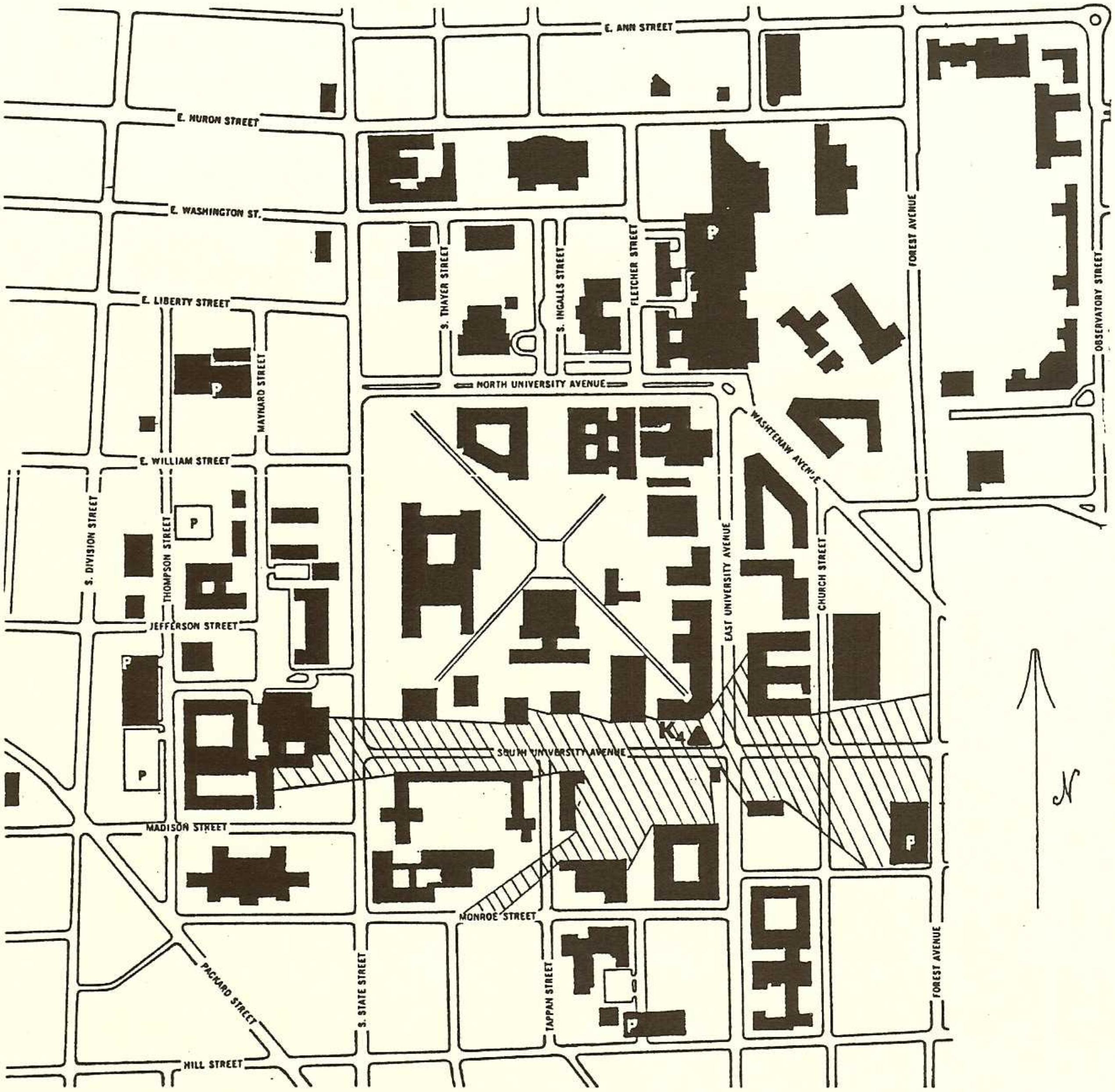
Source: same

Neighborhood of K_3

0 200 400
feet

Map 4

LINE-OF-SIGHT NEIGHBORHOOD K_4



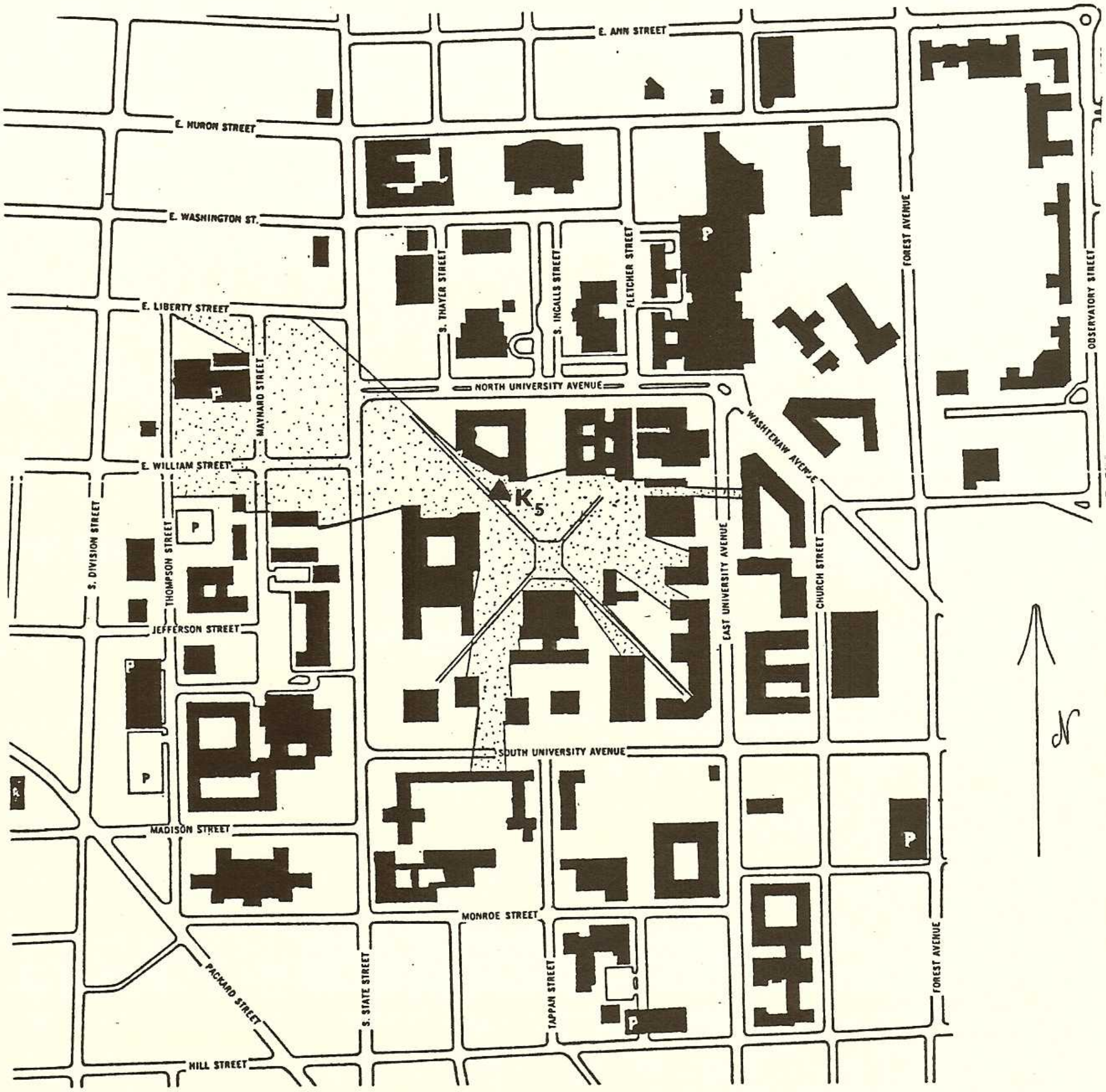
Source: same

Neighborhood of K_4

0 200 400

Map 5

LINE-OF-SIGHT NEIGHBORHOOD OF K_5



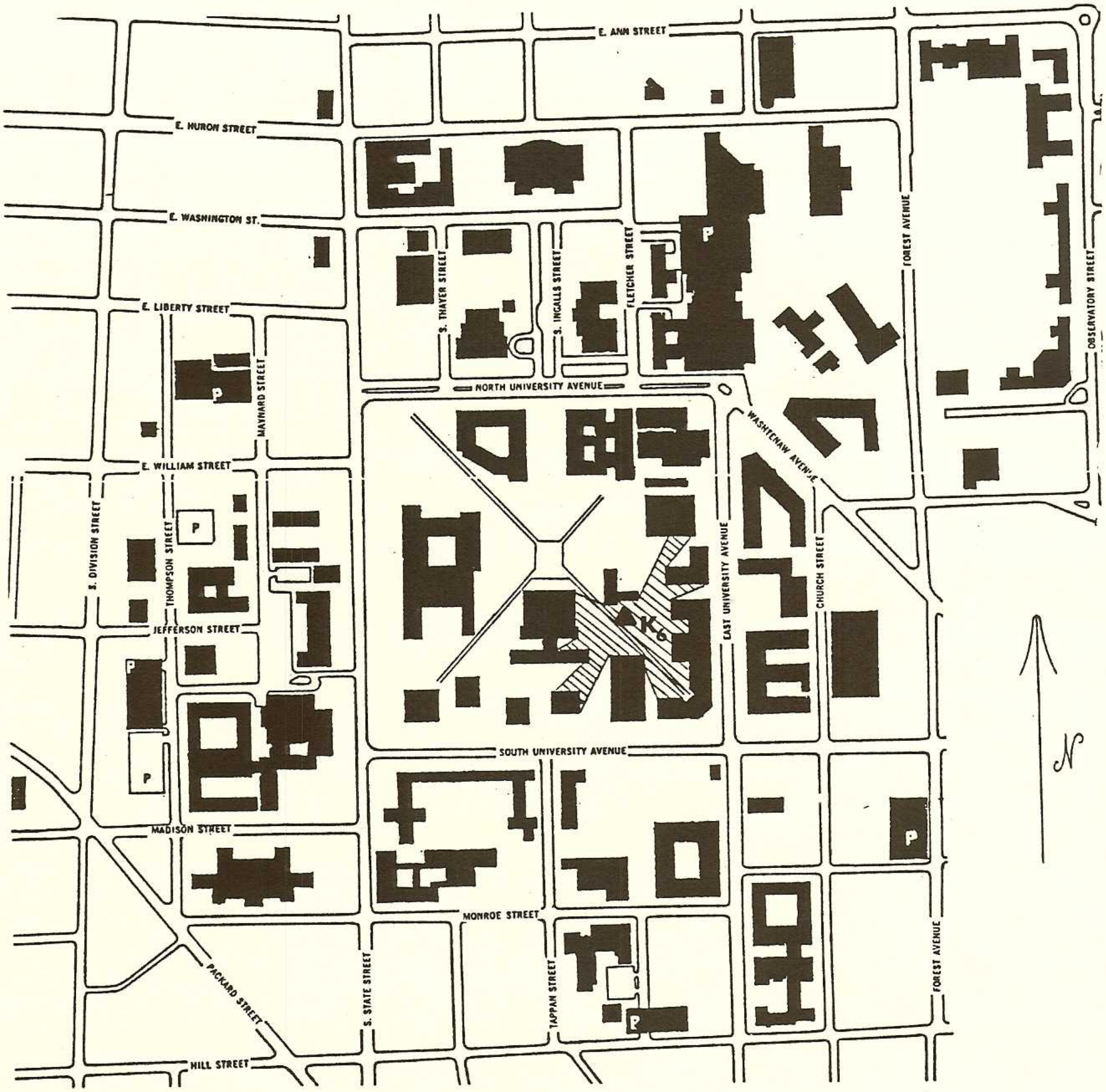
Neighborhood of K_5

Source: same

0 200 400
feet

Map 6

LINE-OF-SIGHT NEIGHBORHOOD OF K_6



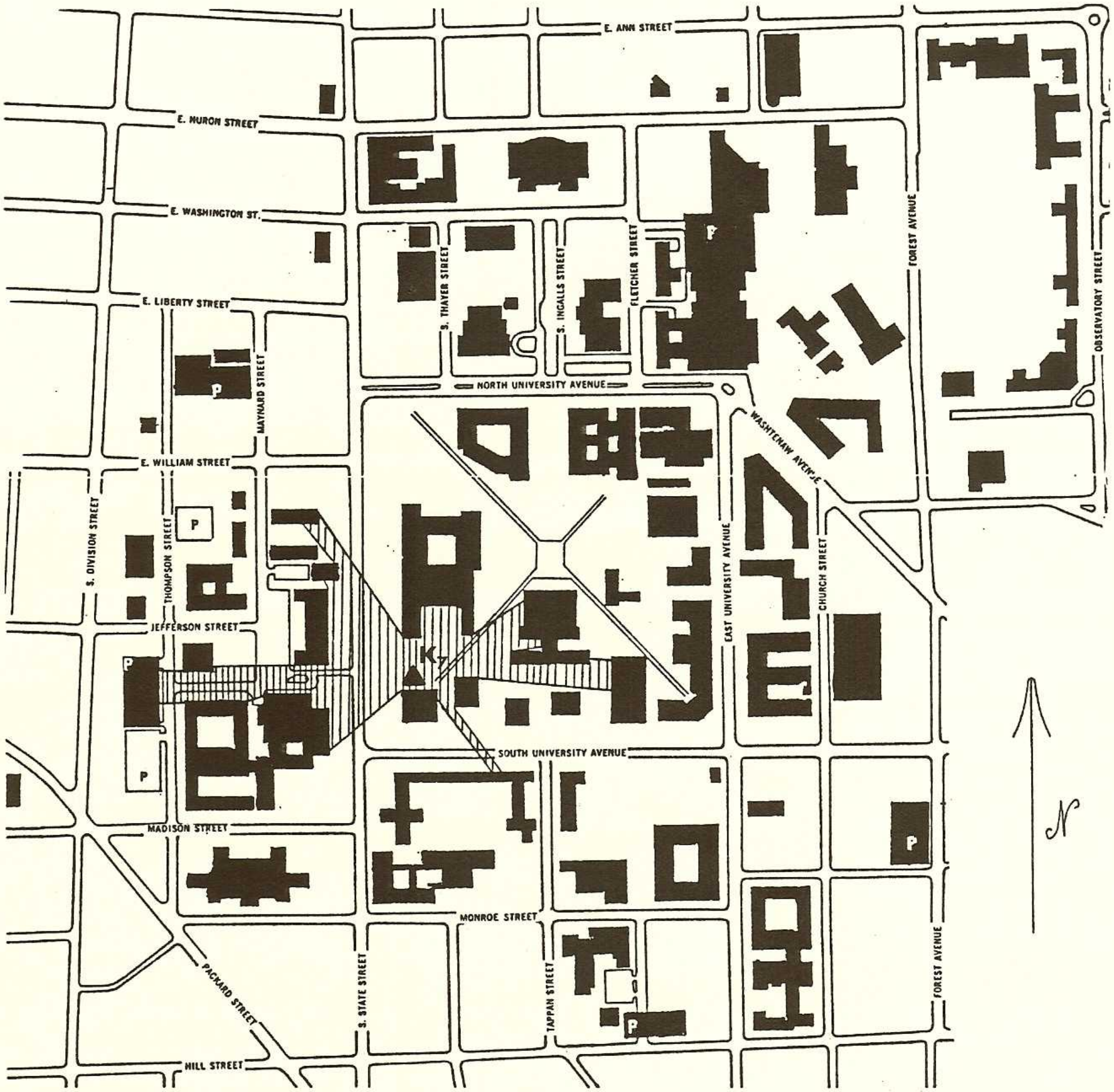
Source: same

Neighborhood of K_6

0 200 400
feet

Map 7

LINE-OF-SIGHT NEIGHBORHOOD OF K_7



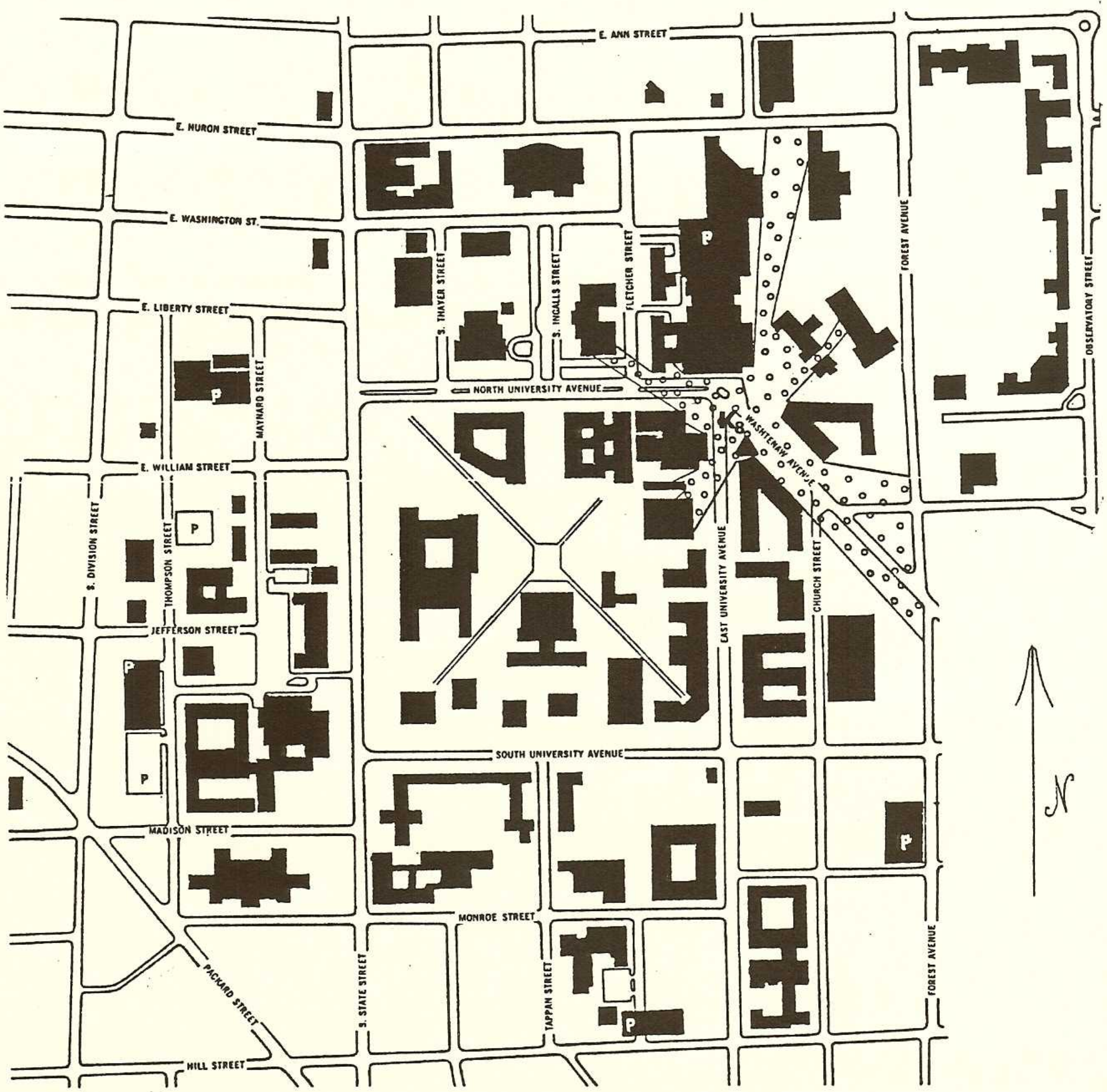
Source: Same

Neighborhood of K_7

0 200 400
feet

Map 8

LINE-OF-SIGHT NEIGHBORHOOD OF K_8



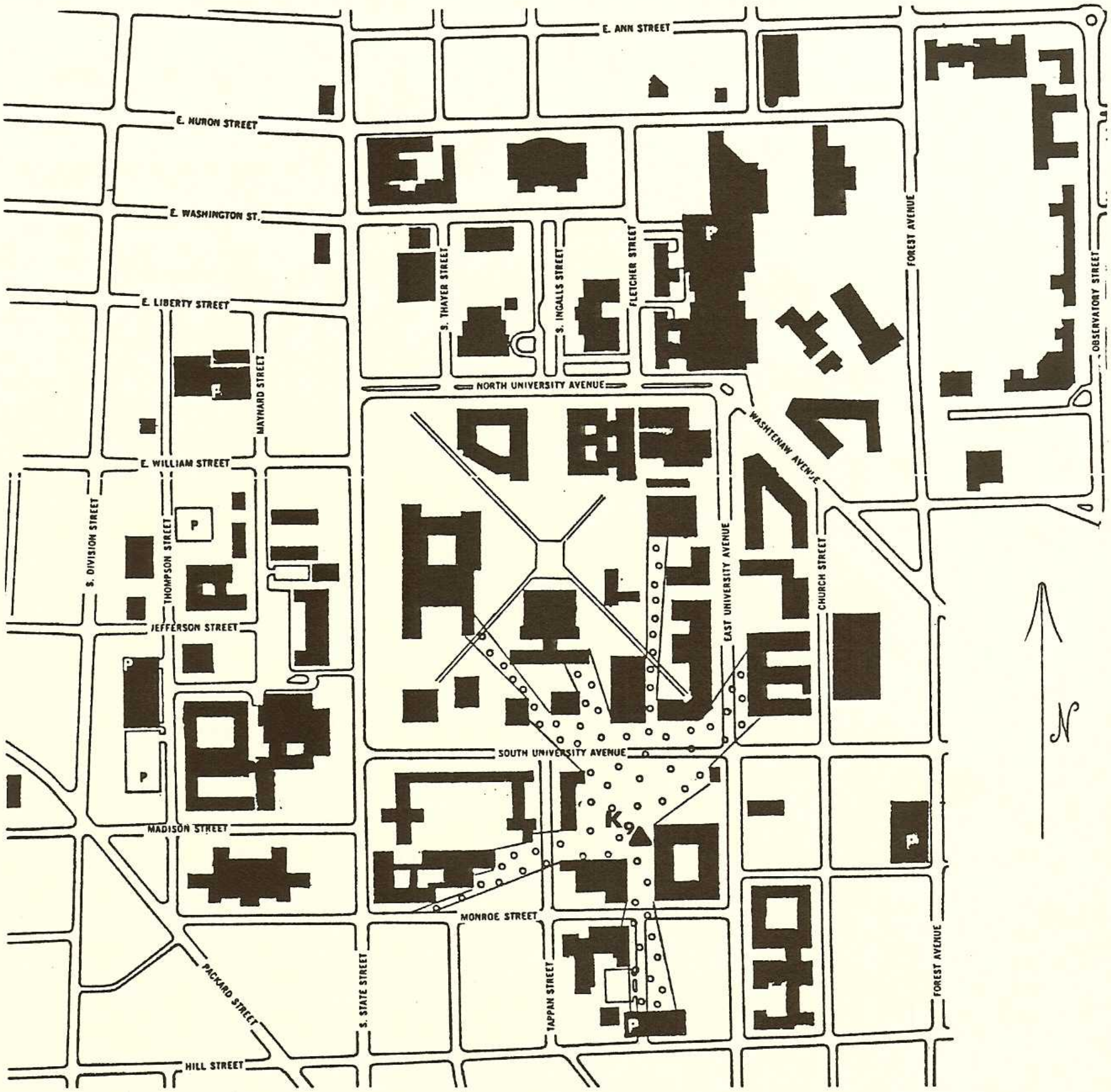
Source: same

Neighborhood of K_8

0 200 400
feet

Map 9

LINE-OF-SIGHT NEIGHBORHOOD OF K_9



Neighborhood of K_9

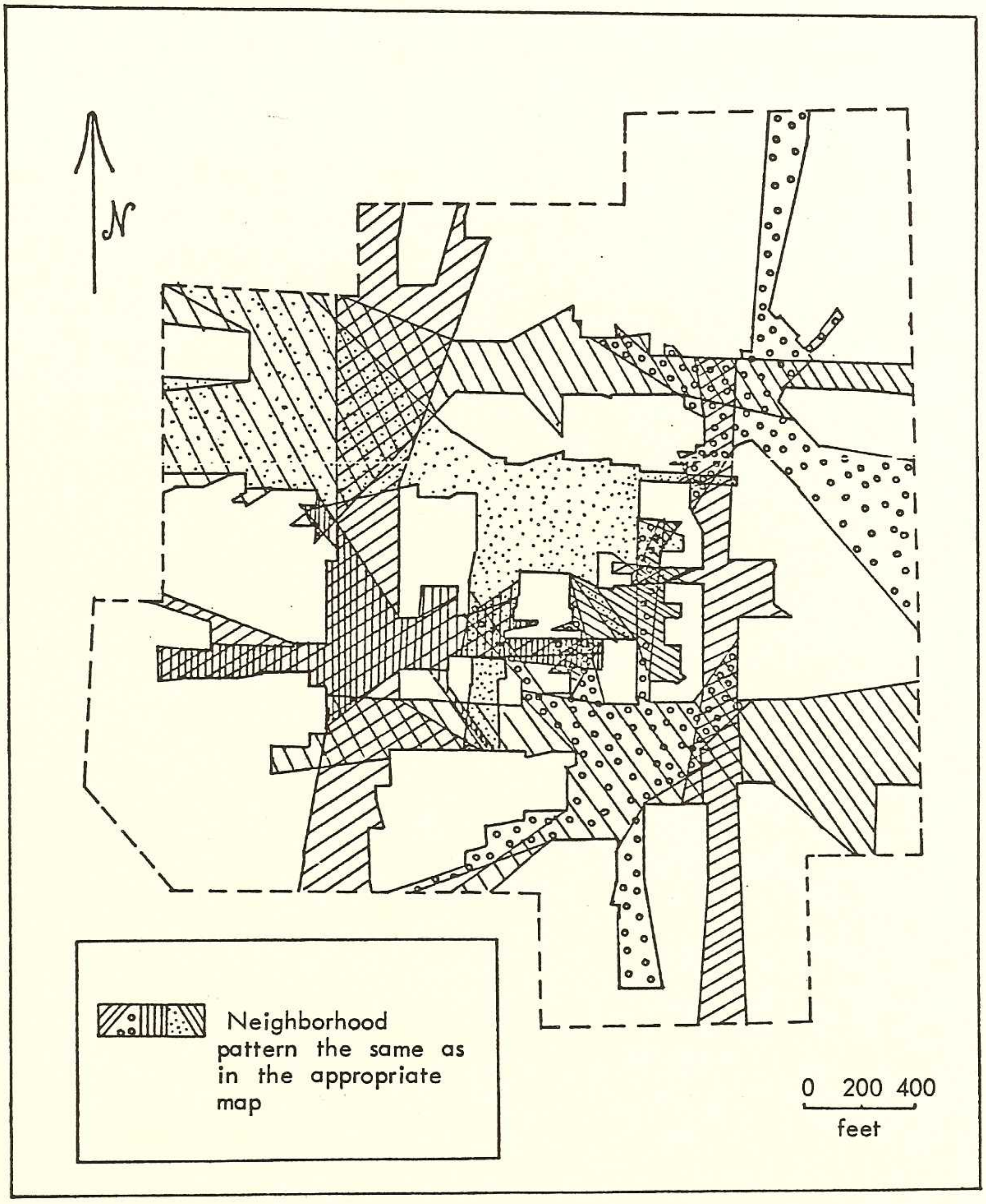
Source: same

0 200 400

feet

Map 10

COMPOSITE OF MAPS



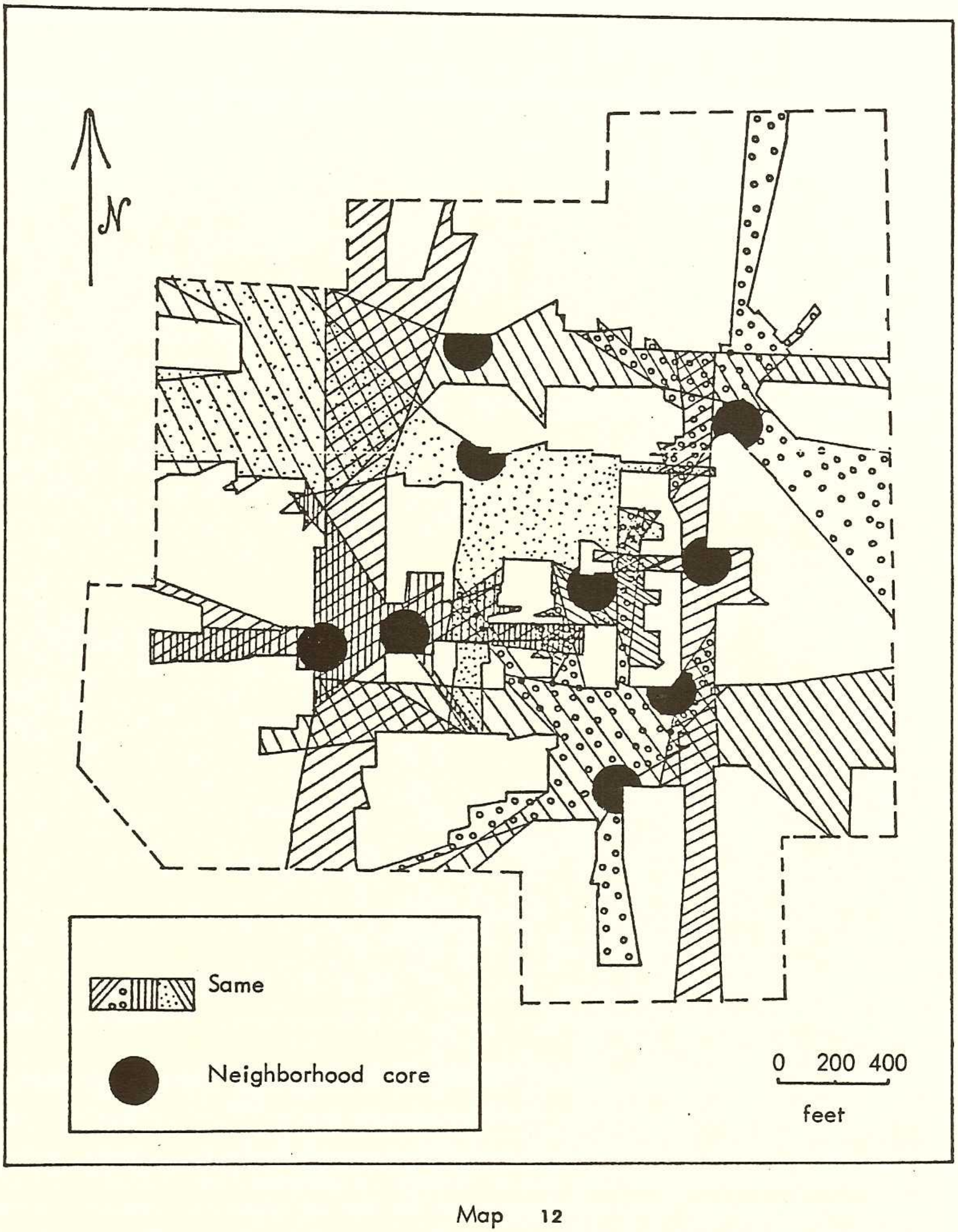
Map 11

To align observed geographical structure with topological structure, we proceed as follows. The set X will consist of message content that has been, or that is to be, transmitted by kiosk (at a current or future location) to any pedestrian that has or will pass through Central Kioskland. To form open sets of T, we use potential or actual direct message transmission as a differentiating characteristic. An open set is a uniform region around an actual or potential Kiosk location from which messages may be read. The collection of open sets T is a topology for X for the union of open sets is again open (within the union the message content of some kiosk can be read) and the intersection of open sets is again an open set (within the intersection, the message content of all kiosks can be read). $X \in T$ by definition of X, and $\emptyset \in T$ when $\emptyset$ represents potential direct interaction.

A kiosk neighborhood around kiosk K within (X,T) can then be defined as a set in which messages may be relayed from the kiosk, but not necessarily through direct interaction of the pedestrian and the kiosk. That is, the margins of the neighborhood coincide with lines-of-sight from the kiosk, or with the hinterland boundaries of the kiosk functional regions. Also, each such set N of K must contain an element $G \in T$ such that $K \in G \subset N$ for the core of the functional region in which the kiosk can be read directly is precisely such a G, and any N defined by lines-of-sight from K must contain such a G. So functional region and topological neighborhood coincide. The functional regions of Map 11 are kiosk neighborhoods. Map 12 shows these kiosk neighborhoods with an open set (the core of the functional region) exhibited as a dark circle.

Choropleth maps and topological mosaics

Kiosk neighborhood mapping, shown in Map 11, suggests that to investigate, rather than to just describe, the configuration determined by these neighborhoods, it would be useful to consolidate the information presented in this map to



Map 12

reduce clutter. One obvious way to do this is to form a choropleth map from this information (Map 13), where areas from which x kiosks can be observed are all covered with the same pattern.

The choropleth map (Map 13) is one way to simplify the content of Map 11. Another way is to form a topological mosaic, as described in Proposition B.1 (Appendix B). Only those sets of interacting neighborhoods in which the hinterland of one overlaps the core of the other are included in this mosaic (Map 14). For example, neighborhood 1 (N_1) based on K_1 intersects core area G_7 based on K_7; $N_1 \cap G_7 = G_7$. So N_1 and N_7 are part of the topological mosaic. Neighborhood and core interaction is as follows.

$$N_1 \cap G_7 = G_7$$

$$N_7 \cap G_1 = G_1$$

$$N_3 \cap G_4 \subset G_4$$

$$N_3 \cap G_9 = \emptyset$$

$$N_4 \cap G_3 = \emptyset$$

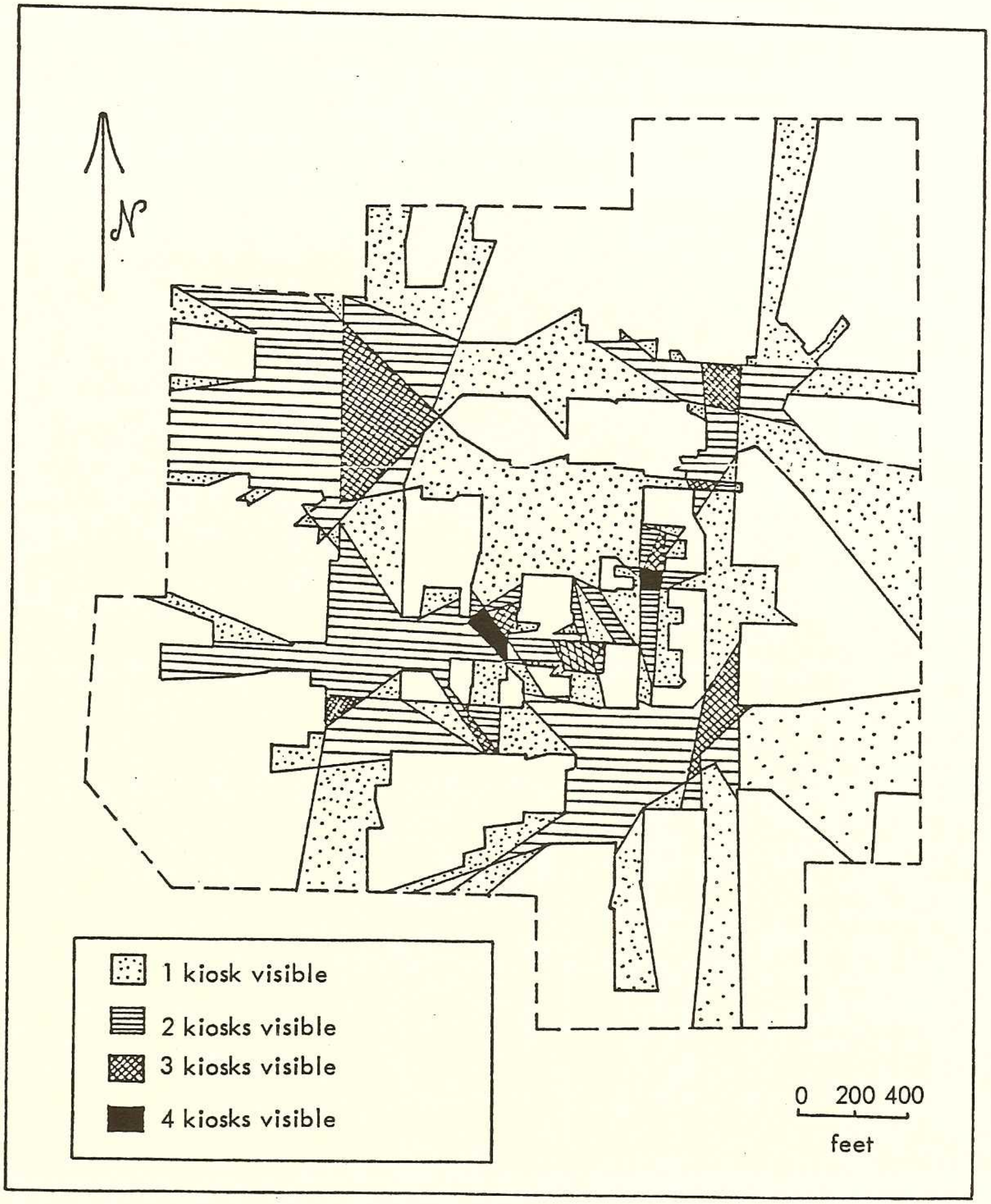
$$N_4 \cap G_9 \subset G_9$$

$$N_9 \cap G_3 = \emptyset$$

$$N_9 \cap G_4 = G_4$$

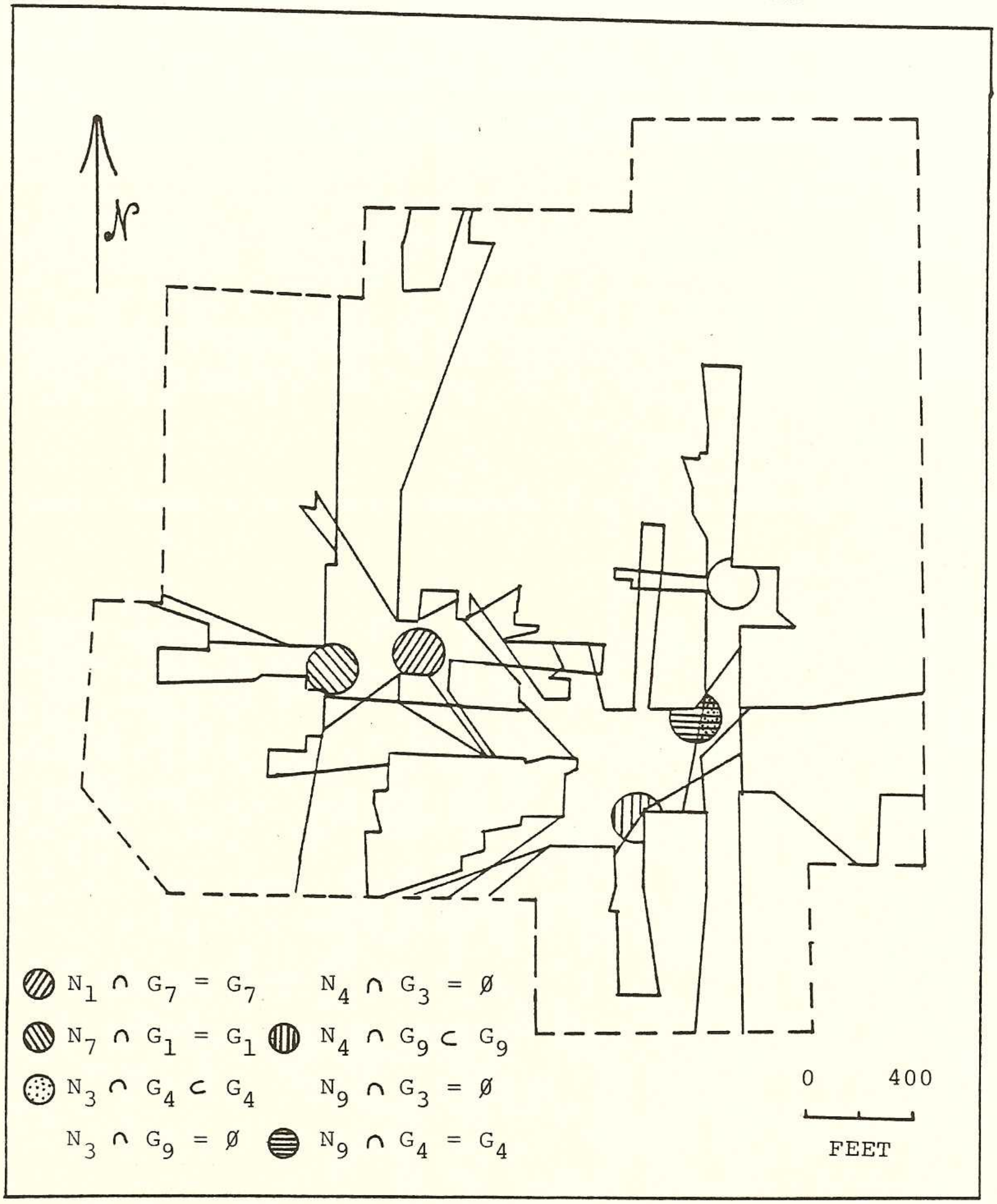
The neighborhoods N_1, N_3, N_4, N_7, N_9 form a topological mosaic in Kioskland. This is in contrast to the choropleth mosaic in which all neighborhood intersections are included. Within the terminology of the taxonomy, we have that a choropleth mosaic M_c in a topological space (X,T) can be expressed as an appropriate union of intersecting neighborhoods, where the level of neighborhood interaction is contained between stages II.A and II.C.ii.b of the neighborhood interaction taxonomy.

CHOROPLETH MAP OF CENTRAL KIOSKLAND



Map 13

TOPOLOGICAL MOSAIC IN CENTRAL KIOSKLAND



Map 14

By not including II.A in the formation of a topological mosaic we have chosen to emphasize more than just superficial (between hinterlands only) contact between regions. The topological mosaic selects those kiosks and neighborhoods among which there is the most interaction according to the original constraints. These regions are not singled out for particular attention using the choropleth map, although from direct observation one might notice that K_1 and K_7 are more closely related than are other subsets of kiosks (and the same for K_3, K_4, and K_9). The use of the topological mosaic lends precision to this intuitive feeling from direct observation that is not evident on the choropleth map.

Checks on the content of these kiosks suggest that if signs were made large enough, the core of K_1 could be extended so that G_1 and G_7 would intersect. In this case K_1 would dominate K_7. An experiment using different sizes and colors of sign could be constructed to see what kinds of signs could be used to achieve this. The set G_1 would then exert pressure in the mosaic (Proposition B.2, Appendix B).

Using both the choropleth mosaic of Map 13 and the topological mosaic of Map 14, as investigative tools for examining Central Kioskland, the following recommendations arise. If a minimal number of signs are to be distributed and maximum coverage of X is desired, then six signs are necessary to cover kiosk neighborhoods: one on each kiosk not in the topological mosaic (there are four of these K_2, K_5, K_6, K_8), and one sign for each set of kiosks where the size of the sign is sufficient to exert pressure on other kiosks within the set (a sign on one of K_1 and K_7, and a sign serving G_3, G_4, G_9 on K_4).

Thus from the point of view of the kiosk user, identification of topological mosaics is desirable for it permits him to cut his expenses. However, from the point of view of the kiosk builder, efficiency of communciation using kiosks is greatest in regions where there are no topological mosaics and so, from that view, location of new kiosks should avoid creating topological mosaics.

Correspondence of geographical and mathematical framework in this case study indicates that the primary difficulty in approaching and executing such work is in creating an object (geographical topological space in this case) that lies between both topological and geographical structures and yet is also one that is systematically developed with respect to both structures.

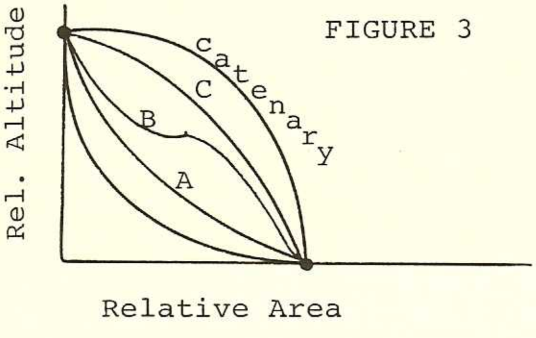
Emergency telephone placement

The Kioskland case study proved useful in analyzing the effectiveness of coverage of Ohio State's Campus by emergency telephones. Students in both mathematics and geography courses were assigned the project of evaluating the extent of coverage provided by the emergency telephone system then in palce, and of determining locations for new telephones. A summary of the results is included in Appendix C.

Urban hypsometry

Identification of curves A, B, and C (Figure 3) that characterize the relations between relative altitude and relative drainage basin area for streams embedded in topographic surface is a geometric realization of hypsometric analysis ("relation of horizontal cross-sectional drainage basin area to elevation"[16]). Curve C characterizes the drainage basin of a young stream, curve A that of an old-age stream, and curve B that of an equilibrium position between the two.[17]

FIGURE 3

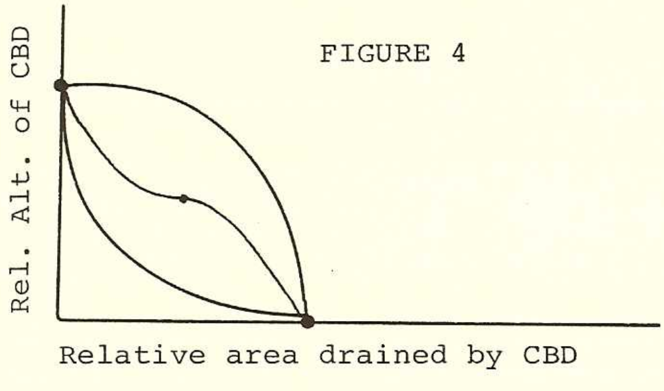


A broader view might suggest that a catenary [y = cosh x] is the limiting form of curve A, and that a reflected catenary is the limiting form of curve C. The presence of a hull of catenaries bounding intermediate positions for hypsometric curves is consistent with geomorphological interpretation of the limiting forms of these curves. For a catenary is the curve that provides maximum support between two extreme positions (as in suspension bridges), and, catenary A is such that the basin supports the maximum amount of sediment deposited from an old-age stream, while catenary C represents a basin that supports the minimum amount of sediment deposited from a youthful stream.

Further reflection suggests that the position of curves may be forced outside the bounding hull; for, in the physical landscape, dam placement could alter deposition of sediment to such an extent that the drainage basin could no longer tolerate its own stream.

Even more generally, transforming these ideas to an entirely symbolic form as a pivot for thought, suggests alternate style of analysis. The hull of catenaries is roughly similar in form to the monad, when catenaries are deformed to semi-circles and the line of stream equilibrium is deformed to the line of separation between Yin and Yang. The inflection point of the stream equilibrium line is the center of the monad.

FIGURE 4



Using the symbol as city leads to the idea of urban hypsometry.[18] In the urban landscape Yin and Yang become human landform and human erosive agent, or CBD and transport line, instead of land and water, or natural landform and natural erosive agent. Transformation of the monad, in the urban landscape, back to the Cartesian plane produces a figure (Figure 4) corresponding to that developed for drainage basins (Figure 3). In this case, faster moving streams of traffic are considered to erode a huge block of cement to steep landforms (high rise buildings). So, caternary A is such that the urban area supports the maximum amount of settled traffic relative to the CBD (the CBD is decaying) while catenary C represents a period of swift building of the CBD and a minimal amount of traffic has settled across the entire urban area.

Between these two positions is an equilibrium position which may be forced outside the bounding hull by severe alteration in traffic patterns or by the introduction of too many high buildings relative to nearby land that can support parking garages. Use of this style of analysis in urban areas, together with line-of-sight neighborhoods at different croos-sectional levels may aid in forming CBD's and urban areas that are arranged harmoniously relative to area/altitude relations.[19]

Street gangs (urban terrorism)

Regions formed by street gangs based on territoriality could lead to the use of the taxonomy to monitor the level of action of street gangs in an urban area; danger occurs in areas of regional overlap, particularly in areas of overlap of core areas (domination), leading to seizure of the focus of a nodal region by one group or another.[26] Assessment of level of activity within the taxonomy could be through current newspaper accounts and word frequency tabulation.

The map that follows shows two such interacting regions derived from data in Ley and Cybriwsky's article.[21] It is intended as representative of the type of map that mightalign nicely with the taxonomy (Figure 5).

OVERLAPPING URBAN STREET-GANG REGIONS

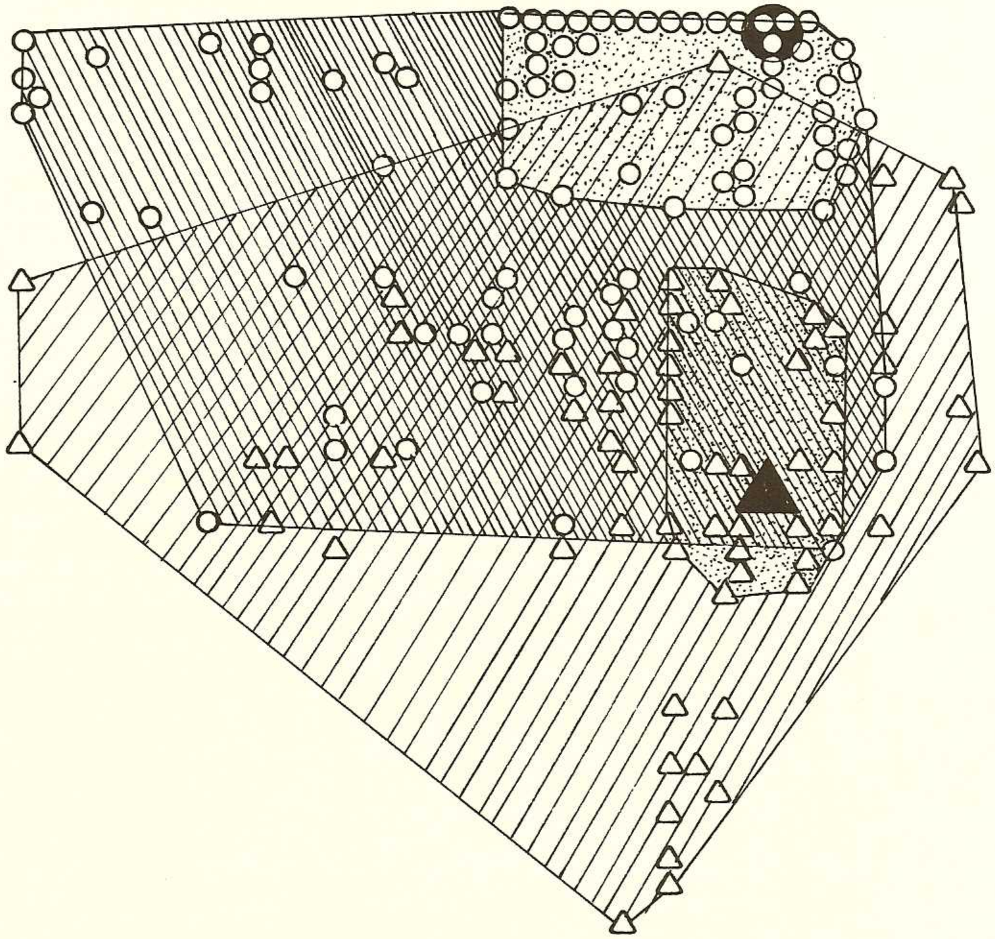


FIGURE 5

APPENDIX A

TRANSFORMATION OF GEOGRAPHICAL TOPOLOGICAL SPACES

Earlier sections examined problems in defining geographical topological spaces and in determining the nature of the internal structure of such spaces. They did not consider comparison of one geographical topological space to another. To accomplish such comparison requires locating structure in both spaces that is invariant with respect to some systematic method of transformation of one space into the other; change is recognized in terms of that which remains fixed. Since point set topology itself already relies heavily on such transformations, initial experiment would invoke those methods. The mathematical development presented below culminates in definition of a homeomorphism, a topological transformation useful in determining similarity of structure.[22]

<u>Definition A.1</u> (definition of a transformation, τ)

A relation τ that relates elements of a set X to elements of a set Y is said to be a transformation of X into Y, $\tau : X \to Y$, if, whenever $x_1\tau \neq x_2\tau$ it follows that $x_1 \neq x_2$ ($x_1, x_2 \in X$). When τ satisfies this definition, it is said to be well-defined.

<u>Definition A.2</u> (definition of a one-to-one transformation)

If $\tau : X \to Y$ is well-defined and if, whenever $x_1 \neq x_2$ it follows that $x_1\tau \neq x_2\tau$ ($x_1, x_2 \in X$), then τ is said to be a transformation from X into Y that is one-to-one.

<u>Definition A.3</u> (definition of onto)

Suppose $\tau : X \to Y$ is well-defined; τ is said to be a transformation of X onto Y if $X\tau = Y$.

<u>Proposition A.1</u> (existence of τ^{-1} as a transformation).

Suppose $\tau : X \to Y$ is well-defined. The relation $\tau^{-1} : Y \to X$ is a transformation if and only if $\tau : X \to Y$ is one-to-one and onto.

Definitions A.1, A.2, and A.3 and Proposition A.1 deal with transformation between sets; the definitions that follow rely on these and provide material that deals with transformation between topological spaces.

<u>Definition A.4</u> (definition of a continuous transformation in terms of open sets).

Suppose $\tau:X \to Y$ is well-defined. Then $\tau:(X,T) \to (Y,U)$ is said to be continuous with respect to the topologies T and U if and only if $R\tau^{-1} \in T$ for each $R \in U$.

<u>Definition A.5</u> (definition of continuous transformation in terms of neighborhoods)

A transformation $\tau:X \to Y$ is said to be continuous at a point $x \in X$ if and only if τ^{-1} of each neighborhood of $x\tau$ is a neighborhood of x. The transformation $\tau:X \to Y$ is said to be continuous if and only if it is continuous at each point $x \in X$. (This is equivalent to Definition A.4.)

<u>Definition A.6</u> (definition of a homeomorphism)

A transformation $\tau:(X,T) \to (Y,U)$ is said to be a homeomorphism if and only if τ is one-to-one, onto, and if both τ and τ^{-1} are continuous transformations.

Homeomorphisms look particularly promising as means of external examination of geographical topological spaces. For the search for properties that are invariant under homeomorphisms is one that could be of interest to physical and human geographers alike, as well as to others in a wide variety of disciplines. Both transformation and set (as process and form, respectively) are essential to such invariance.

APPENDIX B

<u>Proposition B.1</u>

A mosaic M, in (X,T), is a union of neighborhoods N_δ, $\delta \in \Delta$, of (X,T), where the level of neighborhood interaction is contained between stages II.B and II.C.ii.b of the neighborhood taxonomy (Figure 1).

Proof:

A mosaic is "a surface decoration made by inlaying small pieces of variously colored material to form pictures or patterns."[23] There is the interpretation of this definition then that the individual tiles retain their basic characteristic (color) but, when aggregated, work to form an uninterrupted pattern. That is the S_i retain their character; they do not intersect each other, but their hinterlands come into contact, and, pairwise, at least one hinterland has non-empty intersection with another core, as a glue fusing the individual core-tiles with each other, in order to form an uninterrupted pattern. Or, $M = \bigcup_{\delta \varepsilon \Delta} N_\delta$ (where the level of interaction between any two N_δ is contained between levels II.B and II.C.ii.b). Q.E.D.

Proposition B.2:

In a union M* of neighborhoods $N_\delta{}^*$, $\delta \varepsilon \Delta$, of (X,T), a society $S_2{}^*$ will be said to be creating pressure in the civilization M* if there exists $N_0{}^*$ among the $N_\delta{}^*$, $\delta \varepsilon \Delta$, such that $S_2{}^* \cap S_0{}^* \neq \emptyset$ and $N_0{}^*$ and $N_2{}^*$ do not have topologically fused neighborhood content.

Proof:

This is clear from the definition of pressure as "the application of force to something by something else in direct contact with it" when this definition is viewed within the neighborhood taxonomy.[24] Q.E.D.

APPENDIX C

The content of this Appendix consists of the material in Figure 6.

SUN., JAN. 21, 1979 Columbus Dispatch A-7

Practical Approach Counts In Math Class

By Carol Ann Lease
Of The Dispatch Staff

A mathematics lecturer teaching a remedial class at Ohio State University used a practical approach and was pleasantly surprised.

"I got tremendous amounts of work from them," Sandra Arlinghaus said of the students in her fall Introduction to College Mathematics course.

THE STUDENTS were those who scored at the lowest level on a mathematics test given all freshmen entering OSU. There were 1,800 in the course and about 30 in her class.

Arlinghaus

"What we teach is elementary arithmetic," Mrs. Arlinghaus said.

Students come from all mathematics backgrounds. "There are even some who haven't seen fractions," she said.

SHE SOUGHT publicity for what her students accomplished after she saw a report downgrading remedial college classes.

"The assumption was 'If these students haven't learned the material by now, they're ineducable.' I don't think that's true," she said.

"Even though there is some kind of deficiency in the background of some of them, they can do good work."

SHE DECIDED that the students didn't like math because they weren't used to working with numbers, and "it doesn't seem to relate to anything."

She gave them a practical problem: How well does the emergency telephone system at OSU serve the campus?

The system consists of 23 telephones, 11 of which were installed at the end of this past summer, connected directly to the University Police dispatcher.

ANYONE WITH A problem need only pick up one of the phones, which are topped by white lights, and a police dispatcher will answer and send the appropriate aid, be it cruiser, fire truck or ambulance.

Assuming that someone who can see a phone is in a safer area than someone who cannot, the students marked areas on campus maps where phones are visible. They also did the same on another map showing the locations of telephones before the 11 were added this summer.

They then calculated the area of the campus minus the area of the buildings. To do this, they also had to go out and determine the scale of the maps they used because it is not marked.

STUDENTS COULD cooperate in gathering information but had to analyze it themselves, write a report and do their own maps to illustrate how they arrived at their conclusions.

One used a drawing of the cartoon character Snoopy in his work to illustrate how such an attention-getter could make phones more visible.

Another determined that scale varied across the map and used her car to test her initial measurements against distances between various pairs of locations.

MRS. ARLINGHAUS said it was found that there are about 14.7 million square feet of campus not covered with buildings. The old phones could be seen in 41 percent of this territory.

With the 11 new phones, 49 percent is covered for an improvement of 19.5 percent.

"Concensus recommendations were (to put) another phone on north campus and one south of the hospitals," Mrs. Arlinghaus said. Some students also recommended putting emergency phones in the middle of the Oval and near Mirror Lake.

MRS. ARLINGHAUS, who has a doctorate in theoretical geography and is in her second year of teaching at OSU, said she chose the problem because the emergency phone system interested her and "I think the students like to go out on field research."

This quarter she has another introductory math class and has planned a variety of projects.

For example, students will do a project similar to the phone study looking at the distribtution of kiosks or message boards on campus instead.

ALSO, "WE ARE going to measure the rate of discharge of the Olentangy River" if it thaws, Mrs. Arlinghaus said, and do some other things such as throwing snowballs and finding the depth of the depressions they make.

It sounds like play, but all projects have one thing in common, she said — "To do them, you need to know fundamental mathematics."

FIGURE 6

NOTES

*Based on an invited lecture given to the Department of Geography, The University of Chicago, May 2, 1979, and on material developed at The University of Michigan, 1976-77.

1. Maynard J. Mansfield, *Introduction to Topology* (Princeton: D. Van Nostrand Company, Inc., 1963), p. 1.

2. John L. Kelley, *General Topology* (Princeton: D. Van Nostrand, Inc., 1955), p. 37.

3. *Ibid*.

4. *Ibid*.

5. Felix Hausdorff, *Mengenhehre* (Berlin, Leipzig: Walter de Gruyter and Co., 1935), p. 24. Also, Harold T. Betteridge, ed., *The New Cassell's German Dictionary* (New York: Funk and Wagnalls Company, 1965), p. 316.

6. *Ibid*.

7. Kelley, *op*. *cit*., note 2, p. 8.

8. *Webster's Seventh New Collegiate Dictionary* (Springfield, Mass.: G and C. Merriam Company, 1965), p. 828. Definitional problems associated with words such as this are obviously complex; a simple dictionary definition was chosen initially. Specific empirical studies might help to single out definitions. A. L. Kroeber and Clyde Kluckhohn, *Papers of the Peabody Museum of American Archaeology and Ethnology*, Vol. XLVII, No. 1 (1952).

9. Webster, *op*. *cit*., note 8, p. 202.

10. *Ibid*., p. 248.

11. Jan O. M. Broek and John W. Webb, *A Geography of Mankind*, second edition (New York: Mc Graw-Hill, 1968), p. 30. Same strategy here as with the use of the dictionary (see comment in note 8).

12. Paul Alexandroff and Heinz Hopf, *Topologie* (Berlin: Verlag von Julius Springer, 1935), pp. 58-68, *passim*.

13. If one were to admit something similar to reincarnation into the space under consideration, then inheritance might be viewed as a reversible property. In fact, such observation might permit partitioning of human traditions into equivalence classes, where an equivalence relation $\sim$ among elements of a set is such that

a) $a \sim a$ (reflexive)

b) if $a \sim b$ then $b \sim a$ (symmetric)

c) if $a \sim b$ and $b \sim c$, then $a \sim c$ (transitive).

Thus property inheritance is an equivalence relation if, and only if, a concept such as reincarnation is admitted, for otherwise the relation could not be reflexive. Equivalence classes, or almost-equivalence classes (satisfying some, but not all of (a), (b), and (c)) based on a set of fundamental human relations might be used to develop "cultural" plates and to examine zones of contact, zones of generation, and zones of destruction of these plates anchored on a set of human traditions that shift through time and space by the mechanism of diffusion.

14. Territoriality is to be studied topologically in this case. Gerald D. Suttles, The Social Construction of Communities (Chicago: The University of Chicago Press, 1972); Kevin Lynch, The Image of the City (Cambridge, Mass.: M.I.T. Press, 1960).

15. Application of this type would represent an attempt to go beyond the taxonomy. Goals of this sort, as well as creation and modification of taxonomies are stated or reflected in works by Linnaeus, Rashevsky, D'Arcy Thompson, Lewin, Sokal and Sneath, Darwin, Spencer, Tylor, Vallaux, Morgan, Childe, White, Steward and many others.

16. Arthur N. Strahler, "Quantitative Analysis of Watershed Geomorphology," Transactions, American Geophysical Union, Vol. 38, No. 6 (December, 1957), p. 918.

17. Arthur N. Strahler, "Hypsometric (Area-Altitude) Analysis of Erosional Topography," Bulletin of the Geological Society of America, Vol. 63 (November, 1952), p. 1124.

18. Inversion of the usual 'city as symbol'.

19. As discussed with Everette Bannister, Ph.D., (deceased) of the Department of Geography of The University of Michigan; the phrase 'urban hypsometry' is Bannister's suggestion.

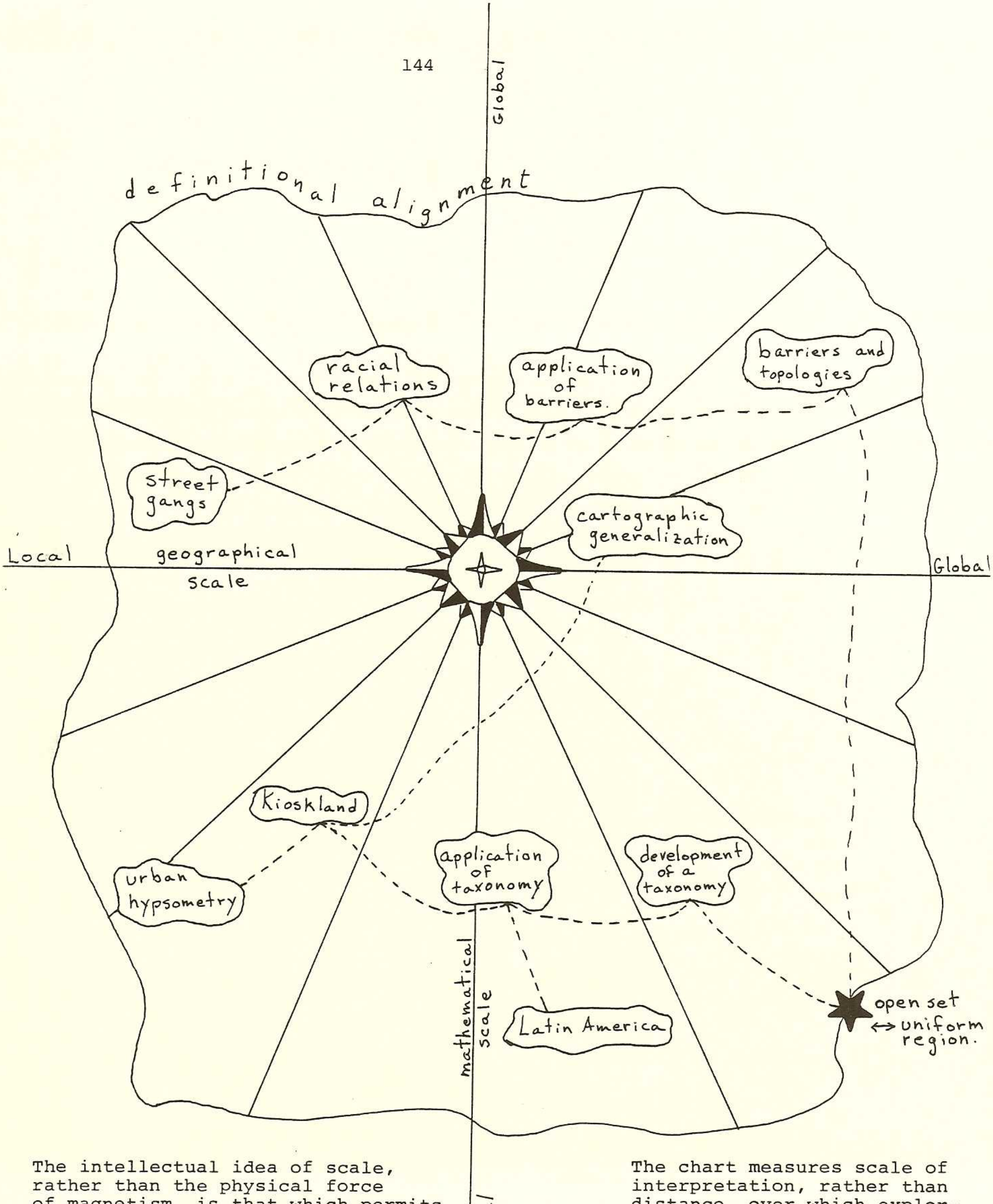
20. David Ley and Roman Cybriwsky, "Urban Graffiti as Territorial Markers," Annals, Association of American Geographers, (December, 1974), p. 498.

21. Ibid.

22. Kelley, op. cit., note 2, pp. 10-11.

23. Webster, op. cit., note 8, p. 552.

24. Ibid., p. 673.



The intellectual idea of scale, rather than the physical force of magnetism, is that which permits identification of cardinal points.

The chart measures scale of interpretation, rather than distance, over which exploration takes place.

A SPACE FOR THOUGHT

I never saw a moor,
I never saw the sea;
Yet know I how the heather looks,
And what a wave must be.

Emily Dickinson

INTRODUCTION

Albert Einstein's comment to Max Wertheimer, that "[my] thoughts did not come in any verbal formulation. I very rarely think in words at all. A thought comes, and I may try to express it in words afterward...I have it in a kind of survey, in a way visually" [Wertheimer, 1959, p. 238], provoked this attempt to describe diffuse thought-processes. Thoughts of this sort might bounce around; in one instant a given thought might be close to another, yet, in the next instant, quite distant from it. Thus the use of metric spaces to describe such thought-processes appears inappropriate unless one allows some sort of folding, to juxtapose "near" and "far," as René Thom does with his use of differential topology (based on the Euclidean metric) to characterize "thought [as] a virtual capture of concepts with a virtual, inhibited, emission of words..." [Thom, 1975, pp. 312-313, 331-332]. Another approach is simply to abandon metric spaces and revert to point-set topology, whose relations form the foundations of differential topology [Auslander and MacKenzie, 1963]. This is the approach taken below; it is more global than was Kurt Lewin's use of point-set topology in 1936 to represent an individual's "life-space" [Lewin, 1936]. Both, however, rest primarily on the set-theoretic notions of inclusion, union, and intersection.

R. H. Atkin also applies topological structure to examine human affairs; throughout, he employs the combinatorial approach to topology [Atkin, 1974]. The combinatorial approach creates the global picture by glueing together small pieces, while the set-theoretical approach begins with the whole and dissects it to look at individual systems and subsystems. The impact that the chosen style of approach has on the final product is as vivid in the application of mathematics as it is in art; from the latter vantage point, the difference in choice of topological approach comes alive as a striking difference in paintings, as between Seurat's pointillistic and Cézanne's impressionistic representations of play, as a walk in the park, or as a hand of cards. As the mathematician Saunders Mac Lane has put it, in commenting on finding combinations of mathematical ideas well-suited to application, "...subtle ideas, fitted by hand to the problem, can lead to triumph" [Mac Lane, 1982, p. 28].

TOPOLOGICAL BACKGROUND

The following sequence of definitions provides material for application; the use of definitions and theorems beyond these would no doubt produce additional insight but is beyond the scope of this article.

Definition 1 [Mansfield, 1963, p. 15; Kelley, 1955, p. 37].

Let X be a non-empty set and let T be any collection of subsets of X. Then T is called a topology for X under the following conditions:

1) futuristic condition (Lemma 1 will motivate this term)
   If $G_\lambda \in T$ for all $\lambda \in \Lambda$, then $\bigcup_{\lambda \in \Lambda} G_\lambda \in T$; this condition admits infinite, as well as finite, unions of sets as members of the topology.
2) historical condition (Lemma 1 will motivate this term)
   If $G_1, G_2, \ldots, G_n \in T$, then $\bigcap_{i=1}^{n} G_i \in T$; this condition admits finite, but not infinite, intersections of sets as members of the topology.
3) $X \in T$ and $\Phi \in T$, where Φ denotes the empty set.

Definition 2 [Mansfield, 1963, p. 15; Kelley, 1955, p. 37].

If T is a topology for the set X, then the couple (X, T) is referred to as a topological space; the sets $G_\lambda \varepsilon T$ are called the open sets of T (G, from the German "gebiet" ("region") is often chosen to denote open sets).

To characterize some of the types of structures that might occur within a topological space, other than the required open sets, Definitions 3, 4, 5, 6, and 7 prove useful.

Definition 3 [Mansfield, 1963, p. 89; Kelley, 1955, pp. 62-63].

A sequence s in a set X is a mapping of the set of non-negative integers, Ω, into X; that is, $s: \Omega \to X$ is defined by $s(\omega) = x_\omega$ for $\omega \varepsilon \Omega$ where x_ω is the value of the sequence s at ω.

Thus, for example, the set X of pages in a book is put into a sequence by the assignment of numerals, beginning with the numeral 1. Here ω represents a numeral while $s(\omega) = x_\omega$ represents the page with that numeral assigned to it.

Definition 4 (Figure 1) [Mansfield, 1963, p. 89; Kelley, 1955, p. 63].

Let s be a sequence in (X, T). Let G be an open set in (X, T) containing the point p. The sequence s is said to be eventually in G if and only if there exists $\omega_1 \varepsilon \Omega$ such that $x_{\omega_2} \varepsilon G$ whenever $\omega_2 \geq \omega_1$, where "greater than or equal to" means is "beyond" in sequential position.

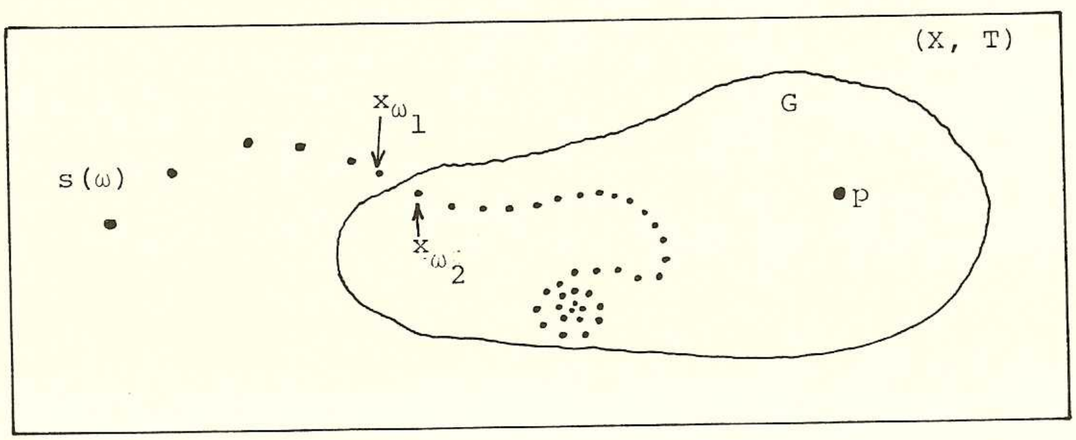


FIGURE 1

Elements of the sequence need only eventually fall into some open set G containing p. A stronger condition for determining how "close" a sequence s is to a point p, chosen a priori, is given in the Definition that follows.

Definition 5 (Figure 2) [Mansfield, 1963, p. 90; Kelley, 1955, p. 63].

Let s be a sequence in (X, T), and let $p \in X$. Then p is a limit point of s if s is eventually in every open set containing p.

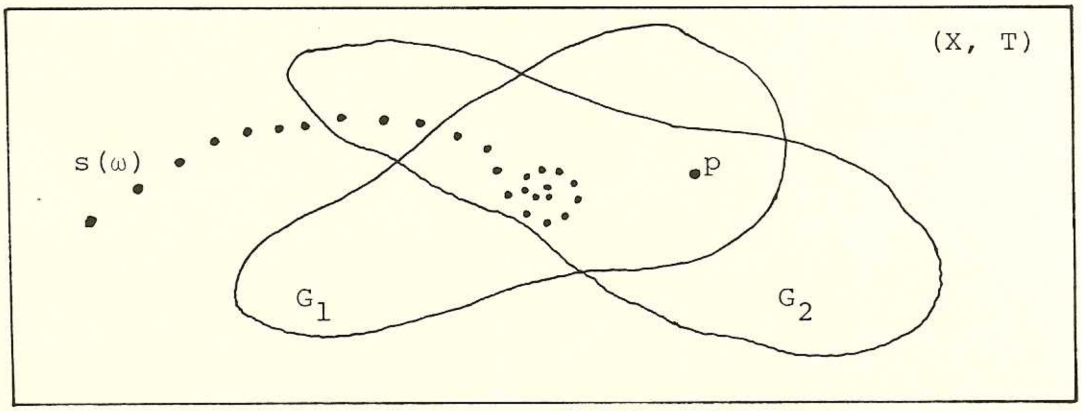


FIGURE 2

Here, s must eventually lie totally within any intersection of open sets of which p is a member. There may, however, be more than one point p in X which satisfies this Definition.

For example, suppose (X, T) is composed of the set X of rational numbers between 0 and 10 (including 0 and excluding 10) and of the collection T formed from left-half closed intervals with integral endpoints (such as [1,2), where the left-hand endpoint is included and the right-hand endpoint is not). Using Definition 3, if Ω = {integral powers of 2} = $\{\omega | \omega = 2^n$, n an integer$\}$, then $s: \Omega \rightarrow X$ produces the sequence {8, 4, 2, 1, 1/2, 1/4, 1/8,...} in X. From Definition 4, this sequence is eventually in G = [0,4), for since 4 is a power of 2, it is an element of T and will serve as ω_1 in Definition 4;

any ω_2 "beyond" 4 is in the interval [0,4). Of course, by similar reasoning, this sequence is also eventually in [0,2), [0,1) as well as in other intervals. Using Definition 5, the point p = 1/2 is a limit point of s since s is eventually in every open set containing p; it is eventually in all left-half closed intervals with 0 as the left-hand endpoint and there are no other left-half closed intervals containing 1/2. Similarly, 1/3 is a limit point for this sequence, as is any other point in [0,1). Clearly, with this choice of topology T for X the sequence s has more than one limit point. The Definition that follows is more restrictive and describes a more highly controlled situation.

Definition 6 [Mansfield, 1963, p. 91; Kelley, 1955, p. 63].

A sequence s in (X, T) is convergent in (X, T) if and only if there exists a unique point $p \in X$ that is the limit point of s.

Using the example above, suppose X is the same and suppose instead that the topology T is formed from left-half closed intervals with rational endpoints. With this topology, that permits "finer" distinctions, 0 is the only limit point of the sequence s described above.

In addition, a given sequence may converge in one topological space and fail to converge in another. For example, consider again the sequence 1, 1/2, 1/4,... . Let X_1 = [0,10), the left-half closed interval of rational numbers from 0 to 10, and let T_1 represent the set of all left-half closed intervals in [0,10). Then this sequence converges to 0 in (X_1, T_1) since $0 \in X_1$ is the unique limit point of s in (X_1, T_1). If, however X_2 = (0,2], the right-half open interval of rational numbers from 0 to 10, and if T_2 represents the set of all right-half open intervals in (0, 10], the sequence s fails to converge in (X_2, T_2) as $0 \notin X_2$ and as there is no other limit point of s in (X_2, T_2) to which the sequence might converge.

To interpret topological structure at a more local scale, the following Definition will prove useful.

Definition 7 [Mansfield, 1963, p. 48; Kelley, 1955, p. 51].

Let (X, T) be a topological space and let Y be a subset of X. The T-relative topology for Y, denoted by S is the collection $S = \{G \cap Y | G \varepsilon T\}$. The topological space (Y, S) is called a subspace of (X, T).

APPLICATION

To align topological material with human communication systems, the following words and the diverse images they might represent are taken as primitive terms.

Primitive Terms:

a) Thought b) Concept c) Conscious.

What these mean will of course vary from individual to individual. Concerns about the interaction of language and thought date from classical Greek philosophy and the liar's paradox of Epimenides the Cretan [Bronowski, 1978, p. 82], to twentieth century logicians Russell, Quine, and Tarski [Bourbaki, 1968, p. 328; Quine, 1960; Tarski, 1956, pp. 154-155], to the linguists Chomsky, Hayakawa, and Whorf [Chomsky, 1968; Hayakawa, 1941; Carroll, 1956], as well as to a host of others in a variety of disciplines. The propositions that follow attempt to capture the diffuse character of thought with "diffuse" mathematics. As the mathematical symbols are, to some extent, free from different interpretations in different languages, the self-reference problem behind these paradoxes is superficially addressed. However, as Tarski puts it, "the language about which we speak need by no means coincide with the language in which we speak," suggesting that any such "capture" of thought needs further evaluation to determine whether or not it is only apparent [Tarski, 1956, p. 402].

For the purpose of a display of mathematical alignment, as opposed to that of an exhibit of linguistic or philosophical underpinnings, "thought" is assumed to be present in the mind of the individual whether or not one contends, as does Whorf, that language shapes one's thoughts and view of the world [Carroll, 1956]. Similarly, "concept" is considered to be an organized grouping of thoughts. There remains with this grouping the difficulty of the Russell Paradox, inherent in any formulation of classes. As Quine comments, however, "...the admission of classes as values of variables of quantification brings power that is not lightly to be surrendered..." [Quine, 1960, p. 266]. Finally, "conscious" is viewed as a state of being that presumes a degree of alertness sufficient for thought and concept formulation. Such alertness requires sufficient imagination to engage in thought and concept formulation, and it thus fits, to some extent, with Bronowski's view that "the central problem of human consciousness depends on the ability to imagine" [Bronowski, 1979, p. 18].

<u>Lemma 1</u>

Let X represent the set of all conscious human thoughts. Let T represent the set of all concepts. Then the pair (X, T) is a topological space, representing collective human thought.

<u>Proof</u>:

To show that (X, T) is a topological space, it is required to show that T meets the three conditions of Definition 1.

1) Show that the union of an arbitrary number of concepts is once again a concept.

   A) Finite unions.

      i) Unions of similar concepts.

         a) The union of two concepts that are similar fuses them as a single concept that might be viewed as a small enlargement of the scope of either original concept.

b) The union of a larger number of similar concepts procedes as in (a), until something akin to a "corporate" image is reached. One would expect this image to be recognizable to those internal to it as well as to those external to it.

ii) Unions of dissimilar concepts

a) The union of two concepts in which the differences between the concepts dominate is yet another concept which focuses on these differences. For example, the radically opposed South African views of "one person--one vote" and "Apartheid" fall together under the broader conceptual umbrella of the role of voting procedures in providing equitable governance.

b) The union of a large number of dissimilar concepts procedes as in (a), until a social structure somewhat like a research university, which thrives on interchange among differing philosophical viewpoints, is reached. Here, the internal view would be opposed to the external view, as in the "town-gown" conflict present in many college communities.

iii) Unions of similar and dissimilar concepts.

The union of a set of related concepts with another set of unrelated concepts would produce a revolution once these sets have merged. For example, the merging of the corporate computer image with the traditionally diverse liberal arts curriculum is creating a technological revolution in printing from word-processing to books printed from camera-ready copy.

B) Infinite unions.

An infinite union of concepts presents itself at a variety of levels in the mind. For example, each of the numbers, "one," "two," "three," ... represents a distinct concept, while the infinite set of all such numbers merges them as the concept of all positive integers. More broadly, the union of an infinite number of concepts permits the indefinite extension of a set of concepts into the future--hence, the label of "futuristic condition" for the arbitrary union property in Definition 1.

2) Show that the intersection of a finite number of concepts is once again a concept.

The intersection of two concepts, similar or not, is the common core of tradition from which they arose. Of course, in the case of highly similar concepts this core would be "larger" than it would be in the case of dissimilar concepts. Any finite set of concepts has a common tradition (including possibly the empty set) around which these concepts have emerged. Since such traditions, be they "great" or "small" [Redfield and Singer, 1954, p. 58], have been established over time, and since the amount of time which has elapsed (in contrast to that which is yet to elapse) is finite, the presence of a finite (rather than an infinite) intersection property is required--hence the label of "historical condition" for the finite intersection property in Definition 1.

3) A) $\Phi \in T$.

The empty set represents the lack of organization of thoughts into a concept; however, this lack is itself a concept and must therefore be included in T.

B) $X \in T$.

The characterization of X as the set of all conscious human thought is itself a concept and must therefore be included in T. QED

Corollary 1

The topological mind-space of an individual, (Y, S), is a subspace of (X, T).

Proof:

Using Definition 7, Y represents the thoughts of a single person and S represents the collection of concepts that his view of the concepts in T admits into his mind.

Theorem 1

Suppose that s_1 represents a sequence of thoughts of individual I_1, in mind-space (X_1, T_1), and that s_2 represents a sequence of thoughts of individual I_2 in mind-space (X_2, T_2). The position of a thought p relative to the thought-sequences s_1 and s_2 may be used to order the level of understanding about p exchanged between I_1 and I_2 as : maximal, significant, moderate, minimal (exposure only), misunderstanding, or bewilderment. It is assumed that such an exchange is communicated via words and that appropriate dictionaries or interpreters are available.

Proof:

It is assumed that I_1 initiates the exchange; dual arguments hold if I_2 is the initiator (Figure 3).

1) Suppose s_1 converges to the unique limit point p in (X_1, T_1);

   A) if s_2 also converges to p in (X_2, T_2), then "maximal" understanding about p is exchanged between I_1 and I_2. This follows since every open set of T_1 containing p eventually contains s_1 and s_2 since every open set of T_2 eventually contains s_2 (Definition 6 and Corollary 1), and there is no room for confusion since no other point p in either X_1 or X_2 has this property.

B) if p is one of several limit points of s_2, then "significant" understanding is exchanged. Again, every open set of T_1 containing p eventually contains s_1, and every open set of T_2 containing p eventually contains s_2 (Definitions 5, 6, and Corollary 1). However, here p is a unique limit point of s_1; if it is also a unique limit point of s_2, then Case 1.A holds. Otherwise, it is one of several limit points within any intersection of open sets in T_2. Thus, there may arise some "incompleteness" in I_2's understanding of the sequence in I_1's mind that is associated with p (as if, in the example following Definition 5, I_1 used 0 as a limit point and I_2 used 1/2).

C) if s_2 is eventually in some open set of T_2 containing p, then I_2 receives exposure to I_1's thought sequence leading to p. Here, elements of s_2 need only eventually fall into some open set in T_2 containing p (Definition 4), so that the exchange linking the two sequences may contain only a very small amount of information.

D) if s_2 is not eventually in any open set of T_2 containing p, then misunderstanding about p results for I_2, and it is recognized immediately by I_1 since there is no other limit point for s_1 in (X_1, T_1).

2) Suppose p is one of several limit points of s_1;

A) if p is also one of several limit points of s_2, then a moderate level of interchange is ensured, since each of s_1 and s_2 must lie eventually in every open set of T_1 and T_2 containing p. The extent of significance in the exchange is reflected in the extent of overlap of the intersections of open sets from T_1 and T_2, which contain p.

B) if s_2 is eventually in some open set of T_2 containing p, then I_2 receives only minimal insight about p (as in 1.C; however, here the degree to which I_2 receives exposure may be even less than in that case).

C) if s_2 is not eventually in any open set of T_2, then misunderstanding about p by I_2 is the result. Such misunderstanding is suspected by I_1 since the number of possible limit points for s_1 is bounded by the size of the intersection of all open sets in T_1 containing p.

3) Suppose s_1 is eventually in some open set of T_1 containing p.

A) Suppose s_2 is eventually in some open set of T_2 containing p. As in 1.C and 2.B, only exposure to each other's sequences results, and in this case that exposure is even more superficial than in the two previous ones.

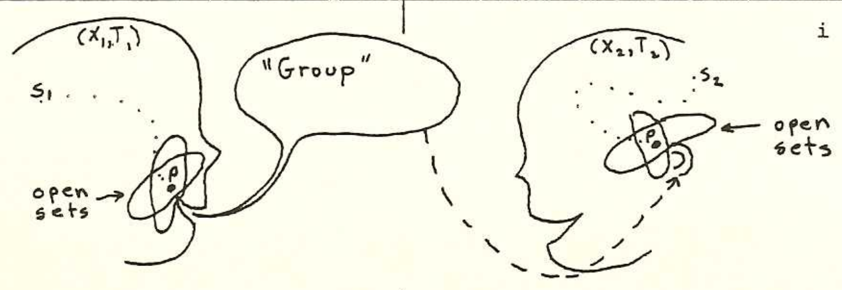
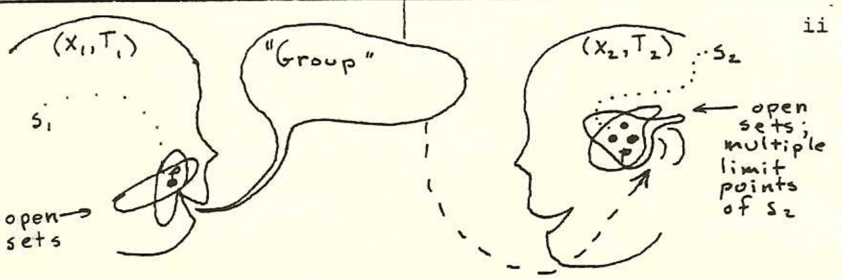
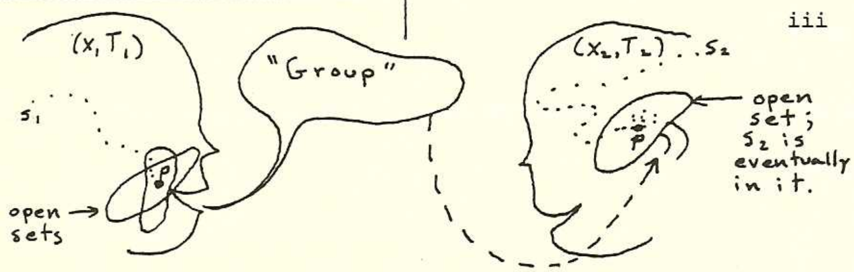
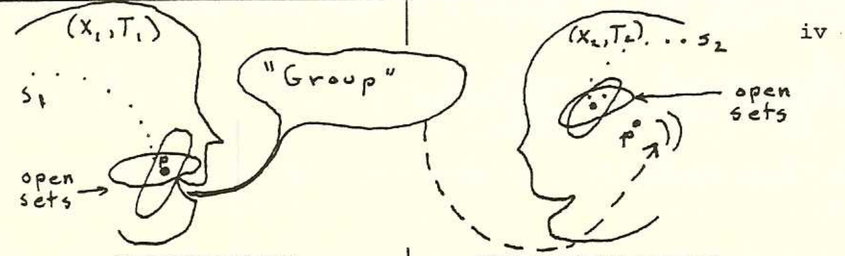
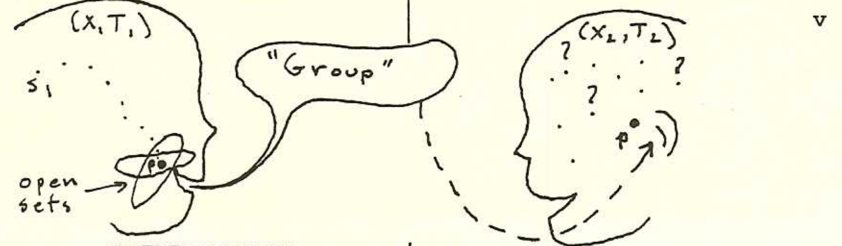
B) If s_2 is not eventually in any open set of p, misunderstanding results. In this case, I_1 likely has only a faint notion that this has happened.

4) Suppose s_1 is not eventually in any open set of T_1.

A) If s_2 is not eventually in any open set of T_2, bewilderment results for both.

As an example, suppose one considers thoughts associated with the word "group." A mathematician's thought-sequence leading to this particular p would contain, in any open set containing the thought, the concepts in the definition of a group as a mathematical system, closed under a single operation, that obeys associative, identity, and inverse properties. Thus, in Figure 3.i, maximal exchange takes place between two mathematicians I_1 and I_2 in which the sequences s_1 and s_2 both converge to p in (X_1, T_1) and (X_2, T_2) respectively. Figure 3.ii illustrates a situation in which s_1 converges to p in (X_1, T_1) and in which s_2 has p as one of several limit points; in this case, significant understanding about groups is exchanged between the mathematician, I_1, and the non-mathematician, I_2. Figure 3.iii shows an exchange in which s_1 converges to p in (X_1, T_1) and in which the sequence s_2 is eventually in some open set in T_2 containing p. Here only a superficial level of communication occurs, resulting

FIGURE 3

SPEAKER, I_1	LISTENER, I_2	POSSIBLE INTERPRETATION BY I_2
 MATHEMATICIAN	i MATHEMATICIAN	MAXIMAL UNDERSTANDING: "Oh yes--a group is a mathematical system closed under a single operation obeying associative, identity,and inverse properties."
 MATHEMATICIAN	ii NON-MATHEMATICIAN	SIGNIFICANT UNDERSTANDING: "Oh yes--we study closed systems too, be the internal linkage through mathematical or genetic relationships."
 MATHEMATICIAN	iii NON-MATHEMATICIAN	EXPOSURE (MINIMAL UNDERSTANDING): "Oh yes--'group', I suppose that must be like 'set,' or 'cluster'."
 MATHEMATICIAN	iv NON-MATHEMATICIAN	MISUNDERSTANDING: "Oh yes--I agree that the peer-group of mathematicians is probably as significant as is any other in determining university policy."
 MATHEMATICIAN	v NON-MATHEMATICIAN	BEWILDERMENT: ??????

in exposure of I_2 to I_1's ideas. Figure 3.iv again shows the mathematician and the non-mathematician, and here it is clear to I_1 that the non-mathematician has completely misunderstood p. The sequence s_2 triggered by the word "group" in I_2 is not eventually in any open set of T_2 that contains p. The final frame, Figure 3.v, represents bewilderment; the word "group" triggers no sequence in (X_2, T_2).

What appears to be critical in this example is not so much that the word "group" has different meanings, but rather that the thought-processes leading to this word were distinct. The non-mathematician required only existence criteria (Definition 5), whereas the mathematician required existence as well as uniqueness criteria (Definitions 5 and 6). This distinction suggests the precision in language-use required in mathematics and mathematical sciences as well as the frustration often felt by non-mathematicians in attempting to communicate with mathematicians. At a deeper level, it suggests yet another approach to expressing the differences between "The Two Cultures," and the implications of these differences for solving problems of global significance, commented on by C. P. Snow [Snow, 1959].

Corollary 2

The way in which individuals choose to partition space may depend on whether or not they expect thoughts to appear as unique limit points of sequences; e.g., mathematicians sit together at one lunch table at the faculty club while humanists congregate at another.

Corollary 3

Thought-sequences that do not have limit points in an individual's topological subspace of (X_1, T_1), might consume much of his thought-space. If the individual figures out a way to find limit points in (X, T) and incorporate them in (X_1, T_1), through further education, then he develops something that, to him, is new.

Certainly Bronowski touches on this idea as he comments that "you have a lot of irrelevant thoughts about what you are going to wear tonight...but you do not put such thoughts into words because they are not specific; you are just conscious that you are thinking..." [Bronowski, 1978, p. 35].

Corollary 4

Thought-sequences that converge to limit points outside of conscious thought-space (X, T) of Lemma 1 might represent dreams. The sequence might eventually lie within an open set of a dream once the boundary between "awake" and "asleep" has been crossed.

Thom, in his catastrophe theory context, sees "a duality between thought and language reminiscent of that...between dreaming and play..." [Thom, 1975, p. 313]. A comprehensive analysis of the role of sequences that converge outside of the thought-space that contains them (as did the sequence 1, 1/2, 1/4,... under the topology of right-half open intervals) is, however, beyond the immediate extension of Theorem 1. It would appear to require additional technical knowledge of verious mental diseases and disorders. (E.g., do multiple personalities always have multiple limit points for even "simple" issues such as physical identity?)

DIRECTIONS FOR FURTHER INVESTIGATION

One of the assumptions of Theorem 1 required adequate translation of language using dictionaries and interpreters where necessary. To conform to Whorfian ideas of linguistic relativity which separates those languages which unify space-time words (such as Hopi) from those which do not, one would need to consider language-specific cases of Theorem 1 [Carroll, 1956].

Further, one might extend this Theorem to include modes of comprehension of sequences other than through language. Certainly insight is often gained through a touch or a glance [Arlinghaus, 1985].

Finally, one might consider, from an experimental vantage point, dualizing the relations within the human thought-space of Theorem 1 as those within an animal instinct-space. In the latter case, however, the issues of whether or not animals can understand parts of "sentences" and of whether or not they have more than one way to say the same thing need to be confronted [Bronowski, 1978, pp. 29, 37].

REFERENCES

Arlinghaus, S. L. Eye-contact graphs. Behavioral Science, 1985, 30, 108-117.

Atkin, R. H. Mathematical Structure in Human Affairs. New York: Crane-Russak, 1974.

Auslander, L. and MacKenzie, R. E. Introduction to Differentiable Manifolds. New York: Mc Graw-Hill, 1963.

Bourbaki, N. Elements of Mathematics: Theory of Sets. Paris: Hermann; Reading, Massachusetts: Addison-Wesley, 1968.

Bronowski, J. The Origins of Knowledge and Imagination. New Haven and London: Yale University Press, 1978.

Carroll, J. B. (ed.). Language, Thought, and Reality: Selected Writings of Benjamin Whorf. Cambridge, Massachusetts: M.I.T., The Technology Press; New York: Harcourt Brace Jovanovich, 1968.

Hayakawa, S. I. Language in Action: A Guide to Accurate Thinking. New York: Harcourt Brace and Company, 1941.

Kelley, J. L. General Topology. Princeton, New Jersey: D. Van Nostrand, 1955.

Lewin, K. Principles of Topological Psychology. New York: Mc Graw-Hill, 1936.

Mac Lane, S. "Proof, truth, and confusion." The 1982 Ryerson Lecture. Chicago: The University of Chicago, 1982.

Mansfield, M. Introduction to Topology. Princeton, New Jersey: D. Van Nostrand, 1963.

Quine, W. V. O. Word and Object. Cambridge, Massachusetts: M.I.T. Press, 1960.

Redfield, R. and Singer, M. "The cultural role of cities." Economic Development and Cultural Change, 1954, 3, #1.

Snow, C. P. The Two Cultures and a Second Look. Cambridge: The University Press, 1959.

Tarski, A. Logic, Semantics and Metamathematics. Trans. J. H. Woodger. Oxford: Clarendon Press, 1956.

Thom, R. Structural Stability and Morphogenesis: An Outline of a General Theory of Models. Trans. D. H. Fowler. Reading, Massachusetts: Benjamin/Cummings, 1975.

Wertheimer, M. Productive Thinking, Enlarged Edition. New York: Harper and Brothers, 1959.

CHAOS IN HUMAN SYSTEMS--THE HEINE-BOREL THEOREM

One theorem, based in topological notions, that is critical in the development of higher mathematical analysis and that might have potential for application to human systems at the global scale is the Heine-Borel Theorem. This theorem holds in all integral dimensions; for ease in visualization, it is realized here along the real number line. Extensions of this Theorem and of the applications suggested below to fractal sets appear promising, but are beyond the scope of this material [Arlinghaus, 1985].

Definition 1 [Taylor, p. 483].

"A point set S on the x-axis is said to be bounded if there is some finite interval which contains all of S; that is, if there exist numbers a and b, $a < b$, such that $a \leq x \leq b$ for all $x \in S$."

Definition 2 [Taylor, p. 491]

"Let S be a point set, and suppose we have a collection of a certain number of open sets such that each point of S belongs to at least one of the open sets. Then we say that S is covered by the collection of open sets."

For example, suppose $S = \{x | 0 < x \leq 1\}$ and suppose T is a collection of open sets (open intervals on the x-axis) with

$$T = \{I_n = (1/(2^n), (n+2)/(2^n)) | n \text{ is an integer}\}$$

[Taylor, p. 491]. Figure 1 displays the approximate positions of the open intervals on the set S. As n becomes large, $1/(2^n)$ becomes small, but it never reaches 0; the sets I_1, I_2, I_3, and I_4 cover everything in S to the right of 1/16. The infinite sequence of open intervals is required to cover $0 < x \leq 1/16$. This set S is covered by the collection T of open intervals, since every x, $0 < x \leq 1$, lies in I_n for some value of n. The set S is bounded, since $0 \leq x \leq 1$ (as in Definition 1); it is not, however, closed--its complement on the real axis is $(-\infty,0] \cup (1,\infty)$ which is not a union of open intervals.

Suppose we make S closed by adding the point 0; $R = S \cup \{0\}$. Does T, as described above, cover this closed and bounded interval R? Clearly it does not, as 0 lies in no I_n. To cover 0, add an interval such as (-1/10, 1/10). Then the infinite collection $U=T \cup \{(-1/10,1/10)\}$ covers R (Figure 2).

The addition of the point 0 to 'close' S, forcing the addition of (-1/10,1/10) to T in order to cover S, has deeper implications. The closed set R may be covered by a finite number of judiciously selected intervals from U; the intervals I_1, I_2, I_3, and I_4 cover everything in R to the right of 1/16, while the added interval, (-1/10,1/10), covers all of R, including 1/16, to the left of 1/10. Thus, five intervals may be used to replace an infinite collection in covering the set R.

Clearly, this is not the case in the situation shown in Figure 1. For, if the intervals $I_1,\ldots,I_k$ (k an arbitrary positive integer) were considered as a finite candidate-set, the interval $0 < x \leq 1/(2^k)$ would remain uncovered. Thus, no finite subset of the infinite collection of open intervals T will cover $S = \{x | 0 < x \leq 1\}$.

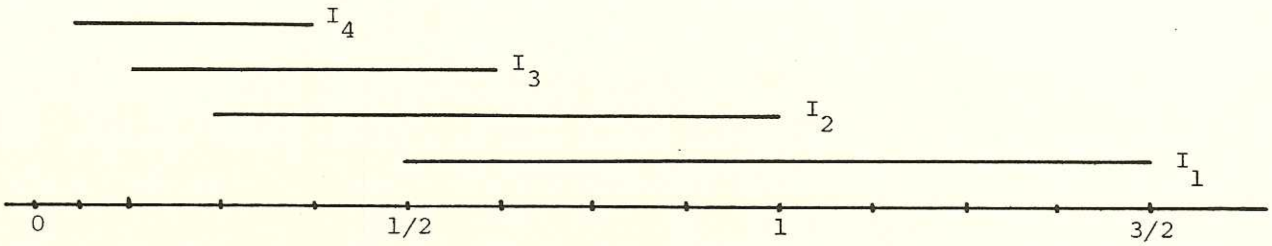
The notion of being able to select a finite subcover from a given covering is the thrust of the Heine-Borel Theorem.

<u>Heine-Borel Theorem</u> [Taylor, p. 493]

"Let S be a bounded and closed point set, and let S be covered by a collection [T] of open sets. Then a finite number of open sets may be chosen from the collection [T] in such a way that S is covered by the new finite collection."

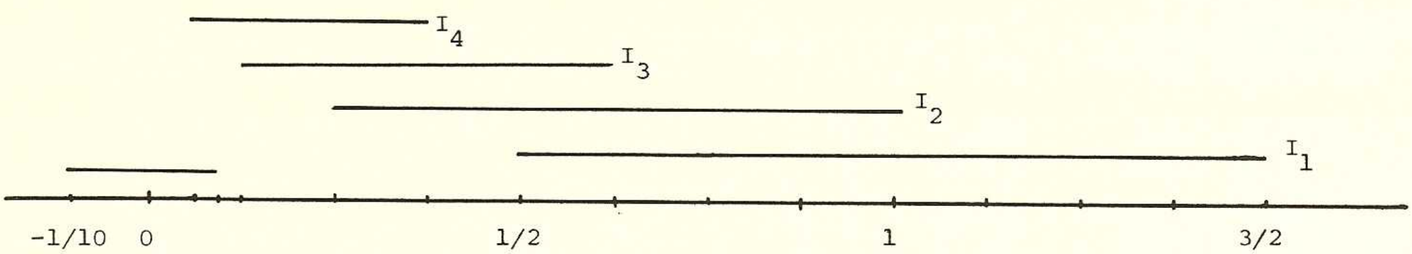
The examples above give the idea of the proof of this theorem; a rigorous proof may be found in Taylor or Rudin.

To apply this Theorem to human systems, suppose that S is a collection of human systems and that T is a collection of interpretations of those systems. Using the Heine-Borel Theorem, we see that if S is closed and bounded, and is covered by T, then from T a finite collection also covering S can be chosen. We might expect S to be bounded by a geographic region (possibly the whole earth), and S might be viewed as closed, if no new input were required from different systems to ensure the functioning of S. The collection T clearly could be finite; however, it rests on belief systems, value systems, and a variety of other social and cultural factors which might be infinite. If the Heine-Borel Theorem holds, a finite number of these views may be chosen that cover, or produce understanding of and rational response to, these systems. In this case, some uniformity in interpretation of the systems is possible in the geographic region containing them. When the number of elements in the finite cover is small, and when the geographic region bounding S is large, global harmony is maximized. Conversely, when the Heine-Borel Theorem does not hold, an infinite collection of interpretations may be required to cover, or understand and respond to, even a fairly small set of human systems (compare to the motivational examples). This suggests cultural chaos; empirical studies along these lines might be drawn from Middle Eastern politics, from English/Irish relations, or from widespread terrorism across the surface of the Earth.



(S,T): $I_1 = (1/2,3/2)$
$I_2 = (1/4,4/4)$
$I_3 = (1/8,5/8)$
$I_4 = (1/16,6/16)$
....

FIGURE 1



(R,U): $I_1 = (1/2,3/2)$
$I_2 = (1/4,4/4)$
$I_3 = (1/8,5/8)$
$I_4 = (1/16,6/16)$
....

and (-1/10,1/10)

FIGURE 2

LITERATURE CITED

1. Arlinghaus, S. L. "Fractals take a central place." Geografiska Annaler, B, 67, (1985), 83-88.

2. Rudin, W. Real and Complex Analysis. New York: McGraw-Hill, 1966.

3. Taylor, A. E. Advanced Calculus. Boston: Ginn and Co., 1955.

INSTITUTE OF MATHEMATICAL GEOGRAPHY (IMaGe)
2790 BRIARCLIFF
ANN ARBOR, MI 48105-1429; U.S.A.
(313) 761-1231; IMaGe@UMICHUM

"Imagination is more important than knowledge"
A. Einstein

MONOGRAPH SERIES

Exclusive of shipping and handling; prices listed and payable in U.S. funds on a U.S. bank, only.

1. Sandra L. Arlinghaus and John D. Nystuen. *Mathematical Geography and Global Art: the Mathematics of David Barr's "Four Corners Project,"* 1986.

This monograph contains Nystuen's calculations, actually used by Barr to position his abstract tetrahedral sculpture within the earth. Placement of the sculpture vertices in Easter Island, South Africa, Greenland, and Indonesia was chronicled in film by The Archives of American Art for The Smithsonian Institution. In addition to the archival material, this monograph also contains Arlinghaus's solutions to broader theoretical questions–was Barr's choice of a tetrahedron unique within his initial constraints, and, within the set of Platonic solids?

2. Sandra L. Arlinghaus. *Down the Mail Tubes: the Pressured Postal Era, 1853-1984*, 1986.

The history of the pneumatic post, in Europe and in the United States, is examined for the lessons it might offer to the technological scenes of the late twentieth century. As Sylvia L. Thrupp, Alice Freeman Palmer Professor Emeritus of History, The University of Michigan, commented in her review of this work "Such brief comment does far less than justice to the intelligence and the stimulating quality of the author's writing, or to the breadth of her reading. The detail of her accounts of the interest of American private enterprise, in New York and other large cities on this continent, in pushing for construction of large tubes in systems to be leased to the government, brings out contrast between American and European views of how the new technology should be managed. This and many other sections of the monograph will set readers on new tracks of thought."

3. Sandra L. Arlinghaus. *Essays on Mathematical Geography*, 1986.

A collection of essays intended to show the range of power in applying pure mathematics to human systems. There are two types of essay: those which employ traditional mathematical proof, and those which do not. As mathematical proof may itself be regarded as art, the former style of essay might represent "traditional" art, and the latter, "surrealist" art. Essay titles are: "The well-tempered map projection," "Antipodal graphs," "Analogue clocks," "Steiner transformations," "Concavity and urban settlement patterns," "Measuring the vertical city," "Fad and permanence in human systems," "Topological exploration in geography," "A space for thought," and "Chaos in human systems–the Heine-Borel Theorem."

4. Robert F. Austin, *A Historical Gazetteer of Southeast Asia*, 1986.

Dr. Austin's Gazetteer draws geographic coordinates of Southeast Asian place-names together with references to these place-names as they have appeared in historical and literary documents. This book is of obvious use to historians and to historical geographers specializing in Southeast Asia. At a deeper level, it might serve as a valuable source in establishing place-name linkages which have remained previously unnoticed, in documents describing trade or other communications connections, because of variation in place-name nomenclature.

5. Sandra L. Arlinghaus, *Essays on Mathematical Geography–II*, 1987.

Written in the same format as IMaGe Monograph #3, that seeks to use "pure" mathematics in real-world settings, this volume contains the following material: "Frontispiece–the Atlantic Drainage Tree," "Getting a Handel on Water-Graphs," "Terror in Transit: A Graph Theoretic Approach to the Passive Defense of Urban Networks," "Terrae Antipodum," "Urban Inversion," "Fractals: Constructions, Speculations, and Concepts," "Solar Woks," "A Pneumatic Postal Plan: The Chambered Interchange and ZIPPR Code," "Endpiece."

6. Pierre Hanjoul, Hubert Beguin, and Jean-Claude Thill, *Theoretical Market Areas Under Euclidean Distance*, 1988. (English language text; Abstracts written in French and in English.)

Though already initiated by Rau in 1841, the economic theory of the shape of two-dimensional market areas has long remained concerned with a representation of transportation costs as linear in distance. In the general gravity model, to which the theory also applies, this corresponds to a decreasing exponential function of distance deterrence. Other transportation cost and distance deterrence functions also appear in the literature, however. They have not always been considered from the viewpoint of the shape of the market areas they generate, and their disparity asks the question whether other types of functions would not be worth being investigated. There is thus a need for a general theory of market areas: the present work aims at filling this gap, in the case of a duopoly competing inside the Euclidean plane endowed with Euclidean distance.

(Bien qu'ébauchée par Rau dès 1841, la théorie économique de la forme des aires de marché planaires s'est longtemps contentée de l'hypothèse de coûts de transport proportionnels à la distance. Dans le modèle gravitaire généralisé, auquel on peut étendre cette théorie, ceci correspond au choix d'une exponentielle décroissante comme fonction de dissuasion de la distance. D'autres fonctions de coût de transport ou de dissuasion de la distance apparaissent cependant dans la littérature. La forme des aires de marché qu'elles engendrent n'a pas toujours été étudiée ; par ailleurs, leur variété amène à se demander si d'autres fonctions encore ne mériteraient pas d'être examinées. Il paraît donc utile de disposer d'une théorie générale des aires de marché : ce à quoi s'attache ce travail en cas de duopole, dans le cadre du plan euclidien muni d'une distance euclidienne.)

7. Keith J. Tinkler, Editor, *Nystuen—Dacey Nodal Analysis*, 1988.

Professor Tinkler's volume displays the use of this graph theoretical tool in geography, from the original Nystuen—Dacey article, to a bibliography of uses, to original uses by Tinkler. Some reprinted material is included, but by far the larger part is of previously unpublished material. (Unless otherwise noted, all items listed below are previously unpublished.) Contents: " 'Foreward' " by Nystuen, 1988; "Preface" by Tinkler, 1988; "Statistics for Nystuen—Dacey Nodal Analysis," by Tinkler, 1979; Review of Nodal Analysis literature by Tinkler (pre–1979, reprinted with permission; post—1979, new as of 1988); FORTRAN program listing for Nodal Analysis by Tinkler; "A graph theory interpretation of nodal regions" by John D. Nystuen and Michael F. Dacey, reprinted with permission, 1961; Nystuen—Dacey data concerning telephone flows in Washington and Missouri, 1958, 1959 with comment by Nystuen, 1988; "The expected distribution of nodality in random (p, q) graphs and multigraphs," by Tinkler, 1976.

8. James W. Fonseca, *The Urban Rank–size Hierarchy: A Mathematical Interpretation*, 1989.

The urban rank–size hierarchy can be characterized as an equiangular spiral of the form $r = ae^{\theta \cot \alpha}$. An equiangular spiral can also be constructed from a Fibonacci sequence. The urban rank–size hierarchy is thus shown to mirror the properties derived from Fibonacci characteristics such as rank–additive properties. A new method of structuring the urban rank–size hierarchy is explored which essentially parallels that of the traditional rank–size hierarchy below rank 11. Above rank 11 this method may help explain the frequently noted concavity of the rank–size distribution at the upper levels. The research suggests that the simple rank–size rule with the exponent equal to 1 is not merely a special case, but rather a theoretically justified norm against which deviant cases may be measured. The spiral distribution model allows conceptualization of a new view of the urban rank–size hierarchy in which the three largest cities share functions in a Fibonacci hierarchy.

9. Sandra L. Arlinghaus, *An Atlas of Steiner Networks*, 1989.

A Steiner network is a tree of minimum total length joining a prescribed, finite, number of locations; often new locations are introduced into the prescribed set to determine the minimum tree. This Atlas explains the mathematical detail behind the Steiner construction for prescribed sets of n locations and displays the steps, visually, in a series of Figures. The proof of the Steiner construction is by mathematical induction, and enough steps in the early part of the induction are displayed completely that the reader who is well–trained in Euclidean geometry, and familiar with concepts from graph theory and elementary number theory, should be able to replicate the constructions for full as well as for degenerate Steiner trees.

10. Daniel A. Griffith, *Simulating $K = 3$ Christaller Central Place Structures: An Algorithm Using A Constant Elasticity of Substitution Consumption Function*, 1989.

An algorithm is presented that uses BASICA or GWBASIC on IBM compatible machines. This algorithm simulates Christaller $K = 3$ central place structures, for a four–level hierarchy. It is based upon earlier published work by the author. A description of the spatial theory, mathematics, and sample output runs appears in the monograph. A digital version is available from the author, free of charge, upon request; this request must be accompanied by a 5.5–inch formatted diskette. This algorithm has been developed for use in Social Science classroom laboratory situations, and is designed to (a) cultivate a deeper understanding of central place theory, (b) allow parameters of a central place system to be altered and then graphic and tabular results attributable to these changes viewed, without experiencing the tedium of massive calculations, and (c) help promote a better comprehension of the complex role distance plays in the space–economy. The algorithm also should facilitate intensive numerical research on central place structures; it is expected that even the sample simulation results will reveal interesting insights into abstract central place theory.

The background spatial theory concerns demand and competition in the space–economy; both linear and non–linear spatial demand functions are discussed. The mathematics is concerned with (a) integration of non–linear spatial demand cones on a continuous demand surface, using a constant elasticity of substitution consumption function, (b) solving for roots of polynomials, (c) numerical approximations to integration and root extraction, and (d) multinomial discriminant function classification of commodities into central place hierarchy levels. Sample output is presented for contrived data sets, constructed from artificial and empirical information, with the wide range of all possible central place structures being generated. These examples should facilitate implementation testing. Students are able to vary single or multiple parameters of the problem, permitting a study of how certain changes manifest themselves within the context of a theoretical central place structure. Hierarchical classification criteria may be changed, demand elasticities may or may not vary and can take on a wide range of non–negative values, the uniform transport cost may be set at any positive level, assorted fixed costs and variable costs may be introduced, again within a rich range of non–negative possibilities, and the number of commodities can be altered. Directions for algorithm execution are summarized. An ASCII version of the algorithm, written directly from GWBASIC, is included in an appendix; hence, it is free of typing errors.

11. Sandra L. Arlinghaus and John D. Nystuen, *Environmental Effects on Bus Durability*, 1990.

This monograph draws on the authors' previous publications on "Climatic" and "Terrain" effects on bus durability. Material on these two topics is selected, and reprinted, from three published papers that appeared in the *Transportation Research Record* and in the *Geographical Review*. New material concerning "congestion" effects is examined at the national level, to determine "dense," "intermediate," and "sparse" classes of congestion, and at the local level of congestion in Ann Arbor (as suggestive of how one might use local data). This material is drawn together in a single volume, along with a summary of the consequences of all three effects simultaneously, in order to suggest direction for more highly automated studies that should follow naturally with the release of the 1990 U. S. Census data.

12. Daniel A. Griffith, Editor. *Spatial Statistics: Past, Present, and Future*, 1990.

Proceedings of a Symposium of the same name held at Syracuse University in Summer, 1989. Content includes a Preface by Griffith and the following papers:

Brian Ripley, "Gibbsian interaction models";
J. Keith Ord, "Statistical methods for point pattern data";
Luc Anselin, "What is special about spatial data";
Robert P. Haining, "Models in human geography:
problems in specifying, estimating, and validating models for spatial data";
R. J. Martin, "The role of spatial statistics in geographic modelling";
Daniel Wartenberg, "Exploratory spatial analyses: outliers, leverage points, and influence functions";
J. H. P. Paelinck, "Some new estimators in spatial econometrics";
Daniel A. Griffith, "A numerical simplification for estimating parameters of spatial autoregressive models";
Kanti V. Mardia "Maximum likelihood estimation for spatial models";
Ashish Sen, "Distribution of spatial correlation statistics";

Sylvia Richardson, "Some remarks on the testing of association between spatial processes";
Graham J. G. Upton, "Information from regional data";
Patrick Doreian, "Network autocorrelation models: problems and prospects."

Each chapter is preceded by an "Editor's Preface" and followed by a Discussion and, in some cases, by an author's Rejoinder to the Discussion.

INSTITUTE OF MATHEMATICAL GEOGRAPHY
2790 Briarcliff
Ann Arbor, MI 48105
U.S.A.

IMaGe MONOGRAPH SERIES–INFORMATION VALID TO 1/1/89

UNIVERSITY LIBRARY STATEMENT

IMaGe Monographs are produced on demand from camera-ready copy, supplied by the author. Guidelines for authors, and a statement of IMaGe review policy appear on a different flier. Price structure is negotiated with Michigan Document Services, Inc., on a regular basis, reflecting changing costs in producing books. Fluctuations might be up or down–up, for example, if equipment maintenance costs rise; down, for example, if the factory location is moved to a cheaper site. Because IMaGe is responsive to the concerns of Acquisitions Departments in university libraries, it has established the following guidelines concerning the frequency of publication of monographs.

1. No single book shall be of a length forcing its cost to exceed $20.00, exclusive of postage and handling.

2. No more than six books shall be issued in a calendar year. If a single monograph number is of a length requiring more than one book for publication, then each of those books counts as one of the six. Therefore, the maximum amount of a basic annual subscription is $120.00.

3. Libraries should specify, at the time the order is placed, the preferred mode of shipping; if left unspecified, books will be shipped via Fourth Class mail, or Library Rate (whatever applies), in the continental United States, and via Air Mail (using a mail company to ascertain the the cheapest mode of transmission) to overseas and foreign libraries.

INSTITUTE OF MATHEMATICAL GEOGRAPHY
2790 Briarcliff
Ann Arbor, MI 48105
U.S.A.

IMaGe MONOGRAPH SERIES–ADVICE TO AUTHORS

ADVICE TO AUTHORS

Manuscripts should be submitted directly to the editor. The monographs in this series are produced on demand from camera-ready copy supplied by the author. All manuscripts will be reviewed by at least two members of the editorial board or by referees suggested by them. Initial submissions should be typed, double-spaced, with one inch margins on all sides. Submit three copies, at least one of which should be copied on one side only. Figures should be in close-to-final form. Once a manuscript has been accepted, the author will be provided with further suggestions for the preparation of the final document. Authors should retain copies of all submissions as the high cost of postage prohibits the return of manuscripts. Royalty checks will be issued periodically, depending on sales. There is no charge to the author for producing these books. Before a book is released, authors will be provided with a transfer of copyright form, and copyright will be procured for each volume in the series by IMaGe.

PURCHASE OF MONOGRAPHS

Monographs may be ordered directly from IMaGe; the price of a monograph does not include postage and handling. A list of monograph titles and abstracts appears on the last page of each book. Because these books are published on demand, they do not go out of print. Standing orders, as well as individual orders, are welcome; the former well be filled ahead of the latter.

EDITORIAL BOARD

INSTITUTE OF MATHEMATICAL GEOGRAPHY
2790 Briarcliff
Ann Arbor, MI 48105
U.S.A.

MICMG DISCUSSION PAPERS, JOHN D. NYSTUEN, EDITOR

The collected work of the Michigan Interuniversity Community of Mathematical Geographers is reprinted in a single essay format (as they were originally reproduced). Royalties from sales will be deposited in the "MICMG" fund for the development of projects in mathematical geography, administered by IMaGe, and disposed of by Nystuen, in collaboration with IMaGe.

Consider ordering one number as reading supplementary to texts in an upper division course. The dates of original release and titles of the individual numbers are listed below. Prices are current as of 1988.

1. Arthur Getis, "Temporal land use pattern analysis with the use of nearest neighbor and quadrat methods." [Re-typed and maps re-drawn.] July, 1963. $1.50

2. Marc Anderson, "A working bibliography of mathematical geography." September, 1963. $3.50

3. William Bunge, "Patterns of location." February, 1964. $2.50

4. Michael F. Dacey, "Imperfections in the uniform plane." June, 1964. $2.50

5. Robert S. Yuill, "A simulation study of barrier effects in spatial diffusion problems." April, 1965. $3.00

6. William Warntz, "A note on surfaces and paths and applications to geographical problems." May, 1965. $2.00

7. Stig Nordbeck, "The law of allometric growth." June, 1965. $2.50

8. Waldo R. Tobler, "Numerical map generalization;" and "Notes on the analysis of geographical distributions." $3.00

9. Peter R. Gould, "On mental maps." September, 1966. $3.50

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. $4.50

11. E. Casetti and R. K. Semple, "A method for the stepwise separation of spatial trends." April, 1968. $2.50

12. W. Bunge, R. Guyot, A. Karlin, R. Martin, W. Pattison, W. Tobler, S. Toulmin, and W. Warntz, "The philosophy of maps." June, 1968. $4.50