Construction Zone

Simple analysis of the logistic function

A derivation supplied by S. Arlinghaus in response to questions from William D. Drake, School of Natural Resources, University of Michigan, concerning aspects of his interest in transition theory. Discussed Tuesday, May 6, 1991, Colloquium in Mathematical Geography, IMaGe. Present: Sandy Arlinghaus, Bill Drake, John Nystuen (this commentary is included in *Solstice* at the request of the latter).

1. The exponential function–unbounded population growth

Assumption: The rate of population growth or decay at any time t is proportional to the size of the population at t.

Let Y_t represent the size of a population at time t. The rate of growth of Y_t is proportional to Y_t;

$$dY_t/dt = kY_t$$

where k is a constant of proportionality.

To solve this differential equation for Y_t, separate the variables.

$$dY_t/Y_t = k\,dt; \int 1/Y_t\,dY_t = \int k\,dt.$$

Therefore,

$$\ln|Y_t| = kt + c_0.$$

Consider only the positive part, so that

$$Y_t = e^{kt+c_0} = e^{c_0}e^{kt}.$$

Let $Y_{t0} = e^{c_0}$. Therefore,

$$Y_t = Y_{t0}\,e^{kt};$$

exponential growth is unbounded as $t \longrightarrow \infty$.

Suppose $t = 0$. Therefore,

$$Y_t = Y_{t0}\,e^0 = Y_{t0}.$$

Thus, Y_{t0} is the size of the population at $t = 0$, under conditions of growth where $k > 0$ (Figure 1).

2. The logistic function–bounded population growth.

Assumption appended to assumption for exponential growth. In reality, when the population gets large, envirnomental factors dampen growth.

The growth rate decreases– dY_t/dt decreases. So, assume the population size is limited to some maximum, q, where $0 < Y_t < q$. As $Y_t \longrightarrow q$ it follows that $dY_t/dt \longrightarrow 0$ so that population size tends to be stable as $t \longrightarrow \infty$. The model is exponential in shape initially and includes effects of environmental resistance in larger populations. One such algebraic expression of this idea is

$$dy_t/dt = kY_t \cdot \frac{q - Y_t}{q}$$

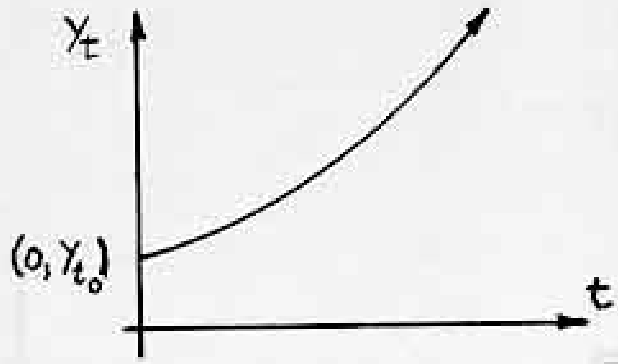


Figure 1

because in the factor $\frac{q-Y_t}{q}$, when Y_t is small, $\frac{q-Y_t}{q}$ is close to 1 (and the growth therefore close to the exponential) and when Y_t is close to q, $\frac{q-Y_t}{q}$ is close to 0, and the growth rate dY_t/dt tapers off. This factor acts as a damper to exponential growth.

Replace k/q by K so that

$$dY_t/dt = KY_t(q - Y_t)$$

and the rate of growth is proportional to the product of the population size and the difference between the maximum size and the population size.

Solve this latter equation for Y_t: (separate the variables)

$$\frac{dY_t}{Y_t(q - Y_t)} = K\,dt; \int \frac{dY_t}{Y_t(q - Y_t)} = \int K\,dt$$

Use a Table of Integrals on the rational form in the left-hand integral:

$$\frac{1}{q}\ln\left|\frac{Y_t}{q - Y_t}\right| = Kt + C$$

$$\ln\left|\frac{Y_t}{q - Y_t}\right| = qKt + qC.$$

Because $Y_t > 0$ and $q - Y_t > 0$,

$$\ln\frac{Y_t}{q - Y_t} = qKt + qC.$$

Therefore,

$$\frac{Y_t}{q - Y_t} = e^{qKt+qC} = e^{qKt}e^{qC}.$$

Replace e^{qC} by A. Therefore,

$$\frac{Y_t}{q - Y_t} = Ae^{qKt};$$

$$Y_t = (q - Y_t)Ae^{qKt};$$

$$Y_t = qAe^{qKt} - Y_tAe^{qKt};$$

$$Y_t(Ae^{qKt} + 1) = qAe^{qKt};$$

$$Y_t = \frac{qAe^{qKt}}{Ae^{qKt} + 1};$$

now divide top and bottom by Ae^{qKt}, equivalent to multiplying the fraction by 1, so that

$$Y_t = \frac{q}{1 + \frac{1}{Ae^{qKt}}} = \frac{q}{1 + \frac{1}{A}e^{-qKt}}.$$

Replace $1/A$ by a and $-qK$ by b producing a common form for the logistic function (Figure 2),

$$Y_t = \frac{q}{1 + ae^{bt}}$$

with $b < 0$ because $b = -qK$, and $q,\ K > 0$.

3. Facts about the graph of the logistic equation.

a. The line $Y_t = q$ is a horizontal asymptote for the graph.

This is so because, for $b < 0$,

$$\lim_{t \to \infty} \frac{q}{1 + ae^{bt}} \longrightarrow \frac{q}{1 + a(0)} = q$$

Can the curve cross this asymptote? Or, can it be that

$$Y_t = \frac{Y_t}{1 + ae^{bt}}?$$

Or,

$$1 = 1 + ae^{bt}?$$

Or,

$$ae^{bt} = 0$$

Or, that $a = 0$? No, because $a = 1/A$. Or, that $e^{bt} = 0$–no.

Thus, the logistic growth curve described above cannot cross the horizontal asymptote so that it approaches it entirely from one side, in this case, from below.

b. Find the coordinates of the inflection point of the logistic curve.

Vertical component:

The equation $dY_t/dt = KY_t(q - Y_t) = KqY_t - KY_t^2$ is a measure of population growth. Find the maximum rate of growth–derivative of previous equation:

$$d^2Y_t/dt^2 = Kq - 2KY_t$$

To find a maximum (min), set this last equation equal to zero.

$$Kq - 2KY_t = 0$$

Therefore, $Y_t = q/2$. This is the vertical coordinate of the inflection point of the curve for Y_t, the logistic curve–dY_t/dt is increasing to the left of $q/2$ ($d^2Y_t/dt > 0$) and dY_t/dt is decreasing to the right of $q/2$ ($d^2Y_t/dt < 0$). So, the maximum rate of growth occurs at $Y_t = q/2$. [The rate at which the rate of growth is changing is a constant since the first differential equation is a quadratic (parabola)].

Horizontal component:

To find t, put $Y_t = q/2$ in the logistic equation and solve:

$$q/2 = \frac{q}{1 + ae^{bt}}$$

Solving,

$$1 + ae^{bt} = 2; e^{bt} = 1/a; e^{-bt} = a; -bt = \ln a,$$

$$t = \frac{\ln a}{-b}$$

Thus, the coordinates of the inflection point of the logistic curve are:

$$(\ln a/(-b), q/2).$$

In order to track changes in transitions, such as demographic transitions, monitoring the position of the inflection point might be of use. To consider feedback in such systems, graphical analysis (Figure 2) of curves representing transitions might be of use.

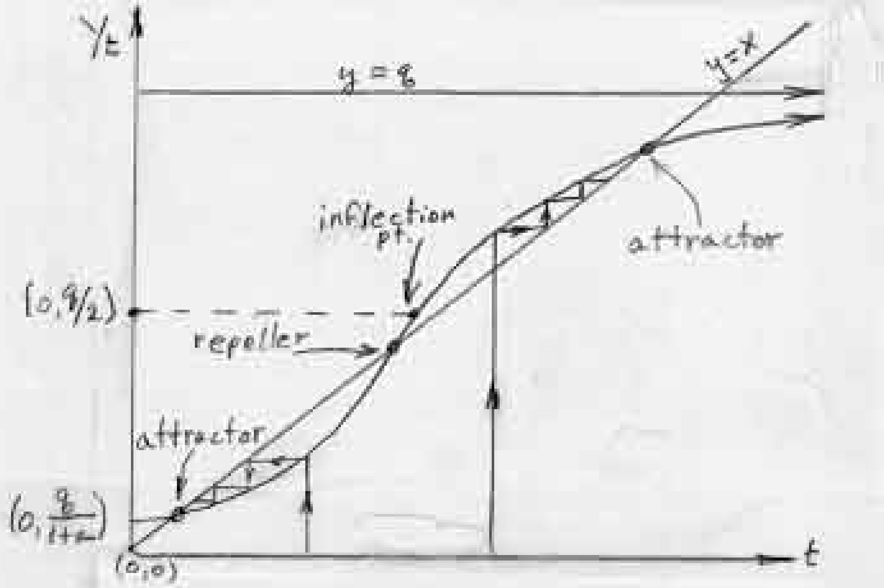


Figure 2. The intersection points of the line $y = x$ with the logistic curve are, using terms from chaos theory, attractors on either end, and a repelling fixed point in the middle, possibly near the inflection point of the curve.

Educational Feature
Topics in Spatial Theory
Based on lectures given by S. Arlinghaus
as a guest speaker in John Nystuen's
Urban Planning, 507, University of Michigan
Feb. 21, 28, 1990; four hours

The people along the sand
All turn and look one way.
They turn their back on the land.
They look at the sea all day.

...

They cannot look out far.
They cannot look in deep.
But when was that ever a bar
To any watch they keep?

Robert Frost *Neither Out Far Nor In Deep*

I. Introduction

Theory guides the direction technology takes; mathematics is the theoretical foundation of technology. To become more than a mere user of various software packages and programming languages, which change rapidly (what is trendy in today's job market may be obsolete tomorrow), it is therefore critical to understand what sorts of decisions can be made at the theoretical level. Underlying theory is "spatial" in character, rather than "temporal," when the objects and processes it deals with are ordered in space rather than in time (most can be done in both–decide which is of greater interest). The focus with GIS is spatial; hence, the theory underlying it is "spatial."

This is not a new idea; D'Arcy Thompson, a biologist, saw (as early as 1917) a need for finding a systematic, theoretical organization of biological species that went beyond the classification of Linnaeus. What he found to be fundamental, to characterization along structural (spatial, morphological) lines (rather than along temporal, evolutionary lines) was the "Theory of Transformations"–in Thompson's words:

"In a very large part of morphology, our essential task lies in the comparison of related forms rather than in the precise definition of each; and the deformation of a complicated figure may be a phenomenon easy of comprehension, though the figure itself have to be left unanalysed and undefined. This process of comparison, of recognising in one form a definite permutation or deformation of another, apart altogether from a precise and adequate understanding of the original 'type' or standard of comparison, lies within the immediate province of mathematics, and finds its solution in the elementary use of certain method of the mathematician. This method is the Method of Co-ordinates, on which is based the Theory of Transformations. [*The mathematical Theory of Transformations is part of the Theory of Groups, of great importance in modern mathematics. A distinction is drawn between Substitution-groups and Transformation-groups, the former being discontinuous, the latter continuous–in such a way that within one and the same group each transformation is infinitely little different from another. The distinction among biologists between a mutation and a variation is curiously analogous.]*

I imagine that when Descartes conceived the method of co-ordinates, as a generalisation from the proportional diagrams of the artist and the architect, and long before the immense possibilities of this analysis could be foreseen, he had in mind a very simple purpose; it was perhaps no more than to find a way of translating the form of a curve (as well as the position of a point) into numbers and into words. This is precisely what we do, by the method of coordinates, every time we study a statistical curve; and conversely translate numbers into form whenever we 'plot a curve', to illustrate a table or mortality, a rate of growth, or the daily variation of temperature or barometric pressure. In precisely the same way it is possible to inscribe in a net of rectangular co-ordinates the outline, for instance, of a fish, and so to translate it into a table of numbers, from which again we may at pleasure reconstruct the curve.

But it is the next step in the employment of co-ordinates which is of special interest and use to the morphologist; and this step consists in the alteration, or deformation, of our system of co-ordinates, and in the study of the corresponding transformation of the curve or figure inscribed in the co-ordinate network.

Let us inscribe in a system of Cartesian co-ordinates the outline of an organism, however complicated, or a part thereof: such as a fish, a crab, or a mammalian skull. We may now treat this complicated figure, in general terms, as a function of x , y *. If we submit our rectangular system to deformation on simple and recognised lines, altering, for instance, the direction of the axes, the ratio of* x/y *, or substituting for* x *and* y *some more complicated expressions, the we obtain a new system of co-ordinates, whose deformation from the original type the inscribed figure will precisely follow. In other words, we obtain a new figure which represents to old figure under a more or less homogeneous strain, and is a function of the new co-ordinates in precisely the same way as the old figure was of the original co-ordinates* x *and* y *.*

The problem is closely akin to that of the cartographer who transfers identical data to one projection or another [reference below]; and whose object is to secure (if it be possible) a complete correspondence, in each small unit of area, between the one representation and the other. The morphologist will not seek to draw his orgainc forms in a new and artificial projection; but, in the converse aspect of the problem, he will enquire whether two different but more or less obviously related forms can be so analysed and interpreted that each may be shown to be a transformed representation of the other. This once demonstrated, it will be a comparatively easy task (in all probability) to postulate the direction and magnitude of the force capable of effecting the required transformation. Again, if such a simple alteration of the system of forces can be proved adequate to meet the case, we may find ourselves able to dispense with many widely current and more complicated hypotheses of biological causation. For it is a maxim in physics that an effect ought not to be ascribed to the joint operation of many causes if few are adequate to the production of it.

Reference: Tissot, Mémoire sur la representation des surfaces, et les projections des cartes géographiques (Paris, 1881)."

Sir D'Arcy Wentworth Thompson, pp. 271-272, in *On Growth and Form.*

Look at Thompson's comments concerning biological structure to see what parallels there are, already, with GIS structure and to see what they might suggest–compare to Tobler's map transformations.

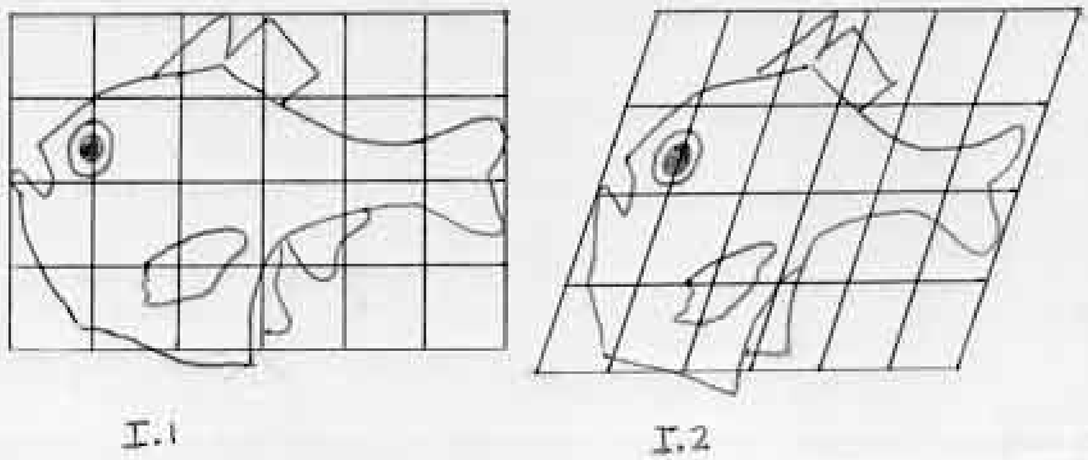


Figure I Sample of Thompson's Transformations. Fig. I.1: Argyropelecus olfersi. Fig. I.2: Sternoptyx diaphana.

1. GIS (the digitizer) uses coordinates to translate forms (maps) into numbers.

2. All GIS software translates numbers into maps, which may then be printed out, parallel to inscribing a fish in a set of coordinates, translating it into a set of numbers, from which the fish may be reproduced at any time (Figure I.1)

3. Thompson's deformations correspond to the ideas of scale shifts on maps. Transformations describe shifts in scale. Figure I.2.

4. Thompson's comments on the distinction between discontinuous and continuous reflects partitioning of mathematics into discrete and continuous. Discrete need not be finite–look at two different types of garbage bag ties–twist ties and slip-through ties, and imagine them to be of infinite extent.

5. We see simple transformations in GIS–maps might be stretched or compressed in the vertical direction. Imagine using a small digitizing table to encode a large map by deliberately recording "wrong" positions—then use a transformation within the computer to correct the "wrong" positions so that the map prints out correctly on the plotter. Large digitizing tables become unnecessary.

6. We look, for future direction, to the Theory of Groups. For today, we confine ourselves to a few simple transformations.

II. Transformations

Transformations can allow you to relate one form to another in a systematic manner allowing retieval of all forms. To do this, you need to know how to define a transformation so that this is possible. Beyond this, one might consider a stripped-down transformation, for even more efficient compression of electronic effort [Mac Lane].

A. Well-defined (single-valued).

Let "tau" be a transformation carrying a set X to a set Y: in notation, $\tau : X \longrightarrow Y$.

Tau is said to be well-defined if each element of X corresponds to exactly one element of Y. Visually, this might be thought of in terms of lists of street addresses: the set X consists of house addresses used as "return" addresses on letters. The set Y consists of other house addresses. The transformation is the postal transmission of a letter from locations in X to locations in Y. A single value of X maps to single value of Y.

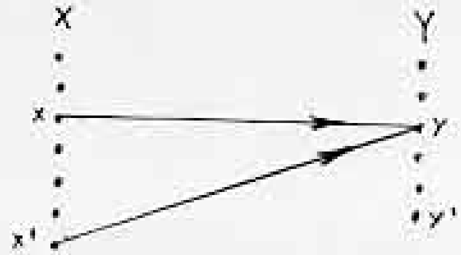


Figure II.1 This is a transformation–two distinct letters (x and x') can be posted to the same address (y). (Many-one map).

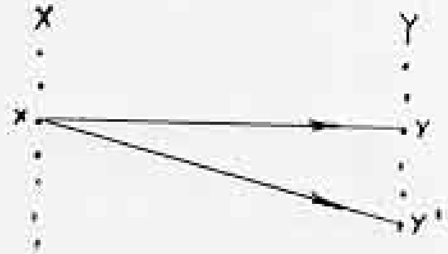


Figure II.2 This is NOT a transformation–one letter (x) cannot, itself, go to two different addresses (y and y') (new technology of e-mail permits this–suggests for possible need for change in fundamental definitions). (One-many map).

B. Reversible

i. One-to-one correspondence.

A one-to-one correspondence is a transformation in which each x in X goes to a distinct y in Y; the situation depicted in Figure II.1 cannot hold. From the standpoint of reversibility, this is important; if the situation in II.1 could hold how would you decide, in reversing, whether to "return" y to x or to x'??

ii. Transformations of X onto Y

A transformation of X onto Y is such that every element in Y comes from some element of X; there are no addresses outside the postal system (Figure II.3).

Figure II.3 This is a transformation–it is neither one-to-one, nor onto (y' is outside the system).

iii. A transformation τ from X to Y is reversible– it has an inverse τ^{-1} from Y to X if τ is one-to-one and onto; it has an inverse from a subset of Y to X if τ is one-to-one (Figure II.4).

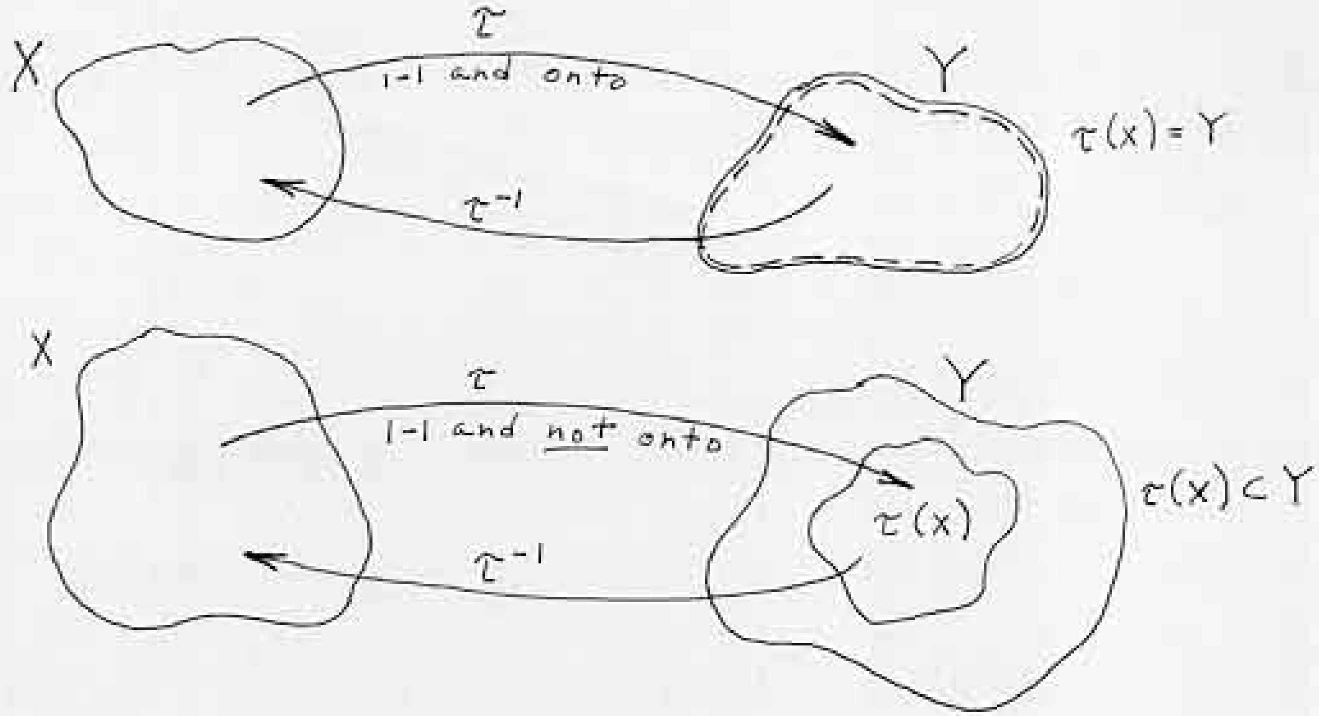


Figure II.4 In the top part, $\tau(X) = Y$. In the bottom part $\tau(X)$ is properly contained in Y; this is like data compression–like ZIP followed by UNZIP.

C. Rubbersheeting

The use of transformations that have inverses is critical in rubbersheeting; associations between data sets must be made in a manner so that correct information can be gained from the process.

III. Types of Transformations

One might consider moving objects within a fixed coordinate system, or holding the objects fixed and moving the coordinate system. Thompson did the latter; rubbersheeting does the latter; NCGIA materials (Lecture 28) comment that the latter approach is particularly well-suited to GIS purposes.

Two major types of transformations:

a. Affine transformations: these are transformations under which parallel lines are preserved as parallel lines. That is, both the concept of "straight line" and "parallel" remain; angles may change, however.

There are four types of affine transformations as noted on suitable NCGIA handout (Figure III.1). Products of affine transformations are themselves affine transformations.

Current technology employs types 1 and 2, quite clearly. CRT allows for translation of maps, and for scale change in y-direction only. Copier also allows for the same, and in addition, permits different shifts in scale along the two axes, allowing maps with different scales along different axes to be brought to the same scale and pieced together. (See output from Canon Color Copier.) On that output, the x-axis if fixed by the transformation and the y-axis is stretched to 200% of the original. Thus, a circle transforms to an ellipse, a rectangle with base parallel to the x-axis transforms to a larger rectangle, and a rectangle with base not parallel to the x-axis transforms to a parallelogram with no right angles (Figure III.2).

B. Curvilinear transformations; neither straightness nor parallelism is necessarily preserved (Thompson fish, Figure III.3).

IV. Exercise, page 5, lecture 28, NCGIA.

V. Steiner networks

If centers of gravity are used as a centering scheme in a triangulated irregular network, then it is desired to have no centroid lie outside a triangular cell. Thus, no cell should have angle greater than 120 degrees, so that the Steiner network (where all angles are exactly 120 degrees) will serve as an outer edge (a limiting position) for the set of acceptable triangulations. Thus, it is important to know how to locate Steiner networks.

VI. Digital Topology

The notion of a "triangulation" is a fundamental concept in topology (sometimes called "rubber sheet" geometry). "Digital" topology is a specialization of "combinatorial" topology in which the fundamental units are pixels. The same "important" theorems underlie each. The Jordan Curve Theorem (which characterizes the difference between the "inside" and the "outside" of a curve, is an example of such a theorem). Using concepts from digital topology, "picture" processing (as a parallel to "data" processing) is possible. There are numerous references in this field; some include works by geographer Waldo Tobler and by mathematician Azriel Rosenfeld. Other key-words to topics of interest in this area include,

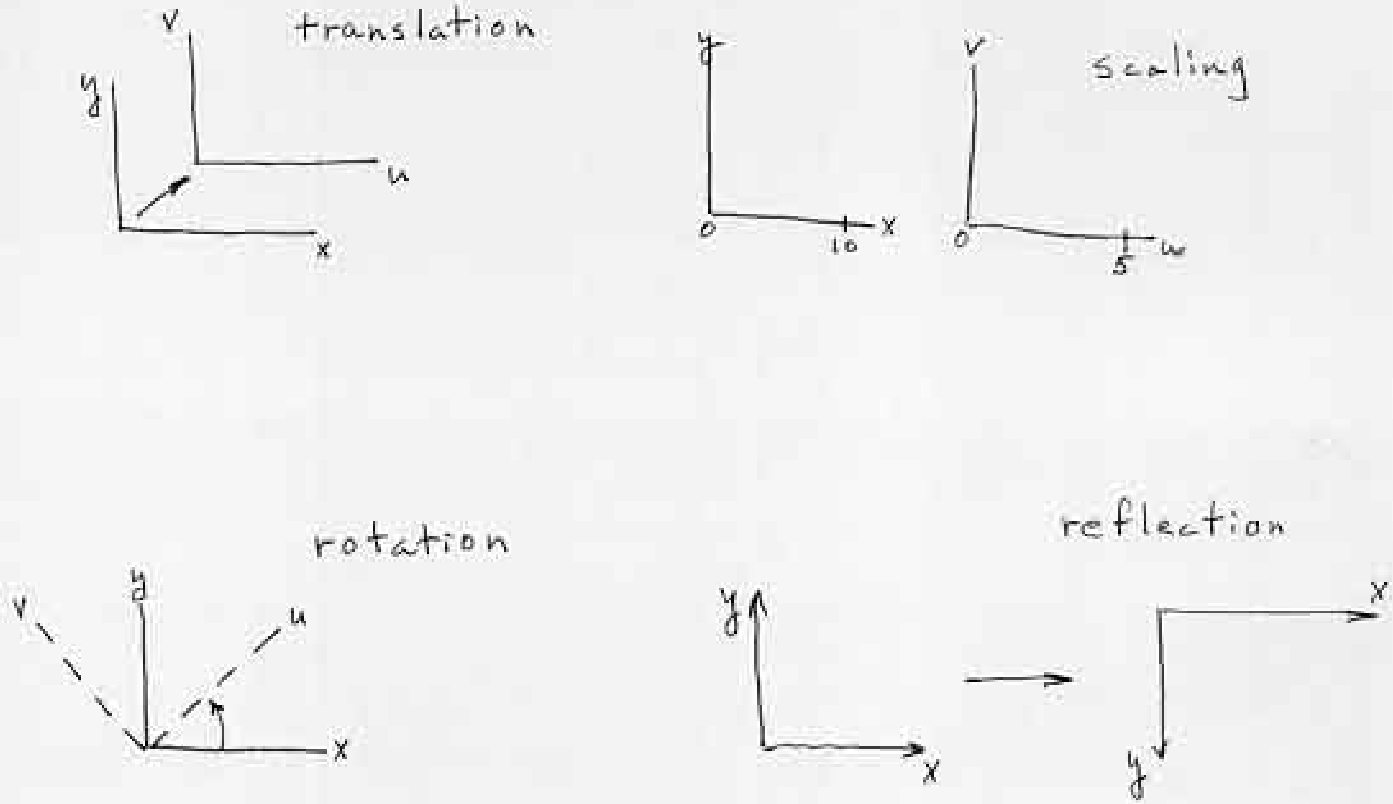


Figure III.1

Jordan Curve Theorem in higher dimensions; quadtrees; scale-free transformations; close-packings of pixels.

VII. The algebra of symmetry–some group theory

D'Arcy Thompson commented that the theory of transformations was tied to the theory of groups. A "group" is a mathematical system whose structure is simpler than that of the number system we customarily use in the "real-world." In our usual number system, we have two distinct operations of "+" and "x"; thus, we have rules on how to use each of these operations, and rules telling us how to link these two operations (distributive law; conventions regarding order of operations).

A group is composed of a finite set of elements, $S = \{a, b, c, \ldots, n\}$ that are related to each other using a single operation of "$*$." Under this operation, the set obeys the following rules (and is, by that fact, a group).

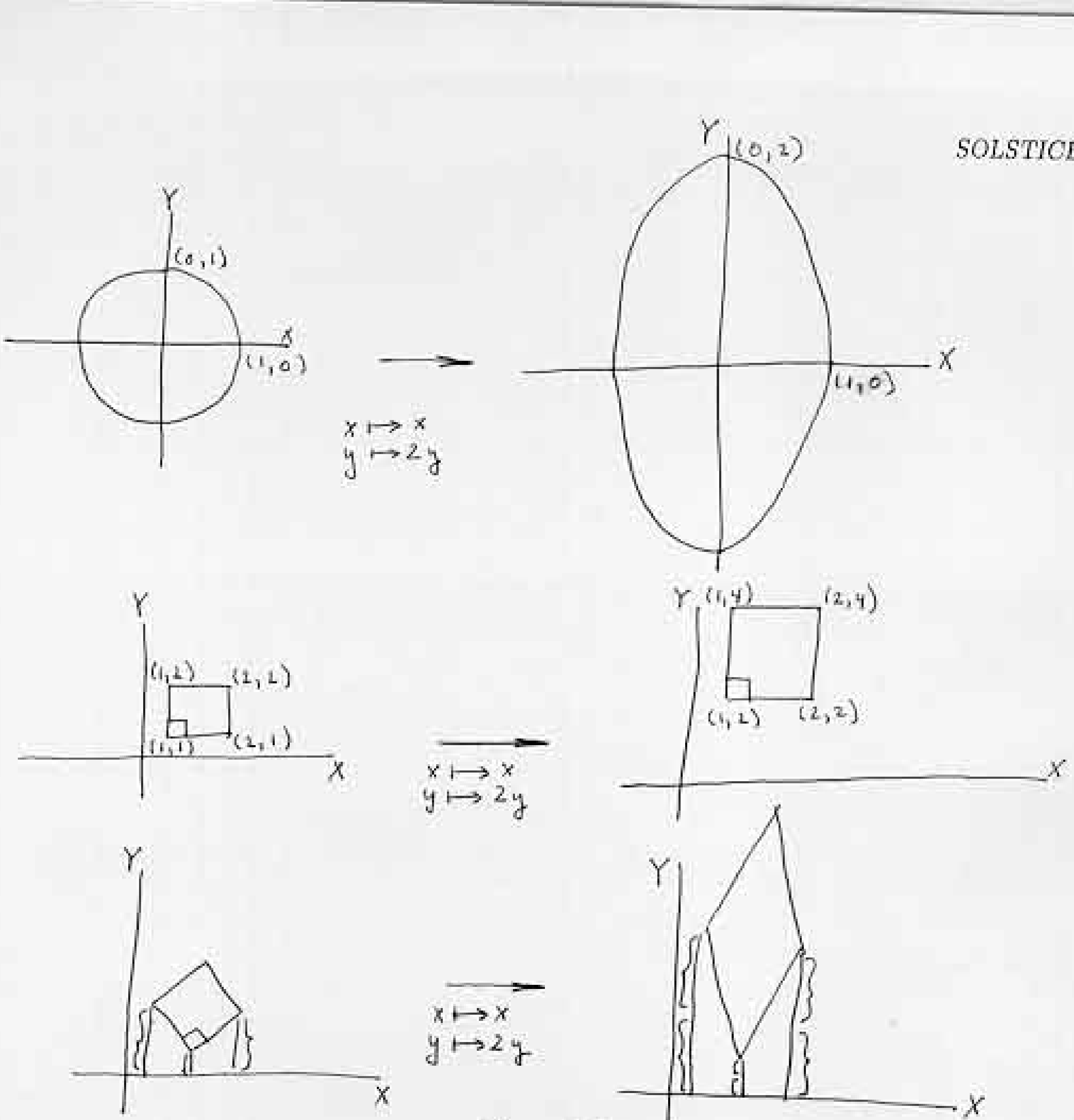


Figure III.2

1. The product, under $\star$, of any two elements of S is once again an element of S—this system is "closed" under the operation of $\star$—no new element (information) is generated.
2. Given a, b, and c in S: $(a \star b) \star c = a \star (b \star c)$. The manner in which parentheses are introduced is not of significance in determining the answer (information content) resulting from a string of operations under $\star$. The operation of $\star$ is said to be associative.
3. There is an identity element, 1, in S such that for any element of S, say a, it follows that

$$a \star 1 = 1 \star a = a.$$

4. Each element of S has an inverse in S; that is, for a typical element a of S, there exists another element, b of S, such that

$$a \star b = b \star a = 1.$$

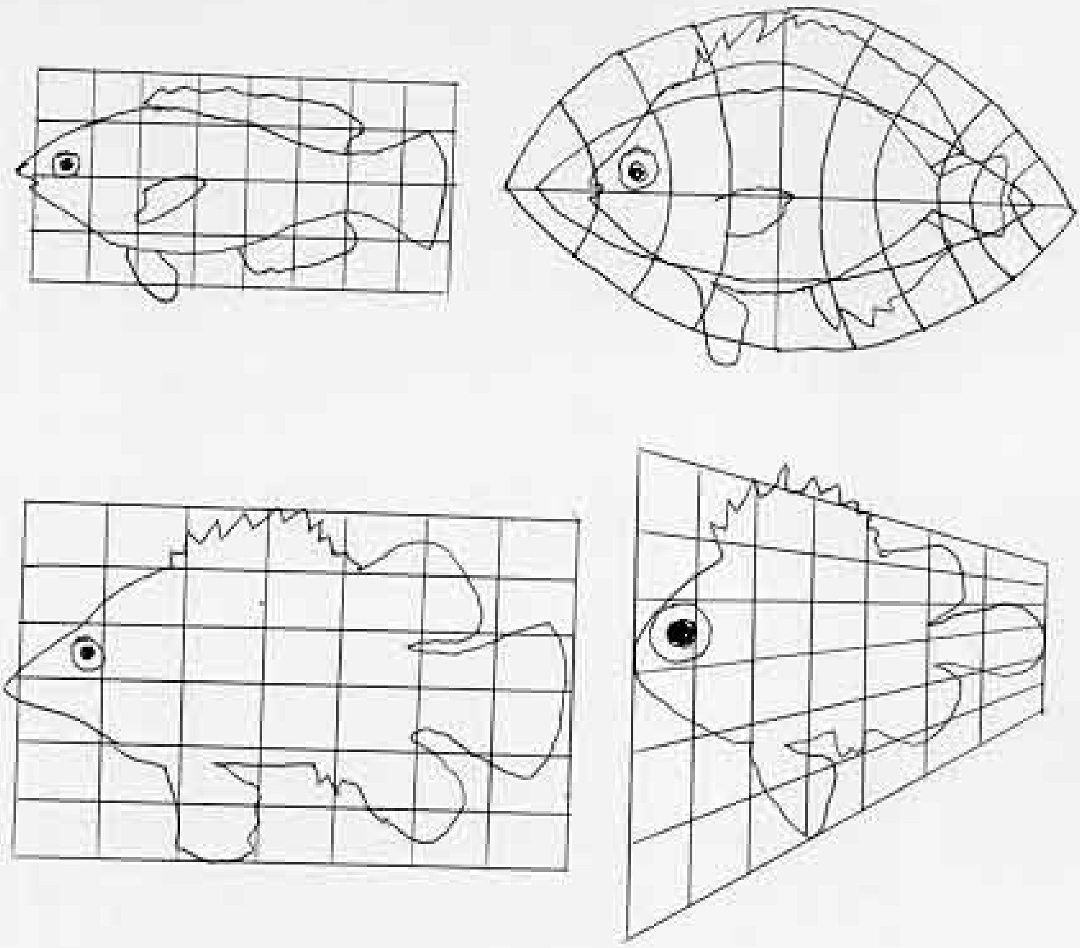


Figure III.3

Denote the inverse of a as a^{-1}. Thus, $a \star a^{-1} = a^{-1} \star a = 1$.

The order in which elements are related to each other, using $\star$, may matter; it need not be true that $a \star b = b \star a$. (Elements of the group do not necessarily "commute" with each other.)

The algebraic idea of "closure" is comparable to the GIS notion of snapping a polygon shut, so that chaining of line segments does not continue forever—the system is "closed."

A. The affine group; affine geometry.

The definition of group given above was to a set of elements and an operation linking them. These elements might be regarded as transformations. In particular, consider the set of all affine transformations of the plane that are one-to-one (translations, scalings, rotations, and reflections). These form a group, when the operation $\star$ is considered as the composition of functions:

i. The product of two affine transformations is itself an affine transformation;

ii. In a sequence of three affine transformations, it does not matter which two are grouped first, as long as the pattern of the three is unchanged—associativity.

iii. The affine transformation which maps the plane to itself serves as an identity element.

iv. Because the affine transformations dealt with here are one-to-one, they have inverses (all translations have inverses; only those linear transformations with inverses are considered here).

Affine geometry is the study of properties of figures that remain invariant under the group of one-to-one affine transformatons. Here are some theorems from affine geometry.

i. Any one-to-one affine transformation maps lines to lines.

ii. Any affine transformation maps parallel sets of lines to parallel sets of lines.

iii. Any two triangles are equivalent with respect to the affine group.

To demonstrate the theorem in iii., consider a fixed triangle with position (OB_0C_0), relative to an x/y coordinate system. Choose an arbitrary triangle, (ABC). Use elements of the affine group to move (ABC) to coincide with (OB_0C_0): a translation slides A to O (Figure VII.1). Two separate scaling operations and rotations slide B to B_0 and C to C_0. This is possible because O, B, and C are not collinear (as vectors, OB and OC are linearly independent).

This is the theoretical origin of the GIS notion that control points must be non-collinear and that there must be at least three of them. From a mathematical standpoint, it does not, therefore, matter whether the control points are chosen close together or far apart; however, from a visual standpoint it does matter. When control points are chosen close together the scaling operation required to transform the control triangle into other triangles is generally enlargement. When the control triangle is chosen with widely spaced vertices, the scaling operations required to transform it into other triangles is generally reduction. Errors are more visible with enlargement. Therefore, it is better, for the sake of visual comfort, to rely on reduction (reducing error size, as well) whenever possible, and therefore, to choose widely-spaced control points.

This is like the exercise above; there are two scalings and another affine transformation (here a translation, in the exercise, a reflection). In either case, the outcome of applying a sequence of affine transformations is still an affine transformation. In this case, it does not matter in what order the scaling operations are executed and in what order, relative to the scaling, the translation is applied. In the case of the exercise, however, this is not the case.

It does not matter in what order the scalings are applied. It is the case that $\tau_1 \circ \tau_2 = \tau_2 \circ \tau_1$. It is also the case that $\tau_1 \circ \tau_3 = \tau_3 \circ \tau_1$. However, it is not the case that

$$\tau_2 \circ \tau_3 = \tau_3 \circ \tau_2 :$$

$$(50,5) \xrightarrow{\tau_2} (50,48) \xrightarrow{\tau_3} (50,432)$$

$$(50,5) \xrightarrow{\tau_3} (50,475) \xrightarrow{\tau_2} (50,4660)$$

Observe, however, that it is possible to solve the problem applying the reflection earlier. Take τ_1 to be the required reflection so that y is sent to $50-y$ (reflection before the scale

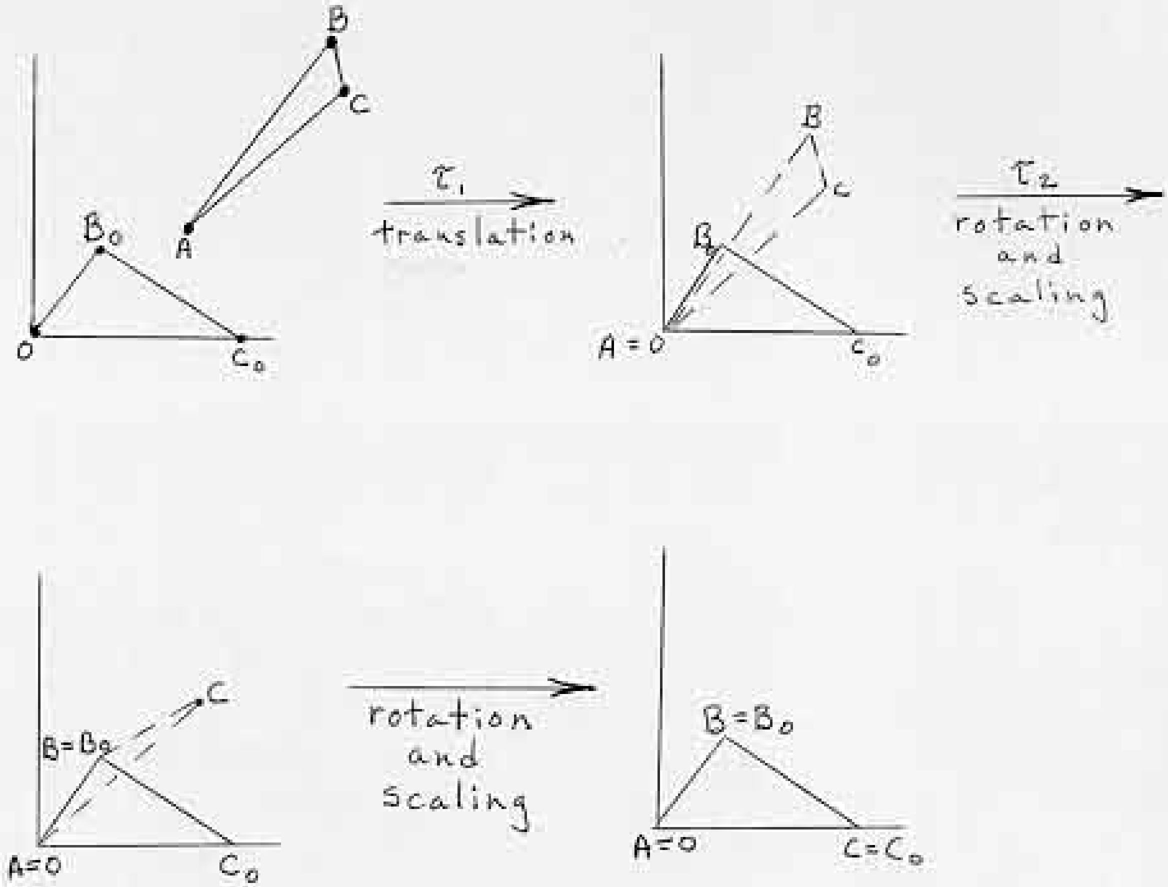


Figure VII.1

change on the y-axis). Figure VII.2 shows the solution here. In the non-commutative case here, there is a sharp difference in the "correct" y-value and the other possible one. In this case, as in the previous one, it does not matter how the application of transformations are separated by parentheses, and it is guaranteed that the product will itself be affine.

Thus, the order of application of affine transformations, within the group (locally), is important. This might cause difficulties (sending you off the screen), or it might be turned to an advantage in zooming-in on something. What caused the problem here was the reflection. Products of rotations of the plane are roatations of the plane; products of translations are translations, and products of scalings are scalings. Here, and as we shall see later, reflections cause non-commutativity (similar problems might have arisen in Figure VII.1, had a reflection been involved).

iv. Any triangle is affine-equivalent to an equilateral triangle (choose whatever control triangle desired—can choose an underlying lattice of regularly spaced triangular points and

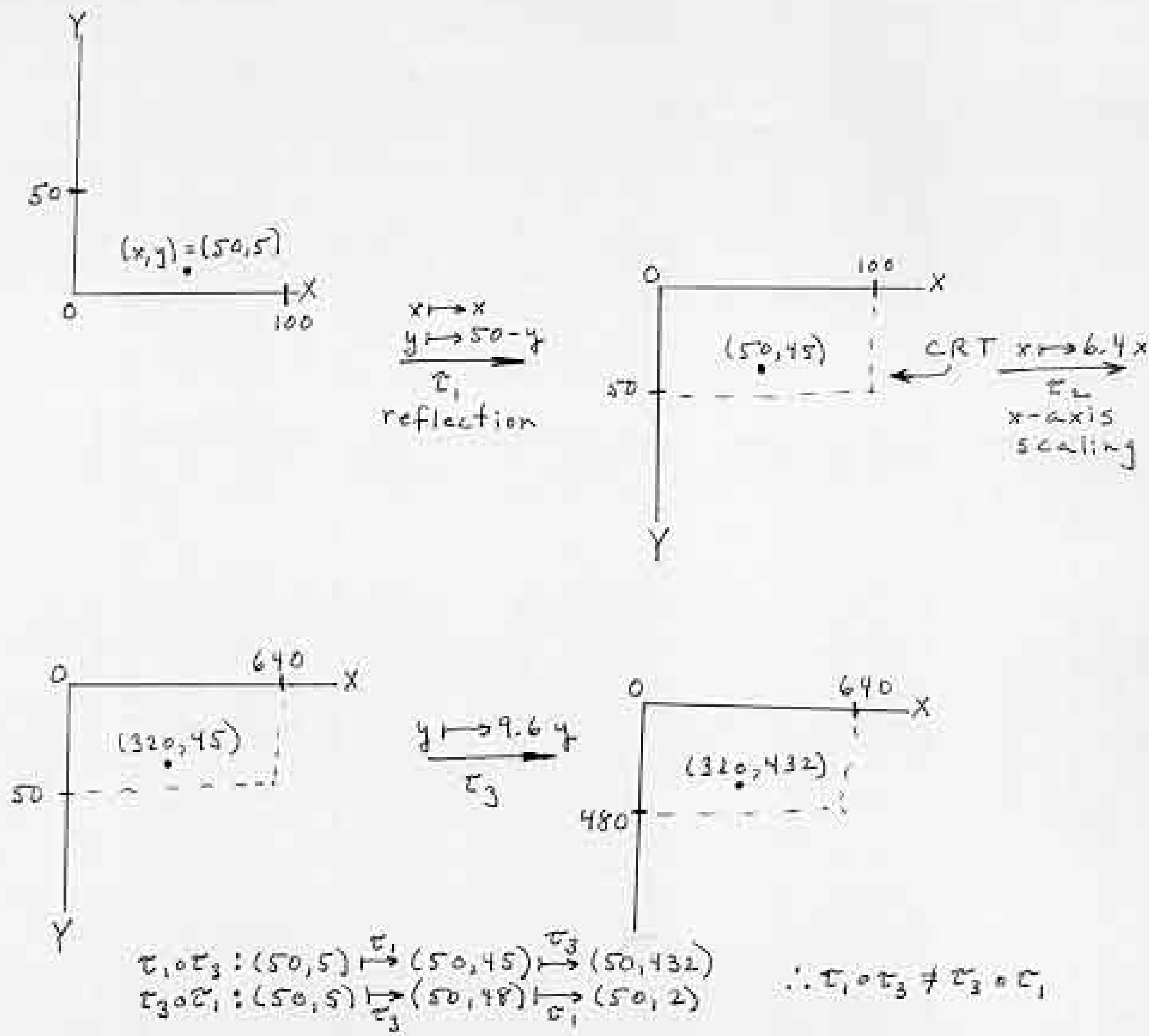


Figure VII.2

rubber sheet them to an irregularly spaced one).

v. Any ellipse is affine equivalent to a circle (demonstrated via copier technology).

a. Parallelism and GIS: crossing lines and polygon area.

Groups suggest how theoretical structure may be built from assembling simple pieces. GIS algorithms for complex processes are also often built from assembling simple pieces.

Straight lines

How can we tell if two lines intersect in a node?

Example from NCGIA Lecture 32: does the line L_1 from (4,2) to (2,0) cross the line L_2 from (0,4) to (4,0)? From a mathematical standpoint, two lines in the Euclidean plane cross if they have different slopes, m_1 and m_2, where the slope m between points (x_1, y_1) and (x_2, y_2) is calculated as $(y_2 - y_1)/(x_2 - x_1)$. In this case, the slope of L_1 is $(0-2)/(2-4) = 1$

and the slope of L_2 is $(0-4)/(4-0) = -1$. The slopes are different, so the lines cross in the plane. However, in the GIS context:

i. Do the lines cross on the computer screen, or is the intersection point outside the bounded Euclidean region of the screen?

ii. Even if the lines cross on the screen, do they intersect at a node of the data base (was that point digitized)?

To answer these questions, it is necessary to determine the intersection point of the two lines.

Equation of L_1: one form for the equation of a line between two points (x_1, y_1) and (x_2, y_2) is

$$y - y_1 = m(x - x_1)$$

where m is the slope and b is the second coordinate of the y-intercept. Thus, L_1 has equation $y - 2 = 1(x-4)$ or $y = x - 2$; L_2 has equation $y - 4 = -1(x-0)$ or $y = -x + 4$.

Solve these equations simultaneously to yield $x = 3$ and $y = 1$.

Thus, if the point $(3,1)$ lies within the boundaries of the screen, the lines intersect on the screen; if the point $(3,1)$ was digitized, then another line might be hooked onto the intersection point. If it was not digitized, then the lines "cross" but do not intersect, much as water pipes might cross but do not necessarily intersect (as in snapping a segment onto the middle of a line on the CRT). This is a graph-theoretic characteristic.

Note that vertical lines are a special case; their slope is undefined because $x_2 - x_1$, the denominator in the slope, is zero. Recognizing vertical lines should not be difficult, but it should be remembered that attempting to calculate slope across an entire set of lines, which might include vertical lines, can produce errors.

Chains of straight line segments.

How can we tell if chains of segments cross?

Because chains are of finite length and are bounded, it is possible to enclose them in a rectangle (no larger than the CRT screen) (Figure VII.3). This is a minimum enclosing rectangle.

Thus, given two chains, C_1 and C_2, if their respective minimum enclosing rectangles do not intersect (as do straight lines) then they do not intersect, and further testing is warranted.

Polygon area:

Calculate polygon area using notion of parallelism (Figure VII.4)

Simple rule, based on vertical lines, to determine if a point is inside or outside a polygon (Figure VII.5)

Centroids of polygons, with attached weights are often used as single values with which to characterize the entire polygon. Centroids are preserved, as centroids, under affine transformations.

These are technical procedures for determining various useful measures and are documented in NCGIA material; all are based in the theory of affine transformations applied to sets of pixels. Move now to consider the mechanics of how sets of affine transformations

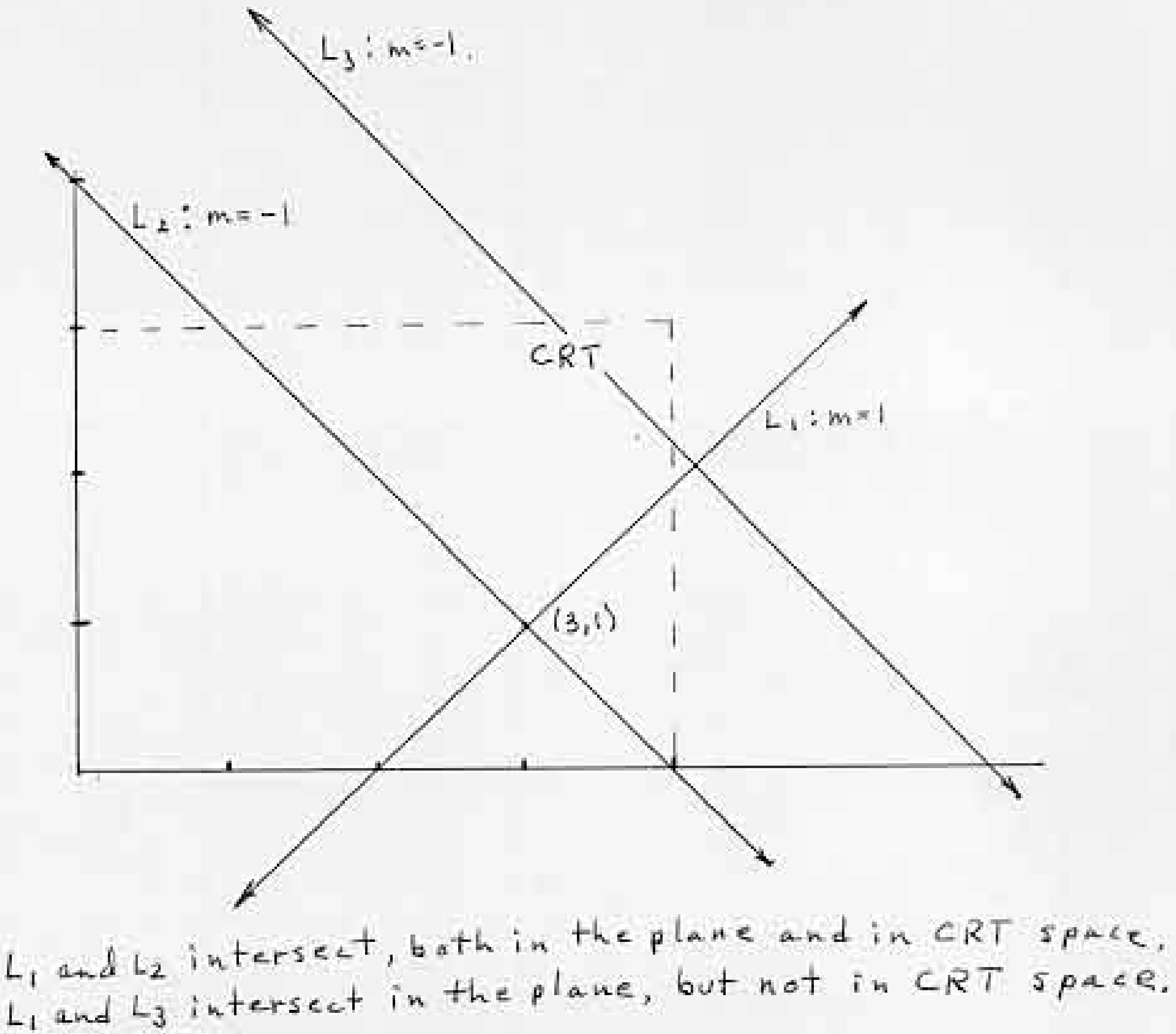


Figure VII.3

might affect a single pixel.

B. Group of symmetries of a square (pixel); the hexagonal pixel.

A square may have a set of rotations and of reflections applied to it as noted in Figure VII.6. Each may be represented as a permutation of the vertices, labelled clockwise. Permutations are multiplied as indicated in the example, below: multiply the permutation (1234) by the permutation (13)(24):

1 goes to 2 (in the left one)

and 2 goes to 4 (in the right one)

so 1 goes to 4 (in the product)

4 goes to 1 (in the left one)

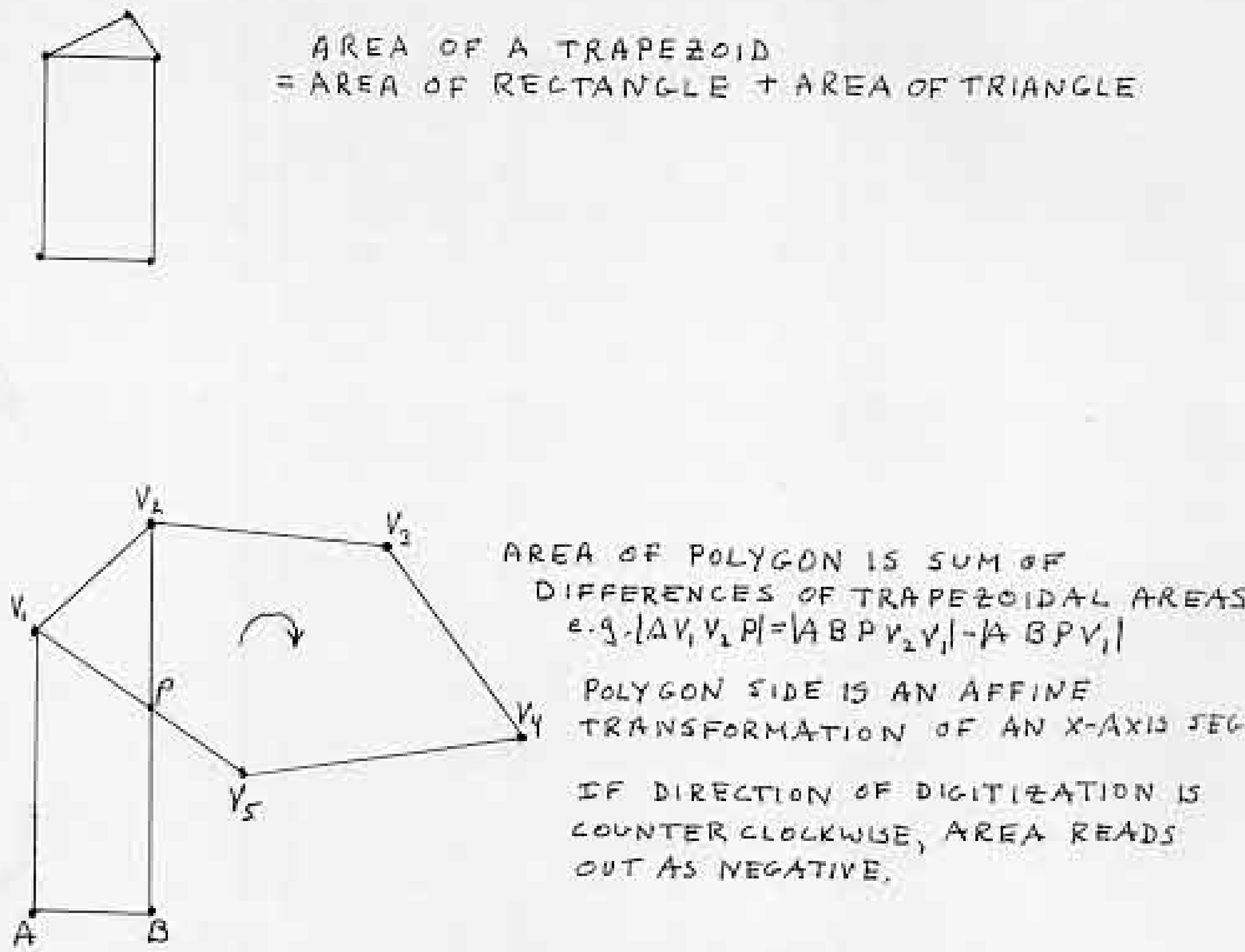


Figure VII.4

and 1 goes to 3 (in the right one)

so 4 goes to 3 (in the product)

3 goes to 4 (in the left one)

and 4 goes to 2 (in the right one)

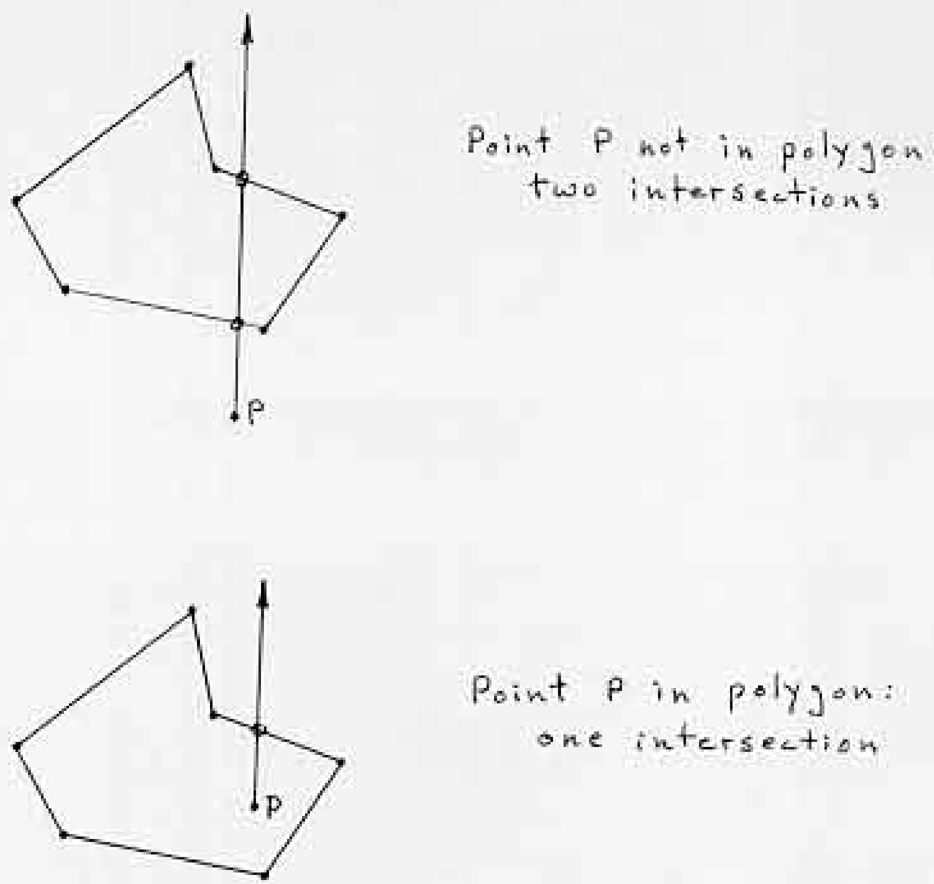
so 3 goes to 2 (in the product)

2 goes to 3 (in the left one)

and 3 goes to 1 (in the right one)

so 2 goes to 1 (in the product)

This last stage is akin to snapping a polygon closed in a GIS environment—here it is a cycle of numbers rather than of vertices. Figure VII.6 shows all the calculations; note, that



Based on NCGIA Lecture 33.

Points P on boundary and self-crossing polygons are a problem.

Figure VII.5

no new permutations ever arise; hence, the system is closed under $\star$; the rotation I serves as the identity transformation; each element has an inverse:

$$I \star I = I; I^{-1} = I$$

$$R_1 \star R_3 = I; R_1^{-1} = R_3$$

$$R_2 \star R_2 = I; R_2^{-1} = R_2$$

$$R_3 \star R_1 = I; R_3^{-1} = R_1$$

$$H \star H = I; V \star V = I; D_1 \star D_1 = I; D_2 \star D_2 = I.$$

So, this system is a "group." It is not, however, a commutative group—for example, $R_1 \star H = D_2$ and $H \star R_1 = D_1$. Once again, a reminder to be careful when combining reflections with affine transformations. Note that the set of rotations (including the identity rotation)

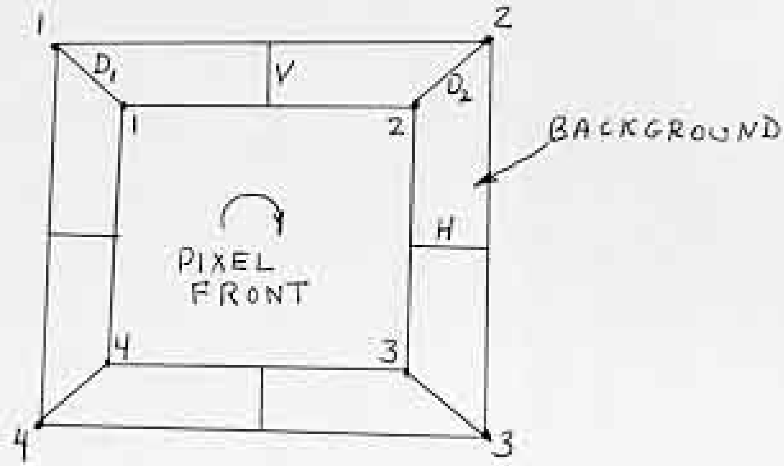


Figure VII.6 Group of symmetries of a square

Rotations:	Permutation representation
I: identity	(1)(2)(3)(4)
R_1: through 90 deg	(1234)
R_2: through 180 deg	(13)(24)
R_3: through 270 deg	(1432)
Reflections:	**Permutation representation**
H: horizontal	(14)(23)
V: vertical	(12)(34)
D_1: diagonal, 1 to 3	(1)(3)(24)
D_2: diagonal, 2 to 4	(2)(4)(13)

Table–operation, $\star$, is multiplication of permutations.

$\star$	I	R_1	R_2	R_3	H	V	D_1	D_2
I	I	R_1	R_2	R_3	H	V	D_1	D_2
R_1	R_1	R_2	R_3	I	D_2	D_1	H	V
R_2	R_2	R_3	I	R_1	V	H	D_2	D_1
R_3	R_3	I	R_1	R_2	D_1	D_2	V	H
H	H	D_1	V	D_2	I	R_2	R_1	R_3
V	V	D_2	H	D_1	R_2	I	R_3	R_1
D_1	D_1	V	D_2	H	R_3	R_1	I	R_2
D_2	D_2	H	D_1	V	R_1	R_3	R_2	I

is itself a group within this group. This is a "subgroup"—it is commutative—the order in which rotations are applied to the square is irrelevant.

Figure VII.6, continued

Permutation	⋆	Permutation	=	Permutation	
(1)(2)(3)(4)		(1)(2)(3)(4)		(1)(2)(3)(4)	I
(1234)		(1)(2)(3)(4)		(1234)	R_1
(13)(24)		(1)(2)(3)(4)		(13)(24)	R_2
(1432)		(1)(2)(3)(4)		(1432)	R_3
(14)(23)		(1)(2)(3)(4)		(14)(23)	H
(12)(34)		(1)(2)(3)(4)		(12)(34)	V
(1)(3)(24)		(1)(2)(3)(4)		(1)(3)(24)	D_1
(2)(4)(13)		(1)(2)(3)(4)		(2)(4)(13)	D_2
(1)(2)(3)(4)		(1234)		(1234)	R_1
(1234)		(1234)		(13)(24)	R_2
(13)(24)		(1234)		(1432)	R_3
(1432)		(1234)		(1)(2)(3)(4)	I
(14)(23)		(1234)		(1)(3)(24)	D_1
(12)(34)		(1234)		(2)(4)(13)	D_2
(1)(3)(24)		(1234)		(12)(34)	V
(2)(4)(13)		(1234)		(14)(23)	H
(1)(2)(3)(4)		(13)(24)		(13)(24)	R_2
(1234)		(13)(24)		(1432)	R_3
(13)(24)		(13)(24)		(1)(2)(3)(4)	I
(1432)		(13)(24)		(1234)	R_1
(14)(23)		(13)(24)		(12)(34)	V
(12)(34)		(13)(24)		(14)(23)	H
(1)(3)(24)		(13)(24)		(2)(4)(13)	D_2
(2)(4)(13)		(13)(24)		(1)(3)(24)	D_1
(1)(2)(3)(4)		(1432)		(1432)	R_3
(1234)		(1432)		(1)(2)(3)(4)	I
(13)(24)		(1432)		(1234)	R_1
(1432)		(1432)		(13)(24)	R_2
(14)(23)		(1432)		(12)(34)	D_2
(12)(34)		(1432)		(14)(23)	D_1
(1)(3)(24)		(1432)		(2)(4)(13)	H
(2)(4)(13)		(1432)		(1)(3)(24)	V
(1)(2)(3)(4)		(14)(23)		(14)(23)	H
(1234)		(14)(23)		(2)(4)(13)	D_2
(13)(24)		(14)(23)		(12)(34)	V
(1432)		(14)(23)		(1)(3)(24)	D_1
(14)(23)		(14)(23)		(1)(2)(3)(4)	I
(12)(34)		(14)(23)		(13)(24)	R_2
(1)(3)(24)		(14)(23)		(1432)	R_3
(2)(4)(13)		(14)(23)		(1234)	R_1
(1)(2)(3)(4)		(12)(34)		(12)(34)	V
(1234)		(12)(34)		(1)(3)(24)	D_1
(13)(24)		(12)(34)		(14)(23)	H

(1432)	(12)(34)	(2)(4)(13)	D_2
(14)(23)	(12)(34)	(13)(24)	R_2
(12)(34)	(12)(34)	(1)(2)(3)(4)	I
(1)(3)(24)	(12)(34)	(1234)	R_1
(2)(4)(13)	(12)(34)	(1432)	R_3
(1)(2)(3)(4)	(1)(3)(24)	(1)(3)(24)	D_1
(1234)	(1)(3)(24)	(14)(23)	H
(13)(24)	(1)(3)(24)	(2)(4)(13)	D_2
(1432)	(1)(3)(24)	(12)(34)	V
(14)(23)	(1)(3)(24)	(1234)	R_1
(12)(34)	(1)(3)(24)	(1432)	R_3
(1)(3)(24)	(1)(3)(24)	(1)(2)(3)(4)	I
(2)(4)(13)	(1)(3)(24)	(13)(24)	R_2
(1)(2)(3)(4)	(2)(4)(13)	(2)(4)(13)	D_2
(1234)	(2)(4)(13)	(12)(34)	V
(13)(24)	(2)(4)(13)	(1)(3)(24)	D_1
(1432)	(2)(4)(13)	(14)(23)	H
(14)(23)	(2)(4)(13)	(1432)	R_3
(12)(34)	(2)(4)(13)	(1234)	R_1
(1)(3)(24)	(2)(4)(13)	(13)(24)	R_2
(2)(4)(13)	(2)(4)(13)	(1)(2)(3)(4)	I

Are there any other subgroups? Yes, I, R_2, H, V also form a commutative subgroup. Note that the product of two reflections is a rotation.

A similar style of analysis might be executed for the pixel viewed as a hexagon. Other theoretical issues arise concerning the possibility of using a crt display with hexagonal pixels.

i. Issues involving centroids

a. Transformation to generate a centrally- symmetric hexagon from an arbitrary (convex) hexagon (rubbersheeting; TIN).

One such issue involves concern for taking a set of irregularly-spaced data points and converting them into some sort of more regular distribution (as with rubbersheeting and a TIN). This procedure illustrates how to transform an arbitrary convex hexagon (V_1, V_2, V_3, V_4, V_5, V_6) into a centrally symmetric hexagon (S_1, S_2, S_3, S_4, S_5, S_6) centered on a point that is easy to find. (See construction in *Solstice I*—Summer, 1990, Vol. I, No. 1., pp. 41-42.) Thus, rubbersheeting would appear possible with an hexagonal pixel.

b. Area algorithm generalizes to hexagons: regular hexagon is two isosceles trapezoids (one on either side of a single diameter of the hexagon).

What else might generalize from the square pixel format to the hexagonal pixel format? A hexagon can be decomposed into two trapezoids; thus one might imagine using an algorithm similar to that for the square pixel to find polygon areas relative to an hexagonal pixel display.

c. Steiner networks as boundaries of sets of hexagonal pixels; given a set of points, find a minimal hexagonal network linking them.

If centers of gravity (centroids) are used as a centering scheme in a triangulated irregular network (or other network of polygons), then it would be nice to have no centroid lie outside a triangular cell (or other polygon). A centroid is the intersection point of medians; it is the balance point on which the figure would rest. Sometimes the centroid lies outside the polygon; Coxeter suggests viewing the centroid as a balance point among electrical charges, thereby allowing for this possibility. Another point that is useful for using as a "central" weight is a Steiner point; in a triangle, it is that point which minimizes total network length joining the three vertices. It is always within the triangle when no angle of the triangle is greater than or equal to 120 degrees. (See *Solstice-I*, Vol. I., no. 2, "Super-definition resolution.")

Assigning point weights to represent polygon values is one way to compare them; another way is to assign centrally-located networks traversing underlying grid lines (Manhattan lines with square pixels, Steiner networks with hexagonal pixels); another way is to overlay the areas—again, a point-line-area classification as mentioned in detail in one of Nystuen's earlier lectures.

ii. Issues involving polygon overlays.

a. Close-packings of hexagons; central place geometry.

b. Fractal approach; space-filling; data compression.

Polygon overlay is familiar from OSUMAP. Look at some abstract geographic/geometric issues that might suggest directions to consider in looking at ideas behind the process of overlays.

Geometry of central place theory–including fractal generation of these layers. Look for a number of issues of this sort, that are theoretical, in using GIS-type equipment. Below is an outline of material in these lectures and of suggestions for future directions in which to look.

I. Introduction: the role of theory. Mathematics is fundamental, and in dealing with spatial phenomena, geometry in particular, is fundamental. Historical precedent from Biology in works of D'Arcy Thompson; Tobler's map transformations.

A. Statement of Thompson regarding the role of theory.

B. Visual evidence: one species of fish is transformed into another actual species by choosing a suitable coordinate transformation.

II. Transformations.

A. Well-defined (single-valued).

B. Reversible

i. One-to-one correspondence

ii. Transformations of X onto Y.

C. "Rubbersheeting"—example from Nystuen lecture, with fire stations. What is involved is creating a transformaton from an irregular scatter of locations to a regular one, locating new points (fire stations) and snapping the surface back to the irregular scatter. This requires transformations that are reversible.

III. Types of transformations and examples.

A. Affine

i. Translation

ii. Scaling

iii. Rotation

iv. Reflection

B. Curvilinear

IV. Exercise—scaling to make digitized map mesh with CRT scale.

V. GIS tie to Steiner networks.

VI. Digital topology. Quadtrees–Rosenfeld, Tobler. Jordan Curve Theorem; American Mathematical Society special sessions on digital topology (run by Rosenfeld). Hexagonal pixels—scanner technology.

VII. Local scale of mathematical extension of the concept of "affine transformation." The algebra of symmetry: definition of a group.

A. The affine group; affine geometry.

i. Parallelism and GIS: crossing lines and polygon area.

ii. Projective geometry; any two lines intersect in a point; no parallels. Here for completeness–not really discussed.

B. Group of symmetries of a square (pixel); the hexagonal pixel.

i. Issues involving centroids.

a. Transformation to generate a centrally-symmetric hexagon from an arbitrary (convex) hexagon (rubbersheeting; TIN).

b. Area algorithm generalizes to hexagons; hexagon is two trapezoids.

c. Steiner networks as boundaries of sets of hexagonal pixels; given a set of points, find a minimal hexagonal network linking them–dealt with in a third lecture, not presented here.

ii. Issues involving polygon overlays.

a. Close-packings of hexagons; central place geometry.

b. Fractal approach; space-filling; data compression.

VIII. Global scale of mathematical extension of the concept of "affine transformation." Topology.

A. Combinatorial topology.

i. Jordan curve theorem. GIS connection, inside and outside of polygons.

ii. Cell complexes; 0, 1, and 2 cells of GIS.

iii. Hexagons derived from barycentric subdivision of a complex.

B. Point-set topology.

i. Definitions.

ii. Consequences of Definitions interpreted in GIS context.

C. Digital topology.

i. Jordan curve theorem–3-dimensions.

ii. Quadtrees.

III. Further extension at different scales. Commutative diagrams—entry to different level of mathematical thought and spatial theory.

Publications of the Institute of Mathematical Geography have been reviewed in

1. *The Professional Geographer* published by the Association of American Geographers;
2. *The Urban Specialty Group Newsletter* of the Association of American Geographers;
3. *Mathematical Reviews* published by the American Mathematical Society;
4. *The American Mathematical Monthly* published by the Mathematical Association of America;
5. *Zentralblatt* Springer-Verlag, Berlin
6. *Mathematics Magazine*