BUREAU OF THE CENSUS
STATISTICAL RESEARCH DIVISION REPORT SERIES
SRD Research Report Number: Census/SRD/RR-87/28

# Shape Representation for Linear Features in Automated Cartography

by

Alan Saalfeld
Statistical Research Division
Bureau of the Census



Recommended by: Tom O'Reagan

Report completed: September, 1986

Report issued: October 1, 1987

# SHAPE REPRESENTATION FOR LINEAR FEATURES IN AUTOMATED CARTOGRAPHY

Alan Saalfeld
Statistical Research Division
Bureau of the Census
Washington, DC 20233
(301) 763-7530

## ABSTRACT

The graphic image produced from a digital file such as a GBF/DIME file may be greatly enhanced by utilizing sufficiently detailed shape information on each line segment. This paper presents a method of structuring shape records to reduce storage requirements and improve appearance of the drawn image. The shape records described here are independent of the map position of the segment. A *standardized* shape is defined and stored as a curve between two (arbitrarily fixed) points in the plane. A standardized shape is moved to any other position in the plane and is scaled up or down prior to drawing by an elementary transformation called a similitude. The operation of transforming a standardized shape to any position is computationally fast and simple. Although the number of standardized shapes is infinite, a small collection of shapes provides good approximations to most shapes encountered in maps. Several important properties of standardized shape representations are examined, including invariance under transformations, independence of topological structure, easy interchangability with other shape representations, and independence of drawing precision.

## INTRODUCTION

To a mathematician, the notion of shape is defined by a family of transformations of space called similitudes. Two figures in space have the same shape if one figure can be transformed into the other by one of these similitudes. In a plane, the family of similitudes consists of translations, scalings, and rotations, and combinations of these three types of movements. Sometimes a mathematician will include reflections in his family of similitudes, but because these transformations reverse orientation, they will not be included in the shape-preserving transformations studied here. Two figures with the same shape are called similar. For example, any two circles are similar because one may be moved into congruence with the other by a scaling followed by a translation; and any two straight line segments have the same shape because either one may be moved into alignment with the other by a scaling followed by a rotation and a translation.

The fact that any two line segments are similar also means that any line segment is similar to or has the same shape as the line segment in the plane from (0,0) to (1,0); and this fact allows us to refer to segments or curves in *standard position*. Every non-closed directed curve (that is, curve whose end points are distinct and ordered) has a straight-line segment associated with it, namely, the directed segment linking its distinct end-points in order. There is a unique similitude of the plane which transforms the first end-point of the segment to (0,0) and also transforms the second end-point of the segment to (1,0). We say that this similitude moves the curve to standard position. Note that two directed curves of the same shape have the same standard position curve and two curves of different shapes have different standard position curves. If the order of the end-points changes, the standard position curve undergoes a rotation of 180°.

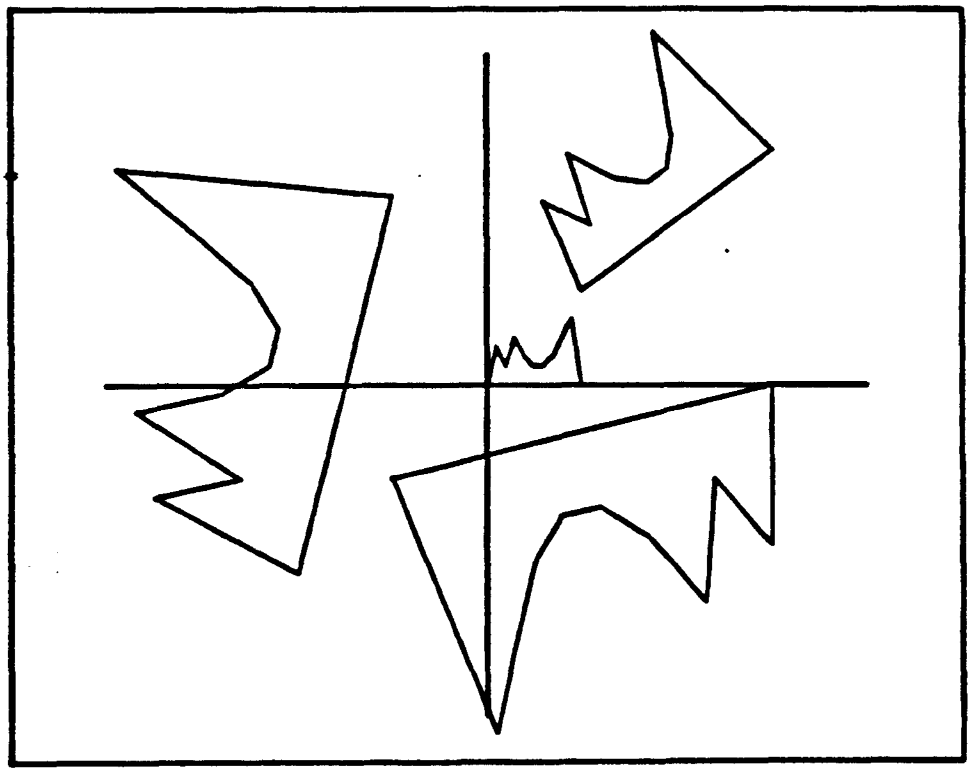

Figure 1. Curves of the same shape and associated segments (one curve in standard position).

After we describe how to transform segments (and, hence, curves) to standard position and from standard position, we will focus on comparing curves in standard position in order to establish a distance measure between curves and, hence, between shapes. Because there is a one-to-one correspondence between shapes and standard-position curves, we can study all shapes simply by examining all curves between (0,0) and (1,0).

## TRANSFORMATIONS TO AND FROM STANDARD POSITION

Arithmetic of complex numbers provides a handy set of tools for describing similitudes or shape-preserving transformations of the plane. We will use the coordinate representation (x,y) and the complex representation x+yi interchangably in the text that follows to describe the transformations of interest to us. Addition of a fixed complex number to all complex numbers transforms the plane of complex numbers by a translation. Multiplication of all numbers by a fixed complex number produces a combined scaling and rotation of the plane. The scaling factor is equal to the magnitude of the fixed complex number; and the angle of rotation is equal to the direction of the vector of the fixed complex number doing the multiplication.

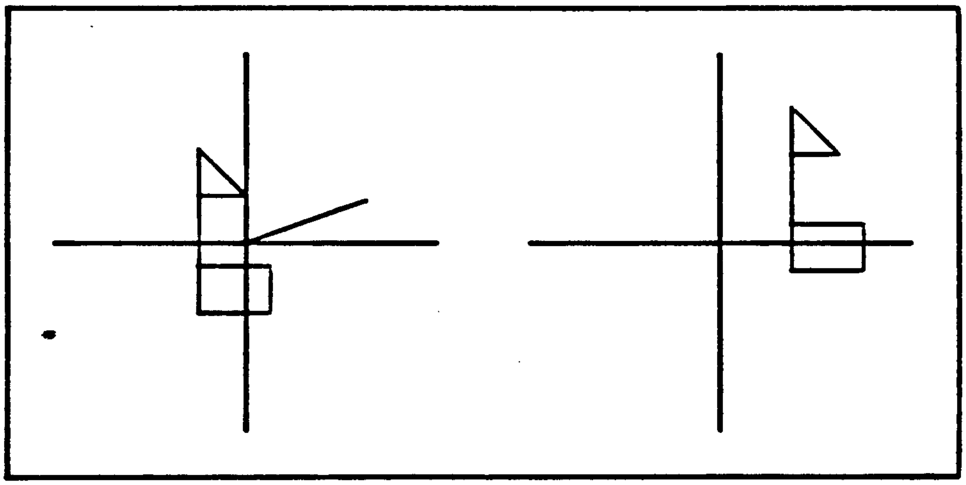

Figure 2. Addition of $\alpha+\beta i$ produces translation.

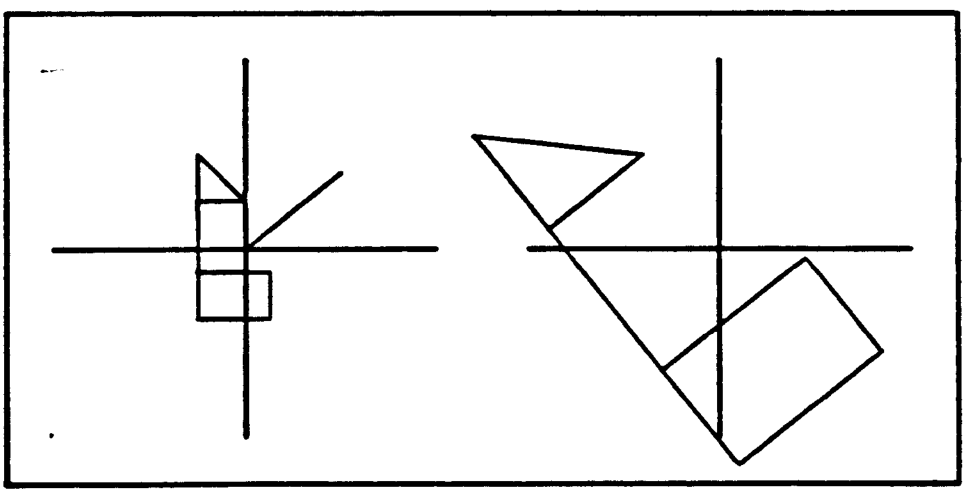

Fig. 3. Multiplying by $\alpha+\beta i$ causes scaling and rotation.

The inverse transformation of addition of $\alpha+\beta i$ is subtraction of $(\alpha+\beta i)$, or addition of $(-\alpha-\beta i)$.

The inverse transformation of multiplication by $\alpha+\beta i$ is division by $\alpha+\beta i$, or multiplication by $(\alpha+\beta i)^{-1}$, or by $(\alpha-\beta i)/(\alpha^2+\beta^2)$, which is the complex conjugate of $\alpha+\beta i$ divided by the norm squared of $\alpha+\beta i$.

The shape-preserving transformation, T, of the segment joining (0,0) and (1,0) to the line segment from $(x_1,y_1)$ to $(x_2,y_2)$ can be described as follows:

$$\text{Let } \Delta = (\Delta_1,\Delta_2) = (x_2-x_1,y_2-y_1).$$

Let $M_\Delta$ be the transformation that corresponds to complex multiplication by $(\Delta_1+\Delta_2 i)$:

$$M_\Delta(x,y) = (x\Delta_1 - y\Delta_2, y\Delta_1 + x\Delta_2).$$

Let $A_1$ be the transformation that corresponds to the addition of the complex number $(x_1+y_1 i)$:

$$A_1(x,y) = (x + x_1, y + y_1).$$

Then $T = A_1 \cdot M_\Delta$, where composition means that the transformation $M_\Delta$ is applied first, then $A_1$ is applied.

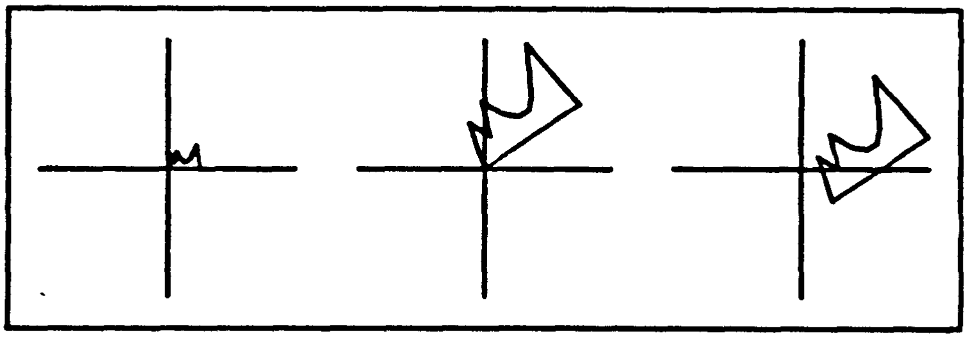

Figure 4. Composition of $M_\Delta$ followed by $A_1$.

Explicitly, in terms of $x_1$, $y_1$, $\Delta_1$, and $\Delta_2$, we have:

$$T(x,y) = (x\Delta_1 - y\Delta_2 + x_1, y\Delta_1 + x\Delta_2 + y_1).$$

The inverse shape-preserving transformation, $T^{-1}$, of the line segment from $(x_1,y_1)$ to $(x_2,y_2)$ to the segment joining (0,0) and (1,0) can similarly be expressed as follows:

$$T^{-1} = M_\Delta^{-1} \cdot A_1^{-1}.$$

As noted above, the inverse of addition is subtraction; and the inverse of multiplication is division or multiplication by the complex conjugate divided by the norm squared.

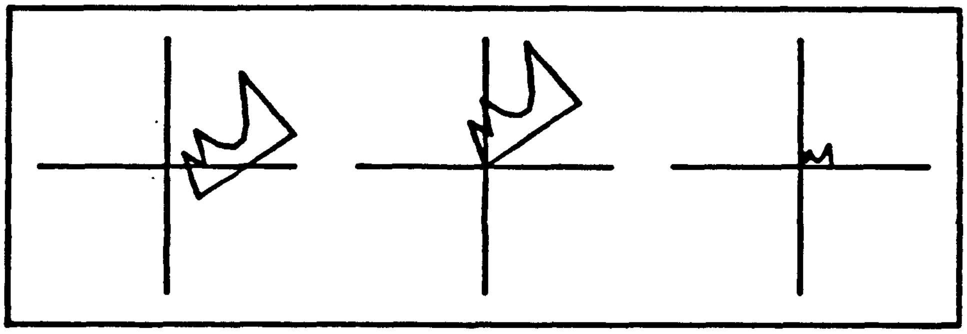

Figure 5. Composition of $A_1^{-1}$ followed by $M_\Delta^{-1}$.

As with T, we have an explicit expression for $T^{-1}(x,y)$:

$$(((x-x_1)\Delta_1-(y-y_1)\Delta_2),((y-y_1)\Delta_1+(x-x_1)\Delta_2))/(\Delta_1^2+\Delta_2^2).$$

## COMPARING CURVES IN STANDARD POSITION

Once two curves have been placed in standard position, they may be compared by some measures of their distance from one another. One simple measure of distance is to compute the area or areas between the curves. If the curves are close, then the area between them will be small. The area between the curves may nevertheless be small if one or both of the curves have spikes; hence, the area measure is not always the best measure of closeness of curves.

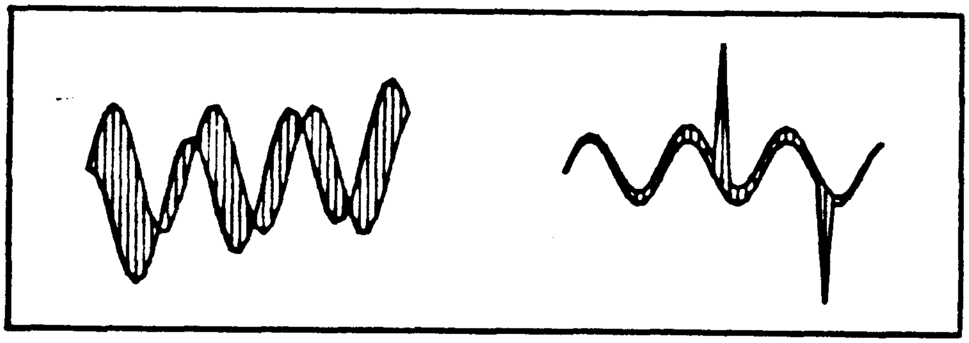

Figure 6. Area difference measures curve closeness.

Another measure of closeness of curves derives from the mathematical notions of closeness of functions. If two curves are expressed as functions $\xi(t)$ and $\zeta(t)$ parametrized by the same normalized parameter t, then there are several fundamental measures for describing distance between the curves which are known in function theory as the $L_p$ measures :

$$\|\xi-\zeta\|_p = [\int\|\xi(t)-\zeta(t)\|^p dt]^{1/p}$$

where the norm ‖ ‖ within the integral refers to a measure of distance between two points in the plane.

The $L_2$ norm (that is $L_p$, when $p = 2$) is easily computed for curves which are polygonal lines and this norm corresponds to averaging the usual Euclidean distance between corresponding points in the plane. A closed formula for calculating the $L_2$ distance between two polygonal lines is included as an appendix.

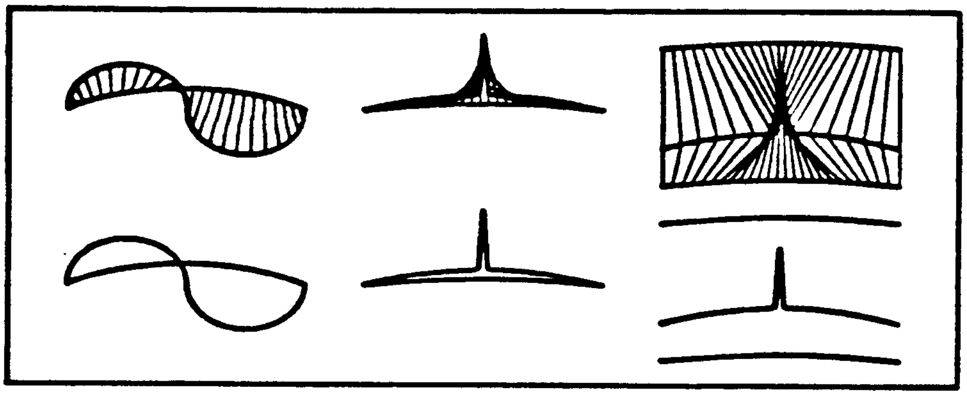

Figure 7. The $L_2$ norm averages distances between corresponding points on two parametrized curves.

The linking lines in Figure 7 illustrate that distances are not always measured to the nearest point on the other curve nor always in a vertical direction.

The primary reason for studying nearness of curves to each other lies in reducing the number of curves used to attain a representative set of curves. If all curves drawn in a particular application are very close to arcs of circles, for example, and if the precision of drawing required is within that closeness tolerance, then the family of arcs of circles will suffice for representing the curves needed for the application.

Map drawing does not require uniform precision in representing curve features. The precision required varies with the scale and the application. Because of data set constraints and equipment limitations such as penplotter operating characteristics, the curves of automated cartography are almost always polygonal lines, often intended to be approximations to smooth curves. For this reason, polygonal lines have a special role in our study of curve types and in our search for a representative family of curves in standard position.

## SMALL REPRESENTATIVE FAMILIES OF CURVES

In a conventional approach to storing curve data used by both the Bureau of the Census and the United States Geological Survey, every different curved line segment has associated with it its own linked list of coordinate pair records (one for every distinguished point on the curve) and possibly a rule or an algorithm for stringing the curve points together (such as a spline fit or some other smooth fit). The number of curve point coordinate pair records is proportional to the number of 1-cell or segment records on the file. Larger maps require proportionately larger curve point files; and the time needed to access each curve list increases with map size.

A small representative curve list eliminates the curve point file and its linkages entirely. A small list of standard-position curves has the same fast access time for large maps as well as small maps. A standard-position curve need not be limited by a particular storage precision. The same standard-position curve may be transformed with different degrees of precision (that is, it may be evaluated at more or fewer points) depending on the size of the transformed image, the mechanical drawing instrument's precision, and the application and appearance requirements of the map being drawn.

Two methods for constructing a small representative curve family for maps are outlined below.

<u>Method I. Statistical Selection</u>
This method will be implemented at the Bureau of the Census as part of an experiment in 1986. This method uses a complete curve file for a map in the current linked-list coordinate pair format. First all of the curve lists are converted to standard-position curves by the appropriate $T^{-1}$ transformation described in an earlier section. After all of the curves have been given a format that makes distance comparisons possible, cluster analysis is performed on the set of standard-position curves to group them into clusters of curves, all of which are close to other members of the same cluster. The number of clusters may be forced (predetermined or adjusted after looking at preliminary results) or the clusters may be self-selecting if the map used in the experiment has a few distinctive types of curves. After clusters have been identified, one centrally located member from each cluster will be selected to become the representative of the cluster; and the map is redrawn using the representative in place of other cluster members. The maps will then be checked for appearance changes and possible inconsistencies due to curve intersections. Computer requirements for the two operations will also be evaluated in the course of the experiment.

<u>Method II. Dense families and approximation theory</u>
The second method chooses a family of representative curves without regard for the particular distribution of curves on a single map. This second approach uses several principles of approximation theory to determine a growing family of curves which may be used to approximate any other curve of a larger family with any precision desired. The growing family of curves used for the approximating is an infinite family, but we may truncate that family at any time and use only a finite subfamily to get within a predefined distance of our desired curve. That distance may be the width of our penplotter, for example.

One simple example of an approximating family is the family of non-intersecting polygonal lines with vertices on a finite grid whose extension is allowed to increase and whose mesh is allowed to grow finer and finer. As the mesh grows finer and finer, the number of possible polygonal lines increases rapidly. Nevertheless, a subfamily of those polygonal lines will be selected to approximate any curve, and as in the experiment describing Method I, we will draw maps with the subfamily and assess the appearance of the results.

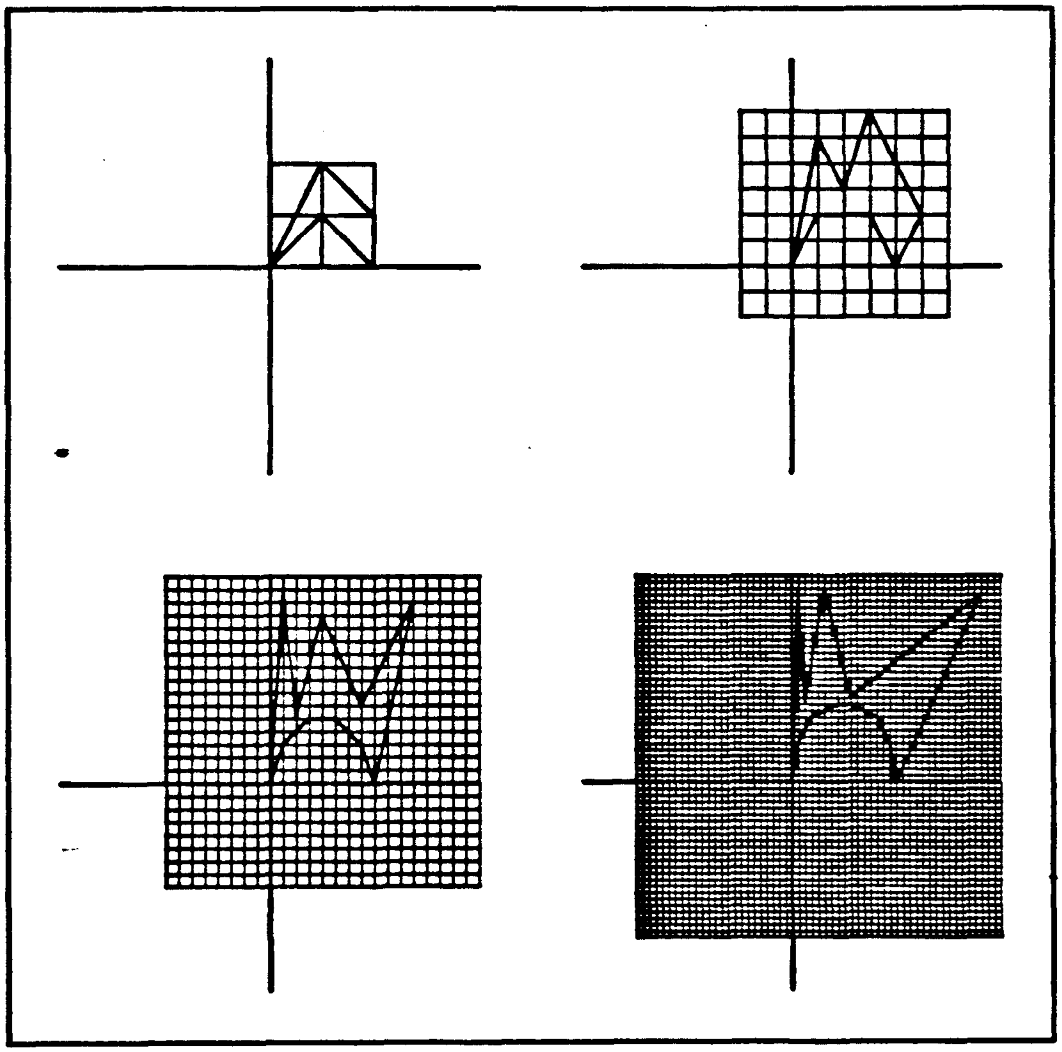

Figure 8. Examples of grids of increasing extent and finer mesh, with illustrative polygonal lines.

Note in Figure 8 that curves beginning at (0,0) and ending at (1,0) do not necessarily remain within one unit or even two units of the base line segment. For this reason, the extent of our mesh must be arbitrarily large. Note also that the finer meshes allow both smoother representations and more jagged curve possibilities because of the sharper angles possible.

## HIGHLIGHTS OF PROPOSED CURVE MANAGEMENT APPROACH

The storage requirements and the shape file access requirements are clearly reduced to a fixed small size instead of being proportional to the map size. Shapes can be reconstructed from the standard-position shape with varying precision requirements. For example, if the shape to be redrawn away from standard position is a semi-circle, and the redrawn size of the semi-circle is to be large, then many points on the semi-circle may be computed and transformed to give the polygonal line approximation to the semi-circle a smooth appearance. However, if the redrawn semi-circle is to be relatively small, then fewer points should be evaluated, transformed, and linked together in a polygonal line approximation.

Because the transformations described above as T and $T^{-1}$ are easy to apply to points and not so easy to apply to arbitrary curves, the approach to transforming an arbitrary curve is to evaluate the curve at a number of points (that number will depend on the quality of the redrawn approximation desired), then transform those evaluated points by the transformation, then rebuild the curve by linking the transformed points in a polygonal line. Some variations to this are possible, such as relinking the transformed points with a spline fit.

With the old method of keeping coordinates of curves as a separate large file, when a map transformation such as an alternative projection was applied to the map, all of the coordinates had to be recomputed. For a family of standard-position curve records, in many cases the new shapes for an alternative projection will not change or may be computed from the standard-positions curves directly.

## CONCLUSION

An alternative approach to storing and representing curves on maps offers potential improvements of efficiency and appearance. Experiments to be conducted in 1986 will determine the feasibility and usefulness of the new approach.

## REFERENCES


Asplund, E., and L. Bungart, 1966, A First Course in Integration, Holt, Rinehart, and Winston, New York.

Broome, F., 1984, "TIGER Preliminary Design and Structure Overview: The Core of the Geographic Support System for 1990," A.A.G. Meeting, Washington, DC.

Foley, J., and A. Van Dam, 1984, Fundamentals of Interactive Computer Graphics, Addison Wesley, Reading, MA.

Pavlidis, T., 1982, Algortihms for Graphics and Image Processing, Computer Science Press, Rockville, MD.

Pennisi, L., 1963, Elements of Complex Analysis, New York, Holt, Rinehart and Winston.

## APPENDIX: AN $L_2$ DISTANCE BETWEEN POLYGONAL LINES

This appendix describes a computationally straightforward formula for computing the $L_2$ distance between polygonal lines parametrized by the same normalized parameter.

Suppose that the curves $\zeta$ and $\xi$ are polygonal lines parametrized by t on [0,1] where $\zeta(t) = (x_\zeta(t), y_\zeta(t))$, and $\xi(t) = (x_\xi(t), y_\xi(t))$; and let the sequence:

$$\{0=t_0, t_1, t_2, \ldots, t_n=1\}$$

consists of all of the parameter values for which either of the two curves changes direction. Suppose that:

$$\zeta(t_i) = (\alpha_{\zeta i}, \beta_{\zeta i}) \text{ and}$$

$$\xi(t_i) = (\alpha_{\xi i}, \beta_{\xi i}), \text{ for } i = 0,1,\ldots,n.$$

Then:

$$\mathrm{Dist}^2(\zeta,\xi) = \int_0^1 \|(x_\zeta(t), y_\zeta(t)) - (x_\xi(t), y_\xi(t))\|^2 \, dt$$

$$= \int_0^1 [(x_\zeta(t)-x_\xi(t))^2 + (y_\zeta(t)-y_\xi(t))^2] \, dt$$

$$= \sum_{i=0}^{n-1} \{(t_{i+1}-t_i)[(\alpha_{\zeta i+1}-\alpha_{\xi i+1})^2$$

$$+(\alpha_{\zeta i+1}-\alpha_{\xi i+1})(\alpha_{\zeta i}-\alpha_{\xi i})+(\alpha_{\zeta i}-\alpha_{\xi i})^2]/3\}$$

$$+\sum_{i=0}^{n-1} \{(t_{i+1}-t_i)[(\beta_{\zeta i+1}-\beta_{\xi i+1})^2$$

$$+(\beta_{\zeta i+1}-\beta_{\xi i+1})(\beta_{\zeta i}-\beta_{\xi i})+(\beta_{\zeta i}-\beta_{\xi i})^2]/3\}.$$

This formula can be used to measure nearness for curves in standard position in order to select an evenly distributed family of such curves and also to study clustering of curve shapes.