A Transformational View of Cartography †

W. R. Tobler

ABSTRACT. Cartographic transformations are applied to locative geographic data and to substantive geographic data. Conversion between locative aliases are between points, lines, and areas. Substantive transformations occur in map interpolation, filtering, and generalization, and in map reading. The theoretical importance of the inverses is in the study of error propagation effects.

Leonard Bernstein, in a recent television lecture, made an exciting, and largely successful, attempt to describe musical concepts in terms of Noam Chomsky's ideas concerning transformational grammars, as originally devised for linguistics.[2, 4] A similar, though less ambitious, attempt is made here to look at a range of cartographic activities from a transformational point of view. The treatment is not particularly Chompskian, although some work on picture languages is available.[7, 11, 21, 23]

The idea of transformations is hardly new to cartography. The well-known example is the procedure by which we associate locations on the two-dimensional surface of the earth with locations on the two-dimensional surface of a piece of paper.[16] The study of the subject of map projections once constituted the bulk of what was known as "mathematical cartography." An extension 'was suggested in 1958 by Julian Perkal:

> *"Cartographic transformations, employed to represent the surface of the earth on maps, are of two simple types: map projections and generalization. Map projections are obtained by an objective mathematical operation. The subject of this paper is the second cartographic transformation — map generalization."*

Perkal then proceeded to derive mathematical rules for a particular method of generalizing lines on maps. As can be seen in his paper the method embodies a notion of spatial resolution.[20] Since that time considerable further work has been done on line generalization, and it is now known that the problem admits of more solutions than were recognized in Perkal's early paper.[9, 25, 28, 41]

At the time that Perkal wrote, computer cartography had not reached the large scale implementation which it now enjoys. Automatic plotters were not commonplace and he did not have available the computer tapes, now in widespread use, of world coastlines and hydrography, of political boundaries, of city street patterns, or of world topography. These fundamental geographical data can now be processed automatically, either to solve geographical problems directly, or to provide geographic illustration in the form of pictures. Thus, a broader view is required of mathematical cartography and has elsewhere been labeled analytical cartography.[35] This is the changing paradigmatic world of cartography.

Perkal recognized only two classes of cartographic transformation: map projections and map generalization. But the entire process of making, and using, a map can be viewed as a sequence of transformations. Original observations are manipulated and digested in various ways to obtain the data going into a map. In the design phase these are converted to a graphic representation, which is then transformed into a fair drawing and printed. Map use requires the assumption that the entire process is of high fidelity. A recent article by Morrison expands on this point.[18]

To give a simple example of how surveying fits into this framework consider a trilateration and its adjustment. Distances (d_{ij}) between places are observed and these are then converted to locations described by a set of position vectors (r_i, r_j), thus (d_{ij}) -> (r_i, r_j). This mapping has an inverse (r_i, r_j) -> (d_{ij}), and the operation is performed in such a manner that the product of the transformation and its inverse is, as nearly as is possible, the identity transformation, $T^{-1} T = I$. Triangulation is similar, but we substitute angles for distances and compute the transformation (${}_i\Theta_{jk}$) -> (r_i, r_j, r_k). For a two-dimensional manifold, such as the surface of the earth, a configuration is defined by 2n numbers, called coordinates, and there are n (n - 1) / 2 possible distances. For such an overdetermined system it suffices to start from a rank ordering of the distances and still obtain a metrically determinate configuration.[15, 24] The usual least squares adjustment algorithms are of course most easily implemented when the observations have the properties of measurement on a ratio scale.[36]

In an abstract study of transformations one would ask whether the inverse exists for a particular transformation, but more general questions are also possible. Does it commute? Is it differentiable? Linear? Closed? And so on. One should keep such questions in mind in the ensuing discussion, even though they are not raised explicitly at all times. Examining cartographic operations from this point of view suggests that they all fall into two classes. The first class consists of those operations that manipulate the locative aspects of the geographical data. These might be called the geometrical transformations. The second class of transformations are those that modify the substantive geographical data. This twofold distinction is not completely satisfactory but does serve as a convenient membrane to separate map projections from, say, map generalization, the two classes recognized by Perkal.

A map projection is defined as an association between places on the earth and places on a piece of paper. This is most easily visualized by imagining sections of string with one end attached to a globe and the other end attached to the paper. A notation for this would be $(\varphi, \lambda) \rightarrow (x, y)$, where (φ, λ) denotes the latitude and longitude of locations on the earth, and (x, y) are the *names* that we give to locations on the paper. We normally place some restrictions on the string attachment procedure and do not allow it to be completely random. The topological requirements of one-to-oneness and neighborhood preservation are the most fundamental.

Sometimes we go further and specify the rules by equations, and these are usually also required to be differentiable. But we do not always have the equations given explicitly. For example, in photogrammetry we *assume* that there is a projective relation between the photograph and the earth. This assumption is then translated into a particular set of parametric equations. These equations are subsequently calibrated, using careful measurements, in the belief that they represent a valid model of the association between the two sets. One of the difficulties of working with sensor imagery obtained by a mechanical scanning process, as is currently used in some earth satellite systems, is that one does not have a completely specified model in the form of parametric equations or else the observations are insufficient to calibrate the model adequately. The same problem occurs when working with old maps, where one attempts to obtain the equations in an *a posteriori* fashion, or when one studies the mental maps which people have of their environment.[33, 34]

The more usual case is that one is interested in a particular aspect of, or feature on, the surface of the earth. One then has a basis for adding further restrictions to the map projection equations. These are normally in the form of one or more partial differential equations, and these constrain the string attachment to within more or less tight bonds. Depending on one's field of interest one thereby obtains a wide variety of map projections, e.g., projections for surveyors are different from those for navigators and these in turn differ from those most useful to geographers.[17] A successful projection in one field may be, and usually is, absurd for other purposes. The cartographer concerned with the tracking of earth satellites may bend the meridians so that the suborbital paths become straight lines; the geographer interested in the flow of ideas may wish a "New Yorker's View of the United States," computed from a map projection which has captured Hägerstrand's mean information field as a differential equation.[14, 33] The administrator or statistician may desire a map on which the territory is depicted in proportion to the population distribution,[27, 31] i.e., satisfying $\partial x \partial y / \partial \varphi \partial \lambda - \partial x \partial y / \partial \lambda \partial \varphi = D(\varphi, \lambda) R^2 \cos \varphi$. The planner may use a map based on the preservation of highway mileages rather than the preservation of great circle distances, or one based on a geometry induced by travel information measured in minutes rather than kilometers. But in each instance one operates on the assumption that a problem solved in the (x, y) coordinates of the map projection also solves a closely related problem on the earth, and it somehow does this more easily than if working in (φ, λ) coordinates directly. Mercator's projection provides the classic example, solving the algebraic problem of finding the intersection angle between a meridian and a logarithmic spiral on a sphere. Whether this is done graphically on a piece of paper or algebraically inside of a computer is of no great import since Descartes showed that there exists a one-to-one correspondence between the geometric points of a sheet of paper and the algebraic set of real number pairs, and between drawn lines and abstract equations.

A map projection is an operation applied to points representing places on the surface of the earth, and these points are usually referred to by their latitude and longitude names. But it is possible to give a single point on the earth's surface several different names.[1, 5, 42] A class of transformations which has become important in today's computer environment is the conversion between these *aliases.* Geographers are familiar with the State Plane Coordinates or their equivalents. But there are many other schemes. Thus in the western United States one can identify locations by their designations in the Public Land Survey; more generally places are located by a mailing address, or by a street intersection, or by a telephone number, etc. There now exist computer programs that will allow conversion between some of these aliases - mailing addresses to State Plane Coordinates, for example.[8] Some of these conversions are performed by explicit equations, some consist of table look-ups, others combine these approaches. Are these transformations continuous? Differentiable? Does logical multiplication with the inverse yield the identity transformation? What are the error propagations, etc.? More theoretical study seems needed in this area.

Cartographers have a penchant for thinking of data in terms of points, lines, and areas. This scheme can also be applied to the locative information. Some places are considered *points.* Others can be conceived of as *lines;* the Mississippi River or U.S. Route 66 are examples. Then there are *areas;* Columbus (Ohio), or the State of Ohio, or zip-code area number 43210, or the telephone area code number 614, or the entire United States. Just as knowledge of the latitude and longitude of places allows specification of some relations between them, so also does knowledge of their other names. For example, telephone area code numbers that are near each other, numerically, represent places that are not near one another, geographically.

The classical cartographic conversions, as represented by plane coordinates, are between point locations. Let us enlarge the set of transformations to include the other ways of specifying locations. Thus, the three-by-three table - of points, lines, and areas versus points, lines, and areas - contains nine possible groups of transforms, with of course many specific instances and subdivisions in each category. An exhaustive examination of each case would be tedious, but a few are worth special mention. All of them raise fascinating questions related to optimal and redundant coding, information content, spatial tolerances, and inverses. One of the simplest, and also most widely used, is the point-name to area-name conversion. A realistic and practical example would be the on-line transformation from street address to police precinct. In one approach one would simply have a directory which associates each street address range with precinct name. Another scheme would be to convert the street address to State Plane Coordinates and the precinct boundary would be defined as an ordered list of numbers in these same coordinates. One then invokes a point-in-polygon algorithm to compute insidedness.

An area-to-area conversion could be by aggregation within a hierarchical system. The City of Columbus is an area contained within the area called Ohio. The aggregation decreases the spatial precision just as does the dropping of a digit in the latitude/longitude place-naming system. More interesting is the conversion between two distinct areal schemes: given that a set of locations is associated with a particular census tract, to which school district do these locations belong? This is the so-called polygon overlay problem.[40] A particularly popular set of areal polygons, in use since at least 1859, consists of quadrilaterals; crudely speaking "squares," on the surface of the earth.[19, 39]. An obvious example is in the recording of world population figures by five degree cells of latitude and longitude. Closely related is the sampling of data at equally spaced intervals from a continuous function by multiplication with a Dirac brush.[12, 22]. The digitized aerial photograph and the digital terrain model are examples.

Imagine that the elevations from a topographical map are stored as an equally spaced array of numbers. This array can be considered a matrix, in the sense of matrix algebra, and one can now use a positional notation for the rows and columns in order to index any location. Suppose that the array, call it **A,** is rectangular of order n by m; thus $\mathbf{A} = (a_{ij})$, i = 1,...,n; j = 1,..., m. Consider now a transformation of the topographic elevations by the matrix equation $\mathbf{A^* = VAH}$, where the square matrices **V** and **H** are chosen to conform to **A.** This cartographic transformation has several interpretations, one of which is map generalization. By an appropriate choice of **V** and **H** the generalization is such that it can be undone, or reversed.[29] The condition is that **V** and **H** have inverses, equivalently that the transfer function does not contain zeros in the frequency domain.[3, 30]. This, then, is an example of a transformation applied to the substantive, rather than the locative, data. The particular example can be extended to the case of randomly located data, i.e., arbitrarily arranged topographic elevations, or to population counts by irregularly shaped census tracts;[32] or to vectorial or tensorial data, and even to the case of non-numerical, categorical data such as land use types.[13] A multivariate instance is the mapping from four channels of Landsat imagery into a classification of cover type. A possible measure of the effectiveness of such a classification is the departure of the inverse from the identity mapping.

Quite similar is the transformation from a sample of data, elevations say, from which one interpolates to a continuous scalar field. The transformation from the field back to the original observation points should yield the identity. This is the interpolation-prediction problem, well known in cartography.[37] Since the transformation from the sample to the spatial field is, literally, one-to-many, it is necessary to invoke a model. This is chosen such that the product of the transformation with its inverse is the identity transformation, as nearly as is possible, and in the averaged least squares sense.

Closely related is the problem of isopleth mapping. Suppose one has counts of the number of individuals by county. From these data one might like to know how the population density, a continuous quantity, varies over the region of concern. What is needed is a procedure for the preparation of a contour map from the aggregate data. The map should clearly be such that the "volume" under the contoured surface within the limits of the irregular polygon should be exactly the same as the observations, the data enumerated for that polygon. This volume-preserving condition is the fundamental inversion requirement that renders the map of sufficient fidelity to be of use. But it is easy to see that there can be more than one volume-preserving contour map for any one set of observations. Additional criteria therefore need to be added to make the transformation from data to isopleths unique. Minimization of some function is mathematically the most natural - the smoothest possible continuous isopleths are desired, of minimal curvature.[38] Alternately, a convolution kernel which most closely approximates (in mean square) the actual probability density function for the locus of activity of individuals could be used to yield a process-oriented contour map, taking into account the spatial non-stationarities of the phenomena.

It is often of interest to bring information stored in a "geographical data bank" into the consciousness of an individual. Frequently this is done by producing pictures, "maps," which are then transmitted via the visual system to the cortex. This is the Pandora's box of cartography as a communication system. Any given set of data can be converted to many possible pictures. Each such transformation may be said to represent some facet of the data, which one really wants to examine as if it were a geological specimen, turning it over in the hand, looking from many points of view, touching and scratching. The effectiveness of the data picture transformation might be measured by how close the inverse comes to reproducing the original, i.e., the extent to which the geographical data can be recreated from the consciousness of the individual. There is, of course, no requirement that the transmission be via visual channels. One can imagine a computer system that examines a digital terrain model and then verbally, via a voice synthesizer, describes the topography in terms relevant to the task at hand.

There are several difficulties here. Virtually nothing is known about what happens to visual information after it reaches the cortex. We have no useful model of the functioning of the human brain, Stockham's work excepted,[26] but this fortunately is an active research area. Furthermore, the facets of the data which should be emphasized are given exogenously and are not in the data. There is some evidence that a "purpose" causes people to think in terms of "structures." A microscopic slide of the liver and an aerial photograph of a Florida swamp are indistinguishable to people if they cannot discern the structure. What is a geographical structure? Presumably it is a transformation of some geographical data, a theory or model of reality. We have returned to the world of Chomsky and Bernstein, where cartographers have yet to tread with assurance.

REFERENCES

1. Barraclough, R, "Geographic Coding," *President's commission on Federal Statistics,* G.P.O., Washington, D.C., (1971), 219-296.

2. Bernstein, L., "The Charles Eliot Norton Lectures, 1973," published with three 33½, rpm records as *The Unanswered Question: Six talks at Harvard,* Harvard Univ. Press, Cambridge, (1976).

3. Burr, E, "Sharpening of Observational Data in Two Dimensions," *Australian Journal of Physics,* 8(1955), 30-53.

4. Chomsky, N., *Reflections on Language,* Pantheon Books, Boston, (1976).

5. Clayton, K., 'Geographical Reference Systems," *The Geographical Journal,* 137:1 (1971), 1-13.

6. Connelly, D., "The Coding and Storage of Terrain Height Data: An Introduction to Numerical Cartography," M.S. thesis. Cornell Univ., Ithaca, (1968).

7. Dacey, M., "A Picture Language for Periodic Strip Patterns," *Geographical Analysis,* 3:4 (1971), 320-333.

8. Dial, R., "Street Address Conversion System," Urban Data Center Report No. 1, Univ. of Washington, Seattle, (1965).

9. Douglas, D., and T. Peucker, "Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature," *The Canadian Cartographer,* 102(1973), 112-122.

10. Ferguson, H. R, *Point in Polygon Algorithms,* Univ. of Washington, Seattle, (no date).

11. Fu, K. S., *Syntactical Method in Pattern Recognition*. Academic Press. New York, (1976).

12. Goodman, J., *Introduction to Fourier Optics,* McGraw-Hill, N.Y., (1968).

13, Guptill, S., "Spatial Filtering of Nominal Data," Ph.D. thesis, Univ. of Michigan, Ann Arbor, (1975).

14. Hägerstrand, T., *Innovationsforloppet ur korologisk synpunkt,* Gleerup, Lund, (1953).

15. Kendall, D., "Construction of Maps from Odd Bits of Information," *Nature,* 231 (1971), 158-159.

16. Lambert, J., "Notes on Maps, 1772," introduced and translated by W. Tobler, *Michigan Geographical Publication No. 8,* Univ. of Michigan, Ann Arbor, (1972).

17. Maling, D., *Coordinate Systems and Map Projections,* G. Philip. London, (1973).

18. Morrison, J., "Towards a functional Definition of the Science of Cartography with Emphasis on Map Reading," Studies *in Theoretical Cartography.* E. Kretschrner, ed., Deuticke, Vienna. (1977) 175-198. Reprinted in *The American Cartographer.* 5:2 (1978) 97-1 10.

19. Nordbeck, S., "Statistiska Kartor Fran 1800-Talets Mitt," *Svensk Geografisk Arsbok,* 40 (1964), 7-18.

20. Perkal, J., "An Attempt at Objective Generalization," translated from "Proba Obiektywnej Generalizacji." *Geodezja I Kartografia,* VII, 2 (1958), 130-142; Michigan Interuniversity Community of Mathematical Geographers Discussion Paper No. 10, (1966), Univ. Microfilms No. OP33067. Ann Arbor.

21. Pfaltz, J., "Web Grammars and Picture Description," *Computer Graphics and Image Processing.* 1, 2 (1972), 193-219.

22. Rosenfeld, A., *Picture Processing By Computer.* Academic Press, N.Y., (1969).

23. Rosenfeld, A. and A. Kak, *Digital Picture Processing.* Academic Press, N.Y., (1976).

24. Shepard, R., "Metric Structures in Ordinal Data," *J. Math. Psychology,* 3 (1966), 287-315.

25. Stegena, L., "Tools for the Automation of Map Generalization," *Automation: The New Trend in Cartography,* E. Csati, ed., Budapest, Int'l Cartographic Assn., (1974), 66-95.

26. Stockham, T., "Image Processing in the Context of a Visual Model," *Proceedings,* IEEE, 60:7 (1972), 828-842.

27. Tobler, W., ‘Geographic Area and Map Projections,” *The Geographical Review,* 53, 1 (1963), 59-78.
28. Tobler, W., “An Experiment in the Computer Generalization of Maps,” Office of Naval Research Report, Task 389-137, Univ. of Michigan, Ann Arbor, (1964).
29. Tobler, W., “Of Maps and Matrices,” *Journal of Regional Science,* 7 (supplement) (1967), 275-290.
30. Tobler, W., “Geographical Filters and their Inverses,” *Geographical Analysis,* 1:3 (1969), 234-253.
31. Tobler, W., “A Continuous Transformation Useful for Districting,” *Annals,* N.Y. Academy of Science, 219 (1973), 215-220.
32. Tobler, W., “Linear Operations Applied to Areal Data,” *Display and Analysis of Spatial Data,* Davis, J. and M. McCullagh, eds., John Wiley & Sons, N.Y., (1975), 14-37.
33. Tobler, W., “Mathematical Map Models,” *Proceedings,* International Symposium on Computer Aided Cartography (Auto-Carto 11), American Congress on Surveying and Mapping and US. Bureau of the Census, (1975), 66-73.
34. Tobler, W., “The Geometry of Mental Maps,” in *Spatial Choice and Spatial Behavior,* Golledge, R. and G. Rushton ,eds, Columbus, Ohio State Univ. Press, (1976), 69-81.
35. Tobler, W., “Analytical Cartography,” *The American Cartographer.* 3:1 (1976), 21-31.
36. Tobler, W., “Numerical Approaches to Map Projections,” *Studies in Theoretical Cartography,* E. Kretschmer, ed., Deuticke, Vienna, (1977), 51-64.
37. Tobler, W., “Lattice Tuning,” *Geographical Analysis,* 11:1 (1979), 36-44.
38. Tobler, W., “Smooth Pycnophylactic Interpolation for Geographical Regions,” *Journal of the American Statistical Association,* to be published Sept. 1979.
39. Tobler, W., “Cellular Geography,” Gale, S. and C. Olsson, eds., *Philosophy in Geography,*(1979).
40. Tomlinson, R., *Computer Handling of Geographical Data,* UNESCO Press, Paris, (1976).
41. Töpfer, F, *Kartographische Generalisierung.* Haack, Gotha, (1974).
42. Werner, P., “National Geocoding,” *Annals,* Assn. Am. Geographers, 64:2 (1974). 310-317.

Dr. Tobler is professor of geography, University of California, Santa Barbara, CA 93106. This article is reprinted, with minor modifications by the author and with the permission of the editor, U. Uotila, from the *Proceedings* of the International Symposium on the Changing World of Geodetic Science held Oct. 6-8, 1976, Ohio State University, Columbus, Ohio.