In the late 1960's a new course, entitled Analytical Cartography, was introduced into the geography program at the University of Michigan. It is the purpose of this talk to explain some of the motivation for this course and to illustrate some of the content by a selection of examples. In such a short lecture it is naturally not possible to cover all of the details, but some of the flavor may be suggested. An understanding of the situation in geographical cartography in the United States in the early fifties of the decade, as I perceived it, is helpful as a background. Cartography has been a university subject in the United States since the late thirties of this century. Of the 3,000 or so colleges in the United States, cartography is taught in most of those which offer geographical instruction, but less than a dozen have developed research specializations in this area. Of these one should add a small number of engineering schools which have professional programs in surveying, photogrammetry, or geodesy. For various reasons there was a gap between official governmental cartography and academic geographical cartography. Geography was once associated with geology, but it has now, at least in North America, moved away from this focus to become a social science. The tradition is thus somewhat different than that of continental Europe. Geography in the late fifties and early sixties went through the so-called "quantitative revolution," in which statistical description replaced verbal description, and formal abstract model building was recognized as superior to anecdotal explanation. Many of the young people entering college teaching in this period soon found themselves teaching cartography. Unfortunately, the production of graduates from schools specializing in cartography at the graduate level seems to have been insufficient to fill all of the undergraduate instructional positions available in this field. As a consequence, the newest staff member was often assigned the introductory cartography course, in which he might not have had a great deal of training and, occasionally, a commensurate interest. But he knew a lot about multiple regression, factor analysis, and related statistical techniques. Thus, courses called Statistical Cartography came into being. Here students learned correlation techniques, trend analysis, and so on, all legitimate topics, but hardly cartography. The subject also seemed sometimes to be treated as a "technique," which could be taught without special qualifications by any staff member, and this attitude certainly could not lead to any expansion of the theoretical core of the field. At the same time it was clear that computers would play a crucial role in the future of cartography. The first computer maps having been produced in 1951, and that many changes were needed in traditional cartography (Tobler, 1959). The whole seemed to be changing so fast that 50 percent of what one learned was obsolete within five years. I had the hope that my lectures would have a half life of 20 years. The problem was to revise drastically the advanced cartography curriculum. The eventual result was Analy-
ical Cartography, as described in the course outline included here. Only minor modification of details has been required in the subject since 1969. It is perhaps also appropriate to draw especial attention to the contrast between this direction of development of the field of cartography with other developing directions. Much of the current literature, for example, emphasizes cartography as geographic illustration, in which communication is paramount (Robinson and Petchenik, 1975).

A popular title would have been Computer Cartography. This did not appeal to me because it is not particularly critical which production technology is used. Such a title would also imply the existence of other courses, perhaps titled Handicraft Cartography, or Pen and Ink Cartography. The substance is the theory which is more or less independent of the particular devices; equipment becomes obsolete rather quickly anyway. Mathematical Cartography could have been used for the title, but this already has a definite meaning (Graur, 1956; Solovyev, 1969), and I had in mind more than is usually covered under this heading. Cartometry, the study of the accuracy of graphical methods, is another available term, but has a rather narrow meaning. One could also speak of Theoretical Cartography. This did not appeal to me on two grounds. It would frighten students, who are always concerned that they learn something practical. Secondly, the precedent is not very attractive. Max Eckert, for example, wrote a great deal about theoretical cartography but did not solve many problems (Eckert, 1921/25). I wished to emphasize that mathematical methods are involved, but also that an objective is the solution of concrete problems.

As regards the substance, rather than the course title, the major difference is perhaps only that a somewhat more general view is taken of the subject. It is appropriate to introduce students to what is already in the literature, to introduce similar concepts which occur in other branches of knowledge, and to suggest new directions. Embarrassingly frequently cartographers claim as unique problems those which also occur in other fields and which may even have been solved there. All professions must constantly fight their myopia. But clearly the application of mathematical methods to cartography is growing rapidly. Thus a simple introduction to Analytical Cartography is through Photogrammetry and Geodesy. These fields have a long mathematical tradition and a healthy literature to which the student need only be referred. The principal equations of the method of least squares and its newer derivatives, of the theory of errors, and some theorems of projective geometry and of potential theory, etc., are all easily reviewed. One can also see tendencies such as the direct production of mosaics from aerial photographs and the computer recognition of objects seen by an imaging system (Duda and Hart, 1973). Another trend is the replacement of triangulation by trilateration.

To see how a more general view is useful, consider this last topic, in which one determines locations from measurements. Suppose that we have identified \( n \) points on the surface of the earth. Between these there exist \( n(n-1)/2 \) distances, \( d_{ij} \). Let us assume that all of these distances have been measured and that the locations of the points are to be found. This means that \( 2n \) coordinates must be determined. In high school one learned that one could compute distances from coordinates by use of the Euclidean formula:

\[
\sqrt{(x_i-x_j)^2 + (y_i-y_j)^2} = d_{ij}.
\]

The surveying problem reverses the arrow, and every surveyor knows how to find the coordinates when given the distances. In this problem the value of \( n(n-1)/2 \) grows much faster (essentially quadratically) than does \( 2n \) and there are thus more equations than unknowns. This has three useful consequences: (1) No solution satisfies all equations, since all empirical measurements have error. This leads naturally to a discussion of least squares methods, error ellipses, and iterative solution procedures. But one also obtains internal checks on the accuracy of the meas-
One only needs to know the ordinal relations of the distances in order to obtain a solution (Shepard, 1966), which comes as quite a shock to the cartographer who strongly believes in numbers. Only 2n measurements are really necessary, and one is then led to consider the “optimal” positioning of surveying measurements, a rather recent development in the literature (Grafarend and Harland, 1973). A cartographic application of this “geodetic” technique might be as follows: Suppose one is successful in getting people to complete a questionnaire in which they are required to estimate the distances between, say, prominent buildings in Vienna. The result of this empirical operation yields data similar to the measurements obtained from a geodetic survey in that it is possible to compute coordinates and their standard errors (Tobler, 1976). One can thus obtain a type of mental map (Gould and White, 1974) and its degree of variance. Comparison with geodetic positions enables one, using Tissot’s Theorem, to measure the amount of distortion of these mental maps. Analyzing such data over time may reveal the rate of spatial learning, which is geographically a very interesting question. A similar technique, incidently, is used in psychology when one attempts to portray separations between personality types. In some respects the psychology literature on trilateration, under another name (Shepard, 1962) is more advanced than the geodetic literature.

Two other examples of old topics in new guises stem from the subject of map projections. In the 1880’s Francis Galton invented the geographical isochrone, a line connecting all points which can be reached in a given time. Isochronic maps are now quite popular, and the concept has been generalized to include travel costs (isotims). These are really geographical circles, recalling that a circle is the locus of points equidistant from a center point. Measure distance in units of time and you have a geographical circle. But what curious circles! They have holes in them, and disjoint pieces, and the ratio of circumference to radius is hardly 2π. What a curious geometry—it makes Einstein seem simple! Or consider a set of concentric isochrones. Now draw in the orthogonal trajectories and one has the equivalent of a set of polar coordinates. Technically these are the polar geodetic coordinates of Gauss, familiar from differential geometry, and for which the metric takes on a particularly simple form. One can draw maps of this geographical geometry by using the usual ideas from the study of map projections, but the reference object is no longer a sphere or spheroid, rather it is more like a pulsating Swiss cheese. As a second map projection example, consider the problem of dividing the United States into compact cells, each of which contains the same number of people, and how it might be approached by a map projection. Figure 1 shows how deliberate distortion can be used to advantage in solving this practical problem. The differential equations covering this situation have been published elsewhere (Tobler, 1973).

Computer graphics are introduced in the second week of the course. One knows, a priori, that all maps which can be drawn by hand can also be drawn by computer-controlled devices. This follows from Turing’s (1936) theorem. Of course can does not imply should. Students are given a short introduction to the equipment which provides a realization of these ideas and to the sources from which they can obtain geographical data tapes and computer programs. Examples are run using the data and equipment available in the local environment. The data tapes include world outlines, county boundaries, street patterns, census tabulations, topographic elevations, etc. Equipment facilities at the University of Michigan include a large computer connected to telephones and tele-type terminals, or graphic devices can be attached to any user telephone for interactive use. The system is outstanding in its ease of access for novices. Students write one program to draw a simple map of their choice in one of the programming languages.
The user of geographical data is, in principle, indifferent as to whether the data are on a geographical map or on magnetic tape. One of the principal uses of geographical maps is that of a graphical data storage device. But tapes are often more convenient than are drawings. These are simply two alternate methods of storing geographical information. We can assert that, when one has enough information on a magnetic tape to be able to draw a geographical map, one also has enough information on the tape to be able to solve all of the problems which could be solved using that map. But to store geographical information in electronic form in such a manner that a geographical map can be drawn requires that the substantive data be given geographical referencing. The usual procedure is to reference points, lines and areas by coordinates: geodetic latitude and longitude, or Gauss-Kruger coordinates, etc. (Maling, 1973). So these must be treated in the analytical cartography course. But this is really too narrow a point of view. Call the fire department and announce that there is a fire in this room, at 48°15'22"N, 16°23'10"E. It would never be done; one would use the street address or the building name. But
there are no cartography books which describe the street naming/numbering system. If I can locate a house using the street address then this label must contain exactly the same amount of information as does the latitude/longitude designation (Huffman, 1952). The telephone area code number for Vienna, 0222, locates this place to circa ±20 km. If I call from the United States to a phone in Vienna, I need to dial 12 digits, and these digits locate an area not much bigger than 1 sq. m. If I know the postal code for Laxenburg, 2361, then I have specified a region to ±5 km.

Equivalently, Gauss-Kruger coordinates can be calculated from latitude and longitude \( \phi, \lambda \rightarrow G, K \) and this is invertible \( G, K \rightarrow \phi, \lambda \). Make a list of as many ways as you can recall of how locations are identified. Some will define points, others will refer to regions (Werner, 1974; Clayton, 1971). Now form a table by repeating this list in the orthogonal direction and consider this table as a transformer; place name to latitude and longitude, and the inverse, might be an example of two transformations. The concern in the literature with computerized address coding, the DIME system (Corbett, 1973), point-in-polygon programs, etc., all relate to these transformations (Barraclough, 1971). More exactly considered, coordinates are a way of naming places which, inter alia, allow all places to be given a unique name, and which allow relations between places to be deduced from their names. The North American telephone area codes, for example, have the property that if two area codes are similar, the places are most likely widely separated, and the converse. Thus the telephone area code scheme implicitly includes a relation between the places. From this relation one can use the trilateration procedure, already described above, to compute and draw a map of North America. In order to do this simply use the implicit relation and then interpret "A is near to (or far from) B" as "adjacent (or non-adjacent)," and then compute \( 2n \) coordinates from the set of these \( n(n-1)/2 \) adjacencies. For details see the delightful paper by Kendall (1971). The point of these remarks is that there is a great deal of theoretical structure buried in the topic of locational coding, and it has remained almost completely unexplored by theoreticians.

When one puts geographical information into a computer, one finds that it is extremely voluminous. A four color map of size \( 10 \times 10 \) cm. contains perhaps \( 100 \times 100 \times 4 = 40,000 \) discrete elements, not a great amount for a computer these days, but this is a small map. One can ask whether all of these elements are necessary. It is clear that not all of the \( 4^{10000} \) possible four color maps of this size can occur, because of the inherent geographical structure. One could throw away a goodly proportion of the \( 4 \times 10^4 \) elements and still have a very useful map. I use the word "goodly" for lack of a numerical estimate. Attempts have been made to apply information theory to this type of cartographic situation, but the bi-dimensional frequency statistics which seem needed are not generally available, except in the case of television pictures (Connelly, 1968). In the lecture sequence these topics lead naturally into the question of geographical map simplification. The condensed book, the overture to a musical work, and a graphical caricature are all somewhat similar modifications of an original. A very related topic is also treated in economics and in sociology under the heading of optimal aggregation (Fisher, 1969; Hannan, 1971). Figure 2 shows an example of a simplification algorithm applied to an outline of Michigan. The method is simple, rapid, and seems effective (Douglas and Peucker, 1973). Figure 3 shows the application of an alternate technique to contours. One of the more appropriate methods for such surfaces seems to me to be that of two-dimensional filtering (Stegena, 1974). This method, although not the only one available, is exactly controllable and invertible (Burr, 1955; Tohler, 1969). Spatial filtering can also be applied to data aggregated by discrete, irregularly-shaped spatial units, e.g., enumeration areas as shown in Fig. 4. A comparable filtering even performs well when the observations are
Fig. 2. Simplification of a polygonal structure by a computer algorithm.

Pedagogically it seems best to introduce students to these methods by considering the field of "picture processing" (Rosenfeld, 1969; Andrews, 1970; Codd, 1968). Here one has data at evenly-spaced spatial intervals, and many important concepts can be explained most easily in this spa-
Fig. 3. Simplification of a topographic surface by spatial filtering. After Tobler, 1967.

Initially homogeneous environment. The notions can then later be generalized to include the more practical case in which the data arrive packaged in irregular spatial polygons or at random point sets. But a critical lack here is a useful model of the functioning of the human brain (but see Arbib, 1972; Stockham, 1972). The design of maps cannot be improved without such a standard against which to test visual effectiveness.

As a simple example of a problem in the processing of geographical data, one can take the case of map overlays, e.g., given a soils map of Michigan and a geological map of Michigan, find the logical intersection of the two. Conceptually this is not a difficult problem, and it can be done using a computer, as is illustrated in Fig. 5. One has several choices in the way in which the regions are stored in the computer. These options become quite technical, but, for example, an area can be described (Meltzer, Searle, and Brown, 1967) as

\[ F(x, y) = \begin{cases} 1 & \text{if in R} \\ 0 & \text{otherwise} \end{cases} \]

or the boundary can be described as an equation \( x(s) + iy(s) = z(s) \), or one can store the skeleton (Blum, 1967) of the region. From all of these representations (and there are more), one can compute the area of a region, calculate whether or not a point lies inside of the region, find whether two regions overlap, and draw maps. It turns out, however, that some representations are more convenient for particular purposes than are others, even though they are algebraically equivalent in the sense that each can be converted into all of the others (Valmer, 1975). A comparison to two methods of solving a pair of simultaneous linear equations may be appropriate. If one remembers a bit of algebra and that linear equations have the form

\[
\begin{align*}
Y_1 &= A_1 + B_1 X \\
Y_2 &= A_2 + B_2 X,
\end{align*}
\]

then the intersection point can be found by the simultaneous solution of these equations. But sometimes it is easier to plot the lines and to read the coordinates of the intersection from the drawing. Many uses of maps are of this nomographic nature. Recall that Mercator's projection, for example, provides a graphic solution to the problem of finding the intersection angle between a North-South great circle and a logarithmic spiral on a sphere. Or consider the problem of finding the nearest gas station when the automobile gauge indicates "nearly empty." With the entire
Fig. 4. Simplification of census data by spatial filtering. After Tobler, 1975.
street system stored in the core of a pocket calculator, and the location of all the gas stations which accept my credit card also stored, and with a minimal path algorithm (Gilsinn and Witzgall, 1973) which is efficient for networks of this size, it seems like a trivial computation.

What is easy, convenient, or difficult depends on the technology, circumstances, and problem. The teaching of cartography must reflect this dynamism, and the student can only remain flexible if he has command of a theoretical structure as well as specific implementations. The spirit of Analytical Cartography is to try to capture this theory, in anticipation of the many technological innovations which can be expected in the future; wrist watch latitude/longitude indicators, for example, and pocket calculators with maps displayed by colored light emitting diodes, do not seem impossible. In a university environment one should not spend too much time in describing how things are being done today. The course outline presented here tries to avoid this bias.

COURSE OUTLINE

Analytical Cartography. Geography 482. 3 credits. Prof. Waldo R. Tobler
University of Michigan, Ann Arbor, Michigan 48109, U.S.A.

Week I. Introduction. Relation to mathematical geography, geodesy, photogrammetry, remote sensing. Replacement of map data storage by computer data storage. Technological change and the need for theoretical approach. Historical perspective.


Week III. Geographical Matrices. Triangular, quadrilateral, hexagonal, and Escher types. Notation, neighborhood property, topological invariance. The varieties of geographical data: nominal, binary, scalar, complex, colored, N-valued, and infinite-valued matrices. Isomorphism to the surface of the earth.

Week IV. Geographical Matrix Operators. Functions of matrices: algebraic, logical, differentiable, invertible; linear, local, spatially invariant (translationally and rotationally). Parallel processing, windows, edge effects. Finite difference calculations.


Week VI. Sampling and Resolution. Fourier interpretations of aliasing, local limited functions, Nyquist limit, comb functions. The sampling theorem, random plane sampling, invisible distributions.

Vol. 3, No. 1, April 1976
Work VII. Quantization and Coding. Analogue and digital processing. Quantization error, reduction of. Information theory: how many aerial photographs are there? Huffman coding, higher order statistics, spatial autocorrelation functions. Television and choropleth maps.


Work IX. Pattern Recognition. Preprocessing, enhancement, feature extraction; discrimination and classification (linear, Gaussian); signal-to-noise ratios; perceptions.


Work XI. Generalized Geographical Operators. Expansion of matrix operators to irregular point sets, to interval data, in such a manner as to include matrix as a special case. Generalized two-dimensional sampling theorem and reconstructions from sampled data.

Work XII. Geographical Coding. Information theoretical content of Latitude/Longitude, street address, l'k code: telephone number, Public Land Survey, and the like. Topological and metrical properties of place naming schemes. Gaussian coordinates. A variety of plane coordinate schemes. Formulas for working on sphere and ellipsoid.


Work XV. Geographical Information Systems. Band width requirements: dollar requirements; hardware and software; input schemes, manipulation algorithms, output schemes. Historical overview and examples: TIROS-ERTS, CATS-PRATS, CLI-MLADS-DIME. Analytical approaches to using geographical data: optimization techniques, sensitivity testing, regionalization, spatial trend analysis, dynamic simulation, growth models, regional forecasting.

REFERENCES


The American Cartographer


---

**Call for Papers**

The Seventh International Conference on the History of Cartography will convene in Washington, D.C., August 7-11, 1977. The Conference will be jointly sponsored by The Library of Congress, Geography and Map Division; The Folger Shakespeare Library; General Services Administration, The National Archives and Records Service, Cartographic Archives Division; The Smithsonian Institution, National Museum of History and Technology.

Papers based on original research in any aspect of cartographical history will be considered for possible presentation in one of the technical sessions scheduled for August 8, 9, and 10. To assure appropriate consideration, send, by July 31, 1976, the title of your proposed paper and an abstract not exceeding 300 words, to: The Technical Program Committee, Seventh International Conference on the History of Cartography, c/o Library of Congress, Geography and Map Division, 815 South Fickett Street, Alexandria, Virginia 22304.