TOBLER, Waldo Rudolph, 1930–
MAP TRANSFORMATIONS OF GEOGRAPHIC
SPACE.

University of Washington, Ph.D., 1961
Geography

University Microfilms, Inc., Ann Arbor, Michigan

MAP TRANSFORMATIONS OF GEOGRAPHIC SPACE

by

WALDO RUDOLPH TOBLER

A thesis submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF WASHINGTON

1961

Approved by________________________________

Department in which degree is granted____________________

Date________________________________

UNIVERSITY OF WASHINGTON

Date: May 9, 1961

We have carefully read the thesis entitled Map Transformations of Geographic Space submitted by Waldo Rudolph Tobler in partial fulfillment of the requirements of the Doctor of Philosophy and recommend its acceptance. In support of this recommendation we present the following joint statement of evaluation to be filed with the thesis.

Mr. Waldo Rudolph Tobler's thesis is an important contribution to the field of geographical cartography. Maps and cartographic representations have been distinctive and much used descriptive and analytical tools of the geography profession. A major and continuing problem facing the geographer is that of devising ways and means of obtaining accurate representation of the earth and selected phenomena upon a plane. Work in map projections is addressed to various facets of this problem. Comparatively recently, cartographic geographers have also become concerned with the use of projections as analytical tools. In this connection, work has been premised on the appreciation of various kinds of spatial relationships which result from differences in the map users' concepts of space in terms of time, distance, cost, convenience, etc. Consequently, in problems where such concepts are important, e.g., industrial location, it should be possible to develop and use suitable cartographic ideas and techniques in their analysis.

In the first portion of Mr. Tobler's study he reviews the more conventional aspects of cartography, namely, the representation of the earth's surface on a plane, and develops an original, still tentative, classification of map projections including a simple graphic technique which is effective in describing and comparing them.

The second and major part of the thesis is concerned with the analytical use of map projections. Distortion inevitably results from the representation of the spheroidal earth on a plane. With great originality, Mr. Tobler relates these distortions, cartographically and graphically, with alternative conceptual distortions, or transformations, of various

THESIS READING COMMITTEE:

relationships having spatial qualities which are susceptible to more effective analysis when represented on other than conventional, physically-oriented map projections. He has provided numerous examples of transformations to demonstrate the usefulness of the techniques, particularly as they relate to research carried on in economic geography. He has also indicated further applications of his ideas and techniques to other areas of geographical research as well as ways in which they may be refined and developed.

The members of the reading committee have pleasure in recommending Mr. Tobler's dissertation as an example of the best kind of imaginative, thorough, and scholarly research in geography.

Thesis Reading Committee: ______________________________

____

____

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of Washington I agree that the Library shall make it freely available for inspection. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by my major professor or, in his absence, by the Director of Libraries. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature ______________________

Date ______________________

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

ACKNOWLEDGEMENTS

Numerous persons have contributed to this study, though, of course, the final responsibility remains mine. Dr. John C. Sherman has supervised my entire career as a graduate student--undoubtedly a frustrating task at times--and has also contributed directly through many suggestions, comments, and advice. To him I am most grateful. Drs. William L. Garrison and William W. Bunge, Jr. have offered encouragement and stimulation beyond measure. Drs. Marion E. Marts and Morgan D. Thomas have read the entire work and offered many suggestions. Dr. C. B. Allendoerfer has commented on the mathematical portions of the work. Mr. Jeremy Anderson conducted the tests mentioned in Chapter IV. Miss Jane Peet typed the entire manuscript and attended to all of the details which this entails. Others who have provided valuable comments are A. Getis, W. R. Heath, G. Kakiuchi, R. C. Kao, D. Marble, H. Porter, M. Teitz, and R. Wolfe. Many others to whom credit is due will be found in the Bibliography.

I also wish to express my appreciation to the Cartographic Laboratory of the Department of Geography for the use of its facilities, and to the Interlibrary Loan Division of the University of Washington for its assistance in obtaining many foreign volumes. The National Science Foundation, through its system of Graduate Fellowships, provided indispensable support for the academic years 1959-1961. I am also extremely grateful to my wife for all that she has contributed to the completion of this study.

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS--Continued

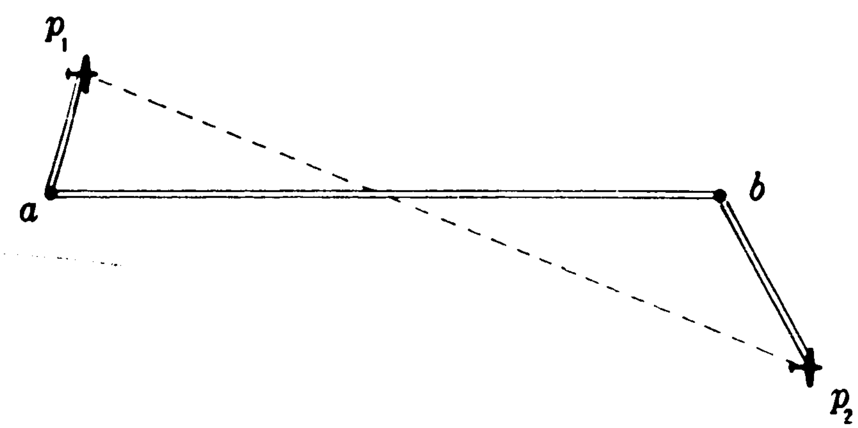
LIST OF ILLUSTRATIONS--Continued

LIST OF ILLUSTRATIONS—Continued

CHAPTER I

INTRODUCTION

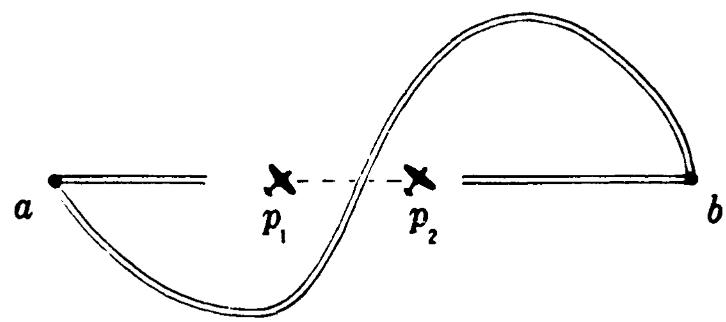
Imagine a simplified geographic arrangement as shown in the accompanying figure (Fig. 1.1). The double lines represent roads; the airplane symbols, P_1 and P_2, are airports; the dashed line is an air route; the dots A and B are two locations.



A Simplified Geographical Arrangement

Figure 1.1

It is clear from this map that the shortest distance from A to B is along the road connecting the two. Suppose, however, that it is desired to travel from A to B in the most rapid manner possible. Under certain conditions it requires less time to travel to the airport (P_1), wait for a plane, fly to P_2, wait for a bus, and then ride to B, than it would require to travel the shortest distance directly. Situations of this general nature are frequent and well known, resulting in expressions such as the phrase "as the crow flies." It would be desirable to be able to draw a map on which the shortest time-distance between two places is the straight line connecting those two places. For the simple situation given, this does not appear difficult. The geographic environment is to be transformed into an arrangement which looks somewhat as follows (Fig. 1.2):

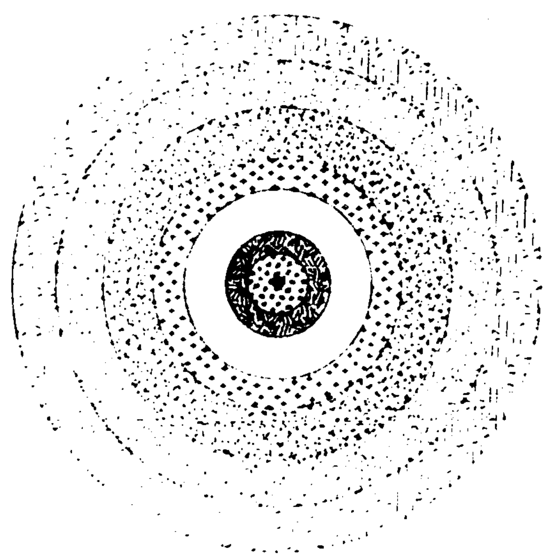


A Transformation of the Geographical Arrangement

Figure 1.2

This simple example represents one type of problem which might be examined. The questions, inter alia, for which answers are desirable would seem to be: Under what conditions are such transformations possible? Is more than one such transformation possible? Can rules be established for performing such transformations in more complex situations?

A second example is to be found in J. H. v. Thünen's theory of agricultural location. The relevant postulates of this theory are a market place located in a uniform region of agriculturally productive land and a method of transport which is free to travel anywhere in this region. Further, the method of transport has the characteristic that freight costs are directly proportional to the distance traveled. The resulting pattern of land use is one in which different agricultural crops are produced in concentric rings around the center of the city.[1] This idealized scheme is depicted in the following figure (Fig. 1.3):



SCHEMATIC OF VON THUNEN LAND USE RINGS

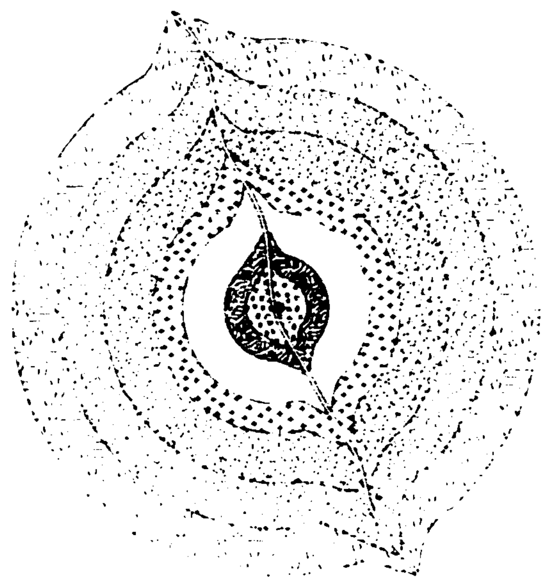
Figure 1.3

Von Thünen now changes his postulates slightly by introducing a river into the region. He allows the transport of agricultural goods to

---

[1]This is a much simplified version of von Thünen's work. See J. H. von Thünen, Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie (Hamburg, 1826). Also A. Lösch, The Economics of Location, W. H. Woglom translation, (New Haven: Yale University Press, 1954), pp. 36-49; and E. S. Dunn, Jr., The Location of Agricultural Production (Gainesville: University of Florida, 1954). The figures (1.3 and 1.4) are modifications of those given in T. Palander, Beitrage zur Standortstheorie (Uppsala: Almqvist and Wiksells, 1935), p. 75.

occur on this river but at a cost which is less than that incurred on land. The concentric pattern of land use is now distorted because of changes in the cost-distance relations (Fig. 1.4).



SCHEMATIC OF VON THUNEN RINGS MODIFIED BY RIVER

Figure 1.4

If it were possible to transform the geographic background in the von Thünen model into a cost-distance space, the land use pattern should again exhibit the concentric ring structure. As is apparent, the earlier example and the von Thünen situation are quite similar and comparable questions are appropriately examined.

The von Thünen model can also be modified so as to relate urban land use, land values, and rents, to distances from the center of cities.[2] Cost-distance concepts are also to be found in discussions of industrial and of retail site location, of the distribution of cities, and of international trade.[3] The application of mapping transformations hence appears to be widespread. In the work which follows, occasional reference to these applications is made from what may be broadly classed as location theory. These brief accounts are not intended by any stretch of the imagination to be comprehensive or complete reviews of the subject. They serve merely to indicate potential uses of the mapping transformations developed. For example, it is assumed that the von Thünen model of land use is generally valid, though the empirical evidence is certainly not conclusive.[4] When the realities of transportation are considered, the von Thünen scheme can be used to account for arrangements of land use which differ quite radically from the concentric zone pattern, but no formulation as simple as this can hope to explain the complexly dynamic interplay of the many

---

[2]W. Alonso, "A Model of the Urban Land Market" (Ph.D. dissertation; University of Pennsylvania, 1960). H. Mohring, et al., The Nature and Measurement of Highway Benefits (Evanston: Northwestern University Transportation Center, 1960). L. Wingo, Transportation and the Utilization of Urban Land (manuscript; Twentieth Century Fund, 1960).

[3]W. Isard, Location and Space Economy (New York: Wiley and Sons, 1956); C. Ponsard, Économie et Espace (Paris: Sedes, 1955); W. L. Garrison, "The Spatial Structure of the Economy," Annals, Association of American Geographers (XLIX, 2, 4, 1959; L, 3, 1960); C. W. Baskin, "W. Christaller's Die Zentralen Orte in Suddeutschland: Critique and Translation" (Ph.D. dissertation, University of Virginia, 1957).

[4]Cf., W. L. Garrison, The Benefits of Rural Roads to Rural Property (Part IV of Allocation of Road and Street Costs; Olympia: Washington State Council for Highway Research, June, 1956); W. L. Garrison and M. E. Marts, Geographic Impact of Highway Improvements (Highway Economic Studies, Seattle: University of Washington, 1958); B. O. Wheeler, Effect of Freeway Access upon Suburban Real Property Values (Seattle: Washington State Council for Highway Research, June, 1956); S. R. Wiley, "The Effect of Limited Access Highways upon Suburban Land Use" (Master's thesis, University of Washington, 1958).

factors involved. Seattle is not simply a homeomorphism of Boston; but such observations are obvious and hardly require special emphasis.

The thought of transforming ground-distances into time- or cost-distances is not original with this writer. It appears in explicit form in Alfred Weber's Theory of the Location of Industries[5] and perhaps earlier. Weber considered that "real" distances should be replaced by "fictitious" distances, composed in part of transport costs. Many subsequent writers have recognized similar concepts. Wingo writes of the "warping of space," Isard of effective distance, others of accessibility.[6]

In a recent work, Bunge presents two maps which further illustrate the concepts.[7] The first map (Fig. 1.5) indicates time-distances from the center of the city of Seattle to various locations by isochrones, the traditional method.[8] The isolines are on a base map which can be said to be

---

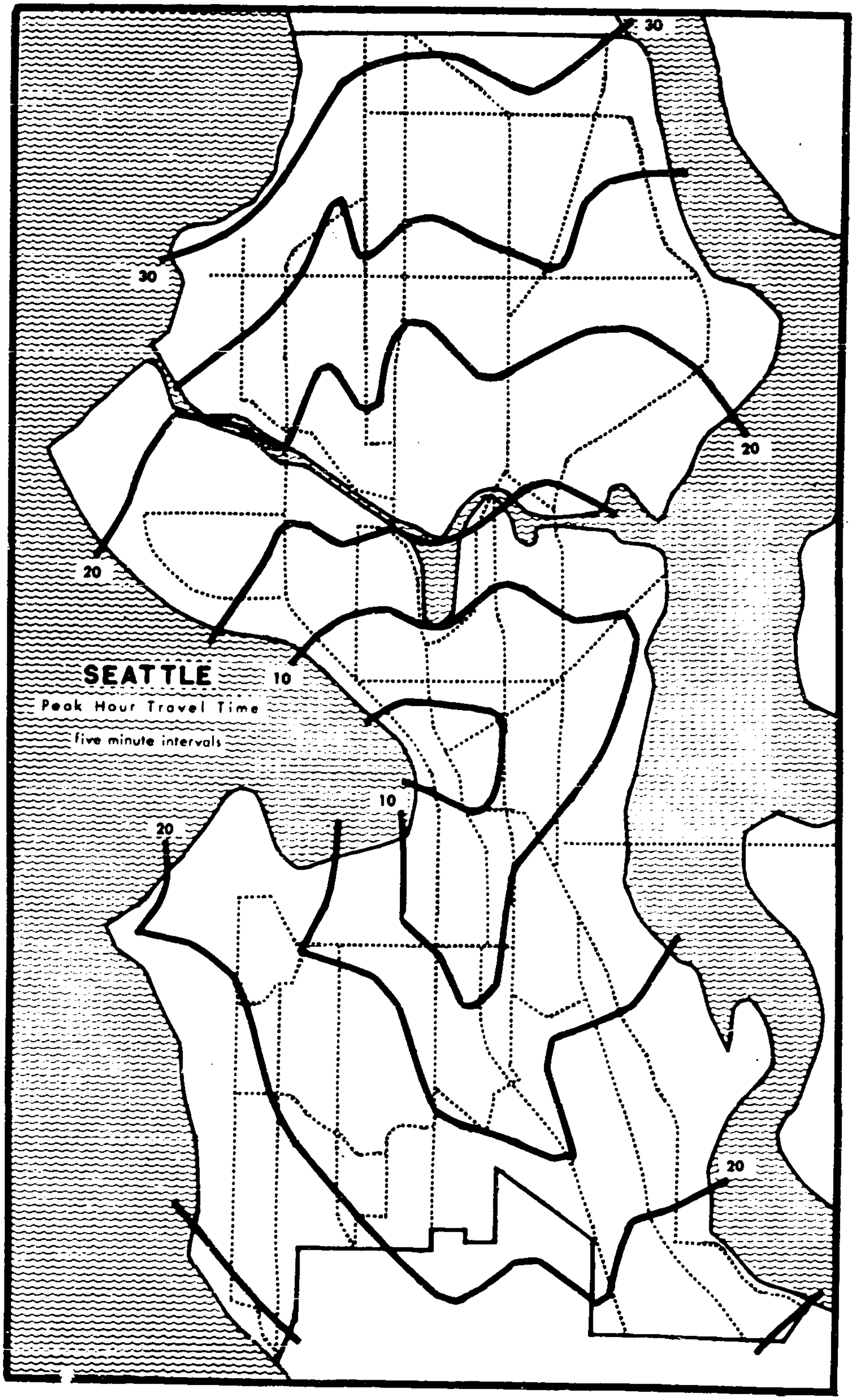
[5]C. J. Freidrich, Alfred Weber's Theory of Location of Industries (Chicago: University of Chicago, 1926), pp. 33-35, 44.

[6]Wingo, op. cit., p. 31; Isard, op. cit., pp. 200-201; W. G. Hansen, "How Accessibility Shapes Land Use," Journal, American Institute of Planners, XXV, 2 (1959), 73-76. Application of these concepts to urban transportation planning is well illustrated in J. D. Carroll, Jr., et al., Final Report: Chicago Area Transportation Study (Vols. I-III; Chicago: Chicago Area Transportation Study, 1960).

[7]W. Bunge, "Theoretical Geography" (Ph.D. dissertation, University of Washington, 1960). The possibility of such maps is also noted in E. Raisz, General Cartography (New York: McGraw-Hill, 1938), p. 266.

[8]An isochrone is a line passing through all places to (or from) which it requires an equal amount of time for a commodity to be transported. Isoline is the general term used to describe any line passing through points having some quantity in common. Occasionally used synonyms are contour and level curve. See J. K. Wright, "The Terminology of Certain Map Symbols," The Geographical Review, XXXIV (1944), 653-654. The first use of the term isochrone is attributed to Francis Galton. Cf., M. Eckert, "Eine Neue Isochronenkarte der Erde," Petermann's Geographische Mitteilungen, LV, 9 (1909), 209.

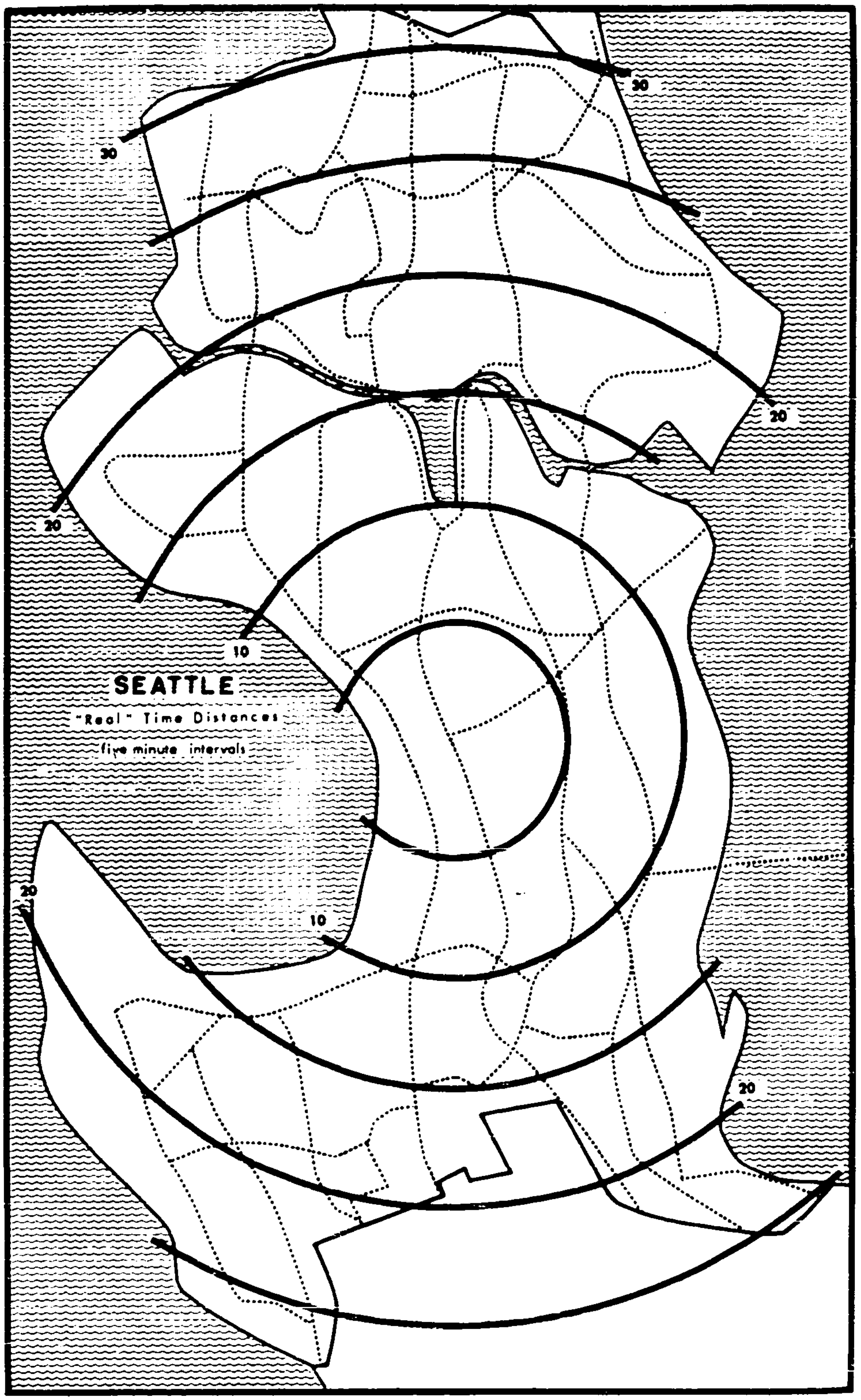
SEATTLE ISOCHRONES



Courtesy of W. Bunge

Figure 1.5

SEATTLE - TIME DISTANCES FROM CBD



Courtesy of W. Bunge

Figure 1.6

true to scale, in the conventional sense of the word. The second map (Fig. 1.6) is a transformation of the first so that distances, as measured in minutes from the center of the city (only), are correct to all other locations. Bunge does not indicate by what method such transformations are to be undertaken. The current work presents an investigation of this problem.

Having indicated a few simple cases in which the geographic notion of distance, measured along a geodesic on a sphere, is to be modified by the substitution of temporal or monetary units for yardsticks, it is not difficult to generalize the concept of distance further. As one example, evidence for a metrogenic substitution of psychological distance for spherical distance will be presented and briefly examined. It is then only a short step from distance to area and an examination of a number of published maps, which in cartographic parlance are referred to as cartograms, leads to what can be considered a generalization of the concept of an equal-area map.

When area is considered it is important to recall that the distribution of resources and human activity over the surface of the earth is quite uneven. The importance of a region can rarely be measured by its size. Theoretical treatises which assume a uniformly fertile plain, or an even distribution of population, etc., are to this extent deficient. As one writer has put it:

> The theoretical conceptions, based on hypotheses of homogeneous distribution must be adapted to geographical reality. This implies, in practice, the introduction of corrections with regard to the existence of blank districts, deserts of phenomenon, massives or special points. That is to say that in practice we have to take into consideration the

anisotropical qualities of the <u>area geographica</u>.[9]
To a certain extent variations in the geographic distribution of phenomena can be eliminated from the maps by appropriate transformations.

The Approach

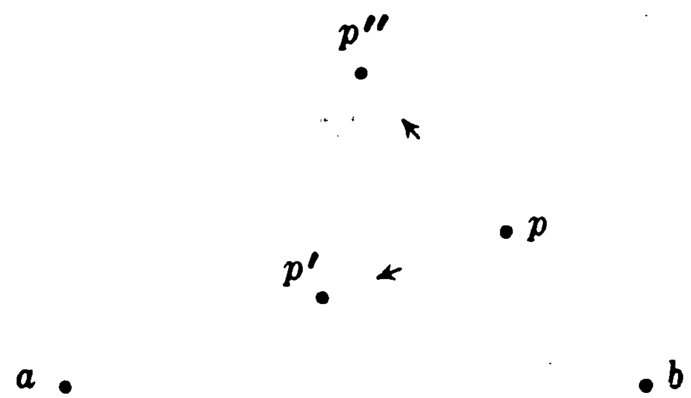
The present approach to these topics is through the theory of map projections, as available to geographers. This is naturally only one of several related but possible approaches. One could begin with the theory of groups or very primitive notions of simple transformations and lead up to transformations of the plane onto itself. The entire topic can, of course, also be considered by making explicit use of Gauss' fundamental forms, as in differential geometry. The approach is hence perhaps backwards but provides a point of departure with which geographers are most familiar. Working from the point of view of map projections has a further advantage; not only is the theory of map projections fairly complete, advantage can also be taken of two thousand years of previous work which has resulted in numerous particular solutions to a wide variety of problems. For example, Alfred Weber has been cited as having been desirous of substituting "fictitious" distances for real distances. This notion has been criticized several times.[10] Take the case of a location at point P in relation to two further locations, A and B (Fig. 1.7). Assume that the fictitious distance from A to P is less than the real distance; P is hence dislocated to P'. Assume further that the fictitious distance

---

[9]E. Kant, "Umland Studies and Sector Analysis," <u>Studies in Rural-Urban Interaction</u> (Lund Studies in Geography; University of Lund, Ser. B, 3, 1951), p. 5.

[10]Isard, <u>op. cit</u>., p. 109; Palander, <u>op. cit</u>., pp. 192-220.

from B to P is greater than the real distance; P is hence dislocated to P". The conclusion is that P must be dislocated to two distinct locations, an apparently absurd situation.



Modification of Distances from Two Points

Figure 1.7

An acquaintance with the literature of map projections, however, immediately indicates how the apparent absurdity is to be resolved. A variant of the doubly equidistant projection is required. The correct relocated position P''' should be at the point of intersection of the circles of radii AP' and BP" centered on A and B respectively. The azimuths are distorted but Weber is not concerned with direction. Chamberlin's trimetric projection can be used to approximate the case of three points, and extensions of the notion for larger numbers of points.[11]

---

[11]In general the fictitious distance AB requires modification to AB' and BA'. If AB' $\neq$ BA', set AB" = BA" = $\frac{1}{2}$AB' + $\frac{1}{2}$BA'. The circles are to be drawn with radii AP' and B"P" intersecting at P'''. These

Similarly, it is known that only special surfaces are isometric with a plane. Complete isometric transformations are therefore not expected to be possible in general. Mappings preserving distances from a point or curve, however, would a priori seem possible from a knowledge of the conventional theory of map projections.

In a broader context, the notion of transformations is of extreme importance to large portions of modern mathematics and has application in almost every field. An attempt is made to point out some geographic applications of this notion. Any pretense to a complete coverage of the topic suggested by the title, however, would be absurd. There is, nevertheless, a departure from the traditional cartographic problem of finding a "best" representation of the surface of a sphere on a flat sheet of paper. The transformations emphasized are not attempts to minimize distortion but contain deliberate distortion of spherical distances. The distortion is chosen to simplify the solution to certain problems or to fit empirical data. The first context is well known from Mercator's projection, which provides nomographic solutions to the navigator's problem by mapping the loxodrome on the sphere into a geodesic on the plane. The second context is less well recognized but is not completely unknown.

The Plan of the Discussion

In Chapter II the concern is with the general theory of map projections. This serves both as a review and as an introduction to the

---

circles often intersect in two points, which causes little difficulty. Yet, under certain conditions the circles may not intersect.

The trimetrigon projection uses the center of the curvilinear polygon defined by the circular arcs as an average distance. Extension to more than three locations can proceed by analogy but becomes increasingly difficult and approximate.

chapters which follow. A complete development of Gaussian surface theory is not presented, but map projections can still be attacked from several directions. Projective constructs are widely available to geographers and in popular accounts. The treatment here given in the initial chapter is largely algebraic rather than projective. While projective models are historically important and occasionally suggest lines of development, they do not always provide a sufficient framework for further extensions of the topic. Chapter II also establishes the foundation for simple graphic techniques which will be used in the ensuing chapters.

The reader is assumed to have a modest, but not detailed, prior knowledge of the subject of map projections.[12] The reader whose acquaintance with map projections is less than that of the average geographer may at times find the terminology slightly obscure but this is not a serious difficulty. A sphere is taken as the general model of the earth. The reader is cautioned that the results do not all apply to a spheroid or other models of the earth.

The next chapter, Chapter III, considers a particular class of map projections, the azimuthal projections, in somewhat greater detail. Here the relation between the results of Chapter II, simple projective models, and graphic procedures is established. Using these methods, a presentation is then given of a few of the traditional problems which have concerned cartographers in their treatment of this class of map projections. Finally,

---

[12] Two excellent non-mathematical introductions are: I. Fisher and O. M. Miller, World Maps and Globes (New York: Essential Books, 1944), and W. Chamberlin, The Round Earth on Flat Paper (Washington: National Geographic Society, 1947). The simplest and most concise discussion of the mathematical theory is probably that in H. Merkel, Grundzüge der Kartenprojektionslehre (München: Deutsche Geodätische Kommission, Reihe A, Heft 17, 1956).

the assumptions of the previous chapter are examined in greater detail and then relaxed. By proceeding in a manner consistent with the earlier established techniques, but with relaxed assumptions, new projections are created. These projections exhibit rather unusual properties but are not entirely new to the literature of cartography. Geographic interpretation and valid uses for projections with some of these properties, however, have hitherto been lacking. In Chapter IV such interpretation is attempted when attacking the problem of metrogenic substitution cited earlier.

The chapter (IV) begins with simple analogues of transportation situations and continues by increasing the complexity of these situations. Examples serve to indicate possible applications of the mappings developed using the concepts elicited from the previous chapter on azimuthal projections. These examples are drawn primarily from the context of what has loosely been referred to as location theory. It is not intended, however, that the application be restricted to this subject.

The transformations themselves can be considered either as being from the surface of the earth to another "sphere," or as mappings of the earth to a plane, or as mappings from one plane to another. The latter interpretation is justifiable and useful for restricted regions of the earth.

The discussion makes use of isolines and is in many respects quite similar to statements by Lösch and particularly Palander. A conscious attempt is made to avoid repetition of Palander's work but a certain amount of overlap appears inevitable. Where possible examples have been chosen so as to extend, rather than duplicate, the notions and excellent isoline maps presented by Palander. Palander's concept of a transport surface is clarified and extended. Certain topological configurations of transport systems are examined and transport characteristics identified. An isomorphism is established between transport isolines and map projections. From the point of view of map projections it is also clear that Tissot's indicatrix (Chapter II) can serve as a measure of the distortion introduced by transport systems. This certainly suggests the possibility of classifying geographic environments on the basis of the amount of departure from an idealized homogeneous geographic space. The final portions of the chapter indicate further geographic applications; distance models and the concept of distance are also briefly examined. The entire discussion is restricted to azimuthal projections.

In Chapter V the inadequacies introduced by consideration of only azimuthal projections are examined. The addition of conditions of equal-area and conformality to transformations preserving distance from a point are discussed.

An attempt is then made to demonstrate that distortion of area on maps can be most meaningful and useful (Chapter VI). The stage is set by an examination of select cartograms and common cartographic problems. The solution is shown to lie in transformations of multiple integrals and the application and relevance of these to geography is explored.

The final chapter reviews a few of the major results and shortcomings.

CHAPTER II

MAP PROJECTIONS

One general notion of a transformation of a two-dimensional surface, such as the surface of a sphere, is given by a pair of equations of the form:

$$u = f_1(\phi, \lambda)$$

$$v = f_2(\phi, \lambda),$$

where u and v are coordinates used to describe the location of positions on the transformed surface, phi (ϕ) and lambda (λ) are coordinates used to describe the location of positions on the original surface, and f_1 and f_2 are any functions whatever.[1] For the discussion of the conventional theory of map projections it will be assumed that f_1 and f_2 are real, single-valued, continuous, and differentiable functions of ϕ and λ in some domain and that the Jacobian determinant does not vanish.[2] This latter condition can be interpreted as requiring that the lines ϕ = constant and λ = constant do not coincide on the transformed surface. The u and v

---

[1]A function is a rule associating elements of one set with elements of another set. This definition of a function is used as we will later include functions whose domain is a finite collection of points. Many-to one and one-to-many mappings will also be encountered. The subscripts in the functional notation will in the future be omitted. It is to be understood that in general $f_1 \neq f_2$; f will be used for both unless special emphasis is desired.

[2]The detailed conditions for an allowable mapping and definitions of surfaces, etc., in terms of equivalence classes have been omitted. The reader is referred to: E. Kreyszig, Differential Geometry (Toronto: University of Toronto, 1959), pp. 3-9, 18-20, 72-79. The Jacobian determinant (J) is given by: $J = \frac{du}{d\phi}\frac{dv}{d\lambda} - \frac{dv}{d\phi}\frac{du}{d\lambda}$.

are to be interpreted as independent coordinates on a plane surface, and ϕ and λ are to be interpreted as geographical coordinates on a sphere. The assumptions are sufficient to ensure a one-to-one mapping of some portion of the surface of the sphere to a plane so that each position (ϕ, λ) on the surface of the sphere (in the domain under consideration) corresponds to one, and only one, position (u,v) on the plane. Conversely, each position (u,v) within some region on the plane corresponds to one, and only one, position (ϕ, λ) on the surface of the sphere. In practice the equations for a map projection are not always given in explicit form; there are also a few map projections whose construction is relatively simple using graphic scaling and plotting techniques but for which equations have never appeared in the literature. The term map projection is used in the conventional cartographic sense, where it is closer to the general concept of a transformation than to that of projective geometry.

Notation

In the following discussion all angles are assumed to be measured in radians; exceptions are specifically mentioned. The latitude (ϕ) has the interval $-\pi/2 < \phi \leq +\pi/2$ as its domain. The longitude (λ) has the interval $-\pi < \lambda \leq +\pi$ as its domain. It will frequently be convenient to use spherical coordinates for positions on the surface of the earth. These are defined by $\rho = \pi/2 - \phi$, $\lambda = \lambda$. It is also convenient to require the origin of the u, v system of coordinates to coincide with the respective origins of the parametric curves of the surface of the sphere. Hence, when the ϕ, λ parameterization is used the u, v coordinates are to be interpreted as rectangular coordinates (x,y) in the plane. When the spherical coordinates ρ , λ are employed, the u, v coordinates are to be considered polar coordinates (r,θ) in the plane. Rectangular coordinates are hence used when the map is centered on the intersection of the equator and the prime meridian. Polar coordinates are used when the map is centered on the north pole. When the map is not centered on either of these two positions but at some arbitrary point (ϕ_0, λ_0), the origins of the system of parametric curves will be moved to this new position. The new values ϕ^*, λ^* and ρ^*, λ^* can be related to the earlier systems by equations given in standard cartographic texts. The notation used will be simply $x = f(\phi, \lambda)$, $y = f(\phi, \lambda)$ and $r = f(\rho, \lambda)$, $\theta = f(\rho, \lambda)$, where it is to be understood that $x = f(\phi, \lambda)$, $y = f(\phi, \lambda)$ may imply $x = f(\phi^*, \lambda^*)$, $y = f(\phi^*, \lambda^*)$ and similarly for r, θ as functions of ρ^*, λ^*. Note that the functional notation ignores scalars. These constants do not in any way affect the classification which follows, though they may alter the properties of individual projections as in the case of the gnomonic and stereographic, or orthographic and Lambert equal-area azimuthal projections. The earth radius (R) is the most frequent value encountered, here assumed equal to unity.

All this juggling is necessary in order to treat the projections in their so-called normal cases. For planar and conic projections this is the polar aspect, and polar coordinates are used. For cylindrical projections it is the equatorial aspect, and rectangular coordinates are used. Thus it is not necessary to distinguish oblique aspects of projections, for the geographical or spherical coordinates are simply shifted to a new origin. The rationale for the entire procedure is best demonstrated by an example. The equations for the polar aspect of the stereographic projection are:

$$r = f(\rho) = 2\,R \tan \rho/2$$

$$\theta = f(\lambda) = \lambda.$$

The general oblique aspect in rectangular coordinates on the other hand requires the more complicated equations:

$$x = f(\phi,\lambda) = 2R\,\frac{\cos\phi\,\sin(\lambda-\lambda_0)}{1+\sin\phi_0\,\sin\phi+\cos\phi_0\,\cos\phi\,\cos(\lambda-\lambda_0)}$$

$$y = f(\phi,\lambda) = 2R\,\frac{\cos\phi_0\,\sin\phi-\sin\phi_0\,\cos\phi\,\cos(\lambda-\lambda_0)}{1+\sin\phi_0\,\sin\phi+\cos\phi_0\,\cos\phi\,\cos(\lambda-\lambda_0)}.$$

To consider oblique or transverse versions of the same projection as entirely different projections when a spherical model is used, as many authors have done, is an unnecessary complication.

In addition to rectangular and polar coordinates more obscure coordinate systems (oblique, parabolic, elliptic) could be used. In fact, any system of coordinates in a plane (meeting rather simple requirements) can be considered to be a map projection. The azimuthal equidistant projection ($r = \rho$, $\theta = \lambda$) is the system of polar coordinates, and the rectangular projection ($x = \lambda$, $y = \phi$) is the system of rectangular coordinates introduced by Descartes. Measurement in these systems, however, no longer has the same meaning as measurement in the plane. Thus, though, the metric must be modified, the u, v coordinates could be interpreted as the normal case of some projection in each instance. Such an interpretation requires as many systems of coordinates as there are projections.[3]

[3] In terms of the classification to be presented this would have a consequence that all map projections belong to category D. This can also be achieved by defining a different set of parametric curves on the original surface for each projection, an approach used by Kreyszig (Ibid., p. 175). Map distance contours as defined in Chapter III are similar.

Employing both polar and rectangular coordinates is clearly not necessary; at times it is awkward; in other instances the treatment is simplified.

The Classification of Map Projections

The desire for a classification of map projections stems from the fact that there are an infinite number of projections, each with different properties. The fundamental problem of classification with regard to map projections hence appears to be the discovery of a partitioning of this infinite set into a comprehensible and useful finite number of all-inclusive and preferably non-overlapping classes.

Several classifications of map projections are to be found in the cartographic literature. The advantages and disadvantages of each, of course, depend on the purpose, just as the property by which classes are to be distinguished depends on the purpose. The classification based on geometric models separating projections into conic, cylindric, planar, polyconic, etc., is convenient and is often used. Another most important method of classification, based on the preservation of certain geometric properties of the surface of the sphere, separates the conformal projections from the equal-area projections, leaving a third class which is neither (Tissot's aphylatic projections). Other classifications are based on the appearance of the meridians and parallels, i.e., whether these consist of circles, ellipses, quartics, etc. Projections can also be classified according to the form of the equations; whether these are algebraic, transcendental, linear, and so forth. Maurer, in his study of map projections, attempts to partition 237 map projections into classes and subclasses based on combinations of these systems.[4] Several of the problems of classification are discussed by Maurer, and a most interesting but very involved Venn diagram is presented showing the overlap and interrelations of the various classes. Tissot is another who recognized fine distinctions and obtained an elaborate classification of map projections.

For the purposes of the present discussion a simple parametric classification of projections is employed. This classification is not common in the cartographic literature but has the advantage of being very simple and including all map projections. In addition to the overview afforded, the parametric classification suggests the simple graphic methods which are to be used in the ensuing chapters. There are, of course, still many alternate methods of classifying map projections.

A Parametric Classification

The parametric classification is based on the fact that the equations for the location of lines of latitude and of longitude on the map in some cases depend on only one of these quantities. For example, in the truly cylindrical projections, the lines of longitude depend on longitude alone, $x = f(\lambda)$, and the lines of latitude depend only on latitude, $y = f(\phi)$. Here it is immediately apparent that the parametric classification tacitly assumes the normal case for each projection. That is, it assumes that centrally symmetric azimuthal projections are centered on a pole, that cylindrical projections are onto cylinders normal to the equatorial plane, and that conic projections have the axis of the cone paral-

---

[4] H. Maurer, "Ebene Kugelbilder, ein Linnesches System der Kartenentwürfe," Ergänzungsheft Nr. 221 zu Petermann's Mitteilungen (Gotha: Justus Perthes, 1935).

lel to the earth axis. Other cases of these projections will fall into different categories of the classificatory scheme.[5] Since the projection properties are invariant under rotations of the reference sphere, this drawback is not important for the present purpose.[6]

Two equations are required to specify the location of both curves of equal latitude and of equal longitude on the map. Since each equation can be a function of two parameters, there are logically only sixteen possible combinations of the general transformation $u = f(\phi,\lambda)$, $v = f(\phi,\lambda)$, eliminating parameters one after another. These special cases are:

$$\begin{array}{llll}
u = f(\phi,\lambda) & u = f(\lambda) & u = f(\phi) & u = f(C) \\
v = f(\phi,\lambda) & v = f(\phi,\lambda) & v = f(\phi,\lambda) & v = f(\phi,\lambda) \\
\\
u = f(\phi,\lambda) & u = f(\lambda) & u = f(\phi) & u = f(C) \\
v = f(\lambda) & v = f(\lambda) & v = f(\lambda) & v = f(\lambda) \\
\\
u = f(\phi,\lambda) & u = f(\lambda) & u = f(\phi) & u = f(C) \\
v = f(\phi) & v = f(\phi) & v = f(\phi) & v = f(\phi) \\
\\
u = f(\phi,\lambda) & u = f(\lambda) & u = f(\phi) & u = f(C) \\
v = f(C) & v = f(C) & v = f(C) & v = f(C),
\end{array}$$

[5]The normal case of the gnomonic projection has the form $\theta = \lambda$, $r = f(\rho)$ but the equatorial aspect has the form $x = f(\lambda)$, $y = f(\phi,\lambda)$. As already mentioned, each projection aspect has a normal form if special coordinates are used.

[6]A rhumb line on the earth (assumed spherical) is a straight line on the equatorial case of the Mercator projection, but not on the oblique (transverse) Mercator. However, the loxodrome in oblique spherical coordinates does become a straight line on the oblique Mercator. As the oblique pole is not used in navigation, the terrestrial rhumb line is no longer a straight line. This point often causes confusion.

where C is a constant. The sixteen cases have been arranged in the form of a matrix which can be said to be interchange-symmetric. The symmetry results in simple interchanges of the u, v coordinates which are indistinguishable from the symmetric cases. The interchange of position a_{23} (second row, third column) is position a_{32} (third row, second column). Thus, if the u and v coordinates are interpreted as polar coordinates, the interchange of

$$a_{23} \qquad \begin{array}{l} \theta = f(\phi) \\ r = f(\lambda) \end{array}$$

becomes

$$a_{32} \qquad \begin{array}{l} \theta = f(\lambda) \\ r = f(\phi), \end{array}$$

which could, of course, have been achieved by giving a different interpretation to the u and v coordinates, interchanging r and θ. A map drawn in the system $\theta = f(\phi)$, $r = f(\lambda)$, position a_{23} with $u = \theta$, $v = r$, in which the parallels are rays from the center and the meridians concentric circles, illustrates the procedure and is given in a paper by Brown.[7] If the u, v coordinates are interpreted as rectangular coordinates, the interchanges amount to rotations of ninety degrees. The Plat Carée projection $x = \rho$, $y = \lambda$, more generally $x = f(\rho)$, $y = f(\lambda)$, is a variant. An alternate variant is $r = \phi$, $\theta = \lambda$. The distinction between sense preserving and sense reversing projections need not be discussed.[8]

---

[7]B. H. Brown, "Conformal and Equiareal World Maps," American Mathematics Monthly, XXXXII (April, 1935), 212-223.

[8]Those which reverse directions, as on a map seen from the reverse side of the paper.

Several of the sixteen cases (nine to be exact) can usually also be eliminated from consideration, for geographic interpretation generally has not been demonstrated. These include the last row and column which map the entire domain (sphere) into a line or point.[9] As Miller and Fisher indicate, the doubly equidistant projection, when centered on a point and its antipodal point, maps the entire sphere onto the straight line of length πR connecting these two points.[10] Such projections have to date found little application in geography.

Elimination of the interchanges and the degenerate cases ($J \equiv 0$) leaves four valid classes of map projections:

A) $u = f(\phi, \lambda)$, $v = f(\phi, \lambda)$ B) $u = f(\lambda)$, $v = f(\phi, \lambda)$

C) $u = f(\phi, \lambda)$, $v = f(\phi)$ D) $u = f(\lambda)$, $v = f(\phi)$

For the sake of brevity these have been labelled A, B, C, and D, respectively. Because of the present dual interpretation of the u and v values as rectangular or polar coordinates, it is convenient to treat two situations in each category, a total of eight:

A) $x = f(\phi, \lambda)$, $y = f(\phi, \lambda)$ $\theta = f(\rho, \lambda)$, $r = f(\rho, \lambda)$

---

[9]However, see note 9, Chapter IV.

[10]O. M. Miller and I. Fisher, World Maps and Globes (New York: Essential Books, 1944), p. 67.

B) $x = f(\lambda)$ $\quad$ $\theta = f(\lambda)$

$y = f(\phi, \lambda)$ $\quad$ $r = f(\rho, \lambda)$

C) $x = f(\phi, \lambda)$ $\quad$ $\theta = f(\rho, \lambda)$

$y = f(\phi)$ $\quad$ $r = f(\rho)$

D) $x = f(\lambda)$ $\quad$ $\theta = f(\lambda)$

$y = f(\phi)$ $\quad$ $r = f(\rho)$.

The coordinates are not to be interpreted as a relabelling, for in this event even the projections of category D, $\theta = f(\lambda)$, $r = f(\rho)$, belong to category A, i.e., $x = f(\rho)\cos f(\lambda) = f(\phi, \lambda)$.

The relation of the parametric classification to the more common classification based on projective models is simple but the nomenclature is somewhat confusing. To establish the connection to the geometric classification a brief overview is presented.

Relation to the Geometric Classification

Category D of the classificatory scheme is clearly the simplest of the four. Most map projections studied in elementary texts cover only this class of projections. To give realization to these projections, the cylindric projections can be obtained by interpreting the u-coordinate as abscissa and the v-coordinate as ordinate in a Cartesian or rectangular coordinate system, $x = f(\lambda)$, $y = f(\phi)$. The meridians and parallels are orthogonal straight lines. The conventional cylindric projections are further restricted to the very special case $x = f(\lambda) = R\lambda$, which implies a right circular cylinder. Elliptic, parabolic, hyperbolic, and more general cylinders never seem to have been considered.[11] If the right circular cylinder is taken as basic, then the spacing of the parallels is all that can be varied on the projection. The most important projections in this category are Mercator's conformal, the equal-area cylindrical, the equi-rectangular, and Miller's cylindrical projection. Symmetry is invariably taken about the equator, obviously not a mathematical requirement.

The conic projections are another realization of category D in polar coordinates, with $\theta = f(\lambda) = C\,\lambda$, where C is a constant restricted to the interval $0 < C < 1$, implying a right circular cone. These projections have straight lines as meridians and arcs of concentric circles as parallels. The maps are pie or fan-shaped. The constant C is often called the constant of the cone; the vertex angle of the fan-shaped map depends on this constant. Non-constant (one parametric) functions for the spacing of the meridians, or the use of constants greater than unity, are only rarely mentioned in the cartographic literature. The center of the concentric circles which define the lines of latitude may be the north pole (normal case), or this pole may itself be one of the circular arcs--a truncated conic projection. The most used of the conic projections are Lambert's conformal and Alber's equal-area, usually with two standard parallels.

The centrally symmetric azimuthal projections also appear in category D, using polar coordinates as the special case $\theta = f(\lambda) = \lambda$. This

---

[11]The Gauss-Kruger projection of the ellipsoid uses an elliptic cylinder; L. Blaser, Einführung in die Kartenlehre (Mathematische-Physikalische Bibliothek, Reihe I, Nr. 81; Leipzig: Teubner, 1951), p. 57.

is clearly a limiting position of the conic projections with $C = 1$. The more important of these azimuthal projections are the gnomonic, stereographic, equidistant, equal-area, and orthographic.

Category C contains some of the so-called oval projections, the pseudoconic projections, and several lesser known varieties. The oval projections with straight parallels and curved meridians (sinusoidal, Mollweide, Eckert) are of the form $x = f(\phi, \lambda)$, $y = f(\phi)$, and, in virtually all cases encountered, the simpler situation $x = \lambda f(\phi)$, $y = f(\phi)$. These can also be shown to be polycylindrical developments. The pseudoconic projections (Bonne, Werner, Weichel) are of the form $\theta = f(\rho, \lambda)$, $r = f(\rho)$ with curved meridians and concentric circular arcs as parallels. Category C also includes the pseudoazimuthal projections, as recognized in the Soviet literature.[12]

Projections belonging to category B are not at all well known. This category does, however, contain azimuthal projections as the special case $\theta = \lambda$. The azimuthal projections used in practice have the simpler form of category D.

Category A includes the polyconic projections, the oval projections with curved parallels, and generally the projections with curved meridians and parallels. These projections are often referred to as conventional projections or unclassified projections. A few of the other projections in this category are Lagrange's projection of the sphere within a circle, Aitoff's projection, Hammer's projection, August's conformal projection

---

[12]See A. V. Graur, Matematischeskaya Kartografia (Leningrad: Leningrad University Press, 1956), pp. 65-66, or D. Maling, "A Review of Some Russian Map Projections," Empire Survey Review, XV, 115, 116, 117, (1960).

within a two-cusped epicycloid, and the elliptic projections of Guyou and Adams.

Graphic Representation of Projections

It is not the purpose of this paper to analyze the classificatory scheme nor to attempt to fit a large number of projections into this scheme, although this can be accomplished without too great difficulty. The parametric classification includes all map projections; this is sufficient for the purpose at hand. The graphic implications of the parametric classification are more valuable. We can now turn to this more stimulating subject with immediate results.

Every map projection can be analyzed or represented by means of at most two diagrams. This follows directly from the fact that the equations for category A, the most involved possibility, can be rewritten in the form:

$$F_1(\phi, \lambda, u) = 0$$

$$F_2(\phi, \lambda, v) = 0$$

which can be interpreted as two separate surfaces in space. Each such surface can be represented diagrammetrically by level curves or by a block diagram. This may appear somewhat involved, only one diagram being required to show the entire projection as a map. However, as an aid to analysis and understanding, the graphs are very useful.

The two diagrams for category A map projections will of necessity be "three-dimensional," or, equivalently, the level curves for these diagrams. Categories B and C will require only one three-dimensional drawing and one graph each. Category D requires only two simple graphs (one of the partial derivatives is zero; the level curves coincide). Map projections which fall into this category can be completely characterized by two curves, one on a u-graph and one on a v-graph. Thus as the projections become less complex in terms of the classificatory scheme, their diagrammatic representation becomes less complex. The converse question, whether any graph can be interpreted as a map projection, is more difficult.

The graphic procedure is most easily demonstrated by projections of category D. The graphs corresponding to Mercator's projection are shown in the accompanying figure (Fig. 2.1). This cylindrical projection has as

GRAPHS FOR THE MERCATOR PROJECTION

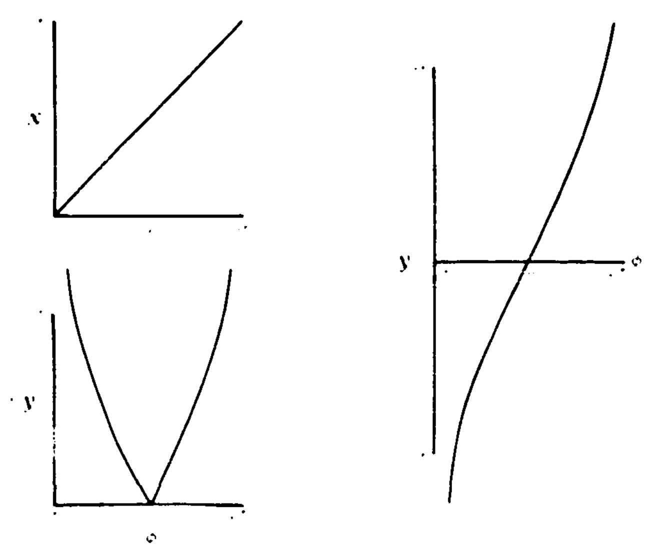


Figure 2.1

its mapping equations (normal or equatorial case):

$$x = f(\lambda) = R\lambda$$

$$y = f(\phi) = R \ln \tan (\pi/4 + \phi/2)$$

where we take R = 1. The one graph, that specifying the location of the lines of latitude, can be given different forms, depending on whether one employs geographical or spherical coordinates for the latitude. The symmetry about the equator can clearly be seen and is typical of cylindrical projections.

The graph corresponding to an equal-area conic projection with two standard parallels is shown in the next figure (Fig. 2.2). The truncated form of this projection appears in an intercept of the axis ($f(\rho)$) which is not at the origin.

GRAPHS FOR ALBER'S EQUAL-AREA PROJECTION

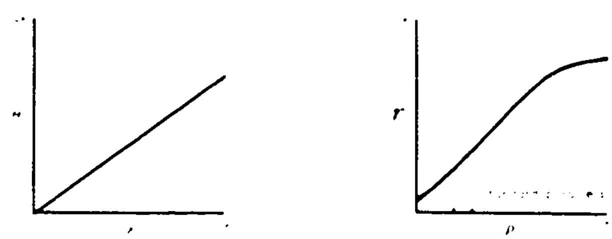


Figure 2.2

Centrally symmetric azimuthal projections are also contained in category D. The two graphs for such a projection, the stereographic, are presented in Figure 2.3. The curves could also have been plotted in polar, rather than rectangular, coordinates.

GRAPHS FOR THE STEREOGRAPHIC PROJECTION

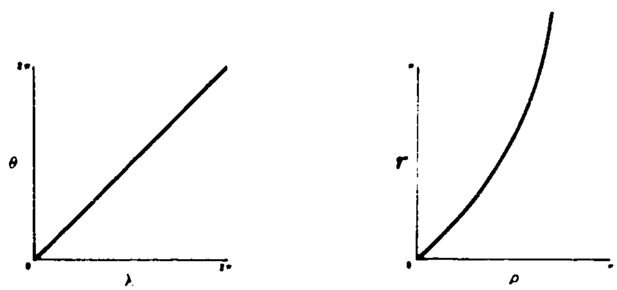


Figure 2.3

The sinusoidal projection has been chosen to treat a more complicated case. This projection belongs to category C and has as its mapping equations:

$$x = f(\phi, \lambda) = R \lambda \cos \phi$$

$$y = f(\phi) = R \phi.$$

The graph (Fig. 2.4) which represents the former of these equations is

GRAPHS FOR THE SINUSOIDAL PROJECTION

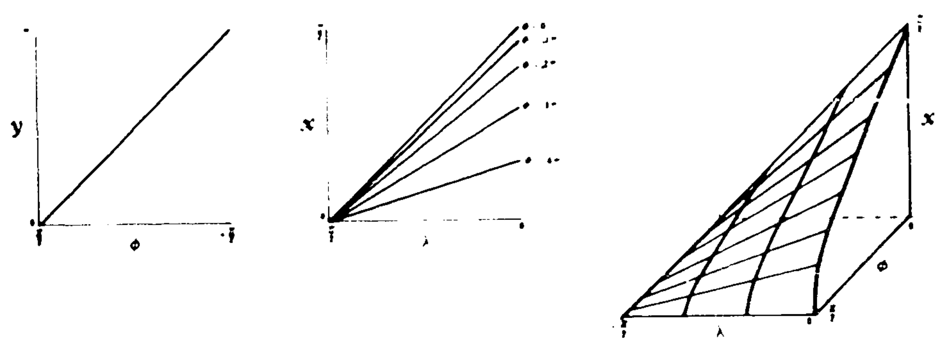


Figure 2.4

symmetric in all four quadrants, and it is therefore only necessary to show one quadrant. This graph is shown in both forms, as a block diagram and as level curves. Since the surface $F(\phi, \lambda, u) = 0$ specifies three sets of level curves, this example can be drawn in a variety of orientations. The pictorial representation of the sinusoidal projection should suffice to illustrate the procedure for extending the graphic method to further cases. In the representation shown the block diagram has been drawn in rectangular coordinates; a representation in cylindrical coordinates may occasionally be more convenient. The level curves are then no longer traces of intersections of parallel planes but profiles along the parametric surfaces.

The usefulness of the graphic representation of the equations for map projections will be more apparent in the ensuing chapters, particularly the chapter on azimuthal projections. Four variables (u, v, ϕ, λ) appear in the pair of equations used to specify a map projection, hence a complete graphic representation would require a four-dimensional diagram.[13] Nevertheless, by assuming one of the equations to be fixed, the effects of changes in the other equation can be studied. Observations regarding the properties of projections are particularly easily deduced from the one-parameter graphs. The block diagrams are less susceptible to rapid visual analysis, at least without practice. The results to be

---

[13]The pair of equations can be reduced to one equation by the use of complex variables and then a diagrammatic representation can be achieved. The traditional method of illustrating map projections by a drawing of the graticule does this by comparing a u,v-plane with a ϕ, λ-sphere. See K. Knopp, Elements of the Theory of Functions, trans. F. Bagemihl (New York: Dover, 1952).

established for azimuthal projections naturally can be modified to consider the cylindric, conic, and more difficult situations.

Distortion on Projections

Certain properties of the surface of a sphere can be preserved when such a surface is transformed to a plane. Incidence relations, for example, are generally invariant. Several other properties cannot be maintained and are said to be distorted. Distance relations between all points on a sphere are among the properties which cannot be shown on a plane map without distortion. Of the inevitable modification of properties recognized as occurring during the transformation from a sphere to a plane several can be measured. Other modifications, such as changes of shape in the large, have not been successfully measured to date.[14] The most discussed distortions concern the modification of distances, areas, and angles. Transformations which preserve angles in the small are called conformal; those which preserve areas are called equal-area.[15] A simple method of measuring several types of distortion at any point (except singular points) on a map is due to Tissot.[16] Only a very brief review of Tissot's results is given here; more detailed treatments are available in the literature.[17]

---

[14]A promising numerical characterization of shapes is given in W. Bunge, op. cit., Chapter III.

[15]For conformal, occasionally also orthomorphic, isogonal, or autogonal. For equal-area, also orthembadic, authalic, equivalent, or homalographic.

[16]M. Auguste Tissot, Mémoire sur la Representation des Surfaces et les Projections des Cartes géographiques (Paris: Gautier-Villars, 1881).

[17]Merkel, op. cit., or L. Driencourt and J. Laborde, Traité des Projections des Cartes geographiques (Vol. I; Paris: Herman, 1932).

Tissot noted that the sphere and map are locally affinely connected so that an infinitesimal circle on the sphere is transformed into an ellipse. This ellipse is now known as Tissot's indicatrix. Tissot further showed that every real, single-valued, continuous transformation preserves the orthogonality of at least one pair of intersecting curves. If the transformation is not conformal, then there is one, and only one, such pair. The orthogonal directions of this intersecting pair of curves are used for the major (b) and minor (a) semi-axes of the indicatrix, for it can be shown that the maximum and minimum scale distortion occurs in these two directions. Tissot's ellipse hence provides a measure of the distortion at any point on the map. The areal distortion is the ratio of the area of the ellipse to the area of the infinitesimal circle. A condition for an equal-area map is therefore that the product of the scale distortion in the two orthogonal directions is unity ($ab = 1$). A conformal map distorts in equal amounts in all directions about a point (but in different amounts at different points), hence the ellipse has zero eccentricity, or $a = b$, a circle.[18] Since conformality and proportionality of area are mutually exclusive properties for plane maps of the surface of a sphere, the area of this circle on a conformal map is of necessity enlarged or diminished (except at singular points). Maps on which are drawn small ellipses have appeared in the literature to illustrate the distortion on specific projections. An alternate presentation has been to draw lines of equal deformation on the maps. Tissot's four measures of distortion can, at most, be functions of both latitude and longitude. A simple corollary

[18]The angular distortion, however, is not the eccentricity.

is then that, as two-parametric functions, each can be illustrated by a diagram similar to those used for the projections themselves. The graphic procedure is very simple and it may be desirable to include such diagrams on published maps. Total distortion, which has occasionally been considered, can be defined as the volume (area) under these graphs, and average distortion by appropriate division of the total distortion.

Latitude and longitude on the sphere form an orthogonal system. If they form such a system on the map, we know immediately that they form the axes for Tissot's indicatrix. Conversely, if the latitude and longitude do not meet at right angles on the map (such a map cannot be conformal), the axes of the ellipse do not coincide with latitude and longitude. These are useful results.

Measurement of the departure from invariance (i.e., distortion) of a property is of course only important if that property is to be used. Only those properties which occur in the problem for which a map is prepared need be invariant. The preparation of maps, however, is a slow process, and it is often desired that a map be used for graphic solutions of several problems. The difficulty here is that distortion of some properties inevitably occurs. Invariably maps are used for purposes for which they are not intended.

Selection of Projections

There are an infinite number of map projections, as can easily be demonstrated. A natural question is to ask how one should select a specific projection from this infinite variety. In a very few cases the geometry of the problem immediately suggests the solution. This, for example, occurs in the photography of the earth from a rocket or satellite and in radar mapping.[19] Usually, however, the question is more difficult to answer.

If one arbitrarily chooses one equation, $u = f(\phi, \lambda)$, to satisfy certain conditions (the general assumptions of differentiability, non-vanishing Jacobian, etc., are normally to be retained), the second equation can still be chosen in a wide variety of ways. If the one equation is specified, then a conformal projection can be obtained only by choosing the second equation in such a manner that the two together satisfy the conditions for conformality.[20] Similarly, equal-area projections are obtained if, and only if, the indicatrix values of Tissot satisfy the relation a times b equals unity.[21] More generally, given some a priori conditions which a map projection is to satisfy, these conditions are translated into differential requirements. The selection of a projection then becomes a problem involving the solution of partial differential equations, often a difficult task. Fortunately, many solutions are already available. Another difficulty arises from the specification of the a priori conditions.[22] Since each of

---

[19] D. Levine, et al., Radargrammetry (New York: McGraw-Hill, 1960); C. H. Barrow, "Very Accurate Correction of Aerial Photographs for the Effects of Atmospheric Refraction and Earth's Curvature," Photogrammetric Engineering, XXVI, 5 (1960), 798-804.

[20] The condition for conformality can be expressed in several ways, e.g., $a = b$, or

$$\left(\frac{dx}{d\phi}^2 + \frac{dy}{d\phi}^2\right)\cos\phi = \frac{dx}{d\lambda}^2 + \frac{dy}{d\lambda}^2, \text{ or } \frac{dy}{d\phi} = \frac{1}{\cos\phi}\frac{dx}{d\lambda} \text{ and } \frac{dx}{d\phi} = -\frac{1}{\cos\phi}\frac{dy}{d\lambda}, \text{ or}$$

using Gauss' first fundamental forms, $E:F:G = E':F':G'$, and several others.

[21] Or $J = R^2\cos\phi$, or $EG - F^2 = E'G' - F'^2$.

[22] There are also more practical problems, e.g., fitting a particular page format. A discussion can be found in A. Robinson, Elements of Cartography (2d ed.; New York: Wiley and Sons, 1960).

several fields has diverse objectives, requirements for projections differ considerably.[23] Requiring too much (e.g., a geodesic mapping of the entire sphere or a projection which is both conformal and equal-area) may lead to no solutions. More commonly, the person requiring the map is not clear or not specific about the types of problems which are to be solved. In such cases it is difficult to find satisfactory solutions.

Tissot defines a method leading to a solution of a portion of the problem by requiring the conformal map chosen (if conformality is required) to be that which has the least areal distortion.[24] Similarly, the equal-area map used should be that which has the least angular distortion. Mackay has recently suggested other solutions by noting that mathematical conformality and equal-area are not directly proportional to the visual conceptualization of these constructs.[25] Miller has used the concept of a tailor-made projection designed specifically for a particular portion of the globe so that the distortion is distributed about the area in question, minimizing the total distortion.[26] Miller attributes the term to Boggs but the concept is well known, in less refined form, from the use of conic projections for areas of latitudinal extent, cylindric projections

---

[23] This aspect is emphasized in F. Reignier, Les Systèmes de Projection et leurs Applications (Vol. I; Paris: Institut Géographique National, 1957).

[24] An oversimplification; see the references cited under note 27.

[25] J. Ross Mackay, "Geographic Cartography," The Canadian Geographer, 4 (1954), 1-14, and J. R. Mackay, "Conformality: Mathematical and Visual," The Professional Geographer, (N.S.), X, 5 (September, 1958), 12-13.

[26] O. M. Miller, "A New Conformal Projection for Europe and Africa," The Geographical Review, XLIII, 3 (1953), 405-409.

for areas of longitudinal extent, and azimuthal projections for approximately circular areas. Tissot, Laborde, and Young have contributed greatly to the clarification of this difficult problem.[27] Somewhat related is the notion of Umbeziffern as recently used in the German cartographic literature.[28] Basically, this consists of defining a small number of equal-area or conformal mappings of the surface of the sphere to a plane, and then arriving at new projections by transformations of the plane onto itself. The approach appears very promising because of its simplicity.

At one time the drawing of a projection or projection grid constituted a major problem. This led to a search for projections on which the meridians and parallels are curves of specific classes. The solutions now available are: all conformal projections with parametric curves (ϕ, λ) as circles (the straight line as a special case), solved by Lagrange; all equal-area projections with parametric curves as circles (Gravé); all conformal projections with parametric curves as conics (the circle being a special case), solved by Von der Mühll; all equal-area projections with parametric curves as conics (Brown).[29]

Definition of the problem from an alternate point of view naturally leads to different results. The technical capabilities of society are also important. Increased air travel brought the long neglected azimuthal

---

[27]Tissot, op. cit.; Driencourt and Laborde, op. cit., particularly Volume IV; A. E. Young, "Conformal Map Projections," The Geographical Journal, LXXVI (October, 1930), 348-351; also, A. R. Hinks, et al., "A New Treatise on Map Projections," The Geographical Journal, LXXXIII (1934), 145-150.

[28]K. Wagner, Kartographische Netzentwürfe (Leipzig: Bibliographisches Institut, 1949).

[29]Brown, loc. cit.

projections into prominence as air age maps. One of the advantages of the American polyconic projection is the convenience by which maps can be plotted from standard tables.[30] Today it is often simpler to dispense with tables completely, generating data as required on a digital computer, with map display if desired.[31] The use of projections in a real time computer system has recently led Kao to an analysis using linear algebra.[32] The simplicity of the perspective projections proves a considerable advantage in these situations.

A complete enumeration of criteria for the choice of a particular projection would be quite lengthy. Equally tedious would be a complete listing of the properties which obtain on the specific individual projections proposed to date. The reader is referred to the literature with a reminder from Robinson: "There are some projections for which no useful purpose is known, but there is no such thing as a bad projection--there are only poor choices."[33]

This rather brief discussion of map projections must suffice. An attempt has been made to emphasize materials which are not readily available in the American literature on cartography. We now turn to a study of a particular class of projections.

---

[30] C. H. Deetz and O. S. Adams, Elements of Map Projection, U. S. Coast and Geodetic Survey Special Publication 68 (5th ed.; Washington: U. S. Government Printing Office, 1945), p. 60.

[31] W. R. Tobler, "Automation and Cartography," The Geographical Review, XLIX, 4 (October, 1959), 526-534.

[32] R. C. Kao, Geometric Projections and Radar Data (Santa Monica: System Development Corp., 1959).

[33] Robinson, op. cit., p. 57.

CHAPTER III

AZIMUTHAL PROJECTIONS

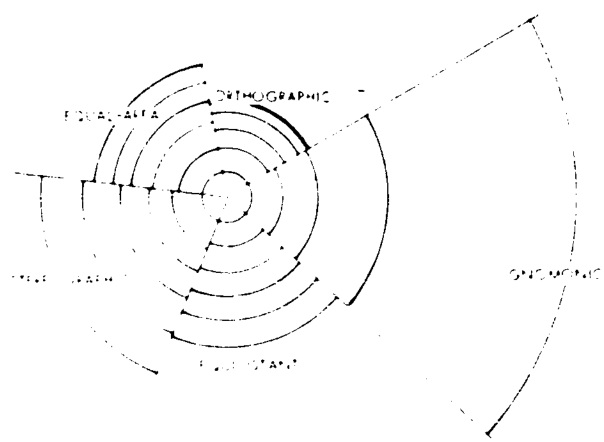
Azimuthal projections are perhaps the simplest of all map projections and are well suited to the present purpose. The graphic methods suggested in the previous chapter can be used effectively in reviewing some of the problems which have concerned past cartographers in dealing with this class of projections. These graphic methods provide a somewhat simpler overview than those available in the cartographic literature and facilitate the transition to the unusual azimuthal projections which are germane to the main problem.

The name azimuthal derives from the characteristic fact that bearings, or azimuths, from the center of the map are correct to all points on the map.[1] The well known azimuthal projections are the stereographic, the orthographic, and the gnomonic. Less well known but still not obscure are the Lambert equal-area and the azimuthal equidistant. Variations of these projections and a considerable number of other azimuthal projections have also been developed. The only difference between the various azimuthal projections is in the spacing of lines of equal spherical distance from the center of the map. This relation is quite clearly shown by

---

[1]In England also zenithal. In German some confusion exists regarding the use of the term. Some authors consider the term zenithal synonymous with centrally symmetric. The term planar occasionally also is used for azimuthal projections. Although azimuthal projections can be considered planar, planar projections are not necessarily azimuthal.

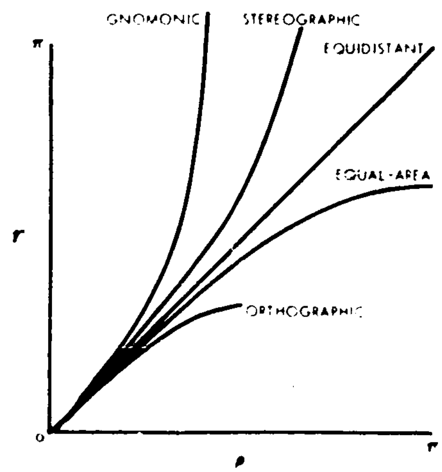
azimuthal maps centered on one of the geographic poles. Although the equatorial and oblique projection aspects retain the same qualities with respect to the center of the map, or point of tangency, the polar aspect is conventionally used. It is simple and well suited to demonstration as the geographic system of latitude and longitude can be employed with only minor modification. This convention is adopted here without loss in generality. A polar or circular diagram (Fig. 3.1) is commonly drawn to illustrate the differences between several azimuthal projections--a difference in the spacing of parallels--and performs this task in an excellent manner.



AZIMUTHAL PROJECTIONS — REPRESENTATION USING CIRCULAR DIAGRAM

Figure 3.1

Another method of illustrating the same relation is to plot spherical distances (in radians) from the point of projection against map distances on a graph (Fig. 3.2). This graph is more suited to the present purpose and has certain pedagogic, as well as analytic, advantages. If one considers the graph as representing a rectangular plot of ground across which one is free to wander in almost any manner one chooses, it is obvious that an infinite number of "paths" may be followed.



AZIMUTHAL PROJECTIONS - REPRESENTATION USING RECTANGULAR GRAPH

Figure 3.2

The five "paths" shown (Fig. 3.2) represent the gnomonic, stereographic, equidistant, equal-area, and orthographic projections, all, of course, azimuthal.

Analytically, azimuthal projections are characterized by equations of the form

$$r = f(\rho, \lambda); \ \theta = \lambda,$$

the latter of this pair being the condition for azimuthality. A special case is represented by centrally symmetric azimuthal projections given by

$$r = f(\rho), \ \theta = \lambda.$$

The cartographic literature, almost without exception, has restricted discussion of azimuthal projections to this special centrally symmetric case. In this instance the circular diagram is appropriate. The rectangular graph is seen to be one of the graphs for projections of category D as established in the previous chapter.

Azimuthal Projections with Central Symmetry

The assumptions concerning differentiability, independence, etc., are initially retained. It is first desirable to establish the simple relation between the rectangular graph and the geometric perspective or projective construction of centrally symmetric azimuthal projections often given in cartographic texts.

The development of a geometric perspective projection onto a tangent plane is shown in the accompanying figure (Fig. 3.3). On the left is an arbitrarily motionless and positioned point of projection. Rays are conceived as being projected from this point through a transparent sphere, on which are drawn lines of latitude and longitude (and geographic data), to a plane tangent to this sphere. Taking the point of tangency as the north pole, the normal case for planar projections, simplifies matters somewhat. The map which is "projected" to the tangent plane can be seen by rotating this plane or can be constructed by drawing the lines of latitude as circles with radii as indicated by the intersection of the rays with the tangent plane. Continuing further to the right (Fig. 3.3), these same radii can be plotted on a graph against their spherical distance (in radians) from the north pole. The graph and the circular diagram, or map, are in a certain sense equivalent. The curve on the graph, of course, shows the so-called radius equation for the particular centrally symmetric azimuthal map chosen. The reader should study this figure carefully and fully understand the transition from the polar case map to the rectangular graph <u>and</u> the reverse procedure of going from the rectangular graph to the

RELATION OF PERSPECTIVE DEVELOPMENT AND GRAPH OF RADIUS EQUATION

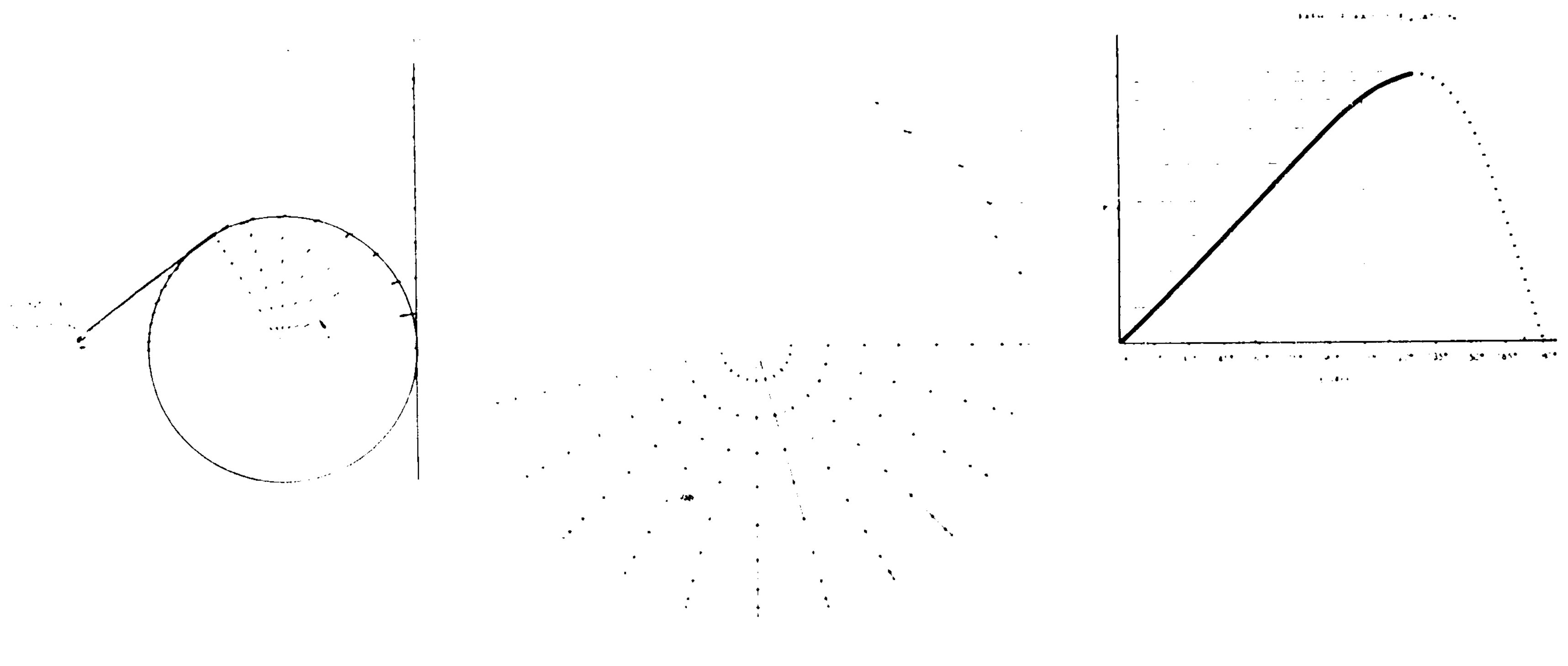


Figure 3.3

creation of a map. The transition from the graph to the map will be extensively referred to later. The transition from map back to the projective situation, however, will not always be meaningful. The dotted line on the graph represents that portion of the sphere which is not represented on the map, a spherical cap excluded because we take only the intersection of the ray and the transparent sphere lying closest to the tangent plane.[2]

Three of the conventional azimuthal projections (the gnomonic, stereographic, and orthographic) can be obtained by this projective method (Fig. 3.4).

PERSPECTIVE DEVELOPMENT OF THE COMMON AZIMUTHAL PROJECTIONS

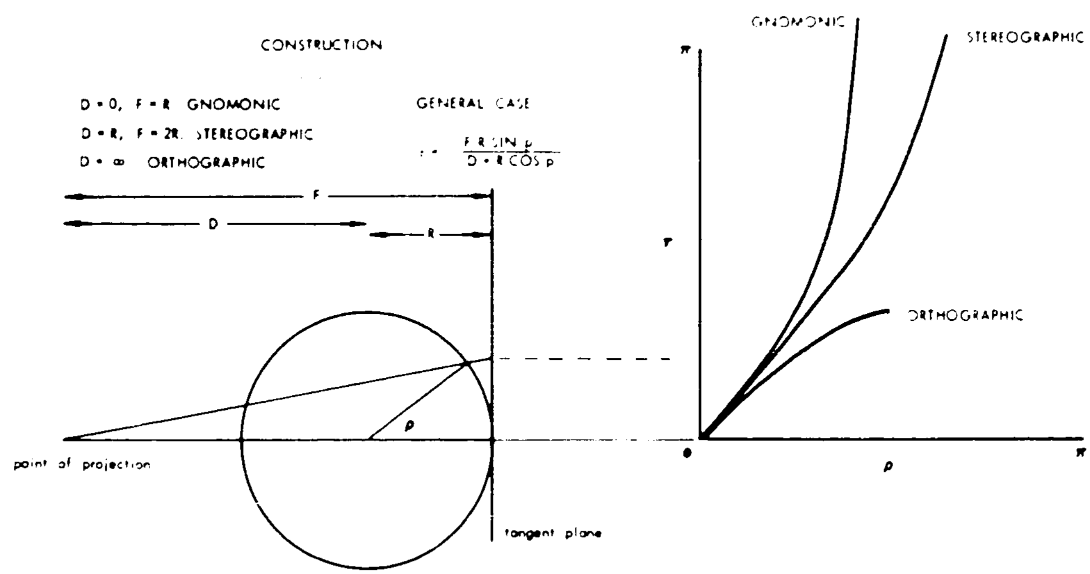


Figure 3.4

The gnomonic, of course, shows all great circles as straight lines but can be used only for areas of less than a hemisphere. The stereographic

---

[2]In practice it is clear which hemisphere is intended. For a more precise definition, see Kao, op. cit., pp. 9-10.

is conformal but will not quite show the entire sphere (on a map of finite extent). The orthographic, limited to a hemisphere, is generally not considered to have any special properties but the polar case shows the parallels in their true length.

In order to restrict the discussion to centrally symmetric azimuthal projections, the point of projection is required to be on an axis of the sphere representing the earth and the projection plane is to lie at right angles to this axis. Having selected a fixed point of projection but varying the position of the plane of projection along this axis changes the scale of the map (except in the case of the orthographic projection). Requiring the projection plane to be tangent to the sphere and changing the position of the point of projection on the axis results in map projections having different properties. As any position, from minus to plus infinity, along the axis can be chosen for the point of projection, there is obviously an infinitude of centrally symmetric azimuthal projections. Clearly, the criteria for the choice of a projection or projection point are ones of usefulness. Projection from select points intermediate to the stereographic and orthographic, with characteristic graphs, are shown in Figure 3.5. Virtually all projections in this category have been selected to balance, distribute, average, or minimize various types of distortion. Clarke's minimum error projections, and the projection of Sir Henry James, are of this type. Airy's projection, though not geometrically perspective, is arrived at by a comparable minimization of errors. Similarly, Breusing produced a projection by using the geometric mean between the conformal and equal-area azimuthal projections. Young later gave evidence that the harmonic mean produces even less errors. Other intermediate azimuthal projections of interest are the approximate equidistant and approximate equal-area suggested by Kellaway. These are typical examples of a search for perspective or projective models which approximate projections derived from analytical considerations. Maurer lists many further centrally symmetric azimuthal projections, some of which are derived from geometric models, others by manipulation of the radius equation. The intermediate projections have occasionally been used for atlas and wall maps; maps of the lunar surface now also are being prepared on an intermediate azimuthal projection.[3]

PERSPECTIVE DEVELOPMENT OF INTERMEDIATE AZIMUTHAL PROJECTIONS

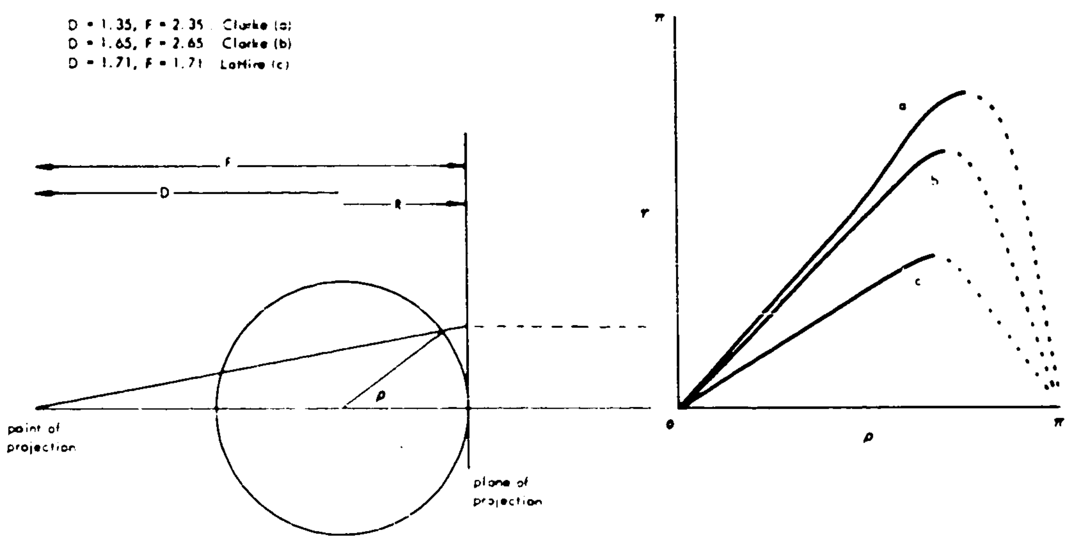


Figure 3.5

[3]L. Y. Dameron, Jr., "Projection for Lunar Map," The Military Engineer, (January-February, 1960), p. 5. Tables for several of these projections are given by Hammer and by Herz: E. Hammer, Ueber die Geographisch Wichtigsten Kartenprojektionen (Stuttgart: Metzler, 1889); N. Herz, Lehrbuch der Landkarten Projektionen (Leipzig: Ta[illegible]er, 1885).

The Soviets are reputed to have objected to the map projection used by R. E. Harrison in his atlas, Look at the World, on the grounds that the projection is "unscientific."[4] Certainly the projection is not one of the standard projections, but it is nevertheless a valid projection. This so-called unscientific and certainly unconventional projection has recently become significant, for it, in fact, shows the earth as it might be seen from an orbiting satellite. A similar geometry obtains in conventional low altitude photogrammetry. As the distortion due to the earth's curvature is in this latter case negligible, it is normally ignored. If one assumes the earth to be a sphere and that the camera is directed in the vertical (the nadir point and the principal point coincide), the equations for this projection are easily derived. The relation to the general equation for perspective azimuthal projections is also quite simple. As the camera distance from the earth becomes infinitely large, the map converges to an orthographic projection. If the camera is not directed in the vertical, the projection is no longer azimuthal and converges on Lecoq's projection.[5] This corresponds to a planar projection with a projection point not on a spherical axis perpendicular to the plane of projection. Planar projections and azimuthal projections are not synonymous, as many authors would have one believe.

---

[4]R. E. Harrison, Look at the World (New York: A. A. Knopf, 1944). Chiao-Min Hsieh, "The Status of Geography in Communist China," The Geographical Review, XLIX (October, 1959), 543.

[5]Reignier, op. cit., Vol. I, p. 259. Equations for a more general case are derived in H. Merkel, "Die Allgemeine Perspektivische Abbildung der Erdkugel," Festschrift Eduard Dolezal (Sonderheft #14 der Oesterreichischen Zeitschrift fuer Vermessungswesen; Vienna, 1952), pp. 169-182.

It is also possible to project points on the surface of a sphere to some other surface and then to the map plane. A centrally symmetric azimuthal projection can be achieved by projecting gnomonically to a tangent cone (with vertex on the axis of projection) and then orthographically to the map plane (Fig. 3.6).

A DOUBLE PERSPECTIVE PROJECTION

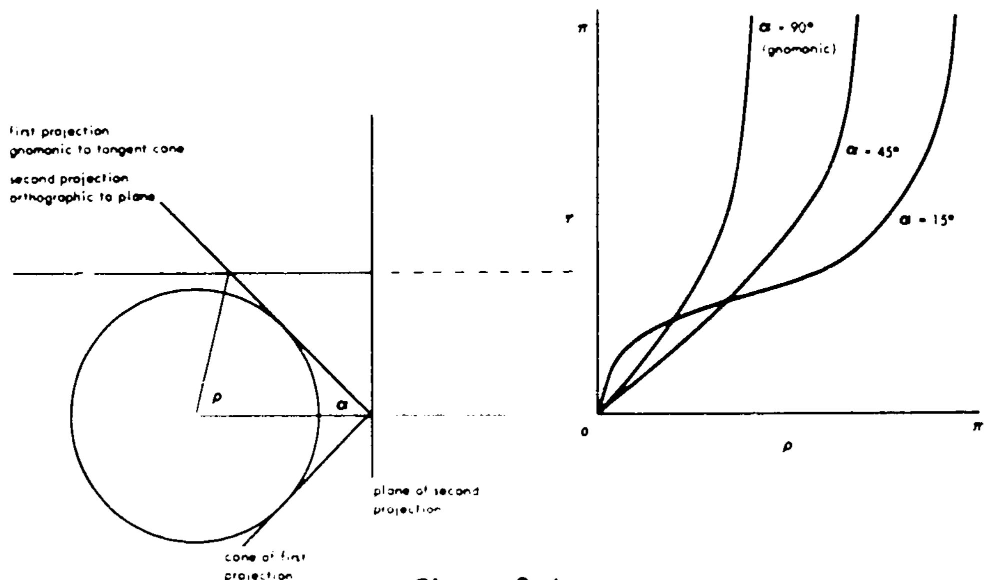


Figure 3.6

Similar projections can be achieved by varying the first and/or second method of projection. Maurer credits the invention of projections onto a cone, then plane, in this fashion to Lidman (1877).[6] Comparable double projections can be developed by using a paraboloid rather than a cone. Projection onto ellipsoids, hyperbaloids, spheres, etc., secant, tangent,

[6]Maurer, op. cit., (System Number 22).

or disjoint, in gnomonic, stereographic, orthographic, or intermediate versions is also quite possible. These projections, of course, are quite similar to Raisz' orthoapsoidal projections.[7] The double projection onto spheres of radii of multiples of that of the reference globe in particular has been investigated by M. Solovyev.[8] One of these cases (Fig. 3.7) is found to be a geometric model for the Lambert equal-area projection. It was previously believed that this projection could not be constructed

LAMBERT'S AZIMUTHAL EQUAL-AREA AS A DOUBLE PERSPECTIVE PROJECTION

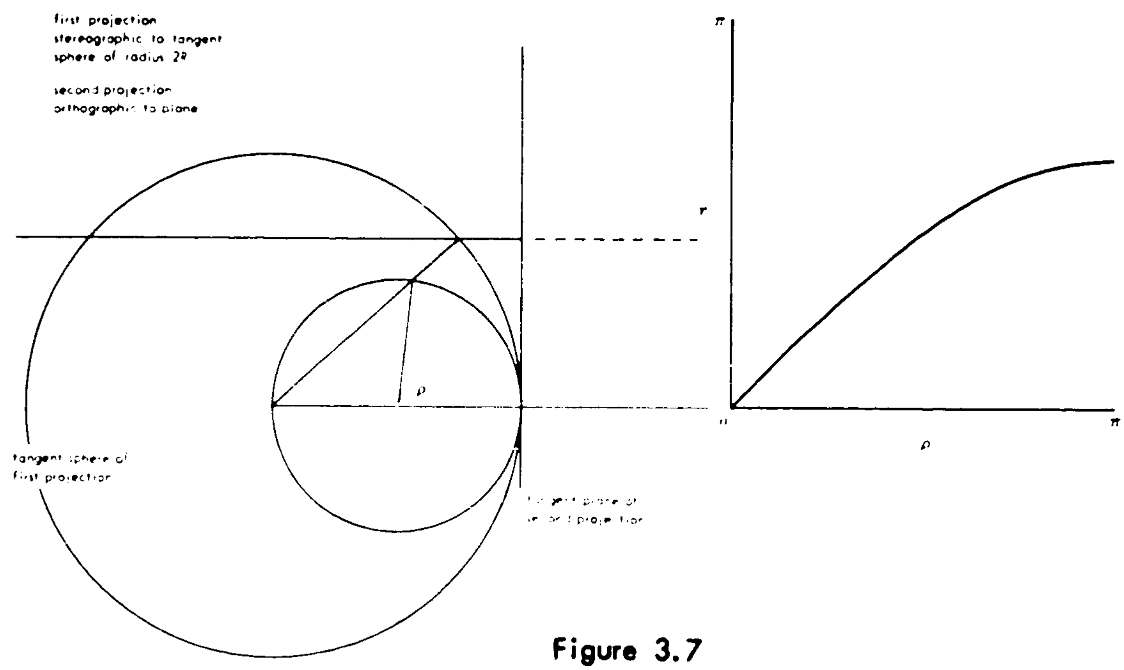


Figure 3.7

---

[7]E. Raisz, "Orthoapsoidal World Maps," The Geographical Review, XXXIII (1943), 132-134.

[8]Noted in Graur, op. cit., pp. 198-199.

from a perspective situation.[9] The Hammer projection (not azimuthal) can be developed by further projecting the Lambert azimuthal equal-area map of a hemisphere onto a new plane tilted at sixty degrees from the first. Aitoff's projection is similarly developed.[10] The two point azimuthal projection, on which all great circles are straight lines, can also be obtained by projecting the gnomonic orthographically onto a tilted plane.[11] Multiple projection (above the second order) will result in many more azimuthal maps. The mapping equations for these projections are normally quite easily derived.

Another simple method of "inventing" azimuthal projections is to space the parallels (in the polar case) arbitrarily or according to some cylindric or conic projection. Selection of radii according to the Mercator projection is possible, for instance. The valuable properties of the Mercator projection are understandably lost in this azimuthal version. This

---

[9]G. P. Kellaway, Map Projections (2d ed.; London: Meuthuen, 1949), pp. 16-21; Deetz and Adams, op. cit., p. 40.
If the general equation for perspective azimuthal projections is solved for a point of projection in the case of the azimuthal equal-area, it is found that the location of the point of projection is a function of the latitude being projected. That is, projection must be considered to be from a moving point. The cartographic literature is not consistent when it accepts moving points of projection for cylindrical projections (Driencourt and Laborde, op. cit., Vol. I, p. 73), but rejects such an infinitude of projections for planar projections. From a different point of view, the question should be whether this projection satisfies the axioms of the projective group (see F. Klein, Elementary Mathematics from an Advanced Standpoint: Geometry (New York: Dover, 1939), pp. 130 et seq.

[10]But from an azimuthal equidistant map. See J. B. Leighly, "Aitoff and Hammer: An Attempt at Clarification," The Geographical Review, XLV, 2 (1955), 240-249.

[11]Miller and Fisher, op. cit., p. 66. This projection is also referred to as the orthodromic (Encyclopaedia Britannica) and is a simple linear transformation of the gnomonic.

would normally be the case. Similarly, the radial distances given for the centrally symmetric azimuthal projections can be interpreted as distances of parallels from the equator on equatorial aspects of cylindrical projections. Thus, the herein suggested modifications can be applied to invent new cylindric or conic projections. A choice must then still be made for the spacing of the meridians. An example of this procedure is given by Adams.[12]

Centrally symmetric projections are also easily and advantageously studied without any reference to perspective models. It is only necessary to refer to the radius equation, or equivalently, to the graph of this function. While it is true that <u>any</u> arbitrary line drawn on the graph can be interpreted as a centrally symmetric azimuthal projection in the manner described, i. e., interpreted by working backwards from the graph to the map, certain curves do not meet the requirements earlier stated and involve difficulties of geographic interpretation. These are omitted for the moment. The graphs presented in Figure 3.8 should cause no trouble and all represent valid map projections. Equivalently, any equation involving two variables can easily be interpreted as a map projection.[13] These results follow rather immediately from the classificatory scheme used here. There is clearly no need to restrict consideration to trigonometric functions,

---

[12]O. S. Adams, <u>General Theory of Equivalent Projections</u>, U. S. Coast and Geodetic Survey Special Publication 236 (Washington: U. S. Government Printing Office, 1945), p. 46 <u>et seq</u>.

[13]Rather as half of a map projection because it represents only one family of parametric curves, just as only the one graph has been illustrated. The restrictions on the equations are discussed below.

as has been fairly common in the cartographic literature.[14]

GRAPHIC EXAMPLES OF TRANSFORMATION EQUATIONS

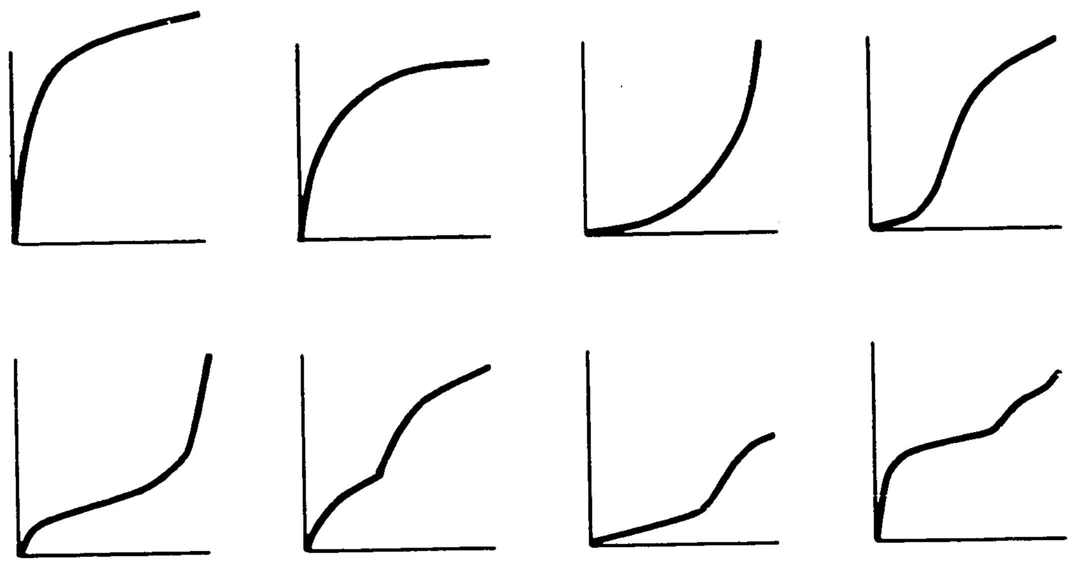
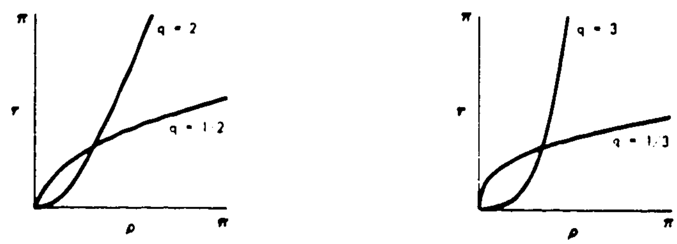


Figure 3.8

If these one parameter equations are to be interpreted as the radius equation of centrally symmetric azimuthal projections, then the other family of lines is fixed and the projection given by: $\theta = \lambda$, $r = f(\phi)$. For example, the generating equations $r = R\phi$, $\theta = \lambda$ for the azimuthal equidistant projection can be generalized to create a larger class of projections, namely, $r = R\phi^q$, $\theta = \lambda$, where the exponent q is an arbitrary positive element from the set of real numbers. In this formulation the azimuthal equidistant projection is the special case, $q = 1$. A few "paths" from this family of curves are shown in Figure 3.9.

[14]However, given objectives of equal-area or conformality, trigonometric functions of necessity occur.



$r = \rho^q$

SAMPLE GRAPHS

Figure 3.9

As an illustration, a map (Fig. 3.10) centered on New York has been prepared using the projection: $r = R\rho^{\frac{1}{2}}$, $\theta = \lambda$, a square-root azimuthal projection. An indication of the distortion which occurs accompanies the map.[15] Thus, map projections are particularly easy to create by arbitrarily writing down equations which satisfy certain simple requirements. Such a procedure may be interesting and occasionally fruitful, but is somewhat foolish as there are infinitely many; it is also necessary to demonstrate

---

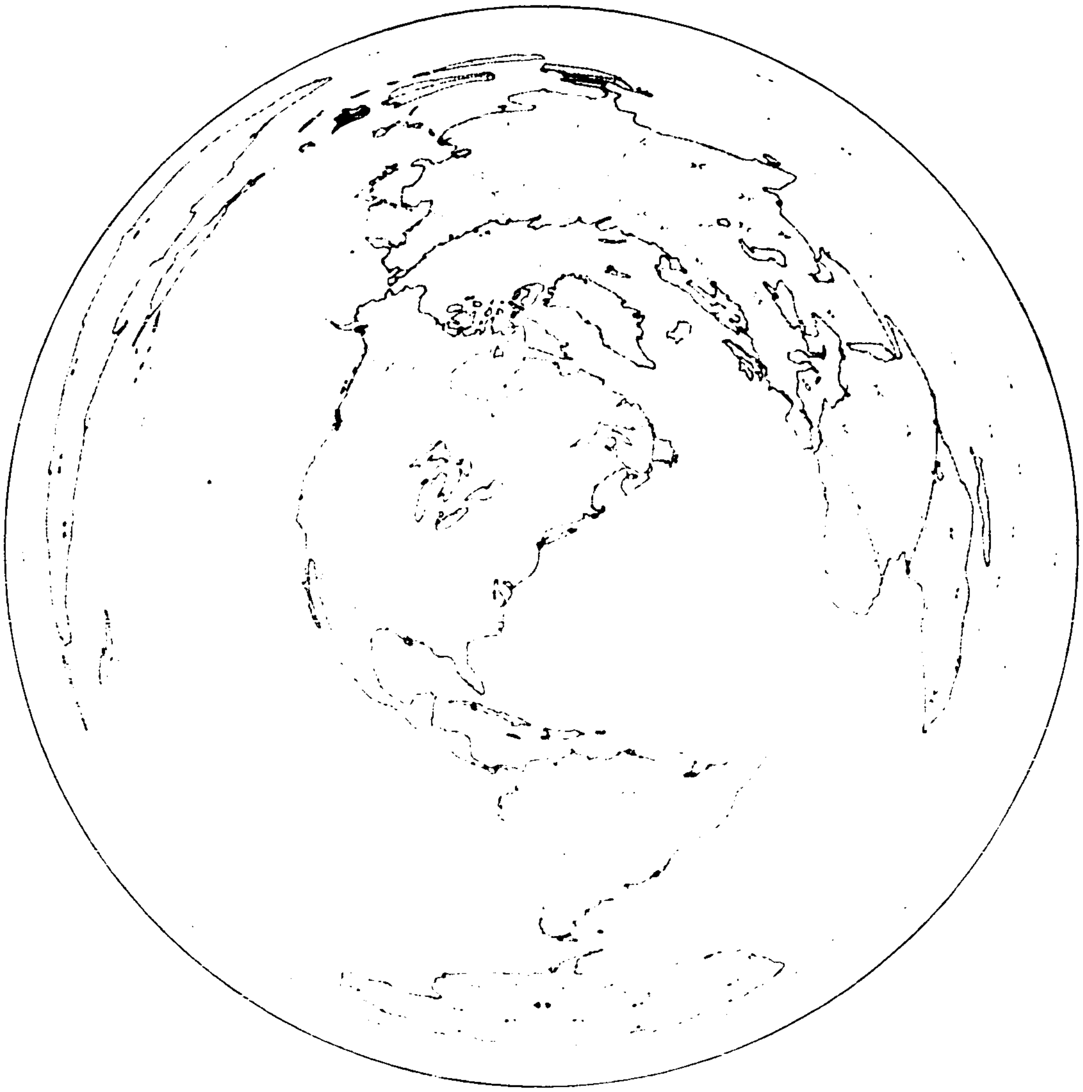
[15]For some projections, including the azimuthal projections with central symmetry, Tissot's measures of distortion are also only functions of one parameter. As the diagrammatic procedure is quite simple, it seems desirable to include such graphs when the maps are published. For azimuthal projections with a one-parametric radius equation, the values are given by:

$$a = f(\phi) = \frac{1}{R}\frac{dr}{d\rho}; \quad b = f(\phi) = \frac{r}{R \sin\rho}; \quad S = f(\phi) = ab$$

$$\Omega = 2 \sin^{-1} \frac{(b - a)}{(b + a)} = f(\phi).$$

55

SQUARE - ROOT PROJECTION

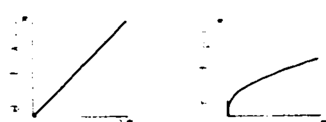


An Azimuthal Projection of the World

CENTERED ON NEW YORK CITY

40° 45' N., 74° W.

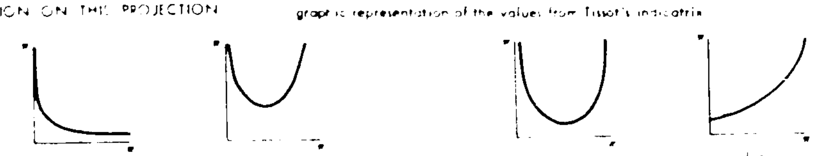
DIAGRAMMATIC REPRESENTATION OF THE EQUATIONS FOR THIS PROJECTION



Azimuth Equation — Radius Equation

The scale of these diagrams corresponds to a globe of one radian. The map scale can also be obtained from this value for any point on the map or can be obtained from the graphs of the distortions.

DISTORTION ON THIS PROJECTION — graphic representation of the values from Tissot's indicatrix



Radial — Circumferential — Area — Angular

Cartographic Laboratories, University of Washington, Seattle

Figure 3.10

some property, relation or use.[16] Further azimuthal maps can be developed by fitting a map to an arbitrary theoretical or empirical curve.

Graphic Methods of Analysis

In addition to simple visualization and comparison of centrally symmetric azimuthal projections the graphs provide a rapid method of analyzing certain properties of these projections. Only a few examples need to be mentioned for extensions of this method are simple. The azimuthal equidistant projection ($r = R\rho$, $\theta = \lambda$), which has gained increased popularity in recent years, is perhaps the most basic, for the radius equation can be represented as a straight line on the graph.[17] If the diagram is scaled in radians, this line is a forty-five degree line (of slope +1). As the name implies, all places are represented on the map at their correct spherical distance from the center. If the curve representing the radius equation of any other centrally symmetric azimuthal projection crosses this line, the intersection point will also lie at its true distance from the center of the map. Similarly, whenever the radius equation of an azimuthal projection of category D intersects the graph of the corresponding equation for the equal-area azimuthal projection twice, the projection in question has an area in some sector which is equal in area to the corresponding sector on the globe. Similar comments can be related to the lengths of the

---

[16] Nehari appropriately remarks that the investigation of particular functions, though yielding much insight, is of necessity a piecemeal procedure. Much more fruitful are investigations akin to those of Bernhard Riemann which examine general existence conditions. Z. Nehari, Conformal Mapping (New York: McGraw-Hill, 1952), p. 173.

[17] For a detailed discussion of this projection, see C. Hagen, "The Azimuthal Equidistant Projection," (Master's thesis, University of Washington, 1957).

parallels and intersections with the graph of the orthographic projection. Projections with some of these properties have been sought from time to time. It is also easy to demonstrate why the most used azimuthal projections can have only one standard parallel and why the conic projections can have two standard parallels.

A change of scale on an azimuthal equidistant projection corresponds to a raising or lowering of the straight line on the graph, keeping the origin intersection fixed. This can also be seen from the equations as a constant change in slope. Changes of scale on other projections are similar.

To create composite projections, one need only join projections where their graphs intersect. This is illustrated for the equal-area and conformal azimuthal projections (Fig. 3.11), the scale of the stereographic having been reduced to achieve intersection (combination of an equal-area and conformal map is not recommended, but used only as an illustration).

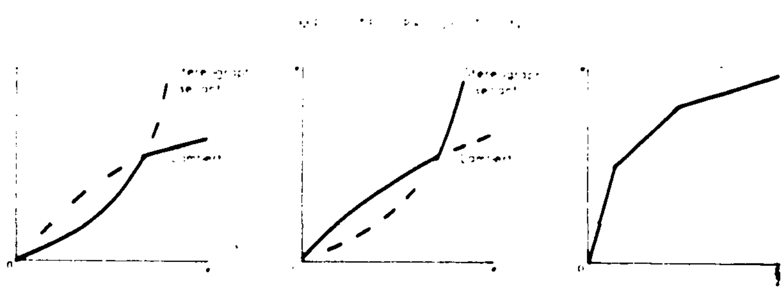


Figure 3.11

A similar, non-azimuthal, composite is Goode's famous homolosine, combining the homolographic (Mollweide) and sinusoidal equal-area projections. Somewhat comparable is Wray's combination of the azimuthal equidistant and the azimuthal equal-area projection into a star-shaped map.[18] A composite map, with large scale near the center, has recently been employed for the use of aviators in the vicinity of airports.[19] The figure (Fig. 3.11c) illustrates the technique, but not the specific projections employed.

The radial scale distortion on an arbitrary azimuthal projection of category D can be estimated by noting the departure of the graph of the radius equation from that of the azimuthal equidistant projection; actually, it is the slope of the radius equation which is important. Curves which are concave downward in general expand the central area of the map and diminish the peripheral areas. The reverse is generally true of curves which are concave upward. This is approximately correct for the radial scale of the map but for areal distortion it is true to only a limited extent due to the spherical model of the earth (compare Fig. 3.10). The graphic analysis, of course, should only be undertaken when the analytic considerations are understood and can be demonstrated. A graphic demonstration of conditions relating to conformality and equal-area, particularly the differential relations between the pairs of equations seems difficult.

Asymmetric Azimuthal Projections

The classificatory scheme employed here makes it obvious that azimuthal projections lacking central symmetry are also possible, though they

---

[18]S. C. Gilfillan, "World Projections for the Air Age," Surveying and Mapping, VI, 1 (January, 1946), 12-18.

[19]L. Y. Dameron, Jr., "Terminal Area Charts for Jet Aircrafts," The Military Engineer (May-June, 1960), p. 6.

can be neither conformal nor equal-area. The condition for azimuthality is $\theta = \lambda$, and the second, or radius, equation has the form $r = f(\rho, \lambda)$. The maps obtained are without doubt somewhat unusual; their relevance to the present problem is indicated in the following chapter. The parallels (polar case) no longer appear as concentric circles. Hammer dismisses these projections in the same sentence in which he recognizes their existence, and there is no cartographic history to review.

The graphic methods can be employed as with all projections of category B. To obtain a map from the graph, one can proceed as before except that a different radius equation applies to each azimuth. A few examples from this infinite set are shown in the accompanying figure, both as map diagrams and as level curves on a graph (Fig. 3.12).

Relaxed Assumptions

To attempt to specify which equations or curves can validly and without difficulty be interpreted as map projections, the reader is again referred to the figure showing the radius curves for the well known azimuthal projections of category D (Fig. 3.2). It is immediately apparent that these curves are of a very special nature. The restrictions can be stated quite explicitly. A specification of the domain is clearly quite important. Concern is only with curves which appear in the first quadrant, i.e., $0 \leq r < \infty$; $0 \leq \rho \leq \pi$. The closed upper end on the abscissa is not necessary, for many projections can cycle endlessly. The orthographic projection, with equations $\theta = \lambda$, $r = R \sin \rho$ is restricted to the interval $[0, \pi/2]$. The equations for the stereographic projection are $\theta = \lambda$, $r = 2R \tan \rho/2$, with domain $[0, \pi)$, a semi-closed interval. The gnomonic is $\theta = \lambda$,

ASYMMETRIC AZIMUTHAL PROJECTIONS

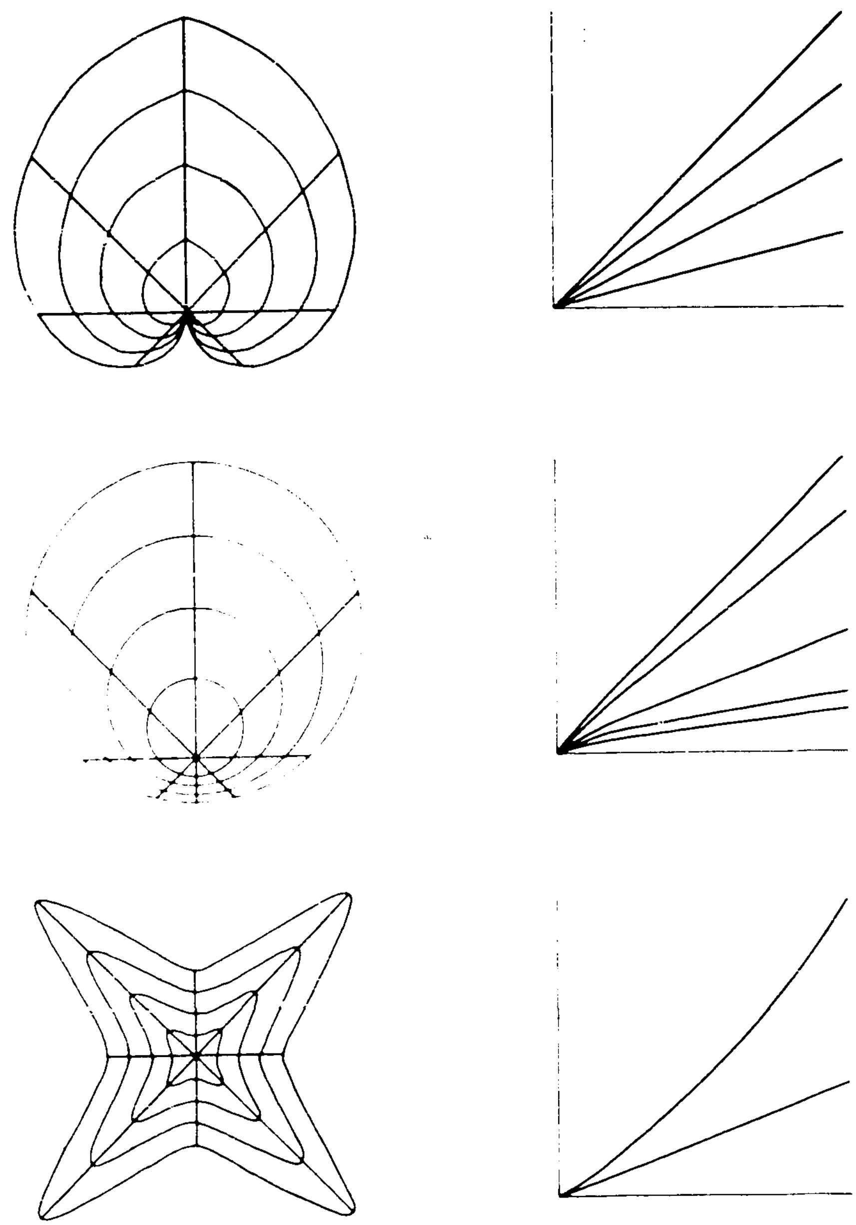


Figure 3.12

$r = R \tan \rho$, and has the semi-closed interval $[0, \pi/2)$ as its domain.

In order to specify more explicitly the important restrictions on the radius equation of conventional azimuthal map projections of category D, the end points of the interval for which the map projection is defined are labelled e^* and e^{**}. The same restrictions apply with slight modifications to the cylindric and conic projections of category D, and, with further modifications, to the projections of categories A, B, and C. The restrictions, in addition to those cited in the previous chapter, are:

a) $dr/d\rho > 0$, except $dr/d\rho \geq 0$ at e^* and e^{**} in the closed interval $[e^*, e^{**}]$. This condition can be restated as $dr/d\rho > 0$ in the open interval (e^*, e^{**}). Although the five azimuthal projections of Figure 3.2 conform to this restriction, it is unnecessarily severe. A requirement that $r = f(\rho)$ be strictly monotone increasing is sufficient for more general azimuthal maps. This allows singular points at which $dr/d\rho = 0$.

b) $d^2r/d\rho^2 > 0$, <u>or</u> $d^2r/d\rho^2 < 0$, except in the case of the azimuthal equidistant projection, all in the interval $[e^*, e^{**}]$. This condition can be dropped completely without leading to difficulties of interpretation.

c) $r(0) = 0$

d) $dr/d\rho$ exists; i.e., $f(\rho)$ is of class c^m, $m \geq 2$, in the open interval (e^*, e^{**}).

Conditions (a), as modified, (c), and (d) are sufficient to insure that $r = f(\rho)$ is continuous, is single-valued, and has a single-valued inverse, all in the domain under consideration. The domain for which a map projection is validly defined is in practice not known in advance; rather, it is determined by the above conditions as applied to the particular function being examined.

A few curves which the restrictions preclude are shown in Figure 3.13.

TRANSFORMATIONS VIOLATING USUAL ASSUMPTIONS

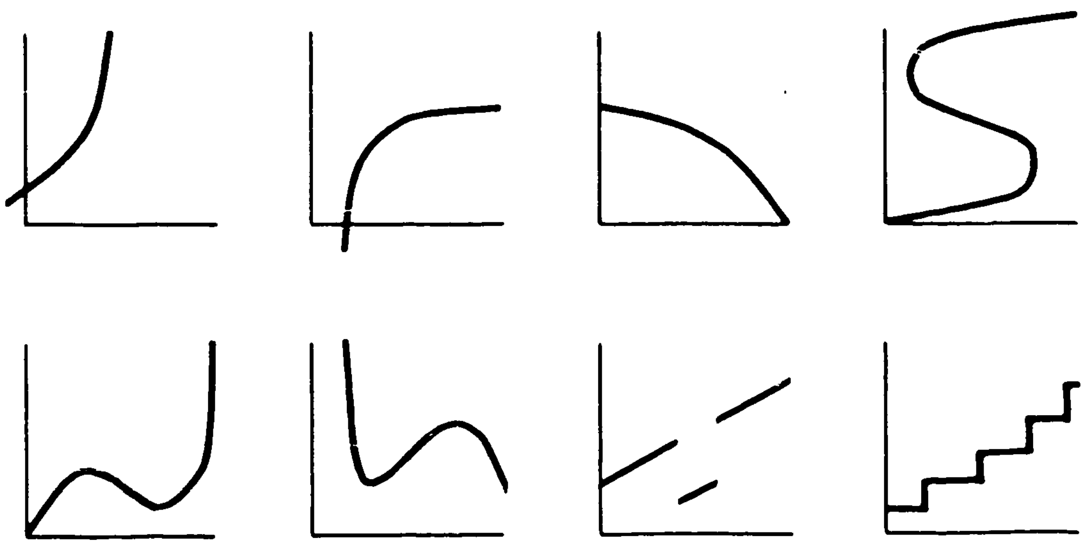


Figure 3.13

The strategy is to examine such transformations, assuming them to be valid map projections, and to proceed graphically by translating the curves into maps in exactly the same manner as has already been demonstrated. The reasons for the restrictions appear rather obviously in this examination and need not be discussed. For the sake of simplicity linear cases are used and generally only one property is examined at a time. The task of assembling the separate and non-linear cases into one map is left to the reader. The discussion focuses on centrally symmetric azimuthal projections as the transition from graph to map is quite easily understood in this case; this device does not diminish the generality of the discussion. Eight specific situations will be recognized. These are not all mutually exclusive and may occur in various combinations. Where well known projections violate some of the restrictions, and this occurs surprisingly often, they will be cited to facilitate understanding.

Particular Cases

Figures 3.14 through 3.21 show individual cases which violate the conventional restrictions. One should note that each position on the abscissa of the graph is interpreted as a circumference on the ground. For the polar case map this is a circle of constant latitude. The ordinate of the graph becomes a radius on the transformed map. Included with some of the figures are graphs showing variants of the specific situation being examined. The reader will be able to imagine the appearance of the maps generated by these variants.

Interruption

Interruption of map projections is fairly common, Goode's interruption of the homolosine projection perhaps being the best known example.[20] The concept has also been applied to other projections such as the star-shaped maps and occurs on Cahill's butterfly projection, Jefferson's six-

---

[20] J. P. Goode, "Studies in Projection: Adapting the Homalographic Projection to the Portrayal of the Earth's Surface Entire," Bulletin of the Geographical Society of Philadelphia, XVII (1919), 103-113.

A. H. Robinson, "Interrupting a Map Projection: A Partial Analysis of its Value," Annals, Association of American Geographers, XXXXIII, 3 (September, 1953), 216-225.

J. W. Stewart, "The Use and Abuse of Map Projections," The Geographical Review, XXXIII (1943), 589-604.

six map, and the polyhedral projections.[21] Interruption also occurs on conic projections, and at the boundary of any map, since the sphere has no boundaries.

The graph $r = a\rho + C$ (Fig. 3.14) violates the third condition for $r(0) \neq 0$.

INTERRUPTION

Figure 3.14

Here C, a constant greater than zero, is the r-intercept of the curve. This yields a ring-shaped azimuthal map with a hole in the center. Such maps have been noted in the cartographic literature, the constant occurring naturally in integration.[22] The origin of the map, a point on the ground, becomes a circle of radius C on the map. A line passing through this point on the ground is interrupted by the open circle on the map.

The variant is simple, the interruption occurring at different places

---

[21] B. J. S. Cahill, "Static and Dynamic World Maps," Bulletin of the American Meteorological Society, XV (1934), 261-265.

M. Jefferson, "The Six-Six World Map," Annals, Association of American Geographers, XX, 1 (March, 1930), 1-6.

Miller and Fisher, op. cit., pp. 92-105.

[22] If it is required that the projection be azimuthal and equidistant, then:

$$a = 1 = \frac{1}{R}\frac{dr}{d\rho}, \text{ whence } R\,d\rho = dr \text{ and } r = R\rho + C.$$

on the map. The variant can be written as $r = a\rho$, $0 \geq \rho \geq \rho_1$; $r = a\rho + C$, $\rho_1 > \rho \geq \rho_2$, etc., defining the interval for each equation.

Elimination

A space elimination occurs when some areas of the spherical surface do not appear on the map. This is well known from the Mercator projection which cannot show the poles, or the orthographic and gnomonic projections which exclude at least a hemisphere.

The graph $r = a\rho + C$, $(C < 0)$ again violates the third condition for $r(0) \neq 0$.

ELIMINATION

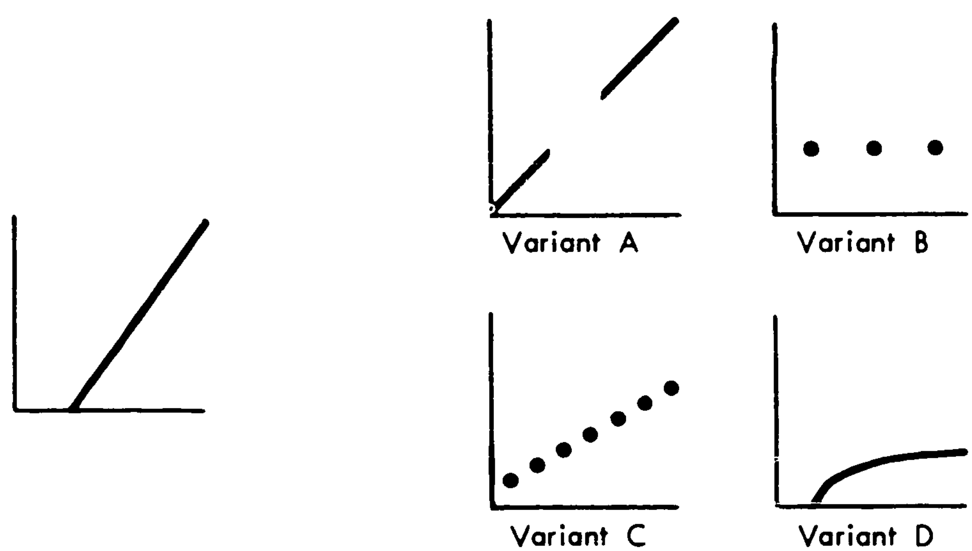


Figure 3.15

The result is a map which cannot show the area in the interval $[0, \rho_0)$, where ρ_0 is the ρ-intercept of the curve. The circumference of the circle of radius ρ_0 on the sphere becomes a point on the map. The variants are somewhat interesting. In each case a vertical extension of some point(s) on the ρ-axis does not intersect the graph of the radius equation. Here again, the equation must be accompanied by a specification of the domain. When the graph consists of lines or line segments the procedure is obvious. When the domain consists of only a finite number of points (Fig. 3.15d), the radius equation takes the appearance of a table:

r	ρ
r_1	ρ_1
r_2	ρ_2
r_3	ρ_3

It is not meaningful to speak of elements which are not in the domain thusly defined. The last variant shows the logarithmic function.

Repetition

A case of space repetition will occur when r is a multivalued relation of ρ. Some places appear on the map in several positions, a one-to-many mapping. It is well known that cylindrical projections can cycle endlessly along the equator. This repetition of positions can occur on many projections, occasionally in several directions. Several cycles of an azimuthal projection are shown as a variant.

REPETITION

Figure 3.16

Collapsing

In this situation areas are collapsed into lines (or points), a many-to-one mapping. The curve shown (Fig. 3.17) violates the first condition for the function, though monotone, is not strictly monotone increasing. Restated, $dr/d\rho = 0$ for an interval rather than only at singular points. In terms of an azimuthal map, the entire area represented by the interval for which $dr/d\rho = 0$ is collapsed into a circular line. The annulus of area maps into the circumference of a circle. The radius

COLLAPSING

Figure 3.17

equation is of the form:

$$r = r_1 \text{ if } 0 \geqslant \rho > \rho_1$$
$$r_2 \text{ if } \rho_1 \geqslant \rho > \rho_2$$
$$r_3 \text{ if } \rho_2 \geqslant \rho > \rho_3, \text{ etc.}$$

Expansion

On many azimuthal projections the antipodal point is expanded into a circumference, and cylindrical projections map the pole into a line. The function shown in Figure 3.18 is somewhat similar. The fourth condition is violated for $dr/d\rho$ is undefined at singular points or becomes infinite. The distance from the center at which this occurs, a circular line of infinitesimal width on the ground, is infinitely distorted to become an area on the map. Other areas may be omitted. Such functions are most easily written in inverse form.

EXPANSION

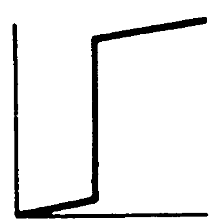


Figure 3.18

Dislocation

Items are said to be dislocated when they do not appear in their proper relative location on the map. In their simplest form such situations can be characterized by a disruption of order relations. The graph is self explanatory.

DISLOCATION

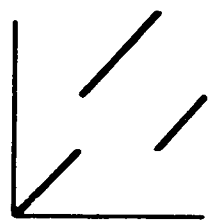


Figure 3.19

Superimposition

An overlapping or superimposition of areas appears on the map when the inverse function is multivalued (Fig. 3.20).

SUPERIMPOSITION

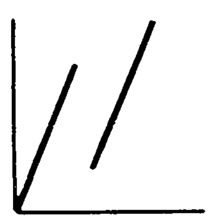


Figure 3.20

One or more positions on the original surface take identical positions on the map. The collapsing of space cited earlier is a similar many-to-one mapping. Conic projections in which the constant of the cone is greater than unity would result in superimposition of areas on a map. It is difficult to draw maps when two (or more) places occur at the same position. Such maps are equally difficult to read or use, being somewhat comparable to multiply exposed photographs. The graph, however, is easily understood.

Inversion

An inversion of space results from an equation which has an interval for which the derivative is negative (Fig. 3.21).

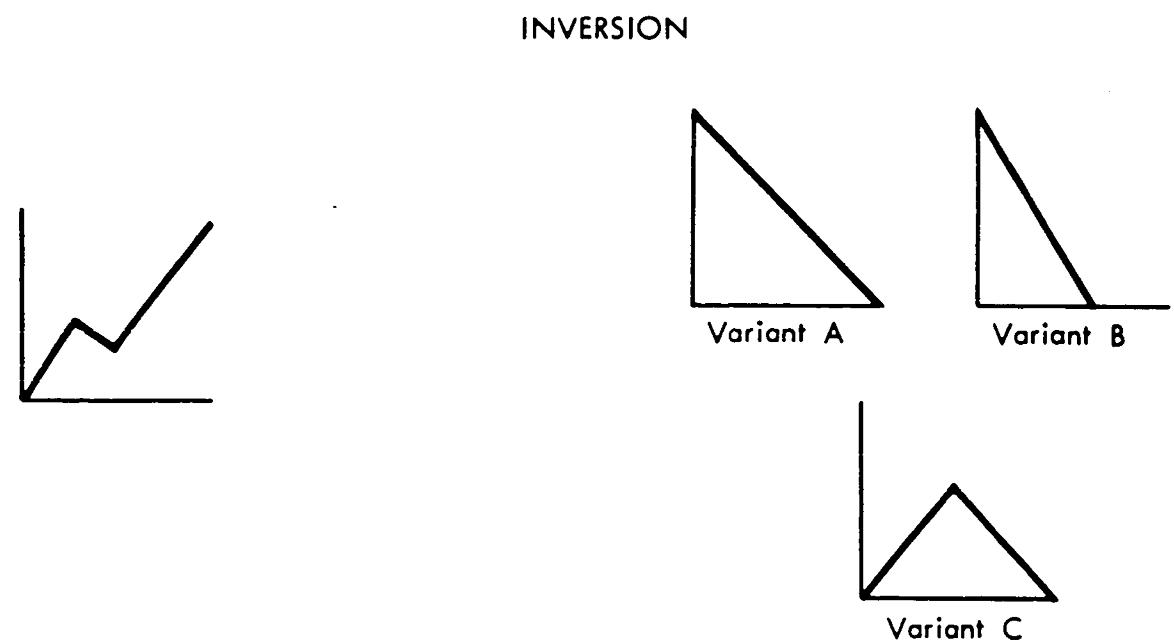


Figure 3.21

This implies that order relations are disrupted in such a way that places which are at greater distances on the ground are closer on the map. The simplest case, $r = -\rho + \pi$, (variant a), can be interpreted as an azimuthal equidistant map with no difficulty except that it is centered on the antipodal point of what was previously the point of tangency. Such a map actually has been used.[23] Similarly, the function $r = a\rho + \pi$, $a < -1$, (variant b), can be interpreted as an interruption (Fig. 3.14), but centered on the antipodal point. This manner of interpretation can be applied to any function which is strictly monotone decreasing.[24] The graphs are again easily understood, but the multivalued relations, combining inversion and superimposition, involve maps which are difficult to draw.

Order Relations

The eight situations which have been identified can all be characterized by rearrangement of the order relationships. Assign numbers to positions on the ρ-axis beginning with those closest to the origin and then examine these order relations on the image or map, as follows:

a)	Interruption	(1234567)	→	(12 345 6 7)
b)	Elimination	(1234567)	→	(1267)
c)	Repetition	(1234567)	→	(12323456657)
d)	Collapsing	(1234567)	→	(1267) 3 4 5
e)	Expansion	(1234567)	→	(123333333334567)
f)	Dislocation	(1234567)	→	(1534726)
g)	Superimposition	(1234567)	→	(12345) 3 6 4 7

---

[23] F. W. Michels, "Drie Nieuwe Kaartvormen," Tijdschrift van het Koninklijk Nederlandsch Aarorijkskundig Genootschap, LXXVI, 2 (1959), 203-209; D. M. Desoutter, "Projection by Introspection," Aeronautics, XL, 2 (April, 1959), 42-44.

[24] Alternate interpretation of negative slope graphs is possible; see Chapter IV. Inversion of a plane $(1/\rho)$ is somewhat different.

h) Inversion (1234567) → (1276543)

It is clear from the order relations, and from the graphs, that the eight cases are not all mutually exclusive and may occur in various combinations. A collapsing, for instance, implies a superimposition, but the converse is not necessary. The figures already presented also show some of the cases combined. The occurrence of many of the situations on conventional projections is at the margins of the maps.[25]

Multivalued Maps

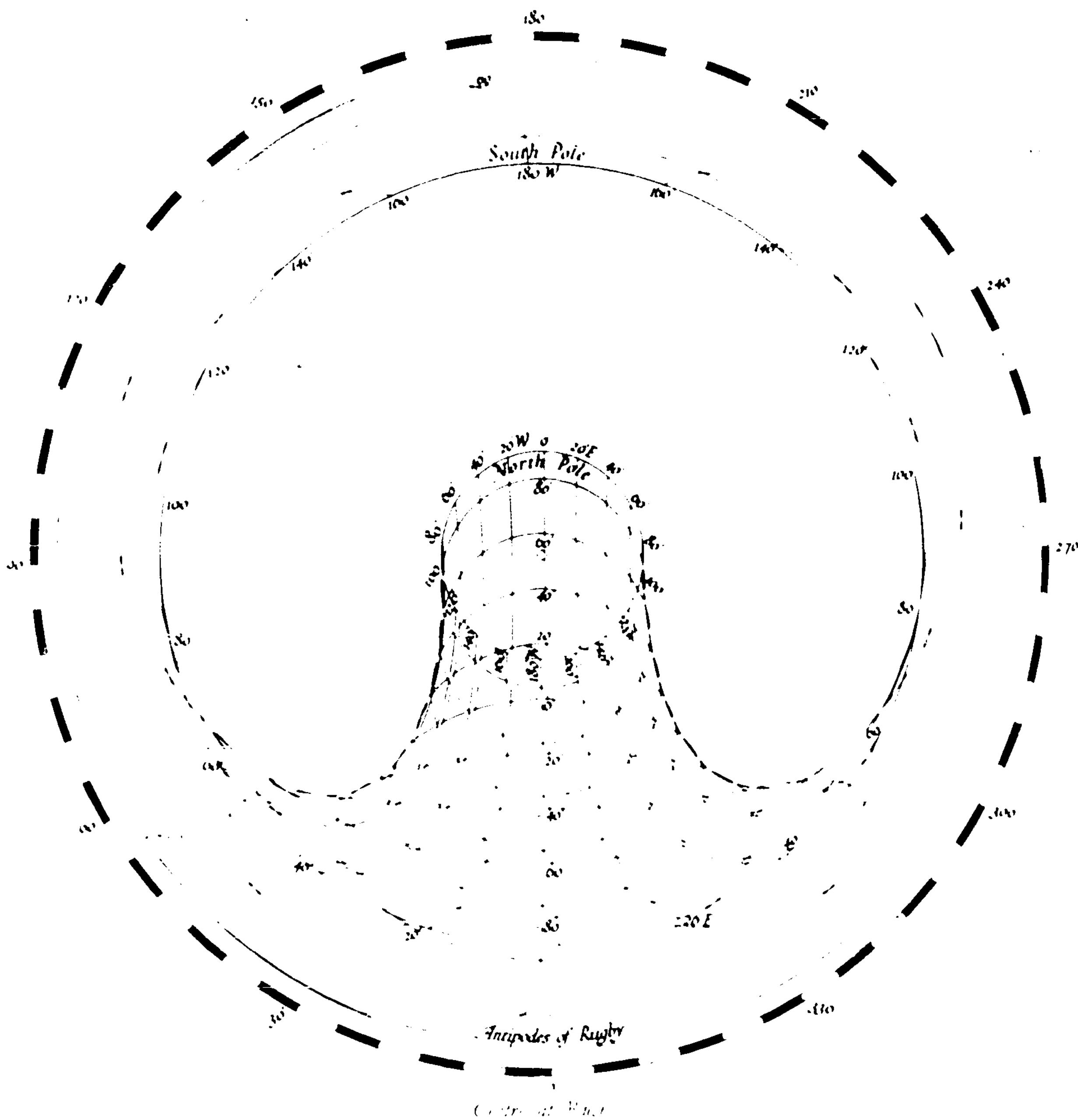
Functions (relations) which are not strictly monotone but have multiple values of r for some ρ, or the reverse, are difficult to conceive as maps. They may result in superimpositions (of points, lines, or areas) on the map. The most striking example in the cartographic literature is A. R. Hinks' retro-azimuthal equidistant projection, which combines interruption and superimposition (Fig. 3.22). The projection was received with some frivolity when first introduced, and one critic wrote: "Mr. Hinks has succeeded in producing a projection on which it is impossible to draw a map."[26]

To clarify the situation, refer to variant c of Figure 3.21. All the difficulty can be eliminated by an appropriate truncation of the domain. This is what in practice is done for the orthographic and other perspective

---

[25]Order relations on a sphere are cyclic; polar coordinates would show this better. Also note that the orders 123, 231, 213, are all correct for a sphere.

[26]"Anon." "Recent Developments in Map Projections," The Geographical Review, XXIX (1920), 687. The uses of this class of projections is given by A. R. Hinks, "A Retro-azimuthal Equidistant Projection of the Whole Sphere," The Geographical Review, LXXIII, 3 (March, 1929), 245-247.



HINKS' RETRO-AZIMUTHAL EQUIDISTANT PROJECTION OF THE ENTIRE SPHERE

From: A.R. Hinks, "A Retro-Azimuthal Equidistant Projection of the Whole Sphere", The Geographical Journal, LXXIII, 3, 1929. Courtesy of the Royal Geographical Society.

Figure 3.22

projections. If, however, it is desired to use the multiplicity of values, they can be shown on maps by the use of halftones or a system of transparent overlays. Occasionally the pasting or folding of separate map pieces into sheets is possible. Strictly speaking, a collapsing would require an infinite number of sheets (or overlays), but the situations are generally not this involved.

Asymmetric Case with Relaxed Assumptions

The azimuthal maps thus far discussed have all assumed central symmetry; that is, they belong to category D and are of the form $\theta = \lambda$, $r = f(\rho)$. By considering individual profiles along particular azimuths, the preceding examination can be extended to the asymmetric azimuthal projections $\theta = \lambda$, $r = f(\rho, \lambda)$. This extension is too simple to require detailed repetition. The eight cases can again be recognized, the possibility of occurrence in this instance being in either of the directions $\lambda = C$ or $\rho = C$. As an example a step function is used. This again corresponds to a collapsing.

It has already been noted that the radius equation of azimuthal projections of category B can be illustrated by a block diagram type drawing. For step functions this appears as a histogram. Figure 3.23 shows the level curves, $r = C$, a plan view of such a drawing with the height of each step indicated, and with a selected profile shown as a graph. The map generated by this equation is shown in the left portion of the figure.

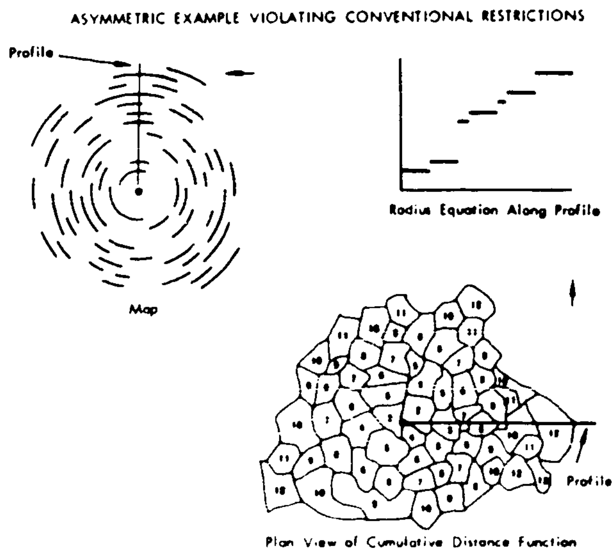


Figure 3.23

Spherical Distances and Map Distances

If equal intervals are marked off on the ρ-axis, these correspond to equal spherical distances from the origin on the globe. If these are mapped onto a plane using a centrally symmetric azimuthal projection, they appear as concentric circles (Fig. 3.24). In the polar case map, these lines can be considered lines of equal latitude and result in the familiar circular diagram. Such lines are conveniently called spherical distance contours.

If equal intervals are taken on the r-axis and are mapped back to corresponding positions on the original surface, we obtain map distance contours, which can be shown as circles on an azimuthal equidistant projection (Fig. 3.24).

SPHERICAL AND MAP DISTANCE CONTOURS

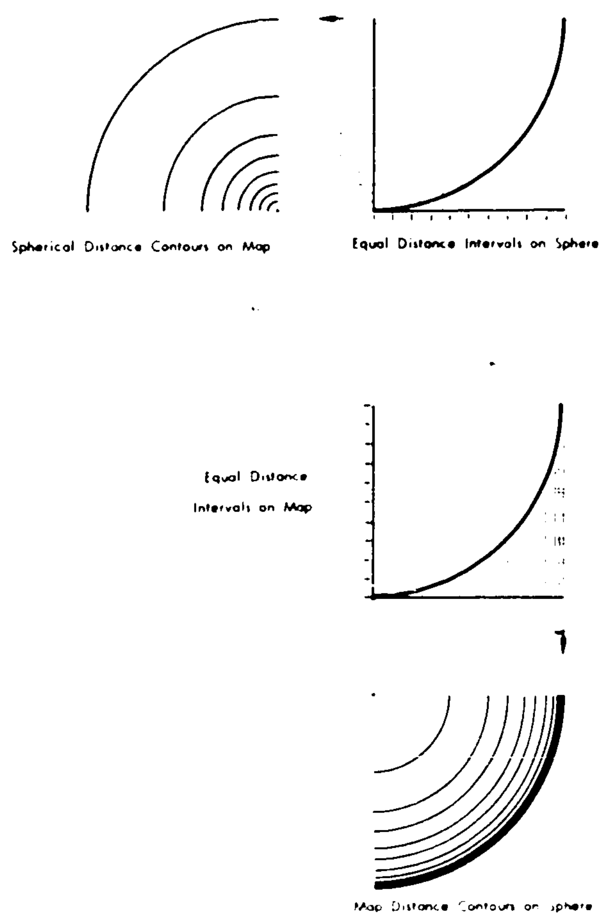


Figure 3.24

Only in the case of this projection do spherical and map distance contours coincide. For the asymmetrical azimuthal projections the spherical distance contours are no longer circles on the map. Similarly, the map distance contours are not circles when examined on the original surface. The concept of spherical and map distance contours is useful in examining relations on the transformations which do not meet the requirements of the conventional theory of projections. For example, on a map which collapses areas (Fig. 3.17) the spherical distance contours coincide. The map distances, on the other hand, may correspond to no positions on the original surface (they do not intersect the function), or they may correspond to an area. However, if the step function is drawn as a connected line (Fig. 3.17a), the map distance contours are superimposed, corresponding to a "cliff." In the case of an inversion the spherical distance contours overlap, which one can think of as an overhang. The drawing of spherical distance contours is generally simpler than the drawing of a complete map with a large amount of geographic data, particularly in the cases of inversion and superimposition.

CHAPTER IV

COST AND TIME DISTANCES

In this chapter the relevance of the transformations obtained in Chapter III by relaxing the assumptions of the conventional theory of map projections is to be demonstrated. Having previously indicated how a wide variety of functions can be translated into azimuthal maps, it is now only necessary to point out situations for which the use of these maps is appropriate. Since the possible examples which might be given are far too numerous, they tend to obscure the presentation and are kept to a minimum. Simple models of transportation systems are used in the discussion, but it is not intended that application be restricted to these situations. The graphic methods and other concepts developed in the previous chapters have already provided the groundwork.

Transport Surfaces

Consider first a uniform surface and a person walking on this surface. It is to be assumed that this person begins walking at a given point and proceeds to each position by the most direct possible route, always walking at a constant rate of speed. This situation can be represented as a line on a graph (Fig. 4.1) where the abscissa is taken to be the distance (ρ) along the ground, the ordinate (r) is the elapsed time, and the slope of the line is the amount of time consumed per unit of distance. The speed is initially taken so that one unit of distance can be

Elapsed Time as a Function of Distance

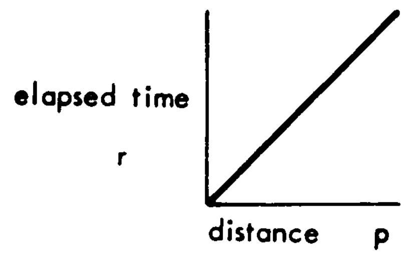


Figure 4.1

covered in one unit of time. It should be quite apparent that this graph can be considered an azimuthal map projection, in particular, the azimuthal equidistant projection. The assumptions regarding a starting point and a uniform surface imply central symmetry and the figure can be regarded as a profile relating points at distance ρ and the time it takes to reach these positions from the center (the origin on the graph). The azimuthal equidistant map hence shows correct time distances under the given conditions.

By changing the assumptions slightly it is possible to investigate the results of innovations in transport technology. In the first instance the uniform surface and the starting point are retained as assumptions (implying central symmetry), but the constant rate of travel is increased. This might correspond to a man running or riding a bicycle. The units of time consumed in covering a unit of distance are hence reduced, which corresponds to a lessening of the slope of the line on the graph. Translated, as before, into an azimuthal map this is simply a change of scale. The effect of an increase in the speed of transport media hence implies a change of scale or a reduction in the total size of the map. The reader may recall representations intended to demonstrate concepts of a "shrinking world" which use this device.[1] Lotka uses the same concept to indicate how the ecological range of a species may be increased by such changes in technology.[2] As is well known, this is too simple a conception for the effects are quite unevenly distributed.

The impact of a non-uniform surface can also be considered. The condition of central symmetry is retained for the moment in the interest of simplicity. A non-uniform transport surface is most easily understood by examining the uniform surface condition in somewhat greater detail. In the previous examples the units of time consumed per unit of distance from the center were constant. A uniform surface hence implies a line of zero slope on a graph as shown in the next figure (Fig. 4.2), where the abscissa

Constant Consumption of Time per Unit of Distance

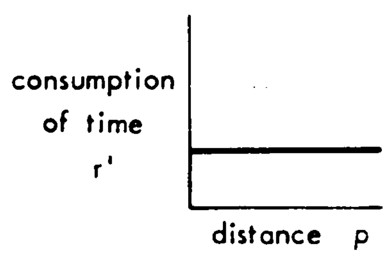


Figure 4.2

again represents distance (ρ) from the origin, and the ordinate (r') repre-

---

[1]For example, Raisz, op. cit., pp. 262-263.

[2]Briefly, one can define the ecological range of a species as the area which is accessible within a fixed time limit. On a uniform plane this increases in proportion to the square of the radius. A. Lotka, Elements of Mathematical Biology (New York: Dover, 1956).

sents units of time. An increase in speed results in a lowering of this line without a change in slope. Elapsed time is the cumulative time from the origin to the point ρ; that is, the area under the curve $r' = C$ (C a constant) from the origin to ρ. This is clearly the integral $\int_0^{\rho} r' \, d\rho$. Using the integral to obtain the graph of elapsed time as a function of distance yields $r = C\rho$, as desired (the constant of integration being taken as zero). A non-uniform surface is thus one which can be represented as having a slope different from zero in at least one instance.[3] Elapsed time as a function of distance is the integral of a profile of this surface.

To give the reader some indication of cases in which a non-uniform surface might occur, the walking man example is again employed. On a surface such as that of our planet earth there are many irregularities. One would encounter hills, swamps and forests, to mention only a few of the obstacles, which would retard progress in walking. Central symmetry has been assumed so that there is no going around these obstacles. A man walking at a rate of five miles per hour on firm ground might be able to progress at only two miles per hour over sandy terrain. The accompanying figure (Fig. 4.3) shows a strip map of an imaginary piece of ground containing typical situations which might be encountered, with estimates of the rate of time consumption over the different portions of the terrain, and with a cumulative time profile along the path. Comparable trafficability maps of different portions of the world have actually been pub-

[3]Perhaps better: a surface of variable curvature.

STRIP MAP WITH ESTIMATED CONSUMPTION OF TIME AND ELAPSED TIME AS FUNCTIONS OF DISTANCE

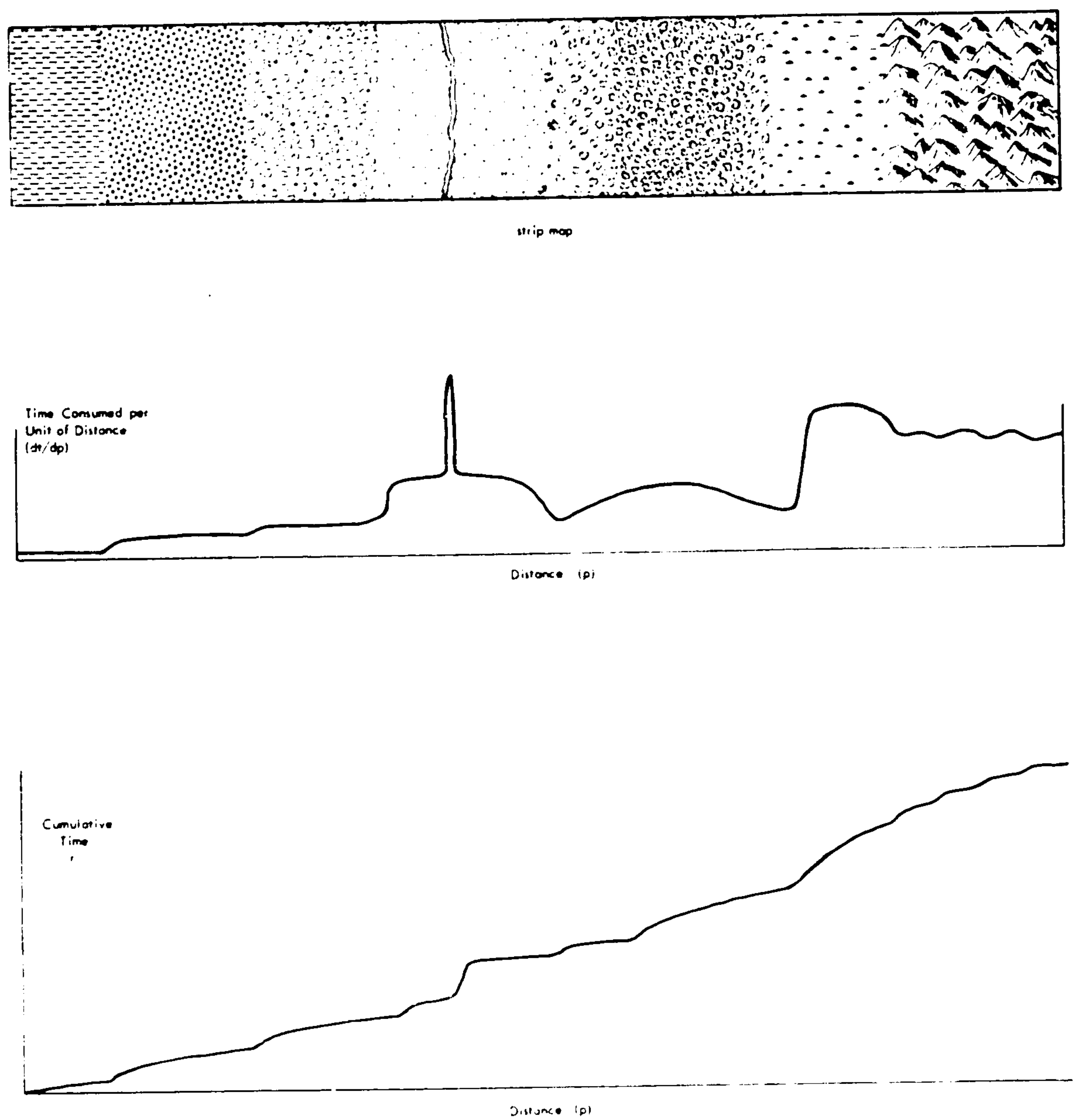


Figure 4.3

lished.[4] The elapsed time profile is a graph similar to that earlier used for azimuthal projections and by using the same procedure as used in the previous chapter--working backwards from the graph to a map--it is possible to obtain a map scaled in units of time so that equal intervals of time are represented by equal units of map distance.

The cost of traveling a particular distance and the time it takes to travel this distance are not unrelated. Rather than attempting to investigate this relation, it will be assumed that they are independent.[5] Nevertheless, costs as a function of distance display characteristics which are quite similar to those indicated for the time and distance relations. The preceding examples of a person walking on a uniform or non-uniform surface can all be rephrased in terms of monetary rather than temporal units simply by substituting the word cost (or some monetary unit) at each occurrence of the word time. Elapsed time is to be translated as cumulative cost, speed as distance rate of consumption, constant rate of speed as equal costs per unit of distance. The illustrations accompanying the strip map (Fig. 4.3) are then to be interpreted as showing the cost of traversing a particular unit of ground and a cumulative cost profile. Using the cumulative cost profile as the radius equation of an azimuthal map, the graphic procedures immediately permit a transformation from ground distance to cost distance so that equal intervals from the center of the map represent equal monetary units. As an indication of reasons for fluc-

---

[4]The classic examples are the German Hindernis-, Befahrbarkeits-, or Gelaendebeurteilungskarten. See W. E. Davies, "Axis War Maps," Surveying and Mapping, VIII, 3 (July-September, 1949), 126-134.

[5]One of the possible relations is postulated in Wingo, op. cit., pp. 50-57.

tuating costs per unit distance, it is clear that the differential natural endowment and characteristics of the environment give rise to differential costs of transportation. Maps attempting to illustrate this notion have been prepared by Boggs, Figure 4.4 being an example.[6]

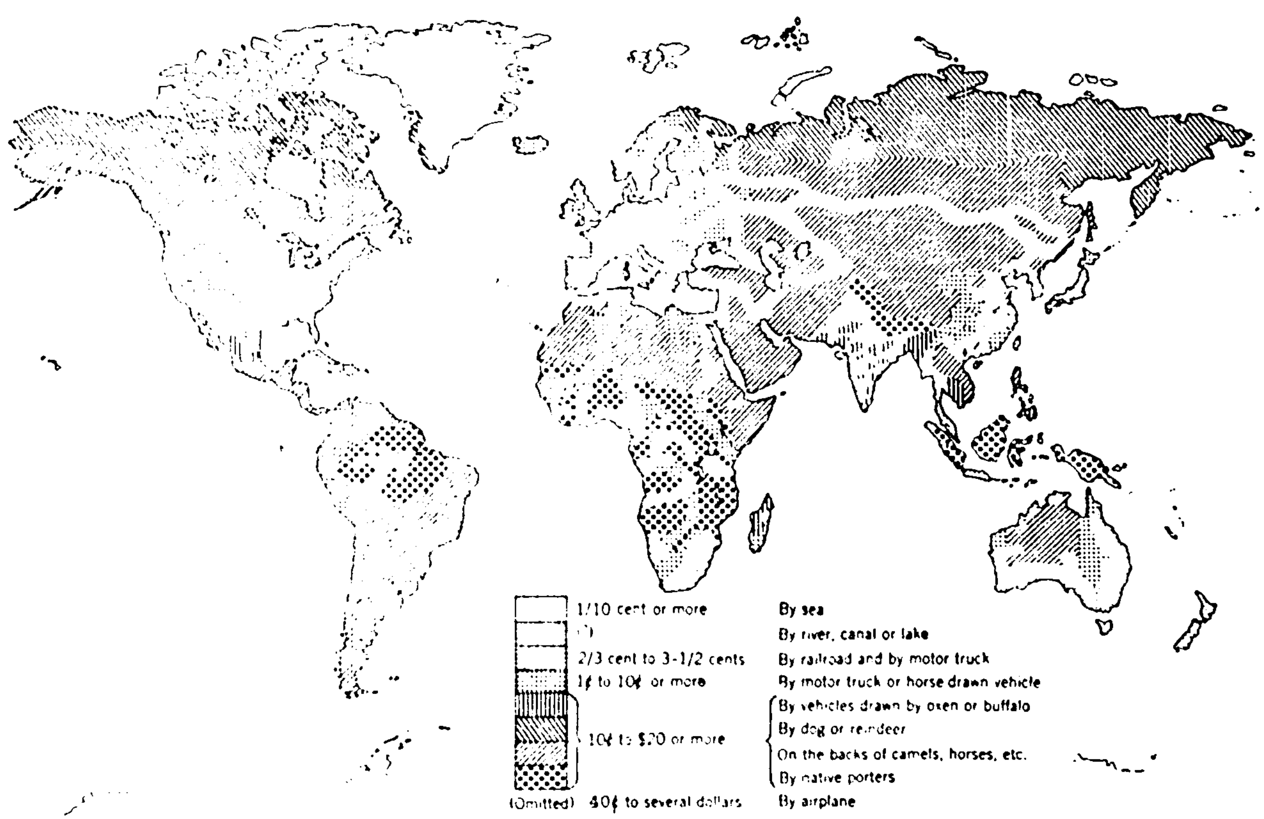
Costs are normally not associated with a person walking but rather with more advanced methods of transportation. In addition to differential over-the-road, construction, and maintenance costs, transport systems are characterized by costs which depend on such factors as frequency of service, amount of use of the route, and on type, size, value, weight, quantity and other characteristics of the commodity being shipped. Competitive, legislative, institutional, and historical factors are also frequently involved. The literature on this difficult topic is extensive and it is not the present intent to delve deeply into the subject.[7] The cumulative time or cost functions will be taken as given without detailed attempt at explanation except as required for reader orientation. These functions can be translated directly into maps by the procedures already given and without integration. In order to maintain a sufficient degree of generality, cumulative time or cost distances will hereafter be used almost interchangeably. The reader is at liberty to substitute other generalized measures of dis-

---

[6] S. W. Boggs, "Mapping the Changing World: Suggested Developments in Maps," Annals, Association of American Geographers, XXXI, 2 (June, 1941), 119-128.

[7] See, inter alia, E. Troxel, Economics of Transport (New York: Rinehart, 1955); M. L. Fair and E. W. Williams, Economics of Transportation (New York: Harper, 1950); S. Daggett, Principles of Inland Transportation (New York: Harper, 1955); P. D. Locklin, Economics of Transportation (Homewood: Irwin, 1954); K. T. Healy, Economics of Transportation in America (New York: Ronald, 1940); National Academy of Sciences, Conference on Transportation Research, National Research Council, publication 840, Washington, 1960.

TRAVEL COSTS THROUGHOUT THE WORLD



From S. W. Boggs, "Mapping the Changing World: Suggested Developments in Maps", Annals, Association of American Geographers, XXXI, 2, 1941. Courtesy of the Association of American Geographers.

Figure 4.4

tance at his pleasure as long as the functions have the same characteristics as those being discussed. Costs may be considered to be operating costs, or charges assessed the user (fares, tariffs), or unspecified utiles. Consideration of utiles has the advantage of including general factors such as convenience or personal preference but is less amenable to empirical measurement.

The simple cases thus far given apply only to distances from one point; they meet several conditions, and, because of central symmetry, correspond to azimuthal maps of category D. Before extending the applications, it is necessary to examine more closely the nature of transportation systems. The terms transport structure and transport characteristic are to be defined, but in a technical sense which only roughly corresponds to the common usage of the words. Transport structure relates to the general arrangement of the transportation system. Several structures are to be identified; the frequently occurring types of transportation in present use can then be classified as belonging to a specific structure. By this device several methods of transportation can be discussed simultaneously. Transport characteristics relate to the measure function associated with distance and are permitted to vary within a structure; the same characteristics may also appear in different structures.

Transport Structures

Transportation systems are structured in such a way that their operation is restricted to particular domains. Thus, ocean vessels are restricted to waters of certain depths; walking is restricted to land areas; airplanes are restricted to the troposphere. A transport medium is free to travel anywhere in its domain. Roadways, railroad tracks, footpaths, rivers, etc., can be considered as being particularly narrow domains but will be called routes, other domains simply areal domains.[8] A route may be directed, as a one-way street, but these will not be considered. A domain may also be a sub-domain of a larger domain. The domain of railroads, for example, consists of railroad tracks; the domain of railroad tracks, however, is generally confined to land areas. For purposes of extension of a railroad system, it is important to consider the domain of land areas. For discussions of travel by rail, however, it is necessary to recognize that the valid domain consists only of the existing tracks. The extent of a domain may fluctuate frequently, of course.

The edge of a domain is referred to as the boundary. An unbounded domain is taken to mean the entire terrestrial surface, a two-dimensional domain. A simply-connected domain is one which is all in one piece and encloses no areas inadmissible to the transport medium. Typical multiply-connected domains might correspond to a body of water containing islands or shallow areas, or land areas containing lakes, swamps, and mountains.

Although it is today true (technologically, at least) that any position on the surface of the earth can be reached, given sufficient expenditure, it is necessary to consider combinations of the domains of several media to achieve this objective. Entry to and exit from domains, therefore, is of particular importance. With each domain is associated a disutility upon crossing the boundary. This disutility can be considered

---

[8]What are here called routed domains are often also called networks; in general, these are multiply-connected. See W. L. Garrison, "Connectivity of the Interstate Highway System," Proceedings, Regional Science Association, VI (1960), 121-137.

a continuous function of arc length ranging from zero to infinity, which also need not be symmetric in direction and fluctuates with time and many other factors. It is simpler, however, to be somewhat less general and to consider only two cases, the first case admitting of entry and/or exit anywhere along the boundary at a finite constant disutility (which may be zero). Such boundaries, for lack of a better term, can be said to be permeable. The second case permits entry and/or exit only at fixed points, the transition points, at some finite disutility. The distinction is easily brought out by an example. A person driving a private automobile is usually permitted to stop and walk away from the road at any point. The boundary of the domain (the edge of the road) is in this sense permeable. A passenger on a bus, subway, ship, airplane, or train, is permitted (or prefers) to leave the vehicle only at select positions, the transition points. Similarly, the shipment of freight by one medium usually terminates only at specific locations and not anywhere in the domain. The disutility incurred upon crossing the boundary in both cases will be taken as a part of the transport characteristic (<u>infra</u>).

The use of airplanes would appear to complicate the classification, for their domains are properly three-dimensional, though the altitudinal limits are of no concern here. Nevertheless, the domain of airplanes can be considered as being a region which touches the earth only at airports--the concern, after all, is only with positions which have terrestrial meaning. In this sense, the case of airplane travel corresponds to an areal domain with transition points. The only portions of the domain which correspond to a ground location are the transition points (airports). To consider any other ground positions as existing on an airplane distance map is usually not meaningful. Also note that when discussing cost-distances, it is occasionally preferable to consider the domain of defined costs rather than the actual domain of the transport medium. To continue the previous example, airplane fares are usually quoted only for airports; the cost domain consists only of this finite collection of locations. When transforming spherical distance into time- or cost-distance, it is important to recall that the domain must be specified.

The foregoing discussion of transport structures has attempted to isolate the topological factors which all transport systems have in common and which appear important for the present purpose. Four types of transport structure have been recognized: areal domains and routed domains, with permeable boundaries or with transition points. The majority of transport media can be assigned to one of these four structures although the actual situation is obviously complicated (Fig. 4.5).

FIGURE 4.5

CLASSIFICATION OF COMMON TRANSPORT SYSTEMS ACCORDING TO STRUCTURE

Vehicle or Mode of Travel	Domain	Boundary or Transition Points
I. Areal Domains		
a. With permeable boundaries		
Small boats	Water areas	Shoreline
Walking	Land areas	Shoreline, other obstructions
b. With transition points		
Commercial vessels	Deep waters	Docks, ports

FIGURE 4.5--Continued

Vehicle or Mode of Travel	Domain	Boundary or Transition Points
Airplanes[a]	Troposphere	Airports, fields
Sound, light, radio television	"Ether" (range often limited)	Receptors
II. Routed Domains		
a. With permeable boundaries		
Automobiles; including trucks, taxicabs, private buses	Roads	Edge of road[b]
b. With transition points		
Railroads; including subways, streetcars	Tracks	Stations
Buses (commercial, public), trolleys	Scheduled route, road, wires	Scheduled stops
Telephones, telegraphs, power lines	Wires	Receptors
Automobiles	Expressway, limited access highway	Expressway exits[c]
Pipe line	Pipe	Terminals

[a]Excludes dropping of materials, and crash landings.

[b]Except where prohibited by congestion or legislative restraints.

[c]The domain then changes from IIb to IIa.

Transport Characteristics

As the reader may have anticipated, transport characteristics are to be identified with cumulative disutilities as a function of distance. Many of the graphs examined in the previous chapter can be shown to be very typical of time or cost distance relations for transport media. Substitution effects, for example, can readily be shown to yield inversions, and multiply-connected domains result in space eliminations. To discuss the relevance of all of the functions illustrated in Chapter III would be tedious and hardly fruitful. Rather than to attempt this in detail a few symmetric examples are given and then substitution effects and asymmetric situations are examined.

Cumulative transport charges are often characterized as increasing at a decreasing rate with distance. The cost curves in this case are such that the costs of transport are less per unit of distance at long distances than at lesser distances. This implies a cumulative cost curve which is concave downward. There are many such functions, one of which has already been demonstrated by the square-root projection (Fig. 3.10). The map, of course, assumes that the domain is unbounded and simply-connected. To this extent the cost-distance map is unrealistic, exactly as is the unqualified statement that cumulative costs as a function of distance are characterized by the form $d^2r/d\rho^2<0$, for transport media with an unbounded simply-connected domain (the entire terrestrial surface) are (at present) rare. Nevertheless, a two-point cost-distance map of the world (no longer azimuthal) using a similar relation might be centered on dual pricing points to illustrate a further concept of freight rates.

Transport systems are also often characterized by terminal charges. To travel any distance, however small, involves an expenditure, usually a fixed charge, which bears no relation to distance and is in addition to the distance rate. Arising from costs of construction and maintenance of terminals, loading costs, or overhead, these terminal charges result in an r-intercept which is different from zero (Fig. 3.14), corresponding to the ring-shaped azimuthal maps. This can be interpreted either as an interruption or expansion. When the transport medium operates on a schedule, the average waiting period may be taken to be comparable to terminal charges. Similar waiting periods (or costs) may occur at intermediate points along a route, or at a destination. At intermediate points one may have to transfer from one transportation system to another, or pass through toll gates, or customs, or wait for refueling, etc. In the movement of freight, break-in-bulk, transshipment, loading and unloading, and the like, have all along been recognized as interruptions incurring costs.

In many transport systems a strictly monotone distance function does not prevail. One often finds what are known as zonal rates. In these instances, contiguous places are lumped into groups and the same rate applies to all places within that rate zone. Fares of this nature are step functions and can be mapped as indicated in the previous chapter. Such rates occur very frequently, step functions being convenient and easily tabulated in the practical process of establishing rates. The parcel post charges in the United States are of this form. A map obtained from the Seattle post office shows the system (Fig. 4.6). The United States is divided into eight zones of varying size, as shown by the concentric arcs on the map. The cost of sending a package from Seattle to any place is the same for all places within each zone, but the cost varies from zone to zone, increasing as the zones increase in distance from Seattle. The actual cost function is tabulated below for a specific parcel weight

Parcel Post Zones

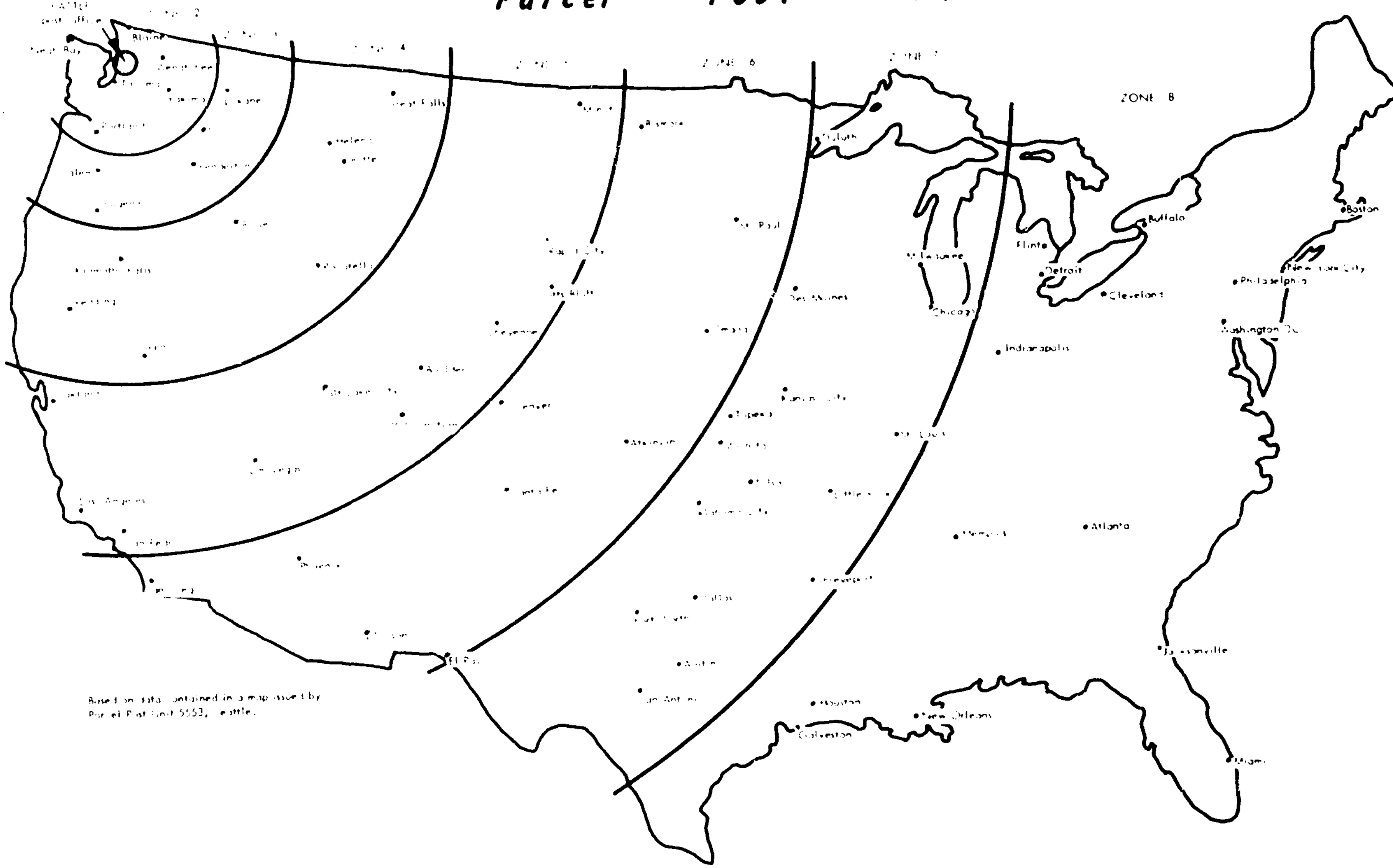


Figure 4.6

(Fig. 4.7).

FIGURE 4.7

PARCEL POST CHARGES IN THE UNITED STATES

For a parcel of more than one but less than two pounds, 1960

Charge in Cents	Distance in Miles	Zone
24	$0 < \rho \leq$ "local"	1
33	"local" $< \rho \leq 150$	2
35	$150 < \rho \leq 300$	3
39	$300 < \rho \leq 600$	4
45	$600 < \rho \leq 1000$	5
51	$1000 < \rho \leq 1400$	6
58	$1400 < \rho \leq 1800$	7
64	$1800 < \rho \leq$ Remainder	8

Source: Rates furnished by the Seattle Post Office. Distances obtained from: 1959 World Almanac, New York World-Telegram, H. Hansen (ed.), New York, 1959, p. 705.

Based on these rates a cost-distance map has been prepared of the United States as it appears to a person wishing to mail a package from Seattle (Fig. 4.8). The scale of the map is in cents. Measurement, of course, is meaningful only from the center of the map, as on an azimuthal equidistant projection; this is more obvious on the post office map. One notices immediately that areas are transformed into arcs of circles corresponding to a collapsing as discussed in the previous chapter. The areas between

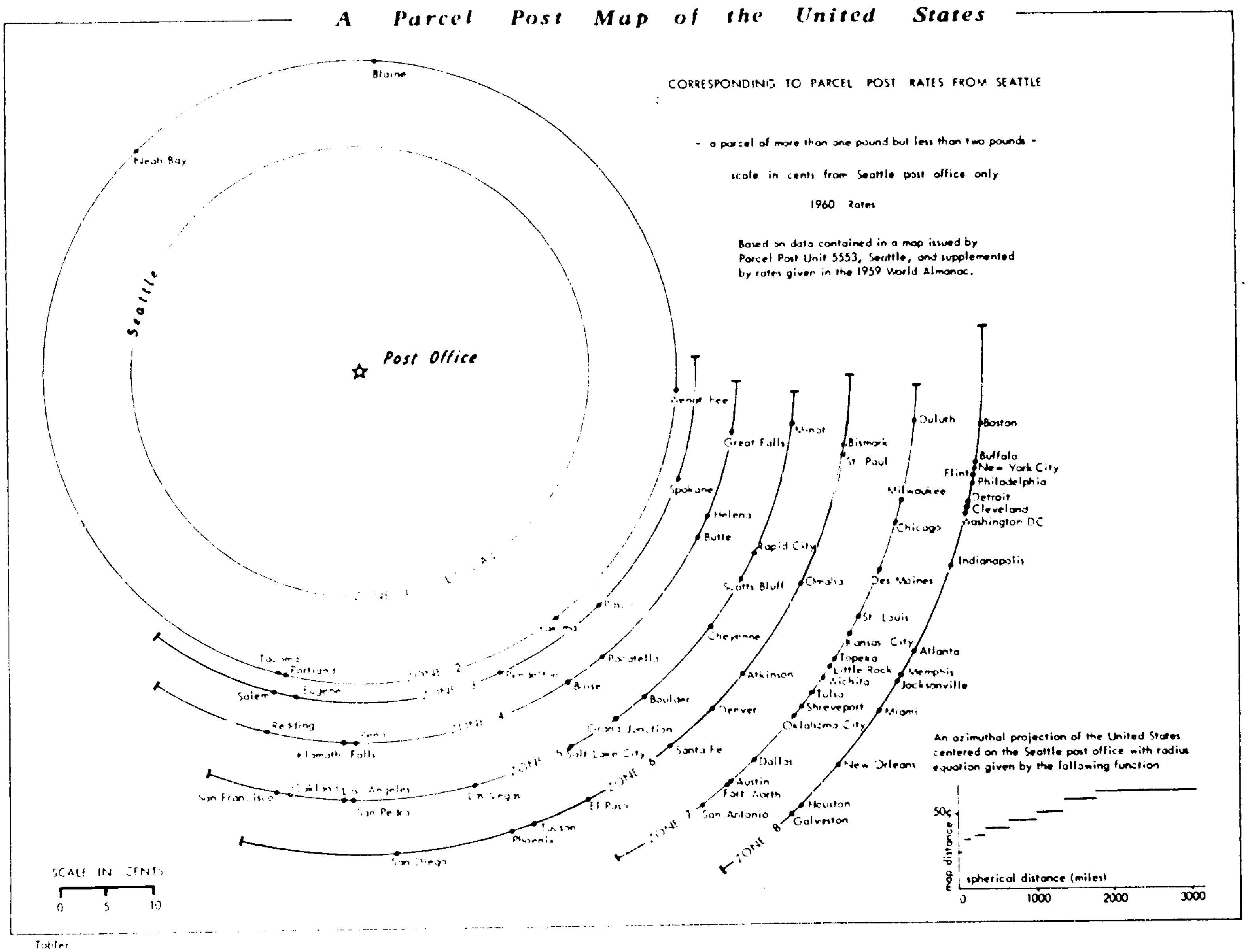

A Parcel Post Map of the United States
CORRESPONDING TO PARCEL POST RATES FROM SEATTLE
- a parcel of more than one pound but less than two pounds -
scale in cents from Seattle post office only
1960 Rates
Based on data contained in a map issued by Parcel Post Unit 5553, Seattle, and supplemented by rates given in the 1959 World Almanac.
Post Office
Seattle
Blaine
Neah Bay
Spokane
Great Falls
Minot
Bismark
St Paul
Duluth
Boston
Buffalo
New York City
Flint
Philadelphia
Detroit
Cleveland
Washington DC
Indianapolis
Milwaukee
Chicago
Helena
Butte
Rapid City
Scotts Bluff
Omaha
Des Moines
St Louis
Kansas City
Atlanta
Topeka
Little Rock
Memphis
Jacksonville
Tulsa
Shreveport
Miami
Oklahoma City
Cheyenne
Pocatello
Atkinson
Boise
Boulder
Denver
Grand Junction
Salt Lake City
Santa Fe
Dallas
New Orleans
Austin
Fort Worth
Houston
San Antonio
Galveston
El Paso
Tucson
Phoenix
San Diego
Portland
Salem
Eugene
Klamath Falls
San Francisco
Los Angeles
San Pedro
ZONE 2
ZONE 3
ZONE 4
ZONE 5
ZONE 6
ZONE 7
ZONE 8
An azimuthal projection of the United States centered on the Seattle post office with radius equation given by the following function
50¢
map distance
spherical distance (miles)
0
1000
2000
3000
SCALE IN CENTS
0
5
10
Tobler

Figure 4.8

arcs do not correspond to any places in the United States and must be left blank. Thus, if one asks: "Where is Lake Michigan?" the answer must be that it lies on a portion of the circular arc between the azimuths of Chicago and Duluth. Postal rates are identical in all directions and the map hence corresponds to an azimuthal projection of category D, with equations $\theta = \lambda^*$, $r = f(\rho^*)$; $f(\rho^*)$ being defined by the table (Fig. 4.7). The domain is the United States, more precisely, the set of all mailing addresses in the continental United States. The map, of course, is commodity specific. A map of the letter rate would be different, the entire United States being mapped onto a circle at a distance of four or seven cents.[9]

The reader will be able to exercise his imagination and, depending on his familiarity with transport characteristics, will note the application of additional functions. The isolated dots (Fig. 3.15d), for example, might correspond to railroad stations, airports, bus stops, expressway exits, and the like. If radio or television communication systems are considered as transporting information with virtually instantaneous velocity, the points can also be thought of as receivers (Fig. 3.15c).

Parcel post rates are unusual in that they are independent of location; i.e., the same characteristic applies no matter where one begins in the United States. By analogy to a sphere or plane the transport surface might be said to be of "constant curvature." Furthermore, the cost-distance for parcel post from Seattle to New York is identical to the cost-

---

[9]The current regular and airmail rates, respectively. Note that this is actually one of the cases excluded in Chapter II ($J = 0$). The map could be further collapsed, setting $\theta = C$.

distance from New York to Seattle. Directional symmetry also occurs only rarely in practice and it is more advantageous to examine the asymmetric situations in greater detail and to establish general procedures. Substitution effects, however, are advantageously inspected along a profile, adding directional asymmetry somewhat later.

Substitution Effects

While it is certainly true that any position on the surface of the earth can be attained with a finite disutility, this in general cannot be achieved without employing several transport media. Only in certain cases does a single transport domain include a large simply-connected portion of the terrestrial surface. The interest, however, usually is in attaining all positions within a given region. The differing media which must be employed will have varying characteristics and structures. The consequences of considering several media simultaneously are referred to as substitution effects.

If one considers all the possible ways in which one position on a sphere can be reached from another, it is clear that there must be infinitely many. In a more practical sense, there are usually several feasible ways to travel to an objective. Each such path will require some effort or time and, in general, the expenditures will vary along each path. This corresponds to a case of space repetition. It is only natural, however, to consider only that time- or cost-distance which is the minimum. In the present context the concern is only with the cumulative time- or cost-distance and there is no necessity to inquire whether the geodesic is unique. It is assumed that the minimum can be found.

Substitution effects are perhaps best explained by an example. The accompanying figure illustrates costs of transporting a commodity over various distances by three different systems (Fig. 4.9). If substitution

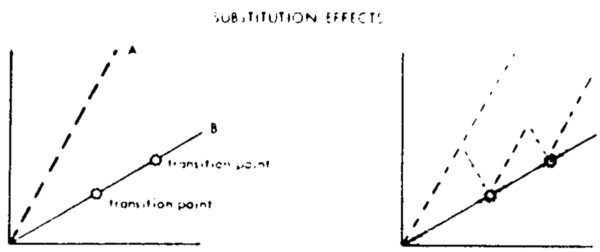
COMBINATION OF THREE TRANSPORT SYSTEMS IN ONE SIMPLY-CONNECTED DOMAIN

shading indicates minimum transport disutility as a function of distance

Figure 4.9

between systems is allowed everywhere and at no cost, the minimum transfer gradient is the piecewise continuous segment shaded in the illustration. The substitution results in a lessening of the total area under the curve, which can be interpreted as a lessening of the total scale on a map. Similar diagrams are occasionally used to illustrate competitive distance relations between truck, rail, and water transport. Extension to non-linear functions is simple. The illustration, however, tacitly and unrealistically assumes that all transport media operate in the same simply-connected domain. If this assumption is removed, the situation changes considerably. The existence of transition points results in particularly striking modifications.

Consider the case of transportation in an open areal domain, on a uniform surface and at a constant rate of speed (system A). Assume also that a further transport system (system B) is available, having a considerably higher rate of speed, but which is structured so that exit is restricted to transition points. Illustrating the situation on a graph, with system A as a dashed line and system B as a solid line, brings out the relations (Fig. 4.10). Up to a certain point it is more advantageous



Shading indicates minimum transport disutility as a function of distance

Figure 4.10

to travel directly by system A. Beyond that point, however, it is more advantageous to travel by system B and then travel in the reverse direction using system A. Emphasizing the minimum time distance to each position on the diagram, it is apparent that this function translates into multiole inversions and superimpositions of space, as well as a lessening of the total scale. If details of the domain are ignored, system A might be identified with automobile or truck travel and system B with travel by rail or air. Or system A might correspond to travel on minor city streets and system B to travel on urban expressways. Decisions whether or not to walk or take a bus are similar. Examples are easily found in both time- and cost-distances. Cognizance of the domain(s) is always appropriate, however. The consequences of substitution between media and domains are considerable. The situation becomes more complicated when there is a cost associated with the crossing of a boundary between domains. The details, however, can all be included in the more general procedure given in the next section.

Asymmetric Situations

Directional symmetry, as exhibited by post office rates, only rarely occurs in practice. More typical are situations in which the distance measure is a function of the direction in which travel is undertaken. In relating these cases to azimuthal projections of category B, advantage can be taken of a correspondence between isolines and map distance contours.

An isoline is a line connecting positions having some assigned number in common, as contours on a topographic map. The isolines of present concern are known as isochrones, connecting places of equal time-distance to or from some point, and isovectures, connecting places at equal cost-distance to or from some point. Similar are isodistantes, lines connecting places at equal ground-distance. To avoid overly complicated terminology only the generic term isoline is employed here.[10] Isolines are constructed by empirical measurement of time- or cost-distances to a finite number of points, with subsequent interpolation and drawing in a manner comparable to the construction of contours on a topographic map. The procedure is well known and the details need not be given here.[11] Maps using this technique are common; a few examples are given in the accom-

---

[10]T. Palander, *op. cit.*, Chapter XII, gives detailed definitions of several types of isolines. Also, W. Horn, "Die Geschichte der Isarithmenkarten," *Petermann's Geographische Mitteilungen*, CIII, 3 (1959) 225-232.

[11]Robinson, *op. cit.*, p. 201; G. Chabot, "La Determination des Courbes Isochrones en géographie Urbaine," *Comptes Rendus du Congress International de Géographie* (Tome 2, Sec. IIIa; Amsterdam, 1938), pp. 110-113; National Committee on Urban Transportation, *Determining Travel Time*, Procedure Manual 3B (Chicago: Public Administration Service, 1958).

panying illustration (Fig. 4.11).[12]

To establish the correspondence between isolines and map distance contours, the reader can refer to the previous chapter and substitute cumulative time or cost on the ordinate for the term map distance (Fig. 3.24). The map distance contours can be identified as isolines and the converse. This simple procedure allows any set of isolines to be translated into an azimuthal map projection, the only restriction obviously being that one consider only isolines which have meaning when so translated. It is also important to take careful note of the domain of the isolines. Many studies implicitly assume the domain to be simply-connected and to have at least the same areal extent as the study area. Substitution and boundary effects also are only rarely recognized, and the isolines are shown as monotonic functions. This is particularly true of the American literature. Certainly there are instances in which such details are unimportant, but it is appropriate to recognize that assumptions are being made.

Examination in greater detail of one map with isolines corresponding to asymmetric cumulative transport disutilities suffices. Map distances on a transformation of this map could be made to correspond to isolines by taking profiles along different azimuths as the radius equation for that azimuth. The sample map (Fig. 4.12) is a portion of a larger map published

---

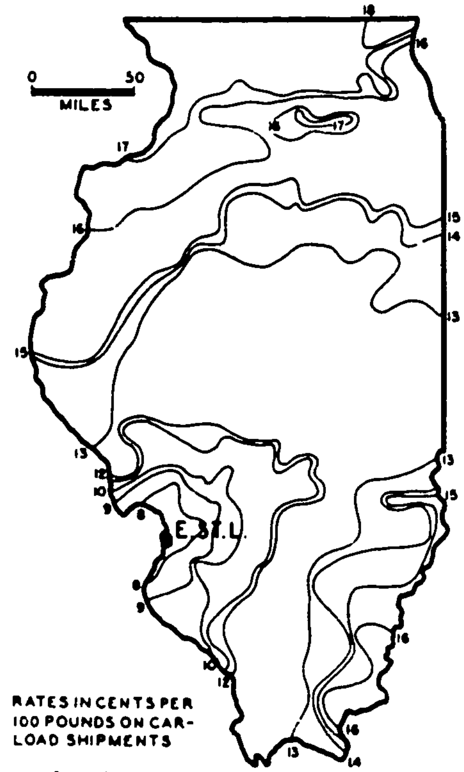
[12]The different assumptions employed in the preparation of each of the maps (Fig. 4.11) are particularly interesting. Additional isoline maps can be found in C. O. Paullin, Atlas of the Historical Geography of the United States, ed. J. K. Wright (Washington: Carnegie Institute and American Geographical Society, 1932), plate 38; State of Illinois, Atlas of Illinois Resources, Part IV, "Transportation" (Urbana: University of Illinois, 1960); E. Fels, Erde und Weltwirtschaft, Vol. V: Der Wirtschaftende Mensch als Gestalter der Erde (Stuttgart, 1956), pp. 223-227.

ISOLINE MAPS

Dallas : Five-Minute Isochrones

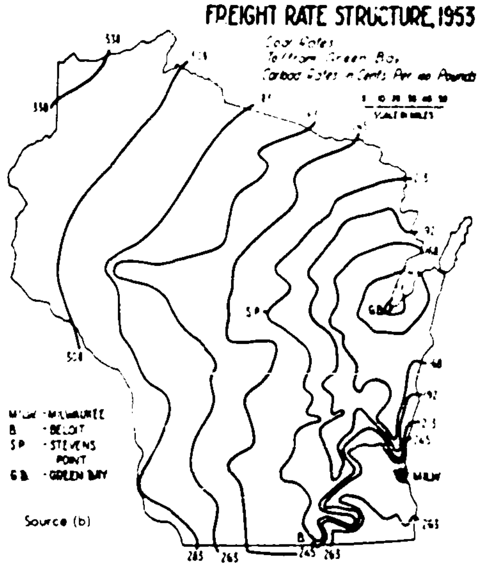
Source Department of Traffic Control, Dallas, Texas

INTERSTATE RATES ON GRAIN TO E. ST. LOUIS



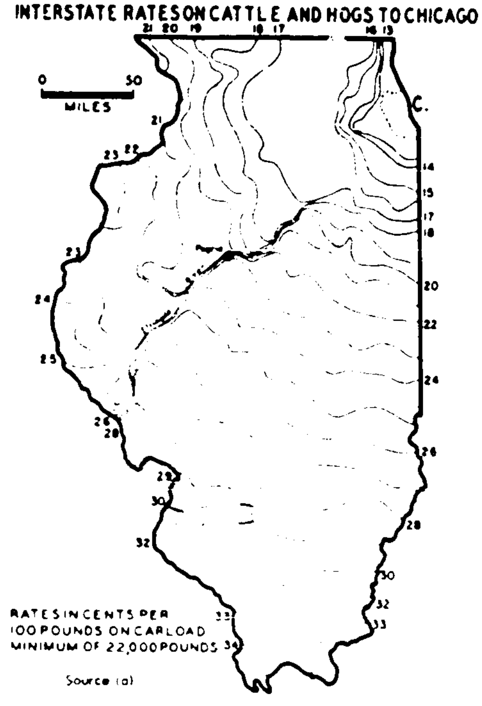
Source (a)

WISCONSIN
FREIGHT RATE STRUCTURE, 1953



Source (b)

INTERSTATE RATES ON CATTLE AND HOGS TO CHICAGO



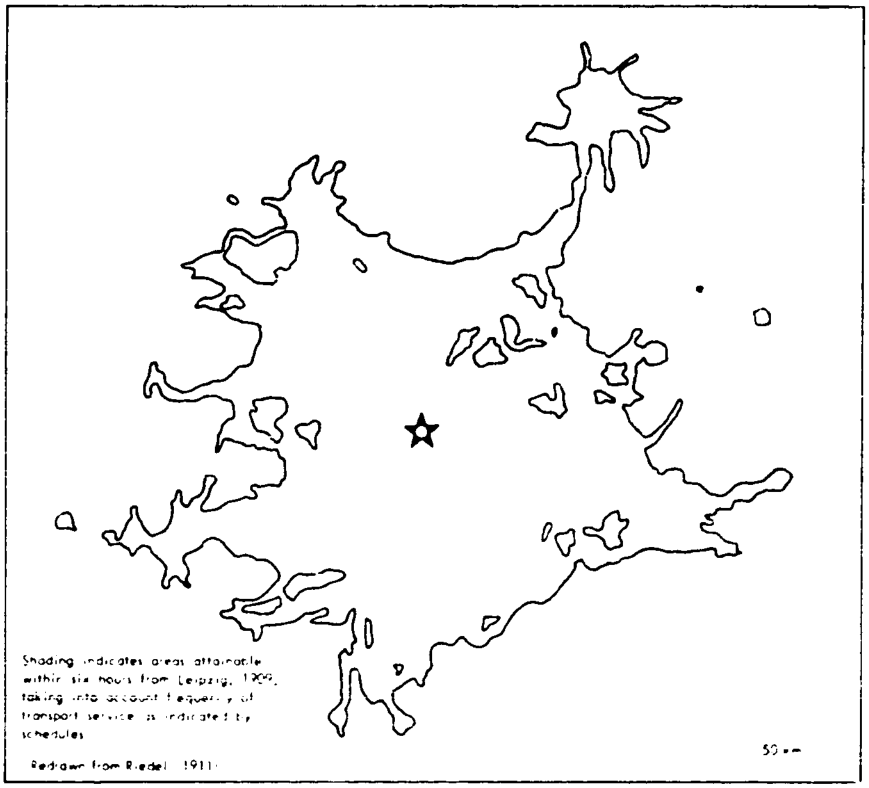
Source (a)

(a) J. W. Alexander, "Freight Rates as a Geographic Factor in Illinois", Economic Geography, XX, 1944. Courtesy of Economic Geography.

(b) J. W. Alexander, S. E. Brown, & R. E. Dahlberg, "Freight Rates Selected Aspects of Uniform and Nodal Regions", Economic Geography, XXXIV, 1, 1958. Courtesy of Economic Geography.

Figure 4.11

Leipzig : the Six - Hour Isochrone



Liège : the One - Hour Isochrone

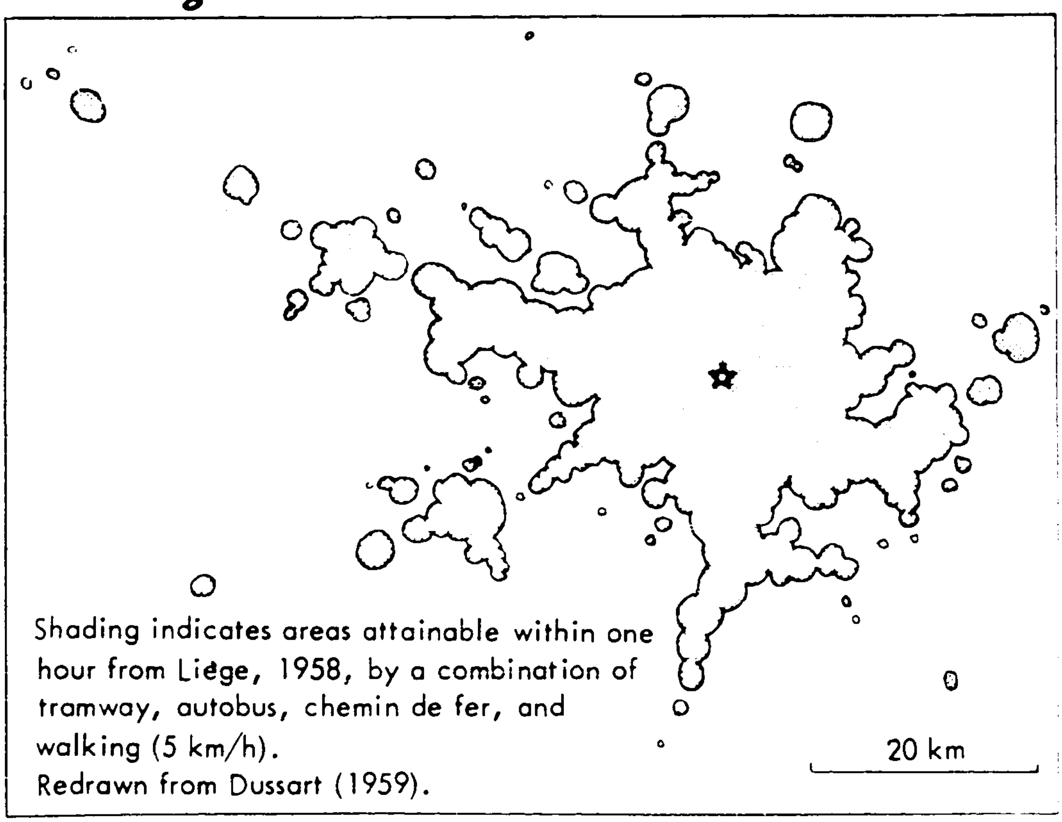


Figure 4.11 (continued)

TRAVEL TIME FROM BERLIN - 1909

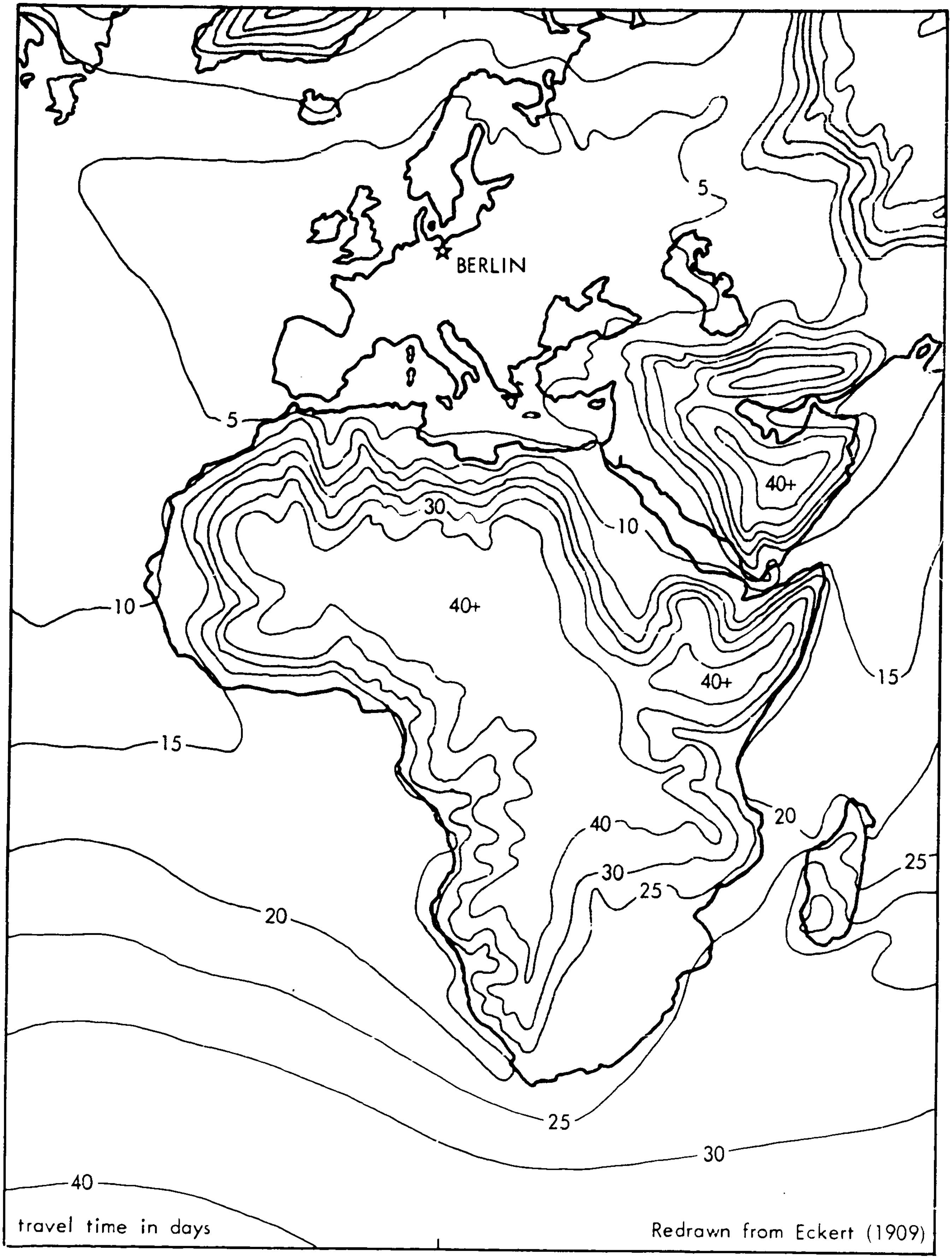


Figure 4.12

in 1909 by the German cartographer Max Eckert.[13] The average travel time from Berlin, using generally available transportation facilities, is indicated on the map by isolines. Eckert recognized that the projection to be used is the azimuthal equidistant and that the isolines may consist of disjoint segments, which, of course, indicate inversions. The accuracy of the data, including travel by camel, native safari, canoe, railroad and steamship, but not airplane or automobile, is not really of concern here. Inaccuracies in the location of the isolines would translate into inaccuracies of position on a transformed map. The isoline interval also may obscure minor inversions.

It is clear from Eckert's map that, from Berlin, the coastal areas of Africa could be achieved in less time than the interior portions of the continent. The interior hence is further away in time-distance. The inversions show clearly on the profile interpolated from the azimuth intersection of isolines (Fig. 4.13). A similar profile today would probably be on the average lower but would also show even greater oscillations.

An azimuthal map, centered on Berlin, with map scale in days from Berlin, is difficult to draw, there being multiple inversions and superimpositions, but could be attempted. Africa would literally be turned inside out with respect to Berlin. More easily drawn and yet indicative of the distortion would be the spherical distance contours.

---

[13]M. Eckert, "Eine Neue Isochronenkarte der Erde," Petermann's Geographische Mitteilungen, LV (1909), 209-216, 256-263.

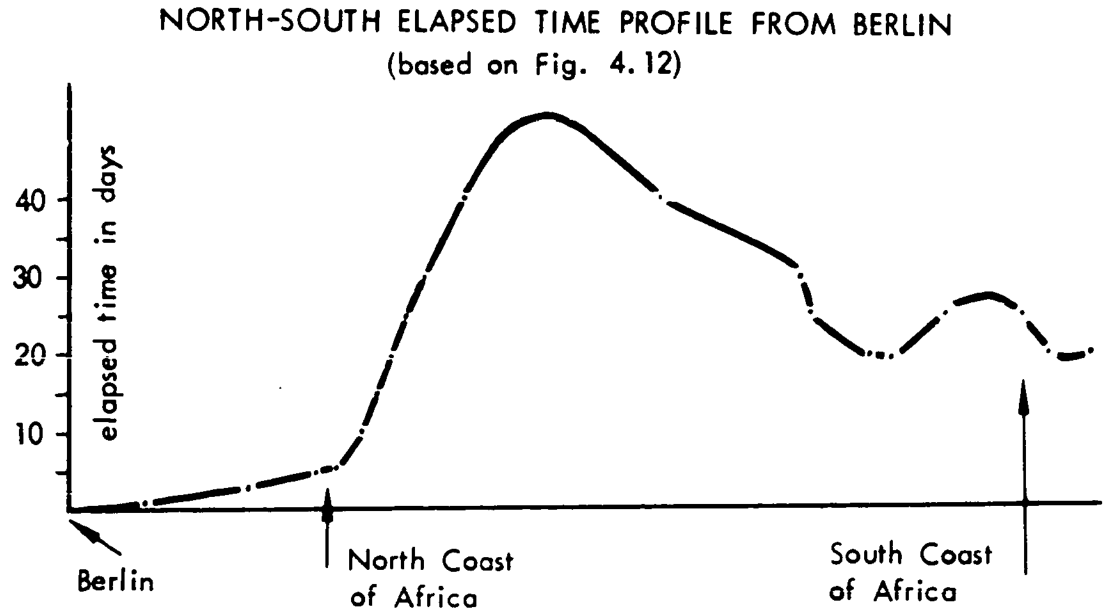


Figure 4.13

Inverses

The inverse of a function $u = f(\phi)$ is given by the function $\phi = F(u)$. In the present context it can be said that locations (u,v) on the map correspond to locations (ϕ,λ) on the surface of the earth, i.e.:

$$\phi = F_1(u,v)$$

$$\lambda = F_2(u,v).$$

Without belaboring the assumptions and restrictions applying to the domain and the functions, the traditional map can be said to be such that (almost) every point on the map corresponds to one, and only one, position on the earth. If the transformations which were examined in the previous chapter are allowed, this is no longer necessarily true. The difficulties to navigation or information storage created by maps which have positions which correspond to several terrestrial positions, or to no terrestrial position, are rather obvious. Hence it is to be expected that the lack of single-valued inverses will cause difficulties. This is not necessarily the case, as can easily be demonstrated.

The graphic method of examining inverses most appropriate to the present context is a movement from a picture of the map at the left of the graph of the radius equation, across to the function and downward to the picture of the ground. That is, the direction of movement is the reverse of that previously used. In the first illustration (Fig. 4.14) a

LAND USE EXAMPLE OF INVERSE TRANSFORMATION

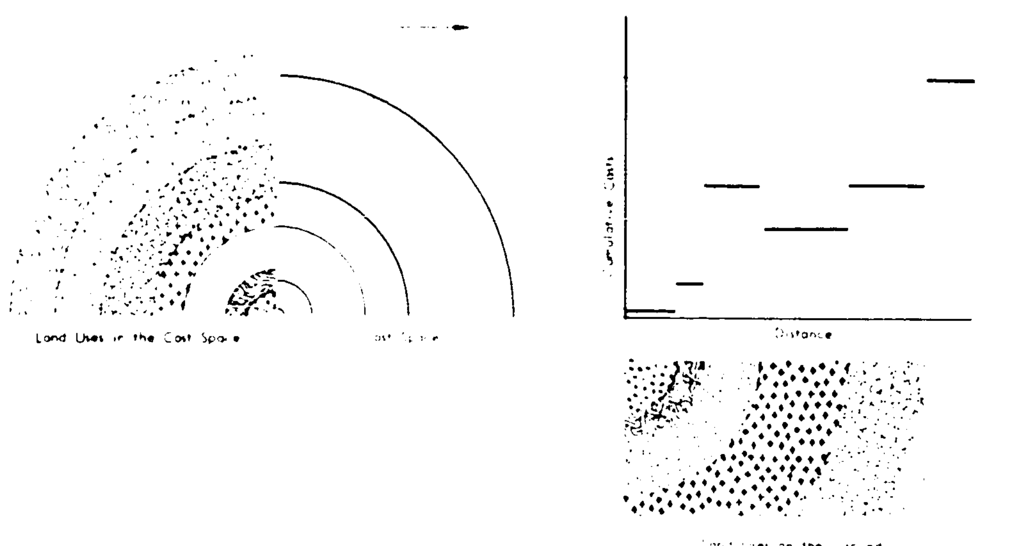


Figure 4.14

von Thünen system of land use is postulated and the cumulative transport cost function is taken to be a step function. When the land use pattern is mapped from the cost surface back to the original surface it no longer has the identical form. The actual form will depend on the specific cumulative transport cost function employed. Any function would serve for purposes of illustration; the centrally symmetric non-monotonic step function has been chosen quite arbitrarily and the reader may substitute other functions. When transport costs are proportional to distance, as in the original von Thünen model, the graph of the cumulative transport costs is a sloping straight line. We note that a highly complicated pattern of land uses may still fit the von Thünen model of circular rings. In the illustration (Fig. 4.14), one land use does not appear on what has been taken to be a picture of the ground. This is because the interruption on the cost distance map does not correspond to any position on the surface of the ground, i.e., a line drawn from the ordinate parallel to the ρ-axis does not intercept the function. This can be shown to be a land use excluded from production because of transport costs. Similarly, the boundary between two land uses may map back into an area of mixed land use. This can also be demonstrated by the von Thünen model as an area where the rate curves coincide. An inversion has also been shown, resulting in the multiple appearance of a land use. An example without central symmetry would yield an even more complicated pattern of land uses.

The limitations of the von Thünen scheme are made clearer by consideration of the inverse. In particular, if the transport cost is such that the cumulative cost curve has zero slope--as with the letter rate--the unrealistic result is that only one type of land use can occur. The assumption of uniform agricultural productivity has also not been relaxed (however, see Chapter VI). Burgess' concentric zone theory, and Hoyt's sector theory of urban land use, nevertheless, can be shown to be compatible by similar consideration of inverses.[14]

Taking another example, population density is postulated--for purposes of illustration only--to be a decreasing linear function of time-distance from the center of a city, with a piecewise continuous cumulative transport characteristic (Fig. 4.15). This transport characteristic may

POPULATION DENSITY EXAMPLE OF INVERSE TRANSFORMATION

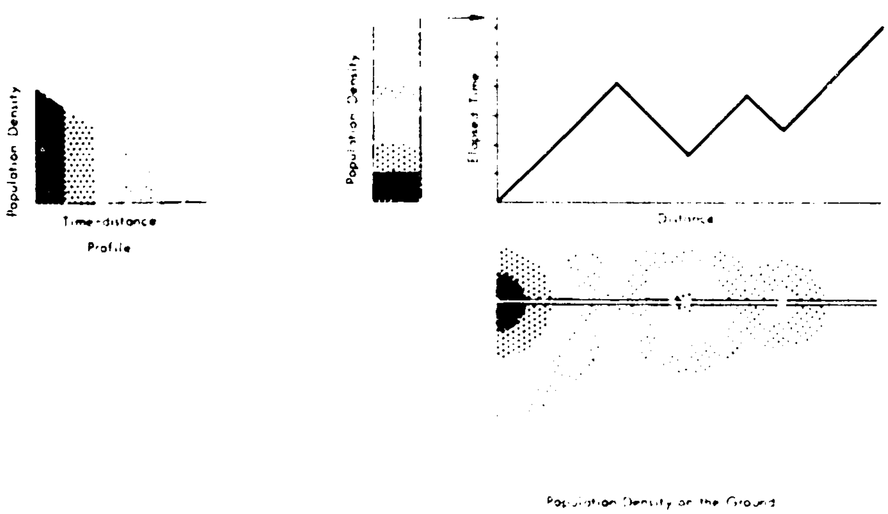


Figure 4.15

be taken to represent a combination of automobile travel on minor streets and on an urban expressway, as previously discussed under substitution.

---

[14]Many authors have recognized this compatibility. E. Kant, for example, remarks:

> The sector theory has been held up beside the circle theory. . . . We may question whether the sector scheme can be regarded as an independent theory. It is rather a case of deviations from the concentric structure.

(E. Kant, op. cit., pp. 7-8). Also see C. D. Harris and E. Ullman, "The Nature of Cities," Readings in Urban Geography, ed. H. Mayer and C. Kohn (Chicago: University of Chicago Press, 1959), pp. 277-286.

As is to be expected, identical results are obtained by considering the isolines only. Both cases are simple manifestations of the function of a function rule.[15] The validity of this method, of course, is identical to that of the postulates employed.

Further Geographic Applications

In the preceding pages the concern has been largely with transport situations and models similar to that of von Thünen. The importance of transport is widely recognized by geographers; transportation, for example, is central to Ullman's concept of spatial interaction.[16] With regard to the notion of intervening opportunity it becomes clear that opportunities may be located at greater spherical distances and still intervene because of inversions.[17] Centrally oriented projections can be expected to be useful in any study of a nodal region.[18] Reilly's law and similar potential or gravity models all incorporate concepts of distance.[19] Political

---

[15]Let L represent land use; c cost; ρ, λ as before. Postulate $L = f_1(c)$, $c = f_2(\rho, \lambda)$, then $L = f_1(f_2(\rho, \lambda)) = f_3(\rho, \lambda)$.

[16]E. L. Ullman, "The Role of Transportation and the Bases for Interaction," Man's Role in Changing the Face of the Earth, ed. W. Thomas (Chicago: University of Chicago Press, 1956), pp. 862-880.

[17]S. Stouffer, "Intervening Opportunities: A Theory Relating Mobility and Distance," Am. Soc. Rev., XV (December, 1940), 845-867.

[18]For definitions of the several types of regions recognized by geographers, see D. Whittlesey, "The Regional Concept and the Regional Method," American Geography: Inventory and Prospect, ed. P. James and C. Jones (Syracuse: Syracuse University, 1954), pp. 19-69.

[19]G. P. Carrothers, "An Historical Review of the Gravity and Potential Concepts of Human Interaction," Journal, American Institute of Planners, XXII, 2 (Spring, 1956); S. C. Dodd, "The Interactance Hypothesis: A Gravity Model Fitting Physical Masses and Human Groups," Am. Soc. Rev., XV (April, 1950), 245-256; J. Q. Stewart, "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Popu-

geography appears extremely rich in possible applications of transformations, with Gottman's concept of circulation, the notion that political control diminishes with distance from a center of power, Jones' circulation fields, and many other stimulating ideas.[20] Hägerstrand's studies of the migration of peoples, ideas, and innovation provide further areas of application.[21]

The following brief paragraphs attempt to suggest implementation of the transformations in a few of these areas. The discussion begins with psychological-distance, which, as a concept, is very similar to cost- or time-distance, and therefore quite simple.

Psychological Distance

If the thesis of cognitive behaviorism is accepted,[22] the notion of psychological distance also seems valid. If a person is asked the distance to some far-off place, his answer, if any, may be couched in terms of minutes, hours, or days, or in dollars, or in miles. The answer may

---

lation," The Geographical Review, XXXVII, 3 (July, 1947), 461-485; W. Isard, Methods of Regional Analysis (New York: Massachusetts Institute of Technology and J. Wiley, 1960), pp. 493-568.

20W.A.D. Jackson, "Whither Political Geography," Annals, Association of American Geographers, XLVIII, 2 (1958), 178-183; F. Ratzel, Politische Geographie (3rd ed.; Berlin: Oldenbourg, 1923), pp. 189-191; S. B. Jones, "A Unified Field Theory of Political Geography," Annals, Association of American Geographers, XLIV, 2 (1954), 111-123; J. R. Mackay, "The Interactance Hypothesis and Boundaries in Canada," The Canadian Geographer, XI (1958), 1-8; A. Lösch, op. cit., pp. 196-214.

21T. Hägerstrand, "Migration and Area," Migration in Sweden, (Lund Studies in Geography; University of Lund, Ser. B, 13, 1957), pp. 27-158.

22H. and M. Sprout, "Environmental Factors in the Study of International Politics," Journal of Conflict Resolution, I, 4 (1957), 309-328.

be based on assumed or known information; it may be an impression; or it may be a wild guess. It is possible to postulate that each person maintains a mental image of the environment and that this mental image can be brought to light by questioning. Interpretation of the results of such questioning is subject to serious debate, but such an experiment has been performed. Forty students were requested to indicate distances (in miles) from their location in Seattle to various cities throughout the world. Two a priori conjectures were entertained: (1) that the psychological distance function would be concave downward, and (2) that it would be asymmetric. The results are shown in Figure 4.13. The sample design is inadequate to comment on the conjectures, but similar experiments could be devised for azimuthality or area. An alternate sample design might make use of a number of maps of varying distortion, perhaps chosen from the series $r = \rho^q$, centered on the location of the experiment. By presenting these maps to children and asking them to select the map which appears most correct, it might be possible to obtain an estimate of their conception of world relations and an estimate of how this conception changes with age or education. An argument might then be made for the use of maps which reverse the average psychological distortion (if such exists) to be used in teaching geographic relationships. Mackay's investigation of the visual interpretation of conformal and equal-area maps is somewhat comparable. Harris and McDowell also have argued for the use of deliberately distorted maps in teaching.[23]

---

[23] C. D. Harris and G. B. McDowell, "Distorted Maps - A Teaching Device," Journal of Geography, LIV, 6 (September, 1955), 286-289. Every map, of course, is a distortion in some ways.

PSYCHOLOGICAL DISTANCE

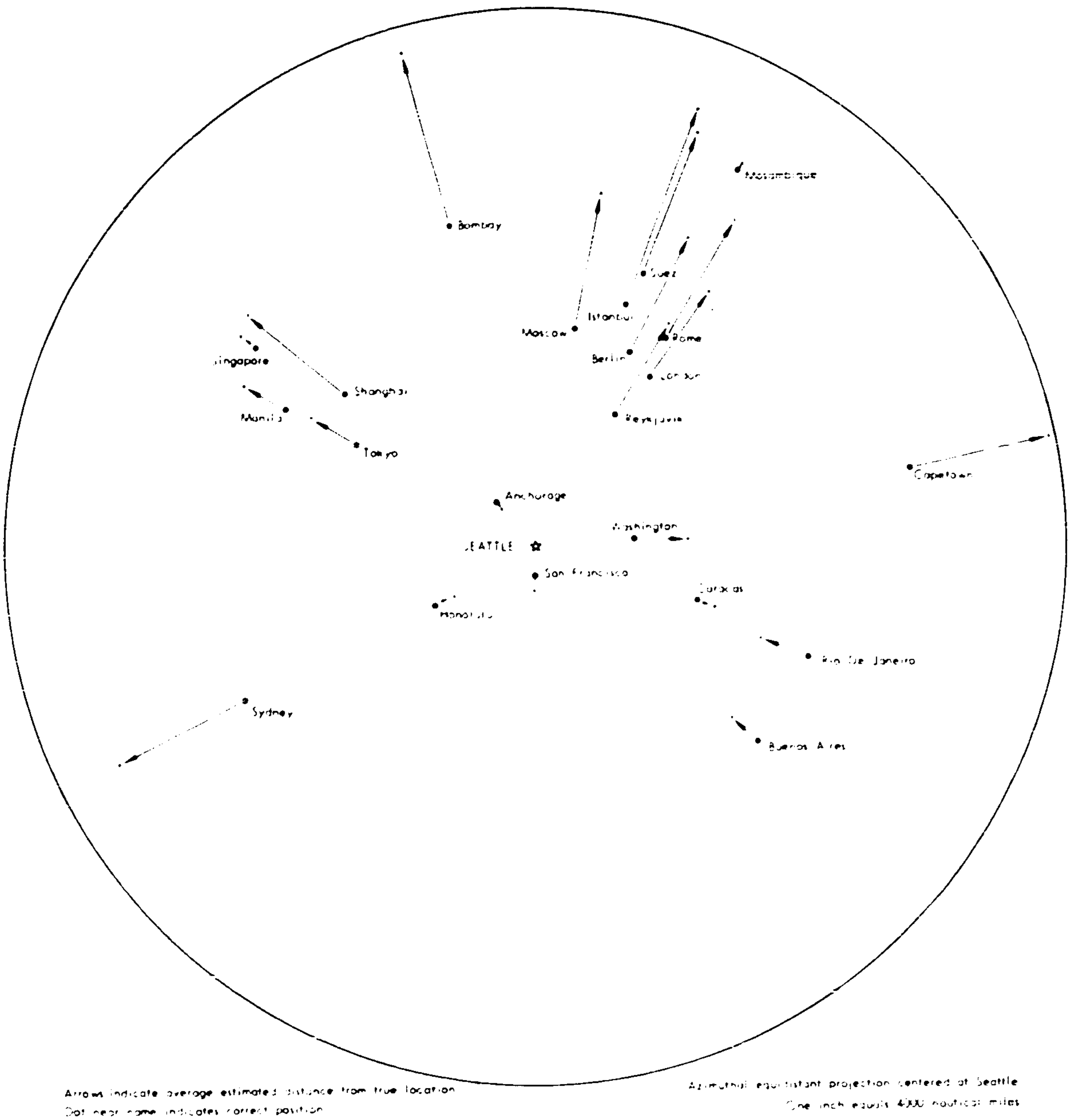


Figure 4.16

Distance Models in Geography

The introduction to this section mentions several geographic concepts and models which, explicitly or implicitly, incorporate notions of distance. This is intimately connected with the geographic importance of relative location. Some authors refer to the friction of distance; others to the attenuating effect of distance; still others to distance as offering a resistance, the overcoming of which requires an expenditure of effort. Resources located at a distance may be of less value because of their relative inaccessibility--one can then speak of a discount rate over space. Watson speaks of geography as a "Discipline in Distance," etc.[24]

The models used or proposed generally refer to a diminution of interaction with increasing spatial separation. Some researchers are not convinced of the validity of this argument, and even if valid, empirical evidence is not conclusive. There is no general agreement as to the form of the relation between distance and human activity, and even less agreement concerning the constants in the equations. The most that can be said is that the relations generally should be monotonic decreasing. The power function $z = a\rho^{b}$, $b > 0 > a$, is frequently used. There are, of course, many

---

[24]J. W. Watson, "Geography - A Discipline in Distance," The Scottish Geographical Magazine, LXXI, 1 (1955), 1-12. An imaginative discussion of the importance of distance in geography can also be found in U. Toschi, Compendio Di Geografia Economica Generale (Firenze: Macri, 1951), pp. 70-88; R. Ajo, "New Aspects of Geographic and Social Patterns of Net Migration Rate," Migration in Sweden (Lund Studies in Geography: University of Lund, Ser. B, 13, 1957), p. 170, remarks:

> As spatial configuration geographic patterns necessarily involve distance in some form or another. . . .Instead of mere distance, it is, evidently, better to try an active, that is functional or, say biologically working live distance - granted that that too must finally be expressed in terms of the common yardstick.

other functions which are monotonic decreasing, some of which have been proposed. Discussions of this topic are given by Hägerstrand, Isard, and others.[25] If the simple power function is accepted, there is still no agreement, empirical or theoretical, as to the value of the exponent. Empirical studies give a value ranging from -0.5 to -3.0, and suggest that the value is a function of the technological level of society.[26] Furthermore, it is not clear whether one value of the exponent should be valid in all situations, or whether distinct types of activity require the use of different exponents, or even forms.

In spite of the aforementioned difficulties, there does appear to be some agreement that utiles (time, cost, convenience) are a more relevant measure of distance than miles or kilometers. In employing distance models it can easily be assumed that distances from one point have been so measured, or, which is almost equivalent, that a mapping transformation $r = f(\rho,\lambda)$ to a cost or time surface has been performed. The power function then becomes $z = a[f(\rho,\lambda)]^b$. By this device it may be possible to account for empirical deviations from the centrally symmetric, strictly monotone decreasing form employed in the distance models.

A further transformation can be performed, however. It has been shown (Chapter III) that curves having a negative slope could be interpreted as an inversion. In the present context there is a more appropri-

---

[25] Hägerstrand, op. cit., part 4; Isard, Methods, loc. cit.; M. Schneider, "Gravity Models and Trip Distribution Theory," Proceedings, Regional Science Association, V (1959), 51-56.

[26] See tables in Hägerstrand, loc. cit.; also, K. L. Heald, Discussion of the Iowa Gravity Model Traffic Distribution Program, Iowa State Highway Commission, mimeographed, Ames, 1960.

ate interpretation of these functions. Thinking in terms of a centrally symmetric azimuthal projection, the radius equation can be obtained by taking the decreasing function of distance as the radial scale distortion. On centrally symmetric azimuthal projections this distortion is a function of only one parameter and can be represented by a graph (cf., Fig. 3.10). For certain of these azimuthal projections the radial scale is a decreasing function of the colatitude. The radius equation is obtained by integration. For the power function, one obtains ($b \neq -1$):

$$\int a\rho^b d\rho = \frac{a}{b+1}\rho^{b+1} + C.$$

This is similar to concepts developed when discussing transport surfaces, and the square-root projection could have been obtained in this manner. The procedure is available for any of the forms proposed for the distance models, though the integrals are not all elementary. To develop projections from each of these models would be little more than an exercise in calculus and is not undertaken here.

The maps achieved by this procedure can be interpreted, in a manner similar to the square-root projection, as showing a type of effective distance. Beyond illustration of concepts, the transformations also have the effect of converting the curvilinear influence of distance to a linear form, so that the maps may more readily be employed as nomographs. For example, the drawing power of a city, store, university, etc., at another location--assuming the validity of the models--can be evaluated by direct measurement of the map distance from the city or store. This is valid only from the center of the map, of course. A scale with non-linear graduations would work equally well on an ordinary (large scale) map. Extension from azimuthal to other types of projections (equal-area, etc.) may also be useful.

Naturally there are limitations to the amount of realism which can be incorporated on the maps by modifications of standard projections. Cost-distance relations may fluctuate over periods of time.[27] The maps, in general, must be commodity specific. Account has not been taken of the amount or importance of flow along routes.[28] These, and other, shortcomings of the material as presented notwithstanding, the transformations appear to clarify, or seem useful in, several aspects of contemporary geography.

The Distance Concept

The ease with which metrogenic substitutions can be performed suggests that an examination of the concept of distance is in order. The examples need not be repeated, but several other measures of distance have been proposed by geographers. Ground-distances and spherical-distances are not synonymous, for the earth is not a perfect sphere. The differences are rather small, of course. The surface of the earth, if bridges and tunnels are excepted, is topologically equivalent to a sphere, and, if cartographers were considerably more adept, this complicated two-dimensional surface could be mapped on a plane preserving areas or angles. The cartographic "problem of the third dimension" (relief representation) would

---

[27] Examples are given by S. W. Boggs, "Mapping Some Effects of Science on Human Relations," The Scientific Monthly, LXI (July, 1945), 45-50 and E. Fels, loc. cit. We can write $r = f(\cdot,\cdot,t)$, but this is an oversimplification.

[28] O. Lindberg, "An Economic-Geographical Study of the Localization of the Swedish Paper Industry," Geografiska Annaler, XXXV (1953), 27-40, attempts to include these factors.

vanish. Distances are more complicated, for geodesics along the ground are not unique, nor are they plane curves.

Social-distance, though difficult to define operationally, has been suggested by Watson and others. Isard suggests that distance might be measured in amounts of fuel or energy expended, or in numbers of intervening opportunities, or in numbers of gear shifts, or stops encountered. Isard's notion of transport inputs also allows substitution from monetary units to their equivalents in commodities (barrels of oil, tons of wheat, steel), or to the factors of production (land, labor). The astronomer scales distances in light years, etc. These substitutions cause little difficulty, for one measure is as arbitrary as the other, convenience, stability, and relevance being the primary considerations. Further investigation is required, however.

It is possible to define the distance from New York to other cities in the world as the number of copies of the New York Times sold in each of these cities during the course of a year. This measure of distance should perhaps be weighted according to population; the reciprocal of sales also seems more meaningful, for cities which receive no copies of the paper can then be considered as being infinitely distant, those receiving many copies as being quite close. Is this a valid definition of distance? Certainly distance, thusly defined, fluctuates from year to year, but there are circumstances under which a change in distances measured in kilometers also could not be detected.[29] A diagram positioning all cities at their correct newspaper-distance from New York could certainly be drawn

[29] H. Reichenbach, The Philosophy of Space and Time, trans. M. Reichenbach and J. Freund (New York: Dover, 1958), pp. 10-14.

and could be called a map. But then how are distances to be defined between cities except by reference to New York? Can any collection of numbers similar to newspaper sales (e.g., telephone calls, other interaction) be taken as a measure of distance? Distance defined in this manner does not appear to have the properties which one desires of distance. The concept of distance is more subtle than is commonly recognized.

The subject of map projections generally applies only to metric spaces. We can review this subject only briefly without attempting to eliminate all of the difficulties. A metric space is any collection of objects which satisfies the following postulates:[30]

1) To every pair of elements (a,b,c,...) of the set there corresponds a unique real number $D(a,b)$.
2) $D(a,a) = 0$
3) $D(a,b) > 0$, if $a \neq b$.
4) $D(a,b) = D(b,a)$
5) $D(a,b) + D(b,c) \geq D(a,c)$

The number $D(a,b)$ is called the distance from a to b. How well does geographic space satisfy these postulates? Equivalently, are the postulates satisfied if we call $D(a,b)$ the travel time or the transfer cost from a to b? Distances are not unique but our consideration of only minimum distances results in a unique real number. The second postulate does not appear to be satisfied by newspaper-distance (supra), nor when terminal costs are included in cost-distances. Postulate four is more difficult; time- or cost-distances are rarely symmetric. There is a distinct grain to

---

[30] L. M. Blumenthal, Theory and Application of Distance Geometry (Oxford: Clarendon, 1953).

the environment. As a consequence, a map such as that prepared by Boggs (Fig. 4.4) can only rarely be correct. Isochrones from a center, and isochrones to a center, are in general distinct. A simple case is the time of travel uphill as opposed to travel downhill. Here travel takes place in a force field (gravity or uniform acceleration) and perhaps the asymmetry of cost-distances at a point can be eliminated by the introduction of mathematically similar fields. However, there are some cases in which $D(a,b) = D(b,a)$; parcel post rates within the United States, for example. If distances between only a few locations are considered, the fact that distances are not symmetric causes less difficulty than if distances between an infinite number of locations in a simply-connected region are to be considered. The difficulties here are not unlike those involved in attempting to include economies of scale in the cost of transport.[31]

Spaces for which the triangular inequality (postulate five) does not hold are known as semi-metric spaces. This postulate would not appear to be satisfied, for inversions have been allowed. These inversions, however, are with respect to spherical distance and occur, in part, because the cumulative transport disutilities have been examined along azimuths, and not along the actual routes by which the various positions have been attained. To this extent the inversions are spurious. A position which is two hours away cannot be closer than one which is only an hour away, for this is an inconsistent use of the measure of distance. Similarly with cost-distances; if $D(a,b) + D(b,c) < D(a,c)$, we should choose a route from a to c via b. Considerations similar to these are examined in the

[31]See M. J. Beckmann, "A Continuous Model of Transportation," *Econometrica*, XX, 4 (October, 1952), 643-660.

following chapter. Nevertheless, it is certainly true that cumulative transport charges are not always strictly monotonic increasing along any minimum cost route (recent legislation seems directed toward forcing this condition, however[32]).

The specific distance function D(a,b) in a metric space can take many forms.[33] Terrestrial minimum time- or cost-distances are usually upper-bounded, and examples of distance functions have been implied (graphically) in the preceding pages, though explicit metrics have not been obtained. To do this for only one commodity, considering combinations of several transport domains, would be quite difficult and goes beyond the objectives of the present work.

The conditions under which the geographically relevant space can appropriately be considered a metric space are difficult to determine. When the assumption of a metric space seems too unrealistic, we should either attempt to ascertain which set of postulates are satisfied[34] or introduce mathematical conditions which can be used to account for deviations from a metric space. The latter course seems simpler, though neither is possible here. In the chapter which follows, distances from only one point are considered and the questions regarding metric spaces are less important.

---

[32]See Healy, op. cit., p. 489.

[33]In Euclidean geometry the familiar form $ds^2 = dx^2 + dy^2$ obtains; on a sphere $ds^2 = R^2 d\rho^2 + R^2 \sin^2\rho \, d\lambda^2$; more generally for locally Euclidean continua $ds^2 = g_{\kappa\rho} du^\kappa du^\rho$, etc. See any text on differential geometry.

[34]As, for example, in E. M. Zaustinsky, Spaces with Non-Symmetric Distance (Memoir 34; Providence: American Mathematical Society, 1959), or R. H. Bing, "Elementary Point Set Topology," American Mathematical Monthly, LXVII, 7, Part II (August, 1960), 42-44.

CHAPTER V

NON-AZIMUTHAL TRANSFORMATIONS

It has been demonstrated that maps can be prepared which are distance-preserving from one point, whatever the metric, by examination of distances along an azimuth. Such transformations will be referred to as polar isometric, in particular, azimuthal polar isometric. Although very simple and useful, these transformations are in many respects unsatisfactory. The azimuthal polar isometric transformations have been achieved at the occasional expense of one-to-oneness. The concept of primary concern, however, is distance--and not direction--so that azimuthality can readily be abandoned. It is desirable to exchange, so to speak, azimuthality for one-to-oneness. In addition, it does not seem entirely reasonable to measure distances in, say, monetary units and yet measure azimuths by reference to angles on the terrestrial sphere. Furthermore, it is not sufficiently clear what is intended by the term azimuth for general surfaces of highly variable curvature.

This latter point can be clarified by reference to the space in which the surface is embedded. We can assume that the embedding space is Euclidean and introduce spherical coordinates on the tangent plane to the surface, taking the origin to be the point of tangency and the equator to be the tangent plane. The lines of constant azimuth on the surface can then be defined as the intersections of the coordinate planes (meridians) with the surface in question. This is hardly a convenient family of one-parametric curves to have introduced on a surface of complicated curvature, whatever the units of distance measurement. The use of geodesic polar coordinates (*infra*) seems more fruitful.

Polar isometric transformations also are much less restrictive than general isometric transformations and are clearly not unique. Hence, it is natural to inquire which of these transformations might be most valuable and what additional properties might also be preserved. To take a simple case, one can begin with what is generally known as a conic projection: $r = f(\rho)$, $\theta = C\lambda$. Isolines drawn on any equidistant conic, $f(\rho) = \rho$, and transformed graphically as before—along the lines λ = constant—yield a polar isometric map. The entire projection is then of category B: $r = f(\rho,\lambda)$, $\theta = C\lambda$. Similar results can be obtained if $\theta = f(\lambda) \neq C\lambda$, as is obvious. More generally, it is possible to take as a basic map any of Schjerning's equidistant projections,[1] $r = \rho$, $\theta = f(\rho,\lambda)$, and obtain the desired polar isometric transformation. As a further example, the accompanying figure (Fig. 5.1) illustrates a simple case in which superimposition can be avoided by a non-azimuthal polar isometric transformation. The map is one-to-one except at certain locations, where it is one-to-many rather than many-to-one as superimposition requires. As a limiting case, the entire surface can be mapped onto the bounded straight line, collapsing all azimuths. This is a valid polar isometric transformation, but would generally be rejected.

---

[1]W. Schjerning, "Ueber Mittabstandstreue Karten," Abhandlungen K. K. Geographischen Gesellschaft, V, 4 (1904). Also in K. Zöppritz and A. Bludau, Leitfaden der Kartenentwurfslehre (Berlin: Teubner, 1912).

A NON-AZIMUTHAL POLAR ISOMETRIC TRANSFORMATION

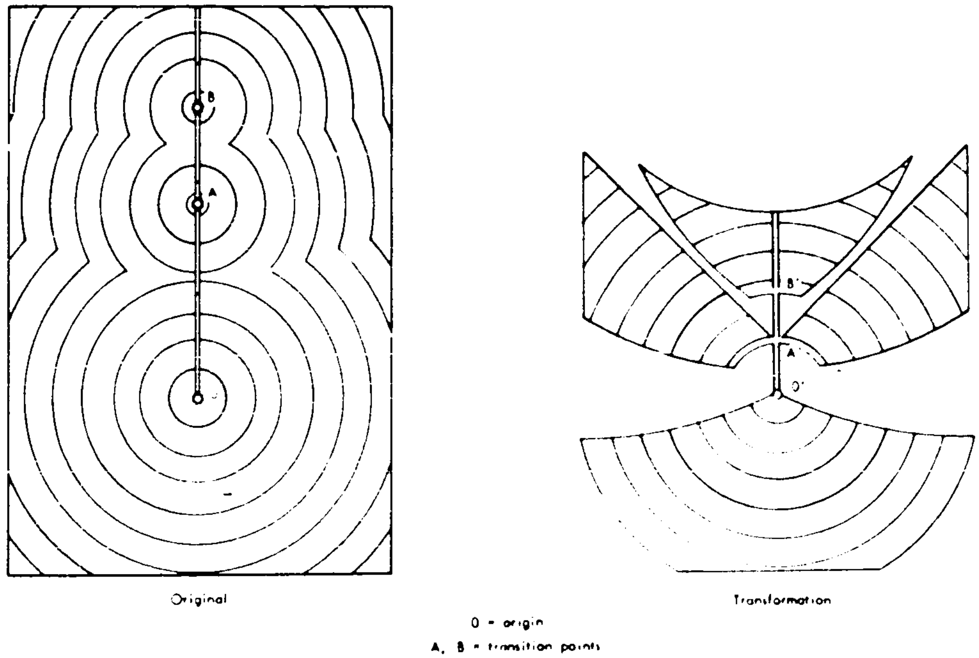


Figure 5.1

Polar Geodesic Transformations

Geodesic transformations of surfaces of variable curvature to a plane are known to be impossible. A transformation which takes geodesics from one point on a surface into geodesics (straight lines) from the corresponding point on a plane, however, is possible--at least in certain instances. The subject of geodesics is a difficult one and a short review is appropriate.[2] A common misconception is that geodesics are curves of shortest distance. On a sphere the great circle connecting two points has two arcs, only one of which is a path of least spherical distance. With

[2]See works by Kreyszig, Morse, Struik, or Willmore, listed in the Bibliography.

suitable restrictions, however, curves of shortest distance can be shown to coincide with geodesics; such restrictions are here assumed, and geodesics will be used synonymously with curves of shortest distance. It is also necessary to assume that geodesics exist. Further, geodesics in general are not unique. On a sphere there are an infinite number of shortest routes from the north pole to the south pole. As a consequence, the azimuthal equidistant projection--when it includes the antipodal point--is a one-to-many transformation at that point. We will not always assume geodesics to be unique.

Geodesic polar coordinates are based on the following results. Through an arbitrary point P of a surface S (of class $\geqslant$ 3) there passes exactly one geodesic in every direction. In a sufficiently small neighborhood U of P this geodesic is the only geodesic from P which passes through an arbitrary point Q of U. The geodesics from P and their orthogonal trajectories are called geodesic polar coordinates on S with center at P and are allowable in U with P deleted. The orthogonal trajectories define geodesic circles, with radius given by the distance along the geodesic. Our objective, and definition, of a polar geodesic transformation is one which maps the geodesics from P on S into geodesics from the image of P on a plane. This transformation is clearly not unique. For a sphere the polar geodesic coordinates are the same as spherical coordinates (poles deleted), and any transformation which sends meridians into straight lines radiating from a center is a polar geodesic transformation. Preservation of order relations and continuity are appropriately assumed here. All the azimuthal projections of a sphere are polar geodesic transformations, to which have been added the requirement that the directions of the geodesics at P are to correspond to the directions through the image of P. From the point of view of the intrinsic geometry of surfaces this is a more satisfactory definition of an azimuthal projection.

The lack of uniqueness of polar geodesic transformations suggests that further conditions may fruitfully be applied. When the angles at which the geodesics intersect (supra) are maintained, the transformation can be referred to as polar azimuthal geodesic (not to be confused with the extrinsic definition of azimuthal). By analogy, when the angles are all reduced in constant proportion, we can speak of polar conic geodesic transformations, etc. When the lengths of the geodesic rays also are preserved, the transformation is polar isometric-geodesic, although polar isometry does not imply a polar geodesic transformation, and vice versa. Polar azimuthal-isometric-geodesic is possible; recall the azimuthal equidistant of the sphere. The transformation presented in Figures 1.1 and 1.2 is of this type, though somewhat more complicated because the geodesics are not unique. The two-point equidistant projection of a portion of the sphere is polar isometric-geodesic from two points (a linear transformation of the gnomonic). When the two points are opposite poles, this polar isometric-geodesic transformation yields the straight line of length πR (see Chapter II), no longer one-to-one and not polar azimuthal. Conditions of conformality or equal-area also might fruitfully be applied to polar geodesic transformations (infra). For a sphere these are possible, even with further conditions.

It seems unsatisfactory to restrict discussion to neighborhoods of P which are sufficiently small to ensure the uniqueness of the geodesic rays, particularly when the actual situations encountered are such that this is a severe limitation. There seem to be three important cases: (a) the geodesic from P to Q is unique; (b) there are some finite number (> 1) of geodesics from P to Q; and (c) there are infinitely many geodesics from P to Q. These situations can be represented diagrammatically (Fig. 5.2).

GEODESICS

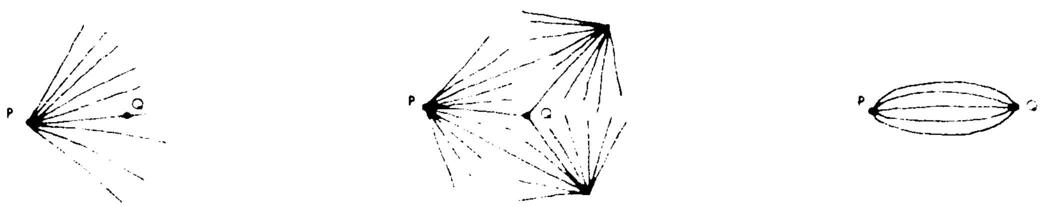


Figure 5.2

The third case illustrates the difficulty. It will be impossible to transform (polar geodesically) a situation of this type to a plane in a one-to-one fashion. In choosing between a truncation of the domain, a many-to-one, and a one-to-many polar geodesic transformation the latter seems preferable (Fig. 5.3). Compare the antipodal point on an azimuthal equal-area or equidistant projection.

Though both polar isometric and polar geodesic transformations are achieved by the application of rather weak conditions, such transformations are useful. The foregoing considerations, informal as they are, help to clarify many of the inadequacies of the azimuthal transformations presented in the earlier chapters. The discussion can be continued with simple graphic methods similar to those used in Chapters III and IV.

POLAR GEODESIC TRANSFORMATIONS

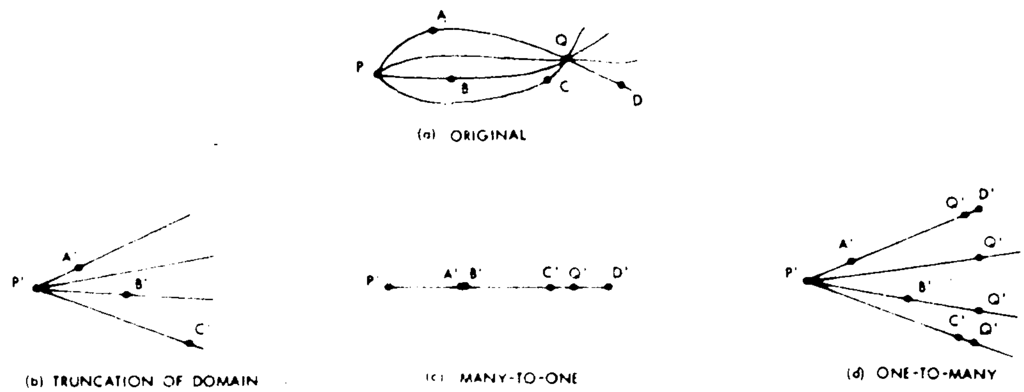


Figure 5.3

<u>The Gradient Method</u>

The gradient method provides a simple technique incorporating many of the foregoing considerations. For azimuthal projections the cumulative distance functions were to be examined along the azimuths λ = constant. This results in profiles which can be taken as graphic representations of the radius equation of an azimuthal projection of category B. This is not always appropriate, as should be clear from the foregoing discussion.

Consider the isolines shown in the accompanying illustration (Fig. 5.4). All the isolines consist of simply-connected curves. All except the innermost isolines, however, are looped with respect to the center.[3] The domain is considered to be bounded by the outermost isoline. A profile along an azimuth may intersect an isoline twice. The result, of course, is an inversion, as is easily verified by drawing the profile.

[3]Not-looped, for simply-connected isolines, is equivalent to the term star-shaped as employed by Nehari, <u>op. cit</u>., p. 220.

LOOPED ISOLINES

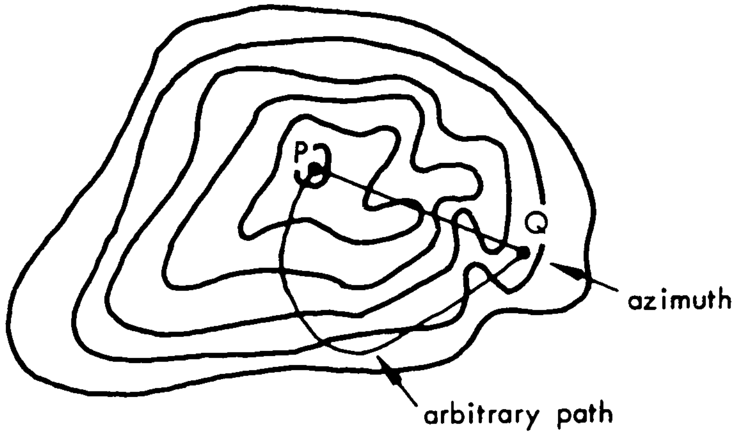


Figure 5.4

However, if any arbitrary path from the origin to a particular position is drawn through the family of isolines and if a cross section along this path is compared with the profile along the azimuth, it is seen that the lengths of the paths vary but that the height of the point (Q) is the same in all cases (Fig. 5.5a). In other words, if the isolines represent cumulative time or cost values, either path could be chosen as the radius equation for a map, and each allows the point to be placed at its correct distance on a time- or cost-distance map (i.e., a polar isometric transformation). The map will no longer be azimuthal and corresponds, in general, to a projection of category A. This arbitrary procedure is permissible because cumulative isolines, representing minimum travel times or costs, have been employed.

Having many arbitrary paths from which to choose, it is natural to pick a path along the gradient, for, of all possible paths, the area under

CROSS SECTION ALONG DISTINCT PATHS

(based on Figure 5.4)

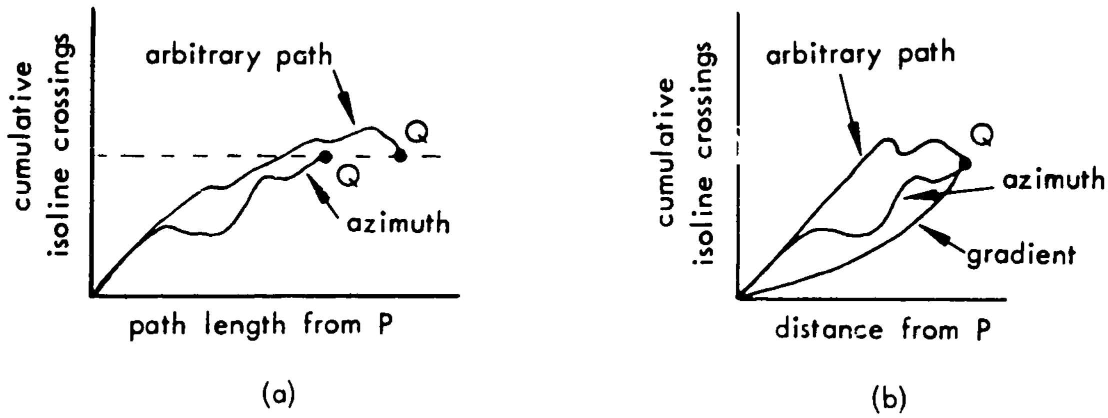


Figure 5.5

the cross section is least of all in the case of the gradient (Fig. 5.5b). To create maps scaled in time- or cost-distances from (or to) the center, the orthogonal trajectories to the family of isolines are used as the lines along which cross sections are taken as transformation equations. It is natural to map positions along these lines into positions along the rays θ = constant, at their desired distance on the map plane, each ray corresponding to one gradient. This avoids the inversion; in the symmetric case (only) the gradients and azimuths coincide, but the isolines need not be looped before the gradient method can be used. If the isolines consist of disjoint segments, the procedure is more difficult, though extensions of the gradient method to these situations is possible; the previous illustration (Fig. 5.1) can be redrawn using an asymmetric and non-linear distance function, transformed along the gradients.

The gradient method clearly yields a polar isometric transformation.

It also allows removal of superimpositions and inversions in certain instances. The gradients, however, have a further significance, for they in fact lie along the geodesics. The gradient method yields polar isometric-geodesic transformations. The apparent simplicity of this method of finding geodesics is somewhat illusionary, for, by assumption, the cumulative isolines were prepared by selecting the minimum disutility associated with reaching particular locations from the center of the map. Once the appropriate isolines (geodesic circles) have been determined the geodesics can then also be ascertained.

There is one major difficulty, however, The gradients have been taken as orthogonal trajectories to the cumulative isolines on a plane diagram. To do this, it is necessary that the cumulative isolines somehow be mapped from a surface of variable curvature onto the flat paper in a manner which preserves the orthogonality of the gradients to the isolines and yet also preserves lengths from the center. It is not clear how (or if) this can be done. Nevertheless, the difficulty can be circumvented. To do this, the notion of a transport surface of variable curvature is abandoned. Instead, it is useful to conceive of the two-dimensional surface of the earth as being covered by a very thin sheet of material through which transport takes place.[4] We imagine this sheet to be of extremely variable permeability. It is not distance which offers resistance to movement but less permeable portions of our imaginary sheet. The disutilities incurred when transport takes place can be interpreted as being the amount of "work" required to transport items through this generally uncooperative

---

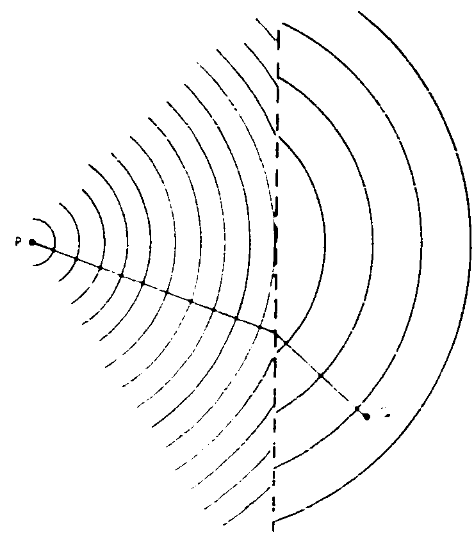
[4]This is a world not unlike Abbott's Flatland, but more complicated. E. A. Abbott, Flatland (New York: Dover, 1952).

sheet. Fortunately, all of the previous materials remain valid. In particular, the comments regarding the domain fit easily into this format, for there are areas (areal domains) of variable permeability, boundaries which may be impossible of penetration, routes of greater permeability which can be entered only at particular points, and so on. In fact, the scheme permits an improvement with regard to airplane and submarine travel. The troposphere, not unrealistically, can be considered a second sheet overlaid on the first, but which can be penetrated from below only at select points. This second thin sheet is also taken to be of variable (on the average greater) permeability, but is insulated from the first by an impermeable boundary except at isolated transition points.

Transport characteristics again can be taken as cumulative amounts of expenditure required to penetrate particular lengths of the sheet(s). Certain lengths may be of perfect permeability requiring no expenditures (step functions); others may require a disproportionate effort for any movement to occur (interruptions). It is still necessary to be commodity specific, etc., but isolines measuring the cumulative disutility required to reach particular locations from one point may be drawn on the transport sheet. As this very thin sheet is in one-to-one contact with the surface of the earth, the isolines can be considered to be drawn on this surface (taken to be a sphere). The gradients to the cumulative isolines are lines of least resistance, or greatest permeability. Not only do all the previous results hold, but simplifications appear. The content of the following is still concerned with isolines on a sphere; the diagrams, however, are no longer to be considered as being drawn on an azimuthal equidistant projection but rather on a stereographic projection. The orthogonal trajectories then still yield the gradients, as the stereographic is conformal; the lengths of gradients are distorted, but the metric is given by the isolines themselves. We can apply this immediately.

It has been known to economic geographers for many years that the least-cost route from an island port to an inland location, when the ocean rate differs from that on land, assuming both rates to be constant, satisfies the law of refraction.[5] This is a special problem of finding geodesics through two areal domains with linear transport characteristics and with a costless permeable boundary separating the domains (Fig. 5.6). Examination of the orthogonal trajectories to the cumulative isolines yields

GRADIENT AND ISOLINES



Gradient and isolines in two areal domains with linear characteristics separated by a costless permeable boundary

Figure 5.6

[5]The most complete review is in C. Ponsard, op. cit., pp. 172-176. Also Lösch, op. cit., pp. 184-186; W. Warntz, "Transportation, Social Physics, and the Law of Refraction," The Professional Geographer, (NS), IX, 4 (1957), pp. 2-7; Palander, op. cit., pp. 337-338.

the least-cost route in the non-linear (differentiable) cases and combinations of arbitrary domains. The gradient method can then be utilized to transform these routes into straight lines on a polar-isometric-geodesic map. When the gradients are not unique, or when the transport characteristics are discontinuous, the difficulties are obvious—on the map of postal zones (Fig. 4.6) the gradients do not yield the routes by which mail is ever sent, but the actual routes are of little concern, etc.

Conformal Transformations

The gradient method immediately suggests further considerations of conformal transformations. Translating isolines into concentric circles and gradients into rays from the center obviously preserves the orthogonality of intersection of these curves. This is a necessary, but not sufficient, condition for a conformal map.

The subject of conformal maps, though much investigated and of importance in many fields, is rather difficult. Only a few general comments on this topic are to be presented here, without going deeply into the details.[6] The geographic relevance of conformal transformations has always been somewhat obscure. Although conformal projections (the stereographic, Lambert conformal conic, Mercator, etc.) are recorded in the geographic-cartographic literature, the major users of such projections are not geographers, but geodesists, meteorologists, oceanographers, navigators, etc. The reason is that angles have not been particularly important to geographers. The concepts most frequently encountered in the geographic litera-

[6]See Nehari, op. cit., or works by Betz, Bieberbach, Knopp, and Thomas listed in the Bibliography.

ture are distance and area. Angles and directions are less explicitly emphasized. This can be contrasted with meteorology where winds are assumed to flow at right angles to isobars (neglecting Coriolis effects, etc.). The intention is not to imply that angles are unimportant for geography, but only that they have not been emphasized. The stereographic projection has been employed here (supra) because of a relation between geodesics and isolines. Any further conformal transformation of the stereographic projection will maintain this relation. It is also relatively easy to calculate the area between given limits after a conformal transformation. Another indication of a justifiable geographic use of conformal transformations is in the potential models suggested by Stewart, for here the orthogonal trajectories to the equipotential lines are important.

We have seen that polar-geodesic-isometries are not unique. Can conformality be added as a further condition? For a sphere this would require straight meridians of correct length and conformality. Such a map is impossible when going from a sphere to a plane, as it implies a complete absence of distortion. In the event that the cumulative transport costs through the transport sheet on the sphere are strictly monotone increasing, not centrally symmetric, and are differentiable, however, the scale distortion of the plane map will vary from point to point and it may be possible to obtain a conformal map of the sphere which also yields valid cost distance relations from some one point.[7]

---

[7]A definitive statement in this respect cannot be given without an analysis in depth of the differential equations involved (assuming differentiability, etc.). We have not undertaken the detailed analysis which would be required. While reference to known map projections may suggest the existence or impossibility of specific combinations of properties, it is not a certain procedure for the general case. A proof of the existence

Lambert's conformal conic projection of the sphere is polar-geodesic in addition to being conformal. It is also known, from the Riemann mapping theorem, that general bounded, simply-connected, two-dimensional domains can be conformally mapped onto a circle. It is too much to expect that this will automatically yield a polar-geodesic map. Nevertheless, these geographic applications of conformal transformations seem to warrant further investigation.

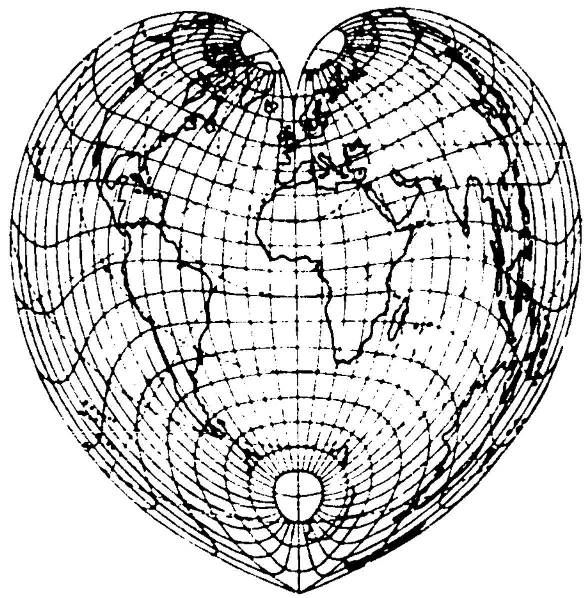
Equal-area Transformations

The von Thünen model is concerned with land use. In this situation it is appropriate to inquire as to the amount of surface area devoted to a particular type of land use. The question then becomes: is it possible to transform geographic space in such a manner that equal cost intervals from some point go into equal map distances and that equal surface areas retain their proportionality of size on the map? For the sphere this is a requirement that $r = \rho$ and that the condition for equality of area be met.[8] The result is Werner's heart-shaped map, no longer azimuthal as was to be expected (Fig. 5.7). The directions are distorted and the map contains an interruption near the center. This causes no problem as distance and area, but not direction, are the concepts of primary concern. The isolines can be plotted on this map as before. We could, in fact, have

(or impossibility) of solutions, in general or under restricted conditions, also often does not tell one how these solutions can be obtained.

[8] $r = R\rho$, $\theta = \frac{\lambda \sin \rho}{\rho}$. The derivation is given by O. S. Adams, op. cit., pp. 31-33. Devised in the fourteenth century, this projection is but rarely employed. Further details can be found in J. Keunig, "The History of Geographical Map Projections Until 1600," Imago Mundi, XII (1955), pp. 1-24.

WERNER'S EQUIDISTANT EQUAL-AREA PROJECTION



From K. Zöppritz and A. Bludau, Leitfaden der Kartenentwurfslehre, Berlin, Teubner, 1912

Figure 5.7

used Werner's projection rather than the azimuthal equidistant for the drawing of isolines all along. Transformation of distances along the lines θ = constant will yield a polar isometry. This transformation does not yield an equal-area polar isometry, however. To retain equality of spherical area and cost-distances from one point, it appears necessary to describe the isolines in mathematical form and to apply the additional conditions for equality of area, unless, of course, graphic methods can be devised. This corresponds to the difficulties encountered in the examination of conformal maps. For certain simple geometric configurations, mathematical description of the isolines can be achieved without excessive effort (Fig. 5.8).

Lambert's azimuthal equal-area projection of the sphere is, of course, polar-azimuthal-geodesic in addition to being equal-area. We cannot, however, draw isolines on either this projection or on Werner's and expect to find that the orthogonal trajectories to the isolines are the gradients (except in very special cases). According to Tissot, there will be only one set of orthogonal trajectories which remains orthogonal after the equal-area transformation. Having (empirically) obtained the isolines, the gradients can be drawn correctly as orthogonal trajectories on a conformal map (e.g., stereographic) or, of course, on the sphere. In the former case the gradients can be transferred to the equal-area map in a manner similar to the plotting of great circle routes on the Mercator projection from a gnomonic map. In the latter case the gradients would be plotted as are other terrestrial positions. The gradient method can then be used to obtain the polar-geodesic isometry; the resulting map will still not be equal-area, however.

Nevertheless, we can conjecture that polar isometries maintaining spherical surface area are possible.[9] The details will depend on the specific distance functions and the resulting maps may appear rather strange. It is also obvious that proportionality of area cannot be maintained when step functions, resulting in collapsing of areas, are involved. The situations with inversions and superimpositions naturally are also more difficult.

Sphere or Plane

Studies which employ isolines are often restricted to areas far

---

[9]See note 7, *supra*.

smaller in extent than the entire sphere. In urban areas in particular it may be appropriate to consider the domains (or transport sheet) as existing on a plane rather than on a sphere. A simple case already treated in the literature is that in which two transport systems exist, one a routed structure radiating from a center, the second an areal structure existing elsewhere.[10] Both systems are taken to have linear characteristics, to have permeable boundaries, and to have no disutility upon crossing the boundary. The routed structure is taken to have a lesser disutility with distance. In this instance it is not difficult to show that the isolines depend solely on the ratio of the coefficients of the two characteristics (Fig. 5.8). The implication is that this model corresponds to urban traffic situations, the routed structure to arterial streets, and the areal structure to the remainder of the urban area. An alternate formulation is to consider the network of streets as consisting of a gridiron pattern.[11] Whether these models are realistic is not really of concern here. The graphic methods previously given for transformations remain valid. When distance from a center is the only concept of concern, the transformations do not differ in any essential points for a bounded plane or a sphere with antipodal point deleted. If areas or angles are also to be considered, the situation naturally changes. The differential conditions for conformality and equality of area are slightly modified.

---

[10] This simple case has been treated a great number of times: Lösch, op. cit., note 26, p. 443; Mohring, et al., op. cit., Appendix; A. Vadnal, La Localisation du Réseau de Routes Collectrices dans le Cercle (Bilbao: Congress European D'Econometrie, 1958); Palander, op. cit., gives other examples. Also compare Figure 3.12.

[11] Wingo, op. cit., p. 39; Lösch, op. cit., p. 442.

LEAST-COST PATH AND ISOLINES

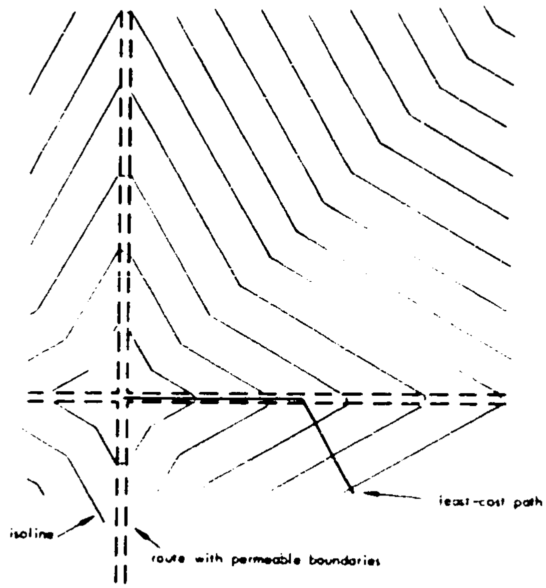


Figure 5.8

If the values for Tissot's indicatrix are to be used in evaluating the distortion introduced by the transformation (or, equivalently, by the transport media), the specific equations are different for a sphere or plane. Whether to consider the transformation as being of a plane or sphere, of course, depends on the accuracy requirements of the particular study. An advantage in beginning with a plane is that it is no longer necessary to employ auxiliary projections such as the azimuthal equidistant, the stereographic, or that of Werner. The properties which we might wish to retain are similar whether we start with a plane or sphere.

As an aside, it is interesting to note that even if the earth were a disk (as some ancients believed) and not sphere-like, the suggested transformations still would be of value. The maps obtained here as transformations also are reminiscent of maps produced in the middle ages. Other equally unusual maps can be considered transformations, as in the chapter which follows.

CHAPTER VI

GEOGRAPHIC AREA

The preceding chapters have attempted to relate distortions of distance on map projections to geographic concepts. It is now postulated that distortions of area can be equally meaningful. A fundamental truism of geography is that the incidence of phenomena differs from place to place on the surface of the earth. It is, therefore, understandable that ceteris paribus assumptions in theoretical descriptions of relations among phenomena which conflict seriously with this fundamental fact are somewhat repugnant to a geographer. Von Thünen, for example, postulates a uniform distribution of agricultural productivity; his economic postulates are no less arbitrary, but they disturb the geographer somewhat less. Christaller's central place theory belongs to a similar category, for the necessary simplifying assumptions, e.g., a uniform distribution of purchasing power, are unsatisfactory from a geographic point of view. In order to test the theory empirically, one must find rather large regions in which the assumptions obtain to a fairly close approximation.[1] The theory, of course, can be made more realistic by relaxing the assumptions, but this generally entails an increase in complexity. An alternate approach, hopefully simpler but equivalent, is to remove the differences in geographic distribution by a modification of the geometry or of the geographic background. This has

---

[1]For an example, see J. Brush, "The Hierarachy of Central Places in Southwestern Wisconsin," The Geographical Review, LXIII, 3 (1953), 380-402.

been attempted by other geographers (*infra*), with some success, but without clear statement of the problem. The graphic methods employed in earlier chapters are no longer as satisfactory, but the topic can be approached by an examination of a number of cartograms. Attention is directed toward those types of cartograms which appear amenable to the metrical concepts of the theory of map projections, without here attempting a definition of the rather vague term cartogram.

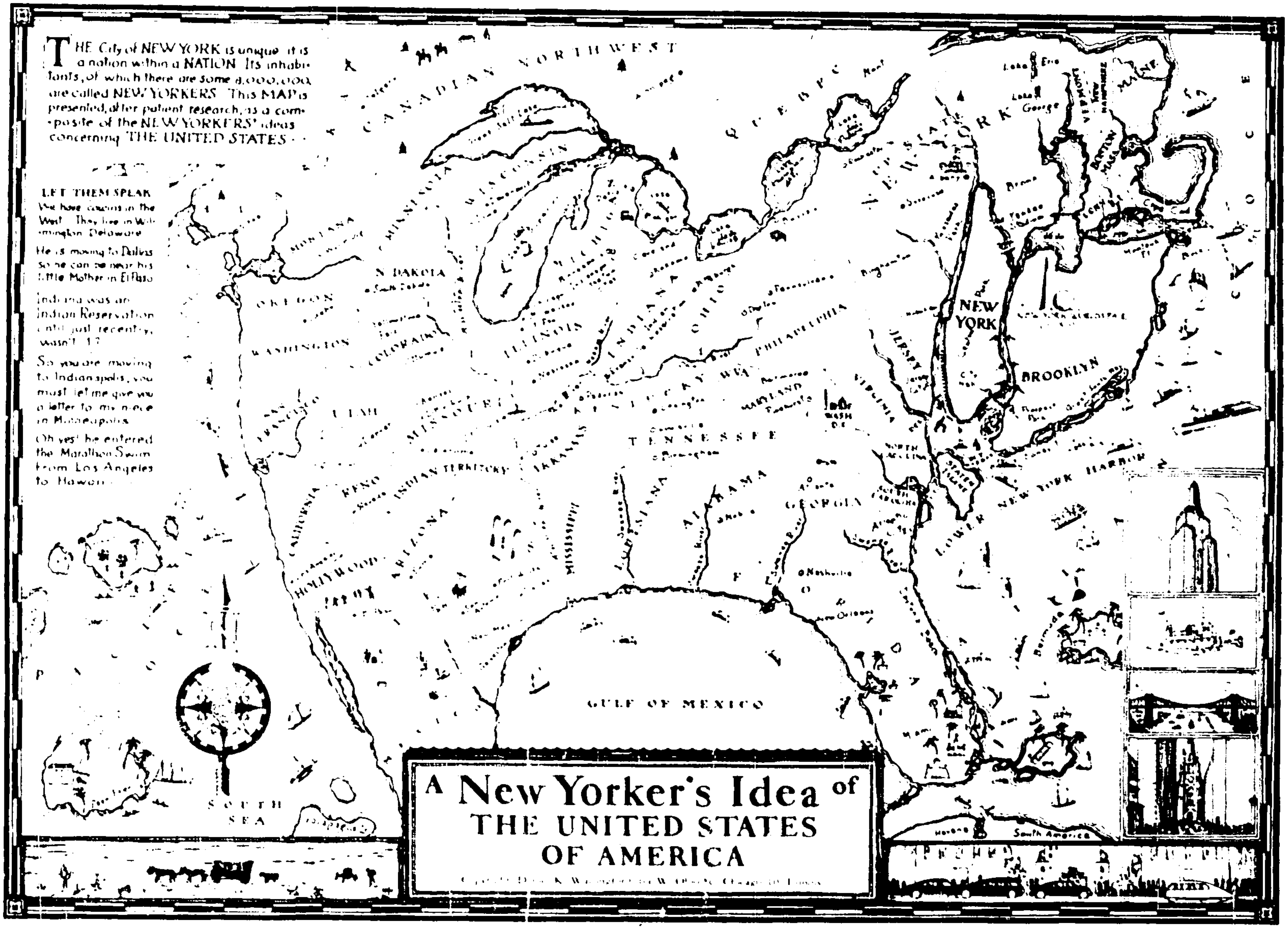
Examples of Cartograms

The accompanying illustration, "A New Yorker's View of the United States," contains several interesting notions (Fig. 6.1). The cartogram reinforces the impression that there is a psychological distortion of the geographic environment in the minds of many persons. It is also clear that this distortion, *inter alia*, is related to distance. Furthermore, the areas of the states are not in correct proportion. Florida, for example, appears inordinately large. Hence, distortion of area can be recognized, though a complete separation of the concepts of distance and area is not possible.

The second illustration is again a distorted view of the United States (Fig. 6.2). The purpose of this cartogram is more serious, however. The area of the states and cities is shown in proportion to their retail sales, rather than in proportion to the spherical surface area enclosed by their boundaries. Harris' point is that the expendable income, not the number of square miles, is a more proper measure of the importance of an area--at least for the purposes of the location of economic activity.[2] Harris also

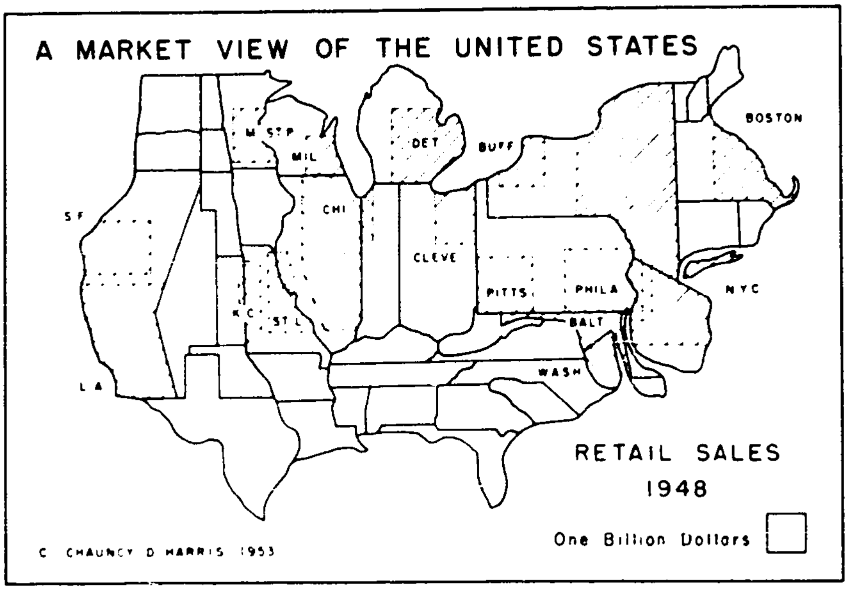
---

[2]C. D. Harris, "The Market as a Factor in the Localization of



Courtesy of Daniel K. Wallingford

Figure 6.1



From C.D. Harris, "The Market as a Factor in the Localization of Industry in the United States", Annals, Association of American Geographers, XLIV, 4, 1954. Courtesy of C.D. Harris.

Figure 6.2

presents cartograms of the United States with map areas of the states in proportion to: (a) the number of tractors on farms, and (b) the number of persons engaged in manufacturing. Raisz presents an alternate cartogram with the area of states in proportion to their population.[3] Hoover stresses a point of view similar to that of Harris and presents a different cartogram of the United States, again with map areas of the cities and states in proportion to their population.[4]

Industry in the United States," Annals, Association of American Geographers, XLIV, 4 (1954), 313-348.

[3]E. Raisz, "The Rectangular Statistical Cartogram," The Geographical Review, XXIV, 2 (1934), 292-295.

[4]E. M. Hoover, The Location of Economic Activity (New York: McGraw-Hill, 1948), p. 88.

Weigert recognizes that the importance of the countries of the world may be more directly proportional to their population than to their surface area and presents a cartogram placing the nations in this perspective.[5] Woytinsky and Woytinsky make extensive use of a similar cartogram (Fig. 6.3).[6] Zimmerman presents further examples: cartograms of world population and output of steel by country.[7]

Whether the cartograms presented are considered maps, based on projections, is a matter of definition and, as such, not really important. Raisz stresses the point that his rectangular statistical cartograms are not map projections. The background of latitude and longitude on Woytinsky's illustration of population (Fig. 6.3) suggests a map projection but is actually spurious. However, as every map contains distortion, the diagrams can be considered maps based on some unknown projection. Certainly the definition which considers an orderly arrangement of positions on the surface of the earth on a plane sheet as a map projection suffices. It also seems adequate to demonstrate that diagrams similar to the foregoing cartograms can be obtained as map projections. But what is the nature of these projections? The question is approached by a detailed examination of a simpler problem posed and solved by Hägerstrand.

Hägerstrand's Problem

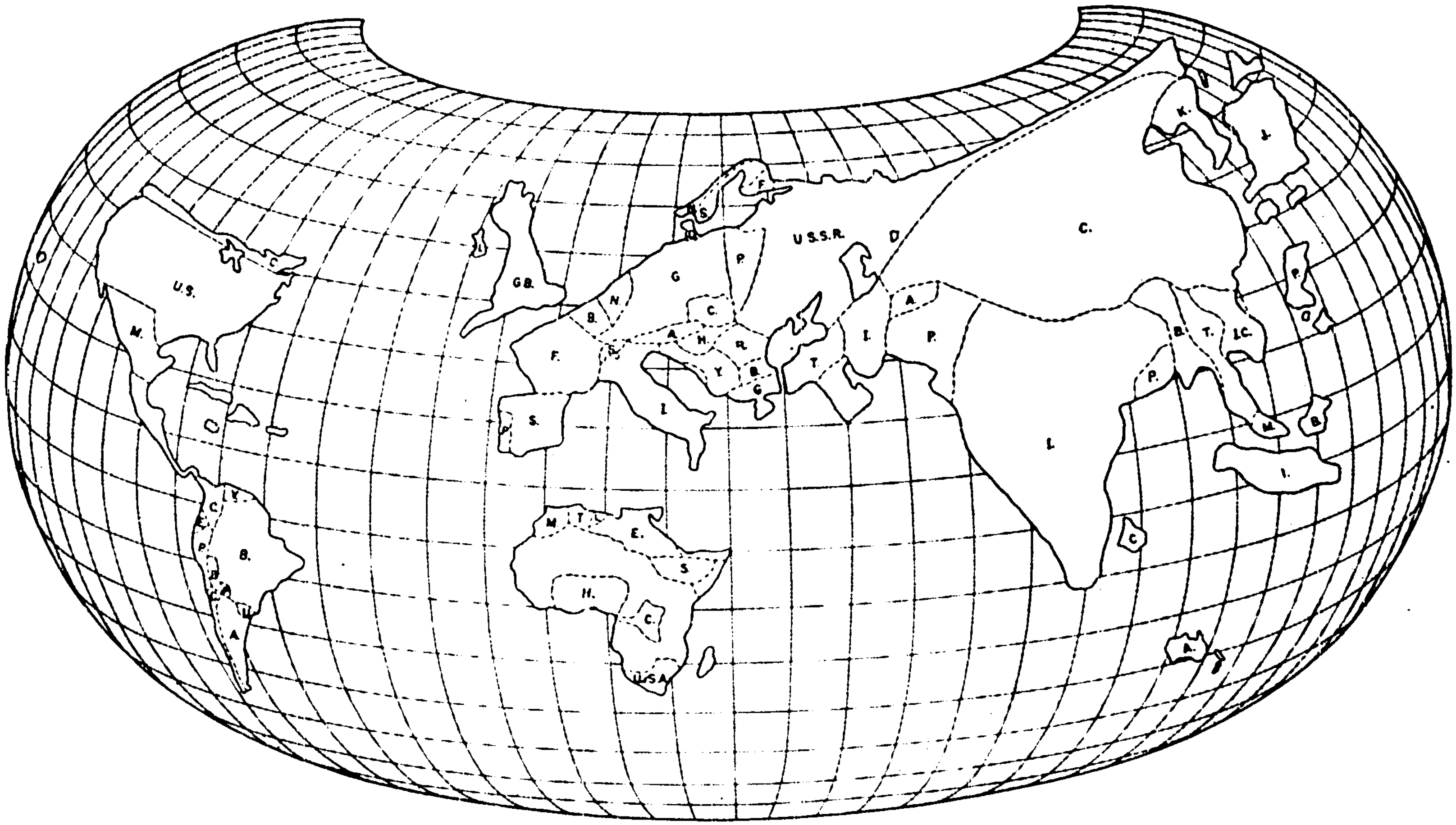
Hägerstrand has been concerned with the study of the migration of

---

[5]H. W. Weigert, et al., Principles of Political Geography (New York: Appleton-Century-Crofts, 1957), p. 296.

[6]W. S. and E. S. Woytinsky, World Population and Production (New York: Twentieth Century Fund, 1953), pp. lxix-lxxii, 42-43, and passim.

[7]E. W. Zimmerman, World Resources and Industries (Rev. ed.; New York: Harper, 1951), p. 97.



Continents and Selected Countries on the Scale of Their Population

From A. S. Woytinsky and E. S. Woytinsky, World Population and Production, Twentieth Century Fund, New York, 1953. Courtesy of the Twentieth Century Fund.

Figure 6.3

people over long periods of time. In discussing the cartographic problem, he states:

> The mapping of migration for so long a period, giving the exchange of one single commune with the whole country in countable detail, cannot be made by ordinary methods. All parts of the country have through the flight of time been influenced by migration. However, different areas have been of very different importance. With the parishes bordering the migrational centre, the exchange has numbered hundreds of individuals a decade. At long distances only a few migrants or small groups are recorded. A map partly allowing a single symbol to be visible at its margin, partly giving space to the many symbols near its centre, calls for a large scale since we wish to be able to count on the map.[8]

Hägerstrand's solution is deferred until some of the implications of the problem are examined.

It is desired to count symbols on the map. This is a clear statement of a common cartographic problem. The situation occurs frequently in the mapping of population, where high concentrations appear in restricted areas and lesser numbers of people are spread more thinly throughout the remainder of the map area. Certainly every cartographer has at some time wished for a distribution of phenomena which does not seem to require that all the symbols overlap. A solution has been the introduction of so-called three-dimensional symbols.[9] An alternate solution is here suggested, based on the theory of map projections. Also note the distinction between the common geographic use of an equal-area map to illustrate the distribution of some phenomena and Hägerstrand's emphasis on the recovery of information recorded on the map.

In the problem as formulated by Hägerstrand, the exchange of migrants

---

[8]Hägerstrand, op. cit., p. 73. Emphasis is in the original.

[9]Cf., A. Robinson, Elements, pp. 169-171.

is known not to be distributed arbitrarily but is a function of distance from a center--the commune being studied. This special situation can again be equated with an azimuthal projection. More commonly, differences from one area to another vary much more drastically, as for example the distribution of population throughout the world.

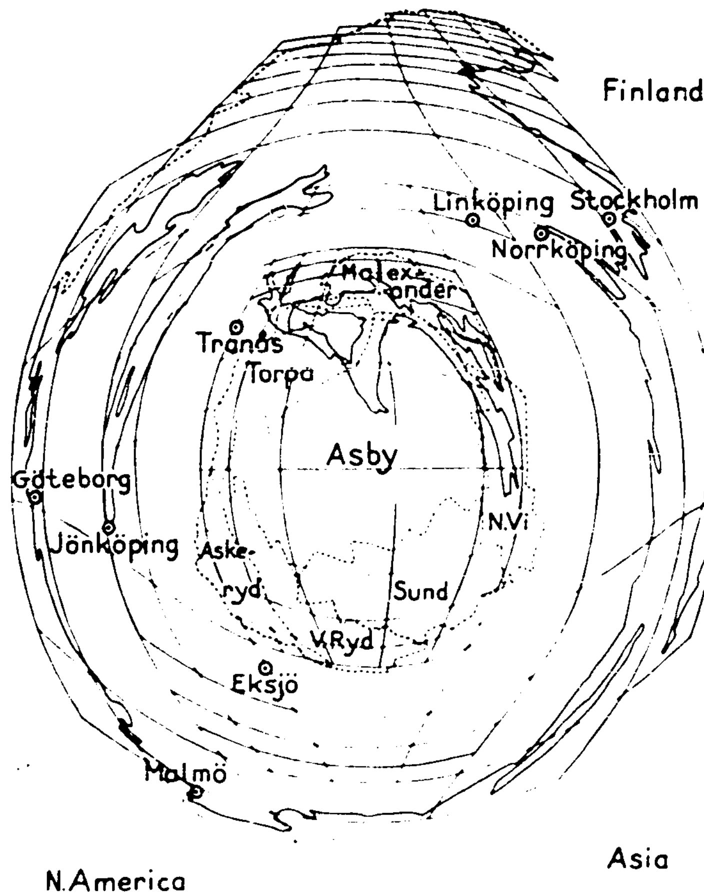
Careful reading of Hägerstrand's comments suggests that the functional dependence is one of decreasing migratory exchange with increasing distance from the center. This can be recognized as a loose formulation of a distance model, as discussed in a previous chapter (IV). In particular, the suspected function of distance can be postulated to be continuous and differentiable, strictly monotone decreasing, and independent of direction. If these postulates are accepted, the functional dependence can be shown as a continuous curve on a graph, in this instance, a curve of negative slope. This curve can be considered a profile along an azimuth, and the expected incidence of migration could be shown on a map by isolines. This suggests that variants of the solution to Hägerstrand's problem can be applied to many isoline maps. Population density, for example, is often illustrated by isolines drawn on maps and an approach to the population cartograms is suggested. This interpretation of isolines is explained in greater detail below and is distinct from that given in Chapter IV.

Hägerstrand's solution is quoted as follows:

> The problem is solved by the aid of a map-projection in which distance from the centre shrinks proportionally to the logarithm of the real distance. (The method was suggested to the author by Prof. Edg. Kant. Maps of a similar kind are used for the treatise "Paris et l'agglomeration Parisienne" 1952). The rule obviously cannot be applied to the shortest distances. Thus the area within a circle of one km radius has been reduced to a dot.
>
> The distortion in relation to the conventional map is of course

considerable.[10]

HÄGERSTRAND'S LOGARITHMIC MAP



From T. Hägerstrand, "Migration and Area", Migration in Sweden, Lund Studies in Geography, Series B, No. 13, 1957.

Figure 6.4

There is no real clue as to how this solution was obtained. A projection which yields the desired result seems to have been plucked out of thin air. Working backwards, however, the radial scale distortion is seen to be ρ^{-1} and it can be inferred that the method used was identical to that

---

[10] Hägerstrand, op. cit., p. 74.

previously given for more general distance models as employed in geography.[11] The space-elimination at the origin is appropriate for it excludes the commune being studied (which does not belong to the domain of migration). But is Hägerstrand's solution the most valid solution to the problem as formulated? The concept of primary concern is not distance, but area. This is implicit in the statement that it is desired to be able to count symbols on the map. The suggestion is that the map show the areas near the center at large scale and those at the periphery at small scale. Such maps would be useful in any study of a nodal region. Hägerstrand's solution achieves this objective, as can be verified by calculation of the areal distortion, at least for areas near the center of the map. But so does the orthographic projection, the square-root projection (Fig. 3.10), and many others. The azimuthal equidistant centered on the antipodal point also yields the desired solution and has been used for this purpose by Michels.[12] Kagami suggests another solution when faced with an almost identical problem.[13]

Cartograms as Projections with Areal Distortion

To clarify the situation it is natural to require that the areal distortion be <u>exactly</u> the same as the expected or observed distribution. Somewhat more precisely, Hägerstrand's problem can be generalized in the following manner. In the domain under consideration there are locations

---

[11]See Chapter IV.

[12]Michels, loc. cit.

[13]K. Kagami, "The Distribution Map by the Method of Aeroview," The Geographical Review of Japan, XXVI, 10 (1953), 463-468.

from which migration to the center originates. Considering the beginning point of each migration to be an "event," each small region (or element of area) will contain a certain number, or incidence, of events. Hence, with each proper partition of the domain there is associated a number, and the area contained within the boundaries of corresponding partitions on the map is to be proportional to this number. In the case of an equal-area projection, the number associated with each partition is the spherical surface area. The similarity to the cartograms previously presented is now clearer. In each instance a set of non-negative numbers (people, tractors, etc.) has been associated with a set of bounded regions (cities, states, nations). The objective is to display the regions on a diagram in such a manner that the areas within the boundaries of the regions on the diagram are proportional to the corresponding number associated with the particular region. Harris recognizes the similarity of the concepts, for his cartogram "A Farm View of the United States" is accompanied in the original article by a histogram of the number of tractors by states.

There seem to be two methods of attacking the details of the transformations in question; one assumes differentiability; the other is an analogue of the first but employs what might be called rule of thumb procedures. Each method has advantages and disadvantages. The differentiable cases display the similarity to equal-area map projections somewhat better, whereas the approximation methods are simpler to use with empirically obtained data.

The data are somewhat difficult to manipulate when the partitions of area are large. It, therefore, is convenient to reduce the values to density form and to think in terms of a continuous distribution which can be represented by isolines. The details of this device are well known and can be omitted here.[14] The map area between given limits is then to be proportional to the volume under the density surface between corresponding limits. The density distribution is assumed to have been described by an equation. For equal-area projections the density of spherical surface area is always unity, so that correct values are also obtained in this special situation. The use of density values facilitates the further objective that common boundaries between regions should also coincide on the final map.

The simplest demonstration of the procedures to be used is afforded by azimuthal projections. The area, as a function of the radius, of a circle on the Euclidean plane is given by the integral $\int_0^{2\pi}\int_0^{r} r\,dr\,d\theta$, which reduces to πr^2. This is the map area. As a simple illustration of the general method, we can take the special case of a distribution of phenomena on a plane which is symmetric about the origin, i.e., the density is taken to be a surface of revolution $z = f(\delta)$, using polar coordinates δ, γ. The volume under this surface is given by the integral $\int_0^{2\pi}\int_0^{\delta} \delta f(\delta)\,d\delta\,d\gamma$, or $2\pi\int_0^{\delta} \delta f(\delta)\,d\delta$. This integral is to be equated to the area of a circle of radius r in the map plane, i.e., $r^2 = 2\int_0^{\delta} \delta f(\delta)\,d\delta$.

For a centrally symmetric density on a sphere write $z = f(\rho)$ for the density and the appropriate integral is $\int_0^{2\pi}\int_0^{\rho} R^2 f(\rho)\sin\rho\,d\rho\,d\lambda$, or $2\pi\int R^2 f(\rho)\sin\rho\,d\rho$, which is again to be equated to the area of a circle in the plane. The resulting equation is $r^2 = 2\int_0^{\rho} R^2 f(\rho)\sin\rho\,d\rho$. If this

[14] C. Brooks and N. Carruthers, Handbook of Statistical Methods in Meteorology (London: H. M. Stationery Office, 1953), pp. 161-165. C. F. Schmid and E. A. MacCannell, "Basic Problems, Techniques, and Theory of Isopleth Mapping," Journal, American Statistical Association, L (1955), 220-239.

equation can be integrated, we can solve for r as a function of ρ, obtaining the radius equation for the desired azimuthal projection. The reader with a cartographic background will recognize the above equation as that yielding the Lambert equal-area azimuthal projection in the special case in which $f(\rho) = 1$.

It has been conjectured (supra) that these results can be obtained by choosing the areal distortion in such a manner that it exactly matches the expected or known density distribution. If true, this would yield a relatively simple method of obtaining the desired transformations. The areal distortion of many map projections also is already known. The proof, for centrally symmetric densities and azimuthal projections of category D, is as follows.

The areal distortion (S) for this class of projections is known (from Tissot) to be $S = \frac{r}{R^2 \sin\rho} \frac{dr}{d\rho}$. Setting $S = f(\rho)$ and using the method of separation yields $R^2 f(\rho) \sin\rho \; d\rho = r \, dr$, whence $\int R^2 f(\rho) \sin\rho \; d\rho = \frac{r^2}{2}$, which reduces to an equation identical to that previously obtained. This proves the conjecture for the special case of centrally symmetric density distributions and azimuthal projections of category D. Lambert's equal-area azimuthal projection is usually obtained in this manner, setting $S = 1$ and taking the constant of integration to be $2R^2$. The proof for more general distributions $z = f(\phi, \lambda)$ can be obtained in a similar manner as a transformation of a surface integral. The resulting partial differential equation to be solved is: $f(\phi,\lambda)R^2 \cos\phi = |J|$. The difficulty of an explicit solution to this equation will depend on the specific form of the function $f(\phi,\lambda)$ and the additional conditions applied. As is typical of differential equations, in general there will be an infinitude of particular solutions, which suggests that additional conditions be applied. These conditions, of course, should be suggested by the applications of the projections (*infra*).

As a concrete example, the distribution of population in an urban area can be described as a density function $f(\delta,\gamma)$ on a plane. Horwood has suggested one such distribution in which density decreases from the center but also varies from one direction to the next.[15] The specific function taken by Horwood is such that density is highest along symmetrically spaced radial streets (n in number) and less in the interstitial areas (compare Fig. 3.12), which is not unrealistic and easily described by trigonometric functions. The population would then be given by the integral $\iint \delta f(\delta,\gamma)\, d\delta\, d\gamma$. To transform this to the map plane set $\iint r\, dr\, d\theta = \iint \delta f(\delta,\gamma)\, d\delta\, d\gamma$, or $\iint r\,|J|\, d\delta\, d\gamma = \iint \delta f(\delta,\gamma)\, d\delta\, d\gamma$, or $r\,|J| = \delta f(\gamma,\delta)$, where $J = \frac{dr}{d\delta}\frac{d\theta}{d\gamma} - \frac{d\theta}{d\delta}\frac{dr}{d\gamma}$. For one solution, not necessarily the most appropriate but simple, stipulate that the transformation is to be azimuthal, i.e., that $\theta = \gamma$. Then $\frac{d\theta}{d\delta} = 0$ and $\frac{d\theta}{d\gamma} = 1$. Hence, $J = \frac{dr}{d\delta}$. The equation to be solved for r is then $r\frac{dr}{d\delta} = \delta f(\delta,\gamma)$ or $r\, dr = \delta f(\delta,\gamma) d\delta$, whence $r^2 = 2\int \delta f(\delta,\gamma)\, d\delta$, and the remaining details are matters of integration and root extraction. This example could be extended to the sphere but for an urban area there is no point in such extension.

The assumption of continuity of a distribution is often not warranted.[16] The data are often in the form of discrete locations, as on a

---

[15] E. M. Horwood, "A Three-Dimensional Calculus Model of Urban Settlement," Regional Science Association Symposium, Stockholm, August, 1960.

[16] This assumes a definition of continuity, which has not been given. The separation of methods used here can be avoided by appropriate neighbor-

population dot map, or grouped into areal units, such as census tracts, or refer to areal units rather than infinitesimal locations, as land values which refer to specific parcels of land. In these cases an analytic solution is usually not feasible and rule of thumb approximations are useful. Even in the case of continuous distributions, descriptive equations are difficult to obtain and, at present, are not available for geographical data, though theoretically possible.[17] Approximation methods, hence, are very useful. They can also be used to demonstrate some of the different types of particular solutions available and some of the additional conditions which may be applied.

The only known description of the actual method used in the preparation of the cartograms previously presented is that given by Raisz;[18] the methods used by others is presumably similar. The population of the states is taken as given, and rectangles proportional to population are drawn on a sheet of paper, adjusting adjacent rectangles until neighbor relations and overall shape are approximately correct. This is illustrated in the accompanying figure. Though the example is very simple, there are still an infinite number of solutions, but some seem more appropriate than others (see Fig. 5.6). Preservation of the internal topology is one con-

---

hood definitions of continuity, and the transformations discussed in this section can be considered continuous. To establish the necessary definitions and apparatus would be too much of a deviation for the primary purpose. See N. Rashevsky, Mathematical Biophysics (3rd ed., Vol. II; New York: Dover, 1960), Chapter 18 and pp. 317-319.

[17]Geodesists (notably Prey) and meteorologists (Haurwitz, Petterssen, Bryson) have employed such descriptions, and geographically useful attempts are now being made.

[18]E. Raisz, "Rectangular Statistical Cartograms of the World," Journal of Geography, XXXV (1936), 8-10.

SAMPLE TRANSFORMATION

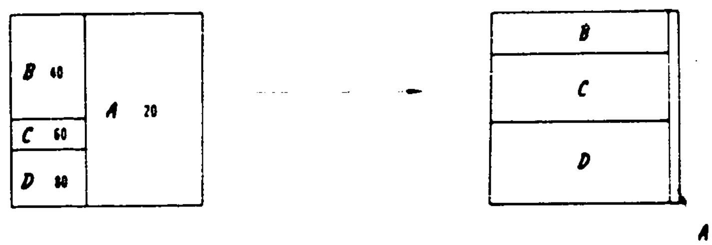


Figure 6.5

dition which might be applied; preservation of the shape of the external boundary is another, etc.

SAMPLE TRANSFORMATIONS

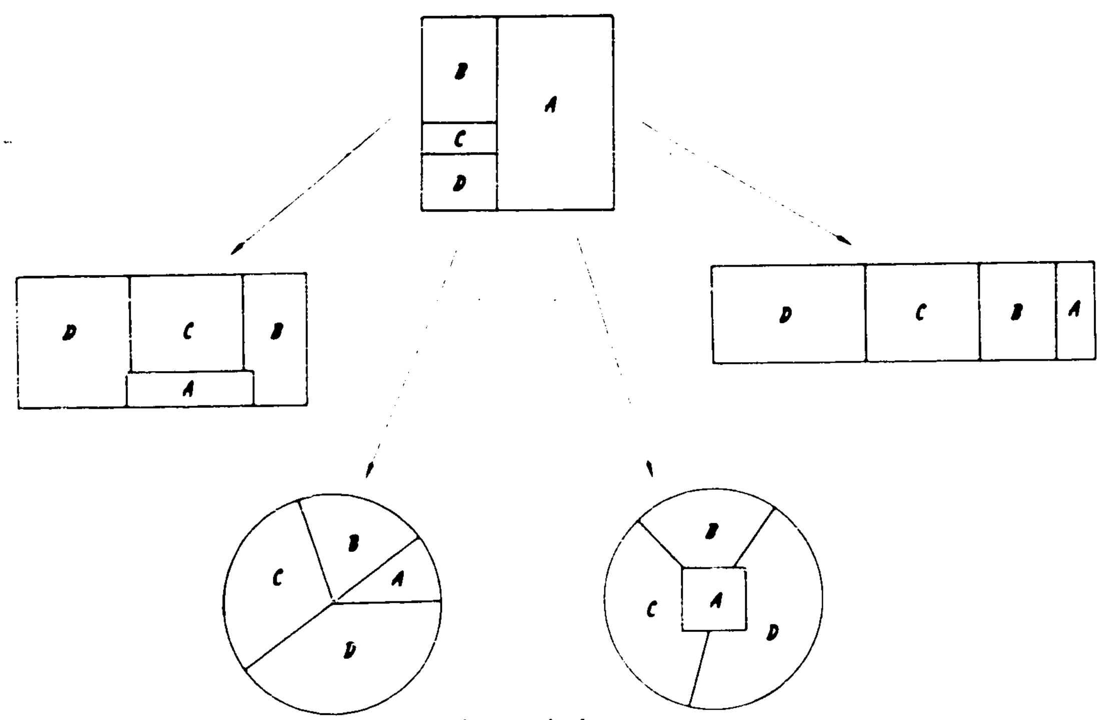


Figure 6.6

If we think in terms of a map of a portion of the earth's surface, an obvious difficulty is that the foregoing examples do not indicate where positions within the original areal units are to be placed within the corresponding partitions of the transformed image. Stated in another way, if positions in the original are described by latitude and longitude, where are the images of these lines in the transformed image? If the partitions represent states, the placement of cities is rather arbitrary, etc. Here the differentiable cases have a distinct advantage. If a coordinate system is introduced in the original, and an assumption of uniform density within each partition (e.g., states) is made, these difficulties can be avoided by estimating lines of equal increments of density on the original. These lines then correspond to an equal-area grid system on a plane (and the converse). A similar method can be employed when the original data are given in the form of a dot map. However, if a partition has no entries, the map area should vanish, a collapsing of space. The population cartogram previously presented (Fig. 6.3) actually consists of several domains, for otherwise ocean areas would be eliminated (lines of latitude and/or of longitude coincide), just as Greenland and Antartica do not appear on the map. Although there is some density of population in these areas, it is of such a negligible nature that the concept of density has little meaning.

The approximation methods are no less valid than the methods used in the differentiable cases; they could be formalized further, but are more akin to topological transformations than those associated with cartography. These transformations do not teach us anything which is not obvious from examination of the differentiable cases, however. More interesting, and more difficult to evaluate, are the geographic uses of maps obtained by the foregoing types of projections or transformations. These applications should also suggest the additional conditions to be applied in selecting a specific transformation from the infinite variety of particular solutions available.

Geographic Applications

Obviously, the maps obtained can be used as were the cartograms previously presented, for they were derived by consideration of such cartograms. These many applications need not be repeated. Further, any distribution plotted on a map using such a transformation shows a ratio, just as a dot map of population on an equal-area projection yields a visual impression of population density. Income symbolized on a map preserving population distribution shows per capita income, etc.

It is also clear that any grid system which partitions the area of the plane map into units of equal size will yield a partitioning of the basic data into regions containing an equal number of elements when mapped back to the original domain. This might be useful, for example, in partitioning states into electoral districts in such a manner that each district contains an equal number of voters, etc. The specific equal-area grids on a plane are infinite in number so that this procedure is not really of much assistance. Equal-area grids are also difficult to define along irregular boundaries, and partitionings (electoral districts, etc.) are usually required to satisfy numerous additional conditions (coincide with city and county boundaries, etc.).

More interesting applications can perhaps be found in the theories of von Thünen and Christaller. It is in this context that Harris and

Hoover attempted to use their cartograms. Von Thünen assumes a uniform fertility of agricultural land, and Christaller a uniform distribution of rural population or income, though both attempt to relax these unrealistic assumptions somewhat. If we postulate that agricultural fertility varies from place to place, i.e., that fertility can be described by a relation $F = f(\rho, \lambda)$, and then apply a transformation of the type described, areas of low fertility will appear compressed on the map and areas of above average fertility will appear enlarged. We then plot an even yield (e.g., in bushels) per unit of map area and, using the inverse transformation, return to the original domain. The even distribution of yields now will be quite uneven, and in fact corresponds to the distribution of fertility. This becomes more interesting if we add the condition that cost-distances from a market place appear as map distances from the center of the map[19] and that the intensity of use (yields) decreases with cost-distance. That is, on the map transformed so that all areas appear of equal fertility, yields are to be plotted as decreasing from the center of the map, as in the von Thünen model. The inverse transformation will then display a distribution of intensity of use which takes into account fertility and cost-distance from the market place.[20]

Just as the von Thünen model can be applied to cities, the foregoing discussion can be rephrased using "suitability-for-construction"

---

[19] See note 7, Chapter V. Isolines of distance from one point, if known, could certainly be drawn on the map when an analytic solution is possible.

[20] The measurement of agricultural fertility is by no means easy. Dunn (op. cit., pp. 67-69) doubts that such measurement can be achieved, but the U. S. Department of Agriculture publishes detailed maps containing a ranked classification of rural land based on its economic value.

instead of fertility. Many urban areas are already built up and construction is no longer feasible; other areas are blighted and have but little appeal; some locations have high prestige value; site and topographic factors vary, and so on. Undoubtedly, measurement of these values is difficult.[21] Nevertheless, the transformation and its inverse can be used as before. Such a transformation takes into account only two factors and is, hence, only of somewhat limited assistance in explaining the totality of urban land uses. The currently available models of urban structure are not outstandingly more successful, however.[22]

Christaller in his work on location theory assumes a uniform distribution of the underlying rural population and then obtains hexagonal service areas and a hierarchy of cities regularly spaced throughout the landscape. We have shown how a distribution may be made to appear uniformly distributed, and the pertinent question is whether Christaller's resulting pattern will now be observed. This would not appear to be the case, for several reasons. There is no guarantee that hexagons will be preserved in the transformation, as would be required. It is also not clear how such a condition might be added as a requirement to choose a particular transformation from the infinite set. Christaller obtains hexagons from consideration of circular service areas, and it is known that the stereographic projection sends circles into circles. The stereographic projection, however, will certainly not result from the density preserving transforma-

[21]Requirements for different classes of land use differ and some measure of intensity of use seems required. Land costs (e.g., in dollars) are biased, as they already reflect accessibility and an estimate of potential returns.

[22]See note 3, Chapter I.

tion in the general case. Conformal projections preserve circles as circles but only locally and would require satisfying both conditions of conformality and a specific areal distortion.[23] For relatively small service areas conformal transformations may be suitable. Christaller is also concerned with distances; his circular service areas are more akin to geodesic circles using a subjectively-valued time-cost distance, and his spacing of cities stipulates some distance between cities. Yet distances are not preserved by the transformations; preservation of all distances is certainly not possible if densities also are to be preserved on a plane map. We might, however, draw hexagons on a uniform density map, and using the inverse transformation examine the resulting pattern of curvilinear polygons on the original domain. Some such procedure is implied by Isard's diagrams of hypothetical landscapes.[24]

We can conclude that the attempt to apply the transformations suggested in this chapter to theories similar to those of von Thünen and Christaller is difficult and only partially successful, though promising and capable of improvement. The deficiencies are to a certain extent due to the inadequacies of the theories themselves, for, at present, they are not sufficiently general nor explicitly formulated.

As map projections, the transformations used in this chapter do not conform to the traditional geographic emphasis on the preservation of spherical surface area. Maps prepared using these transformations,

---

[23]Laborde's surface indicatrice may be of assistance in finding such projections (Driencourt and Laborde, op. cit., Vol. 4, p. 39, et seq. Also, see note 7, Chapter V.

[24]Isard, Location and Space Economy, Chapter XI, Figures 52, 53, and 54.

however, from many points of view are more realistic than the conventional maps used by geographers.

CHAPTER VII

OVERVIEW AND SUMMARY

It is now necessary to take stock and to review briefly what has been accomplished. A basic notion is that the measuring rod of the geodesist or surveyor is less relevant to social behavior in a spatial context than is a scaling of distances in temporal or monetary units. Hence, it is necessary to take into account not only the shape of the earth, but also the realities of transportation on this surface. Automobiles, trains, airplanes, and other media of transport can be considered to have the effect of modifying distances--measured in temporal or monetary units--in a complicated manner. Different distance relations, however, can be interpreted as different types of geometry. A geographically natural approach is to attempt to map this geometry to a plane, in a manner similar to the preparation of maps of the terrestrial sphere. The geometry with which we must deal is rarely Euclidean, and it is, in general, not possible to obtain completely isometric transformations. However, maps preserving distance from one point are easily achieved, whatever the units of measurement, and these have been discussed in some detail. The maps at first may appear strange, but this is only because we have a strong bias towards more traditional diagrams of our surroundings and we tend to regard conventional maps as being realistic or correct.

The geographer's most distinctive tool has long been the map, and the subject at hand has been approached from the point of view of map projections. For the present purpose, however, the cartographic literature on map projections, especially in English, is somewhat obsolete and deficient in generality. It has here been modified slightly to correspond somewhat more closely with the modern notion of transformations.[1] The parametric classification employed in the second chapter cannot be considered more than an ad hoc device which allows a very rapid overview, as it depends too strongly on the particular types of coordinates employed; the interest is certainly not in coordinates. This classification also ignores the important notion of invariance of properties. Similarly, though the transformations presented can be considered to be distortions of distance or area, this misses the point, for what is really important is that certain properties (time-distance, density, etc.) have been preserved. The graphic methods employed for projections will perhaps be useful in teaching. These methods are not used extensively in the cartographic literature, but are otherwise not uncommon. Their use is, of course, not necessary. The difficulties of analytic solutions, which are very real but have not been emphasized, also suggest the use of graphic techniques.

The requirement that time- or cost-distances be diagrammatically represented on a plane sheet is a handicap, for it allows preservation of distances from only select points. An attempt could have been made to represent transport cost relations between all locations mathematically, after the manner of differential geometry. This has not been attempted—though a local metric and curvature might be deduced from the isolines, at least

---

[1] This could be considerably improved. Map projections might better be treated by outlining the modern subject of transformations, then considering the special case in which the terrestrial surface is the referent object.

in the differentiable cases--for it is a highly complex geometry. The rather simple methods used here and the partial approach of considering transport distances from only one point and neglecting long run (temporal) changes in rates, however, have yielded geographically interesting results, for there are many instances in which relations to (or from) only one central location need be considered. This can also be regarded as a preliminary to a more detailed study.

The actual transport structures and characteristics are quite complicated, and are so for even more complicated reasons, which we have made no attempt to discuss. Transfer costs represent only a portion of the many factors in location theory, and they have perhaps been overemphasized. Minimization of transport costs may certainly present a case of sub-optimization. We have not attempted to explain traffic flows, nor to optimize the routing of shipments. Normative opinions and matters relating to desirable transportation policies also have been avoided.[2] Nevertheless, it seems appropriate to consider substitution between all feasible transport media, rather than to study one medium--say roads or railroads--in isolation, and to consider the total trip, not only that portion using a particular medium.

The possible combinations of media and routes between two terrestrial locations are usually infinite in number, and the most obvious procedure is to allow only those combinations which minimize some disutility or maximize some utility. The large majority of studies make such assumptions. This seems to lead immediately to methods of classical physics--

---

[2]See E. F. Penrose, "The Place of Transport in Economics and Political Geography," Transport and Communications Review, United Nations, V, 2 (June, 1952), 1-8.

going back as far as Fermat[3]--which has certain advantages but also disadvantages. The advantage, of course, is that we can use results which have accumulated over many years. The disadvantage is that extreme caution is required to avoid the implicit introduction of postulates which belong to physics and not to the subject at hand.[4] It can be argued that competition, social or economic, tends to require a minimization of costs, but this is really a matter to be postulated and then tested empirically. That travel takes place along a cost-distance geodesic assumes at least that all transport costs are known correctly, that geodesics have been found, that action is taken in accordance with the objective of minimizing costs, and that the user cannot affect the level of transport costs. Any discussion of this subject can become quite involved; we need only mention that considerations other than transport costs are often relevant, that perhaps geodesics are estimated only up to certain marginal levels, or that costs are subjectively evaluated.[5] Such rationalizations often seem designed simply to thwart any attempts at empirical evaluation. More reasonable postulates, perhaps

---

[3]Ca. 1657-1662. See J. Bernoulli, "On the Brachistochrone Problem," A Source Book in Mathematics, Vol. II, ed. D. E. Smith (New York: Dover, 1959), pp. 644-655. Bernoulli's synchrone and Huygens' wave front correspond to our isolines. Further references are given in W. Warntz, "Transatlantic Flights and Pressure Patterns," The Geographical Review, LI, 2 (April, 1961), 187-212.

[4]For similar comments by economists, see M. Beckmann, C. McGuire, and C. Winsten, Studies in the Economics of Transportation (New Haven: Yale University, 1956).

[5]Compare Christaller's subjectively-valued time-cost distances, Zipf's principle of least effort (G. K. Zipf, Human Behavior and the Principle of Least Effort; Cambridge: Addison-Wesley, 1949), or the thesis of cognitive behaviorism--and, for comments, A. Rapaport, "The Stochastic and Teleological Rationales of Certain Distributions and the so-called Principle of Least Effort," Behavioral Science, II, 2 (April, 1957), 147-161.

stochastic, seem required.

The foregoing considerations are necessary because the validity of any application of the transformations suggested here is dependent on the validity of the underlying hypotheses. This is true whether distance or area are the concepts involved. The pair of equations, per se, interpreted as transformations, are devoid of any particular connotations. On the other hand, the application of transformations to geographic purposes has hardly been exhausted. Imaginative reading of D'Arcy W. Thompson's On Growth and Form, for example, suggests many an approach to dynamic geographic phenomena through the use of transformations.[6] Many extensions and simplifications of the notion of a transformation have occurred during the previous 100 years. The majority of these materials have been ignored by the geographic-cartographic literature. For example, the many conformal transformations of a plane that are available can all be applied to obtain conformal maps of the surface of the earth by first using the stereographic (or other conformal) projection. The validity of the geographic uses of the transformations will depend on the accuracy of the underlying hypotheses, just as the validity of a map depends, in part, on the terrestrial model employed, whether this is a sphere, an ellipsoid of revolution, a tri-axial ellipsoid, etc. Many concepts in location theory (and geography in general) are today but loosely formulated; the concept of accessibility, for example. To map theories from a time-cost space to the surface of the earth is even

---

[6] D. W. Thompson, On Growth and Form (New York: MacMillan, 1948), especially Chapter XVII ("On the Theory of Transformations"). The point of view is that growth and evolution are but transformations of related forms. Also, N. Rashevsky, loc. cit. The employment of isolines of any type can be interpreted as the introduction of new coordinates on a sphere (or plane) and may suggest a transformation.

more difficult under these circumstances. A major objective of this study will have been accomplished if some of these problems have been brought into sharper focus.

BIBLIOGRAPHY

Books and Monographs

Abbott, E. A. Flatland. New York: Dover, 1952, 103 pp.

Adams, Oscar S. General Theory of Equivalent Projections. (U. S. Coast and Geodetic Survey Special Publication 236.) Washington: Government Printing Office, 1945, 74 pp.

Beckmann, M., McGuire, C. B., and Winsten, C. B. Studies in the Economics of Transportation. New Haven: Yale University Press, 1956, 152 pp.

Betz, J. Albert. Konforme Abbildung. Berlin: Springer, 1948, 359 pp.

Bieberbach, L. Conformal Mapping. Translated by F. Steinhardt. New York: Chelsea, 1953, 234 pp.

Bing, R. H. Elementary Point Set Topology. (H. E. Slaught Memorial Paper No. 8) Buffalo: Mathematical Association of America, 1960, 78 pp.

Blaser, L. Einführung in die Kartenlehre. (Mathematisch-Physikalische Bibliothek, Reihe I, Nr. 81.) Leipzig: Teubner, 1951, 64 pp.

Blumenthal, Leonard M. Theory and Application of Distance Geometry. Oxford: Clarendon Press, 1953, 347 pp.

Brooks, C., and Carruthers, N. Handbook of Statistical Methods in Meteorology. London: H. M. Stationery Office, 1953, 412 pp.

Bryson, R., and Kuhn, P. M. Half-Hemispheric 500 MB Topography Description by Means of Orthogonal Polynomials. (Report No. 4.) Madison: Department of Meteorology, University of Wisconsin, February, 1956, 20 pp.

Carroll, Jr., J. D., *et al.* *Final Report.* 3 vols. Chicago: Chicago Area Transportation Study, 1960, 252 pp.

Chamberlin, Wellman. *The Round Earth on Flat Paper.* Washington: National Geographic Society, 1947, 126 pp.

Daggett, Stuart. *Principles of Inland Transportation.* 4th ed. New York: Harper & Bros., 1955, 788 pp.

Deetz, C. H., and Adams, O. S. *Elements of Map Projection.* 5th ed. (U. S. Coast and Geodetic Survey Special Publication 68.) Washington: Government Printing Office, 1945, 226 pp.

Driencourt, L., and Laborde, J. *Traité des projections des Cartes Géographiques a l'Usage des Cartographes et des Géodesiens.* Paris: Herman, 1932, 394 pp.

Dunn, Jr., E. C. *The Location of Agricultural Production.* Gainesville: University of Florida, 1954, 76 pp.

Fair, M. L., and Williams, Jr., E. W. *Economics of Transportation.* New York: Harper & Bros., 1950, 757 pp.

Fels, Edwin. *Der Wirtschaftende Mensch als Gestalter der Erde.* Vol. V of *Erde und Weltwirtschaft.* Edited by R. Lütgens. Stuttgart: Franckh'sche Verlagshandlung, 1956, 321 pp.

Fisher, I., and Miller, O. M. *World Maps and Globes.* New York: Essential Books, 1944, 168 pp.

Freidrich, Carl J. *A. Weber's Theory of Location of Industries.* Chicago: University of Chicago, 1929, 256 pp.

Garrison, William L. *The Benefits of Rural Roads to Rural Property.* Olympia: Washington State Council for Highway Research, 1956, 107 pp.

Garrison, William L., and Marts, M. E. Geographic Impact of Highway Improvements. Seattle: University of Washington, 1958, 139 pp.

Graur, A. V. Matematischeskaya Kartografia. Leningrad: Leningrad University, 1956, 364 pp.

Hammer, Ernst. Ueber die Geographisch Wichtigsten Kartenprojektionen, Insbesondere die Zenitalen Entwürfe nebst Tafeln zur Verwandlung von Geographischen Koordinaten in Azimutale. Stuttgart: Metzler, 1889, 148 pp.

Hansen, H. (ed.). 1959 World Almanac. New York: New York World-Telegram, 1959, 896 pp.

Harrison, R. E. Look at the World. New York: A. A. Knopf, 1944, 67 pp.

Haurwitz, B., and Craig, R. A. Atmospheric Flow Patterns and Their Representation by Spherical-Surface Harmonics. (Geophysical Research Paper No. 14.) Cambridge: Cambridge Research Center, 1952, 81 pp.

Healy, Kent T. The Economics of Transportation in America. New York: Ronald Press, 1940, 575 pp.

Herz, Norbert. Lehrbuch der Landkartenprojektionen. Leipzig: Teubner, 1885, 312 pp.

Hoover, Edgar M. The Location of Economic Activity. New York: McGraw-Hill, 1948, 370 pp.

Illinois, State of. Atlas of Illinois Resources: Transportation. Urbana: University of Illinois, 1960, 50 pp.

Isard, Walter. Location and Space Economy. New York: J. Wiley, 1956, 350 pp.

_______. Methods of Regional Analysis. New York: M.I.T. and J. Wiley, 1960, 784 pp.

Kao, Richard C. Geometric Projections and Radar Data. Santa Monica: System Development Corp., 1959, 47 pp.

Kellaway, George P. Map Projections. New York: E. P. Dutton, 1946, 127 pp.

Klein, Felix. Elementary Mathematics from an Advanced Standpoint: Geometry. New York: Dover, 1939, 274 pp.

Knopp, K. Elements of the Theory of Functions. F. Bagemihl translation. New York: Dover, 1952, 140 pp.

Kreyszig, Erwin. Differential Geometry. Toronto: University of Toronto, 1959, 352 pp.

Levine, Daniel, et al. Radargrammetry. New York: McGraw-Hill, 1960, 330 pp.

Lösch, August. The Economics of Location. W. H. Woglom translation. New Haven: Yale University Press, 1954, 520 pp.

Locklin, Philip D. Economics of Transportation. 4th ed. Homewood: Irwin, 1954, 916 pp.

Lotka, Alfred J. Elements of Mathematical Biology. New York: Dover, 1956, 370 pp.

Maurer, H. Ebene Kugelbilder: Ein Linnesches System der Kartenentwürfe. (Ergänzungsheft 221 of Petermann's Geographische Mitteilungen.) Gotha: Justus Perthes, 1935, 87 pp.

Merkel, Heinrich. Grundzüge der Kartenprojektionslehre. Munich: Deutsche Geodätische Kommission, 1958, 120 pp.

Mohring, Herbert, et al. The Nature and Measurement of Highway Benefits: An Analytical Framework. Evanston: Northwestern University Transportation Center, 1960, 165 pp.

National Academy of Sciences. Conference on Transportation Research. (Publication 840.) Washington: National Research Council, 1960, 48 pp.

National Committee on Urban Transportation. Determining Travel Time. (Procedure Manual 3B.) Chicago: Public Administration Service, 1958, 24 pp.

Nehari, Zeev. Conformal Mapping. New York: McGraw-Hill, 1952. 396 pp.

Palander, Tord. Beiträge zur Standortstheorie. Uppsala: Almqvist and Wiksells, 1935, 419 pp.

Paullin, Charles O. Atlas of the Historical Geography of the United States. Edited by J. K. Wright. Washington: American Geographical Society and the Carnegie Institute, 1932, 162 pp.

Petterssen, S. Weather Analysis and Forecasting. Vol. II. New York: McGraw-Hill, 1956, 258 pp.

Ponsard, Claude. Economie et Espace. Paris: Sedes, 1955, 467 pp.

Raisz, Erwin. General Cartography. New York: McGraw-Hill, 1938, 370 pp.

Rashevsky, N. Mathematical Biophysics. 3rd ed., Vol. II. New York: Dover, 1960, 462 pp.

Ratzel, Friedrich. Politische Geographie. 3rd ed. Berlin: Oldenbourg, 1923, 421 pp.

Reichenbach, H. The Philosophy of Space and Time. M. Reichenbach and J. Freund translation. New York: Dover, 1958, 295 pp.

Reignier, Francois. Les Systèmes de Projection et leurs Applications. 2 vols. Paris: Institut Géographique National, 1957, 312 pp.

Robinson, Arthur. Elements of Cartography. 2d ed. New York: J. Wiley, 1960, 343 pp.

Struik, D. J. Lectures on Classical Differential Geometry. Reading: Addison-Wesley, 1950, 221 pp.

Thomas, Paul D. Conformal Projections in Geodesy and Cartography. (U. S. Coast and Geodetic Survey Special Publication 251.) Washington: Government Printing Office, 1952, 142 pp.

Thompson, D'Arcy Wentworth. On Growth and Form. New York: MacMillan, 1948, 1116 pp.

Thünen, Johann Heinrich von. Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. Jena: Gustav Fischer, 1826, 678 pp.

Tissot, M. A. Mémoire sur la Representation des Surfaces et les Projections des Cartes geographiques. Paris: Gautier-Villars, 1881.

Toschi, Umberto. Compendio di Geografia Economica Generale. Firenze: Macri, 1951, 314 pp.

Troxel, Emery. Economics of Transport. New York: Rinehart, 1955, 837 pp.

Wagner, Karlheinz. Kartographische Netzentwürfe. Leipzig: Bibliographisches Institut, 1949, 262 pp.

Weigert, H. W., et al. Principles of Political Geography. New York: Appleton-Century-Crofts, 1957, 723 pp.

Wheeler, B. O. Effect of Freeway Access upon Suburban Real Property Values. Seattle: Washington State Council for Highway Research, 1956, 120 pp.

Willmore, T. J. An Introduction to Differential Geometry. Oxford: Clarendon Press, 1959, 317 pp.

Woytinsky, W. S., and Woytinsky, E. S. World Population and Production. New York: Twentieth Century Fund, 1953, 1268 pp.

Zaustinsky, E. M. Spaces With Non-Symmetric Distance. (Memoir 34.) Providence: American Mathematical Society, 1959, 91 pp.

Zimmerman, Erich W. World Resources and Industries. New York: Harper & Bros., 1951, 833 pp.

Zipf, George K. Human Behavior and the Principle of Least Effort. Cambridge: Addison-Wesley, 1949, 573 pp.

Zöppritz, K., and Bludau, A. Leitfaden der Kartenentwurfslehre. Berlin: Teubner, 1912, 378 pp.

Articles and Periodicals

Ajo, R. "New Aspects of Geographic and Social Patterns of Net Migration Rate," Migration in Sweden, (Lund Studies in Geography), Ser. B, 13 (1957), 170-183.

Alexander, J. W. "Freight Rates as a Geographic Factor in Illinois," Economic Geography, XX (1944), 25-30.

Alexander, J. W., Brown, S. E., and Dahlberg, R. E. "Freight Rates: Selected Aspects of Uniform and Nodal Regions," Economic Geography, XXXIV, 1 (January, 1958), 1-18.

Barrow, C. H. "Very Accurate Correction of Aerial Photographs for the Effects of Atmospheric Refraction and Earth's Curvature," Photogrammetric Engineering, XXVI, 5 (December, 1960), 798-804.

Beckmann, M. J. "A Continuous Model of Transportation," Econometrica, XX, 4 (October, 1952), 643-660.

Bernoulli, J. "On the Brachistochrone Problem," in A Source Book in Mathematics, Vol. II, D. E. Smith, ed., New York: Dover, 1959, pp. 644-655.

Boggs, S. W. "Mapping Some Effects of Science on Human Relations," The Scientific Monthly, LXI (July, 1945), 45-50.

________. "Mapping the Changing World: Suggested Developments in Maps," Annals (Association of American Geographers), XXXI, 2 (June, 1941), 119-128.

Brown, B. H. "Conformal and Equiareal World Maps," American Mathematical Monthly, XLII (April, 1935), 212-223.

Brush, John E. "The Hierarchy of Central Places in Southwestern Wisconsin," The Geographical Review, XLIII, 3 (July, 1953), 380-402.

Cahill, B. J. S. "Static and Dynamic World Maps," Bulletin (American Meteorological Society), XV (1934), 251-255.

Carrothers, Gerald A. "An Historical Review of the Gravity and Potential Concepts of Human Interaction," Journal (American Institute of Planners), XXII (1956), 25-30.

Chabot, G. "La Determination des Courbes Isochrones en Geographie Urbaine," Comptes Rendus du Congres International de Geographie (Amsterdam), II, 3a (1938), 110-113.

Dameron, Leslie Y. "Terminal Area Charts for Jet Aircraft," The Military Engineer (May, 1960), p. 6.

________. "Projection for Lunar Map," The Military Engineer (January, 1960), p. 5.

Davies, William E. "Axis War Maps," Surveying and Mapping, VIII, 3 (July, 1949), 126-134.

Desoutter, D. M. "Projection by Introspection," Aeronautics, XL, 2 (April, 1959), 42-44.

Dodd, Stuart C. "The Interactance Hypothesis: A Gravity Model Fitting Physical Masses and Human Groups," American Sociological Review, XV (April, 1950), 245-256.

Dussart, F. "Les Courbes isochrones de la ville de Liège pour 1958-1959," Bulletin (Societe Belge d'Etudes Géographiques), XXVIII, 1 (1959), 59-68.

Eckert, Max. "Eine Neue Isochronenkarte der Erde," Petermann's Geographische Mitteilungen, LV, 9, 10 (1909), 209-216, 256-263.

Garrison, William L. "The Spatial Structure of the Economy," Annals, (Association of American Geographers), XLIX, 2, 4 (1959), 232-239, 471-482; L, 3 (1960), 357-373.

________. "Connectivity of the Interstate Highway System," Proceedings (Regional Science Association), VI (1960), 121-137.

Gilfallan, S. C. "World Projections for the Air Age," Surveying and Mapping, VI, 1 (January, 1946), 12-18.

Goode, J. Paul. "Studies in Projection: Adapting the Homalographic Projection to the Portrayal of the Earth's Surface Entire," Bulletin (Geographical Society of Philadelphia), XVII (1919), 103-113.

Hägerstrand, Torsten. "Migration and Area," Migration in Sweden (Lund Studies in Geography), Ser. B, 13 (1957), 27-158.

Hansen, Walter G. "How Accessibility Shapes Land Use," Journal (American Institute of Planners), XXV, 2 (May, 1959), 73-76.

Harris, Chauncy D. "The Market as a Factor in the Localization of Industry in the United States," Annals (Association of American Geographers), XLIV, 4 (December, 1954), 313-348.

Harris, Chauncy D., and McDowell, George B. "Distorted Maps - A Teaching Device," Journal of Geography, LIV, 6 (September, 1955), 286-289.

Harris, Chauncy D., and Ullman, Edward L. "The Nature of Cities," in Readings in Urban Geography, H. Mayer and C. Kohn, eds., Chicago: University of Chicago Press, 1959, pp. 277-286.

Hinks, Arthur R. "A Retro-Azimuthal Equidistant Projection of the Whole Sphere," The Geographical Journal, LXXIII, 3 (March, 1929), 245-247.

Hinks, Arthur R., et al. "A New Treatise on Map Projections," The Geographical Journal, LXXXIII (1934), 145-150.

Horn, Werner. "Die Geschichte der Isarithmenkarten," Petermann's Geographische Mitteilungen, CIII, 3 (1959), 225-232.

Hsieh, Chiao-Min. "The Status of Geography in Communist China," The Geographical Review, XLIX (October, 1959), 535-551.

Jackson, W. A. D. "Whither Political Geography," Annals (Association of American Geographers), XLVIII, 2 (June, 1958), 178-183.

Jefferson, Mark. "The Six-Six World Map," Annals (Association of American Geographers), XX, 1 (March, 1930), 1-6.

Jones, Stephen B. "A Unified Field Theory of Political Geography," Annals (Association of American Geographers), XLIV, 2 (June, 1954), 111-123.

Kagami, Kanji. "The Distribution Map by the Method of Aeroview," The Geographical Review of Japan, XXVI, 10 (October, 1953), 463-468.

Kant, E. "Umland Studies and Sector Analysis," Studies in Rural-Urban Interaction (Lund Studies in Geography), Ser. B, 3 (1951), 1-10.

Keunig, Johannes. "The History of Geographical Map Projections Until 1600," Imago Mundi, XII (1955), 1-24.

Leighly, John B. "Aitoff and Hammer: An Attempt at Clarification," The Geographical Review, XLV, 2 (1955), 240-249.

Lindberg, Olof. "An Economic-Geographical Study of the Localization of the Swedish Paper Industry," Geografiska Annaler, XXXV (1953), 27-40.

Mackay, J. Ross. "The Interactance Hypothesis and Boundaries in Canada: A Preliminary Study," The Canadian Geographer, 11 (1958), 1-8.

________. "Conformality: Mathematical and Visual," The Professional Geographer (N.S.), X, 5 (September, 1958), 12-13.

________. "Geographic Cartography," The Canadian Geographer, 4 (1954), 1-14.

Maling, D. "A Review of Some Russian Map Projections," Empire Survey Review, XV, 115, 116, 117 (1960), 203-215, 255-266, 294-303.

Merkel, Heinrich. "Die Allgemeine perspektivische Abbildung der Erdkugel," Festschrift Eduard Dolezal (Sonderheft 14 of the Oesterreichischen Zeitschrift für Vermessungswesen), 1952, pp. 169-182.

Michels, F. W. "Drie Nieuwe Kaartvormen," Tijdschrift (Koninklijk Nederlandsch Aarorijkskundig Genootschap), LXXVI, 2 (1959), 203-209.

Miller, O. M. "A New Conformal Projection for Europe and Africa," The Geographical Review, XLIII, 3 (1953), 405-409.

Morse, M. "Fields of Geodesics issuing from a Point," Proceedings (National Academy of Sciences), XLVI, 1 (January, 1960), 105-111.

Penrose, E. F. "The Place of Transport in Economic and Political Geography," Transport and Communications Review (United Nations), V, 2 (April, 1952), 1-8.

Prey, Adelbert. "Darstellung der Höhen- und Tiefenverhältnisse der Erde durch eine Entwicklung nach Kugelfunktionen bis zur 16. Ordnung," Abhandlungen (Königliche Gesellschaft der Wissenschaften zu Göttingen; Mathematische Physikalische Klasse, Neue Folge), XI, 1 (1922), 1-42.

Raisz, Erwin. "Orthoapsoidal World Maps," The Geographical Review, XXXIII (1943), 132-134.

________. "The Rectangular Statistical Cartogram," The Geographical Review, XXIV, 2 (April, 1934), 292-296.

_______. "Rectangular Statistical Cartograms of the World," Journal of Geography, XXXV (1936), 8-10.

Rapaport, Anatol. "The Stochastic and the 'Teleological' Rationales of Certain Distributions and the So-Called Principle of Least Effort," Behavioral Science, II, 2 (April, 1957), 147-161.

"Recent Developments in Map Projections," The Geographical Review, XIX (1929), 686-688.

Riedel, Johann. "Neue Studien über Isochronenkarten," Petermann's Geographische Mitteilungen, LVII, 1 (1911), 281-284.

Robinson, Arthur. "Interrupting a Map Projection: A Partial Analysis of its Value," Annals (Association of American Geographers), XLIII, 3 (1953), 216-225.

Schjerning, W. "Ueber Mittabstandstreue Karten," Abhandlungen (K. K. Geographische Gesellschaft von Wien), V, 4 (1909).

Schmid, C. F., and MacCannell, E. A. "Basic Problems, Techniques, and Theory of Isopleth Mapping," Journal (American Statistical Association), L (1955), 220-239.

Schneider, Morton. "Gravity Models and Trip Distribution Theory," Proceedings (Regional Science Association), V (1959), 51-5[illegible].

Sprout, Harold, and Sprout, Margaret. "Environmental Factors in the Study of International Politics," Journal of Conflict Resolution, I, 4 (1957), 309-328.

Stewart, John Q. "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population," The Geographical Review, XXXVII, 3 (July, 1947), 461-485.

_______. "The Use and Abuse of Map Projections," The Geographical Review, XXX (1943), 589-604.

Stouffer, Samuel. "Intervening Opportunities: A Theory Relating Mobility to Distances," American Sociological Review, XV (1940), 845-867.

Tobler, Waldo R. "Automation and Cartography," The Geographical Review, XLIX, 4 (October, 1959), 526-534.

Ullman, Edward L. "The Role of Transportation and the Bases for Interaction," in Man's Role in Changing the Face of the Earth, William L. Thomas, Jr., ed. Chicago: University of Chicago Press, 1956, pp. 862-880.

Warntz, William. "Transportation, Social Physics, and the Law of Refraction," The Professional Geographer (N.S.), IX (1957), 2-7.

_______. "Transatlantic Flights and Pressure Patterns," The Geographical Review, LI, 2 (April, 1961), 187-212.

Watson, J. W. "Geography - A Discipline in Distance," The Scottish Geographical Magazine, LXXI, 1 (1955), 1-13.

Whittlesey, Derwent. "The Regional Concept and the Regional Method," in American Geography: Inventory and Prospect, P. E. James and C. F. Jones, eds. Syracuse: Association of American Geographers, 1954, pp. 19-69.

Wright, John K. "The Terminology of Certain Map Symbols," The Geographical Review, XXXIV (1944), 653-654.

Young, A. E. "Conformal Map Projections," The Geographical Journal, LXXVI (October, 1930), 348-351.

Unpublished Material

Alonso, William. "A Model of the Urban Land Market: Locations and Densities of Dwellings and Businesses." Unpublished Ph.D. dissertation, University of Pennsylvania, 1960, 2?6 pp.

Baskin, Carlisle W. "A Critique and Translation of Walter Christaller's Die Zentralen Orte in Suddeutschland." Unpublished Ph.D. dissertation, University of Virginia, 1957, 458 pp.

Bunge, Jr., William W. "Theoretical Geography." Unpublished Ph.D. dissertation, University of Washington, 1960, 189 pp.

Hagen, Carlos B. "The Azimuthal Equidistant Projection." Unpublished Master's thesis, Department of Geography, University of Washington, 1957, 280 pp.

Heald, K. L. "Discussion of the Iowa Gravity Model Traffic Distribution Program." Ames: Iowa State Highway Commission, 1960, 15 pp. (Mimeographed).

Horwood, Edgar M. "A Three-Dimensional Calculus Model of Urban Settlement." Paper presented at the Regional Science Association Symposium, Stockholm, August, 1960.

Vadnal, A. "La Localisation du Reseau de Routes Collectrices Dans Le Cercle." Paper presented at the Congres European D'Econometrie, Bilbao, 1958.

Wiley, Stanley R. "The Effect of Limited Access Highways upon Suburban Land Use." Unpublished Master's thesis, Department of Civil Engineering, University of Washington, 1958, 146 pp.

Wingo, Lowdon. "Transportation and the Utilization of Urban Land: A Theoretical Exploration." New York: Twentieth Century Fund, 193 pp. (Unpublished manuscript).

VITA

Waldo Rudolph Tobler, son of Hanny and Verner Tobler, was born in Portland, Oregon, on November 16th, 1930. He attended elementary schools in Seattle, Washington, and Bethesda, Maryland, and attended high schools in Switzerland. His undergraduate education was obtained at the University of British Columbia and at the University of Washington, where he was granted a Bachelor of Arts degree in 1955. Mr. Tobler received a Master of Arts degree from the University of Washington in 1957. The title of Doctor of Philosophy was conferred on Mr. Tobler by this same institution in 1961.