ICA Chapter

Cartograms as map projections

W. Tobler
University of California
Santa Barbara, USA

Since the time of Ptolemy 2000 years ago the objective of map projections has been to produce maps of high metric fidelity, and this is still the prime objective today. Thus the emphasis on direction, distance, and area preservation. These properties are needed in order to serve a variety of purposes. This way of looking at maps considers them as a type of terrestrial graph paper. This is possible because, paraphrasing John Donne, we use "the net of Meridians and Parallels cast over the world". Different kinds of geographic graph papers are used for different kinds of problems. Thus there are the classic conformal, equidistant, and equal area projections. Particular solutions are provided by the conformal projections, such as the Mercator projection, or the stereographic. Other maps such as the azimuthal projections, including the gnomonic projection, show directions. Equal area projections are most often used for larger regions or countries and for the depiction of statistical information. There are also quite a large variety of lesser known properties for which special projections are used. Still other projections serve for general purpose maps.

Another type of map is referred to as a cartogram. It is the intent of this note to explain how these relate to convention maps of the two dimensional surface of the earth. But these cartograms are often not considered as belonging to the class of map projections. One difference between these and, say topographic maps, is that the phenomena depicted may change more rapidly in time, possibly often even hourly. This suggests that animation is a proper domain for cartograms. However not all history is quite this rapid and these maps can also be useful even if they are based on census information for which the change may be noted only every decade or so. In another respect cartograms have a property in common with the Mediterranean portolan charts of the thirteen to fifteen hundreds in that they are based on empirical observations rather than strictly spherical or ellipsoidal geometrical considerations.

Geographic Graphs:

This notion of map projections as graph paper for spheres can be extended to serve additional, non-traditional uses. A considerable number of these are provided by the class of cartograms, taken to be a special type of map projection. These can be classified in at least two ways. One grouping might be by type of problem or purpose. One common use is simply to present a point of view. This can be as simple as the ego-enhancing "Here's a Representation of My Favorite Region"; these are often intended to be humorous and colorful, and sometimes appear on post cards. Many of these can be referred to as 'Fisheye' maps (Rase, 1997). One more serious recoded use is to aid pilots by enlarging the vicinity of an airport, with a kind of local bubble enlargement. Or they can depict the state of the world from an alternate point of view, as in "The Atlas of the Real World" (Figure 1).

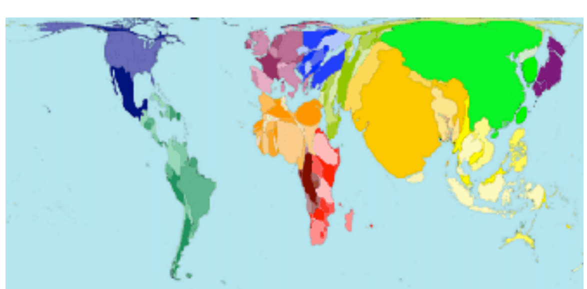

Fig 1: Mathematically computed equal world population projection. Dorling et al, 2007

There is also the problem-solving point of view in which the cartogram, and its inverse, are a way of using an unusual coordinate system that renders a situation more understandable and simpler to manage. In many fields the choice of a proper coordinate system often clarifies a situation. An example of this is to render the earth on a map in geomagnetic, rather than geographic, latitude and longitude coordinates in dealing with terrestrial magnetic problems. Today this might be done inside of a computer, switching between projections as needed, displaying only the final result. This problem-specific group would put the distorted Mercator projection into the category of a warping cartogram that assists in ocean navigation. Cartograms based on movement also seem possible but are rare.

Another approach is to group cartograms into the type of properties that they preserve. In such a classification one might find topological maps in which incidence and adjacency are preserved, but distance, area, and angles are not. The classic case is the well known London Underground map, a style now copied throughout the world. The usefulness of this depiction can hardly be questioned. Another category would be based on the adjustment of distances, using time or cost of travel as the metric instead of kilometers. Typically these cartograms resemble the azimuthal class of map projection in that they are based on a single center. Of course they are not geometric projections, but rather are based on empirical information. A third type of cartogram is related to area. Here the sizes of regions or countries are expanded, or shrunken, according to some numerical quantity, most often population magnitude, but any non-negative measure can be used. This classification of cartograms into three general types seems to cover most examples.

Topological Transforms:

All map projections attempt to preserve neighbor and adjacency properties, and mostly succeed except at the edges of the map. But if this is all that is preserved the maps may appear distorted. "Mental maps" wherein individuals are asked to draw maps of their neighborhoods, or of the world, often preserve these properties but latitude and longitude lines drawn on such maps appear as squiggly lines. The metric properties are not preserved, but sketch maps are still useful! I'm certain that you have made or used them. Comparing the London Underground map from 1910 (at www.ltmuseum.co.uk), metrically correct, with Beck's severely warped map of 1933 (and the modern derivatives in use today) clearly shows what can be done in this respect. The current underground map has also been analyzed for its metric properties using Tissot's results by Jenny (2006). Such analyses are not often performed for cartograms but one can recognize and indicate the angular, area, or distance properties if these are considered relevant. It might be of interest to minimize the distortion of these additional properties in the cartogram since there are always degrees of freedom in the choice of how to represent a specific cartogram – see the equations below. This also applies to the next category.

Distance Transforms:

Constructing maps from measured distances is a well understood problem in surveying known as trilateration. It has been generalized in psychology and there goes under the name of multidimensional scaling. An example is children ranking how much they enjoy playing with other children. Well-liking is considered "closeness", that is, a small distance. In the computer programs relations such as this (the children's locations relative to each other) are converted to coordinates to yield a 'map' of the classroom. The same computer procedures can be used to make geographic maps of travel time or cost, or some other metric. As one example several studies have shown that the interaction between different linguistic groups in bilingual countries, such as Belgium and Canada, are not the same. They are quite 'apart' from each other, and this can be depicted in map form.

These maps seem to be of two types. In one case only isolated locations, for example, airports with travel cost (or time) taken from a schedule, are indicated and the intermediate places are set down between these. An example is Barrett's map of "The world based on airfares from London". The intermediate places are really just arbitrarily inserted and thus misrepresented, and the map needs to be considered disconnected, and really as showing only discreet places. For display in these maps directions from a center are usually retained, i.e., azimuthal. This type of map can be extended, using multidimensional scaling, to represent time, or cost, etc., distance between two places, and then to all places (not only from one center), even using kilometer distances, but the distances will only be approximated, sometimes poorly. This also be done with sociological distances, for example such as differences between languages, customs, gross domestic products, or any differentiating measure. An index of the degree of the fidelity of this representation is available as the 'stress', and this is an overall measure of fit; therefore it is somewhat of an improvement on Tissot's index which measures only local distortion.

The second type is usually implicitly based on continuous contours (isochrones) of travel time or cost from (or to) a location. Even in this case some interpolation is required, but an attempt is made to provide a spatially continuous representation. It could be based on any type of contour map; including, for example, population density, but this is rare. Obviously in these cases there generally results in different distances in different directions – even though azimuths are preserved! This is quite unusual in conventional projections. Isochrones are often amoeba like in shape and even with disjoint pieces. An analogy can be made to pole centered azimuthal projections, with the irregular isochrones (or isotims) corresponding to parallels and the orthogonal trajectories (gradients) similar to meridians emanating from the pole. Directions and angles are clearly not preserved. And the maps generally relate to (or from) only one place, and at only one instance in time.

A related analytical use has been made by Hāgerstrand (1957) of logarithmically-scaled azimuthal maps of migration (Figure 2) in order, as he puts it, to be able to count the symbols indicating the coming and going of people. Most of these individual movements are crowded about the origins and destinations. In other words, this is the resultant effect of the well known distance decay in human affairs. Obviously this logarithmic projection is a proper map to represent the phenomena; angular and area distortions can be calculated. The local area is of course enlarged, and this leads directly to the next category.

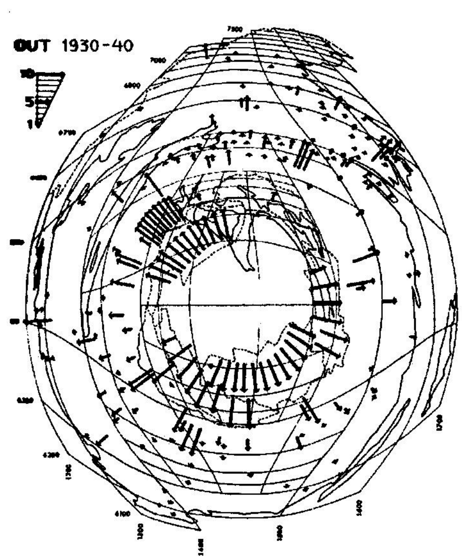

Fig. 40. Out-migration from Asby from 1930 up to 1940.

Figure 2. Logarithmic map of Asby, Sweden, used to study migration. Hägerstrand, 1957.

Area Transforms:

Area cartograms, with region sizes made proportional to some measure such as population or wealth, have recently become common now that a free and fast computer algorithm has become available (Gastner & Newman, 2004). Typical cartography texts include a smattering of such cartograms as examples. But these are generally not treated as map projections but rather as graphical or pictorial illustrations. And they often appear in newspapers, and even as exercises in lower school classes, as can be confirmed by entering the term 'cartogram' into a search engine on the internet. These maps offer a different view of the world and can be based on a variety of topics. (Tobler, 1996, Dorling et al, 2008; Hennig 2013).

The history of these cartograms goes back to the blossoming of statistical graphics in the mid eighteen hundreds (Tobler, 1996), and this is reflected in the terminology applied to these maps. They have been referred to as anamorphoses (France), verzerte Karten (Germany), varivalent maps (Russia), and value-by area maps. Many of the last type of these were prepared by E. Raisz who, starting in the 1930's, presented rectangular depiction of regions and countries with sizes proportional to a variety of phenomena. They have also appeared in other publications as 'statistical maps'. Recently this 'rectangular' type of cartogram has been perfected by researchers in the Netherlands (Van Kreveld & Speckmann, 2007) who have shown that they can be produced automatically by computer (Figure 3). The geographic graticule on such a map projection might have kinks in it and not be smooth, even though the populations are correct.

Figure 3. Mathematically computed equal world population projection. Van Kreveld and Speckmann, 2007

An additional use that has been made of area cartograms is for statistical purposes. When doing area sampling one likes to know that all subjects have an equally likely chance of being chosen. An area cartogram stretches (or shrinks) space and can warp geography so that areas are proportional to the target density (Figure 4). After sampling in this domain - attacking the problem - the inverse is used to apply the solution. This is quite similar to the application and warping of the world by the Mercator projection in connection with the gnomonic projection in ocean navigation. Thus the ‘Transform – Solve – Invert’ procedure is conceptually identical to what is done when using the Mercator projection.

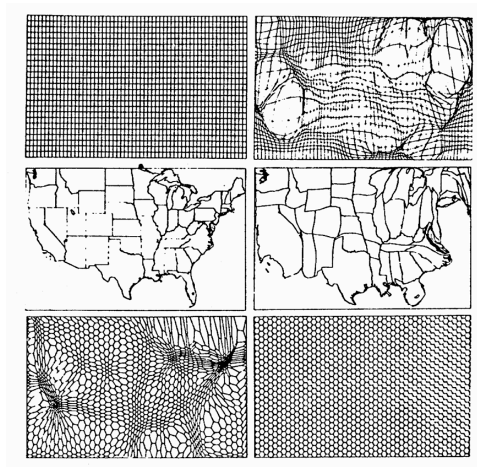

Figure 4: The Transform - Solve - Invert paradigm. Converting the USA to a population density space, overlaying a hexagonal pattern, and then taking the inverse to partition the country into approximately equal population regions. Tobler, 1963

Networks:

When modern transportation situations are considered it is necessary to violate the usual continuous topology. From Los Angeles to New York now costs less than to many intermediate places. Thus the map may need to be turned inside out. Many such conditions occur. Bunge has suggested balloons with the ends of strings glued on to constrain some places to be close together when the balloon is inflated, with the other places bulging out. This is the spherical real world model to be mapped. In another example, parcel postage rates are a step function of distance, as are many other transportation costs. Thus places are lumped together. If one draws postal costs as a function of geographic distance it looks like several places

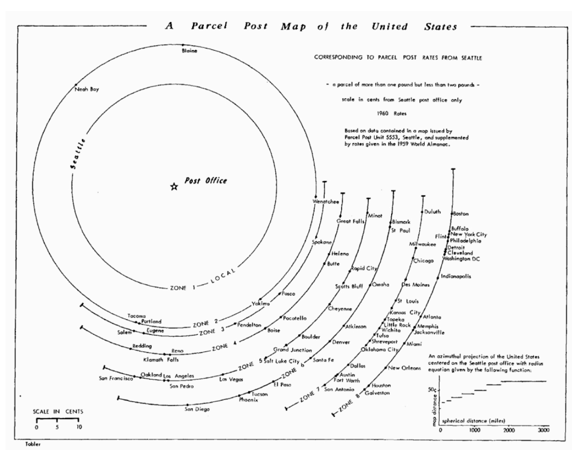

Figure 5: The postal rates are a step function, seen in the lower right. This results in the collapsing of the continuous space into a discrete set of distances. In this particular rendition directions (azimuths) from the Seattle post office are maintained. Tobler 1961

coincide (Figure 5). In the time of the Roman Empire travel by water was more rapid than travel by land. Thus Britain was closer to Rome than was Vienna. Even central Spain was further than Britain. The resulting isochrones are rendered graphically in the Roman History project (orbis.stanford.edu) and depict this, but with inverted positions.

The geographic details in these situations are often best represented as a network. Positioning of the network graphically in a two dimensional diagram must attempt to maintain the measurements and this often distorts the conventional geography. The placement of the isolated locations is possible, most often using a form of multidimensional scaling, but if the geographic coordinates (latitude and longitude) of these locations are known then interpolation between them can be said to be 'difficult'. In a classification of the types of situations that occur one must recognize inversions (geographically far places are closer than conventionally near places), interruptions (nearby places are not together), and many other violations of the traditional assumptions. This is the case even if distances are considered from only one location. But how badly are these maps distorted? Can this disruption be measured? By Tissot's indicatrix or by the stress? Is it possible to detect (that is, calculate) how much a high speed road or railroad system distorts a country? How warped has it become due to an expressway? Perhaps it has been turned almost inside out, with some places very close and others relatively further away! What did the Romans think (metrically) of the shape of the world when land travel was so difficult? These kinds of relations are quickly visible on cartograms. But parts of the maps may overlap, which most (but not all) classic map projections prevent. Another difficulty is that geographic travel times (or costs, etc.) are not symmetric: from A to B is not the same as from B to A. Cartographers most now take into account these global realities and this often requires ingenious cartograms.

The Equations:

Most books on map projections do not consider cartograms. This in spite of the fact that it can be shown, by writing out the equations (see below), that area cartograms are a generalization of equal area map projections; the earth's surface area is the particular measure preserved on an equal area projection. Stretching by population is just a different choice of property to be preserved. And many of the distance oriented cartograms are equivalent to azimuthal projections. A reason that such maps are not frequently discussed in the projection literature might be that they are of more recent origin, and that they are generally produced by individuals rather than government sponsored national mapping agencies. They do appear in some atlases, along with thematic maps. In this respect they are somewhat like the retroazimuthal projections that have very restricted usage. It is certainly the case that Tissot's measures of angular, area, and distance properties can be calculated for these types of maps. Measures can also be devised that indicate whether or not the resulting maps actually match the design objective – i.e., fit what they are intended to show. This may mean going a bit beyond the classical indices. At the moment there do not appear to be evaluation standards for cartograms, at least there is no consensus on this point. The large variety of possible uses makes this difficult.

There are really no books, and few scholarly articles, that consider cartograms as their main subject. The exception is for cartograms that modify areas according to some measure. Here a few recent books can be cited, along with those already mentioned. In dealing with cartograms using metrogenic substitutions (cost or time, etc,) instead of kilometers and based on a single center, the modifications can be evaluated as are azimuthal projections.

Appendix: Equations for an area cartogram

The equal area condition for a map projection in spherical (φ, λ) and plane rectangular coordinates (x, y) is:

$$\partial\varphi/\partial x \;\; \partial\lambda/\partial y \; - \; \partial\lambda/\partial x \;\; \partial\varphi/\partial y \; = \; R^2 \cos(\varphi).$$

The condition equation for a cartogram is:

$$\partial\varphi/\partial x \;\; \partial\lambda/\partial y \; - \; \partial\lambda/\partial x \;\; \partial\varphi/\partial y \; = \; R^2\, D(\varphi, \lambda) \cos(\varphi),$$

where D(φ, λ) is the density distribution on the earth, considered spherical.

Clearly, when the density distribution is constant (unity), the cartogram becomes an equal area projection. There are many solutions to both of the foregoing partial derivative equations. In each case the one condition does not suffice to yield the two equations necessary [x = f (φ, λ), y = g (φ, λ)] to completely define a map projection. Thus some other criteria is applied. The obvious second condition is to require that the angular distortion be minimized, making the image more recognizable. But other conditions, for example rectangular shape or symmetry about the equator, are often used for equal terrestrial area maps. The rectangular Value-by-Area maps of Erwin Raisz come to mind as alternatives.

The inverse of a cartogram of this type is to be found in the usual manner using the Jacobian defined by the partial derivatives, as described in books on advanced calculus, e.g., Kaplan (1952; 96-100). If the defining equations are not given explicitly then a two dimensional finite difference empirical iteration and interpolation needs to be used.

References:

F. Barrett, 1992, "The World based on airfares from London, *The Independent*, 9 Feb, page 6, London.

W, Bunge, `1966, *Theoretical Geography*, 2nd edition, Lund University, Gleerup.

C. Dobbin, 2013, *London Underground Maps*, London.

D. Dorling, et al., 2008, *Atlas of the Real World*, London, Thames & Hudson.www.worldmapper.com

M. Gastner, Newman, M, 2004, "Diffusion-Based method for producing density-equalizing maps", *Proc*., National Academy of Sciences of the United States of America, 101(20):7499-7504.

T. Hāgerstrand, 1957, "Migration and Area", in *Migration in Sweden*, D. Hannerberg, ed., University of Lund, Sweden.

B. Hennig, 2013, *Rediscovering the World: Map Transformations of Human and Physical Space*, Springer, Berlin.

B. Jenny, 2006, "Geometric distortion of schematic network maps", *Bulletin*, Society of Cartographers, 40:15-18.

W. Kaplan, 1952, *Advanced Calculus*, Addison-Wesley, Reading.

M. Van Kreveld, Speckmann, B., 2007, "On Rectangular Cartograms", *Computational Geometry: Theory & Applications*, 37(3):175-187.

E. Raisz, 1934, “The Rectangular Statistical Cartogram” *Geogr. Rev.*, 24: 292-296.

Rase, W., 1997, “Fischauge-Projektionen als kartographische Lupen“ (Fish-eye projections as cartographic lenses), *Salzburger Geographische Materialien*, Heft 25:115-122.

W. Tobler, 1961, "Map Transformations of Geographic Space", Ph.D. dissertation, Department of Geography, University of Washington, Seattle.

W. Tobler, 1973, "A Continuous Transform useful for Districting", *Annals*, New York Academy of Sciences, Vol. 219: 215-220.

W. Tobler, 1996, "Thirty Five Years of Computer Cartograms", *Annals*, Association of American geographers, 94 (1):58-73.