# THE GEOMETRY OF POSSIBLE/IMPOSSIBLE MAPS

Jean-Claude Muller
and
John L. Honsaker
Department of Geography
University of Alberta
Edmonton, Alberta, T6G 2H4

Points separated in space are ordinarily put into relation with one another by the Euclidean distance - measured by kilometers, miles, rods, cubits, or other such units. It may be meaningful in some situations to relate them by other quantities, such as: the time needed to travel between them, the number of messages passing between them, etc. Many conceptual surfaces are overlying the physical surface of the earth and we are reminded that "the role of distance as a dimension of society is to be judged not in physical units of length alone but rather in terms of cost distances, time distances and the like."[1] With proper definition, the quantification of such relational 'distances' may be easy. However, the procedure for presenting the information on a map is by no means as straightforward as for the ordinary map distances.

Imagine a triad of points M, N, O whose time-distances would satisfy the triangular inequality $d(M,N) + d(N,O) < d(M,O)$. The pairs of points (M,N) and (N,O) are connected by a superhighway whereas a footpath is connecting M and O. What is the mapping of the points M, N, and O? Obviously, the geodesics joining the image points cannot be represented by straight lines on the time-surface. Take four points M, N, O and P whose distance $d(M,N)$, $d(M,O)$, $d(M,P)$, $d(N,O)$, $d(N,P)$, and $d(O,P)$, are equal. It is impossible to map those points in two dimensions such that Euclidean distances are all equal.[2]

The cartography of cultural or mental distances would raise similar problems. Many "non-visible" components of the geographic space would not be understood, however, if they were not "materialized" in some kind of map which would connect them to the physical and human aspects of the geographical environment. There are few cartographic theories which deal specifically with the representation of distance-spaces measured in odd units such as dollars, hours, postal charges, frequency of telephone calls, cultural interaction or cognitive distances. The question is the following: can we map a conceptual surface in its own operational terms in order to reveal the potential links between its particular geometry and geography? We assume that a two-dimensional display is the most practical solution for conveying the idea of pattern and spatial relations.[3] Within this limitation one of two types of mapping may be adopted.

## Symbolic Representations

A cartographic symbol is overlayed on the geographical map and shows the distances of a conceptual space with respect to one origin or between places. Isochrone or isocost maps are examples. The geographic space remains invariant. The distance-space superimposed to the geographic area can be thought of as a volume whose heights are simulated by cartographic symbols.

## Equidistant Representations

Imagine a map which is scale calibrated and whose graphic distances replicate some conceptual measure of distance such as time or cost. The mapping usually involves a non-Euclidean transformation of the geographical area. Consider a matrix of distances among n points. A graphic representation of these points in two dimensions could be done by adopting one of the following systems:

1) Polar equidistant. Distances are preserved when measured from a given center-point (Fig. 1). The substantive content of such a map is very limited, since only one row or one column of the distance matrix is shown. In 1941 Boggs already pointed out the limitations and inadequacies of polar maps: "we are interested in [distance] everywhere, in all directions, not only from a single [point]".[4]

2) Multipolar. Theoretically, distances are preserved between all points. We must find a two-dimensional configuration whose graphic distances approximate the $n(n - 1)/2$ conceptual distances. The solution is trivial when the metric of the conceptual surface is Euclidean in two-space. When the distance function is unknown (which is usually the case) one may choose a graphic approximation based on the concept of crow's flight distances. The graphic configuration will be inexact since the distance function implemented in the scaling algorithm is not identical to the distance function

governing the conceptual surface (Fig. 2). Various non-Euclidean and non-metric spaces illustrate this problem.

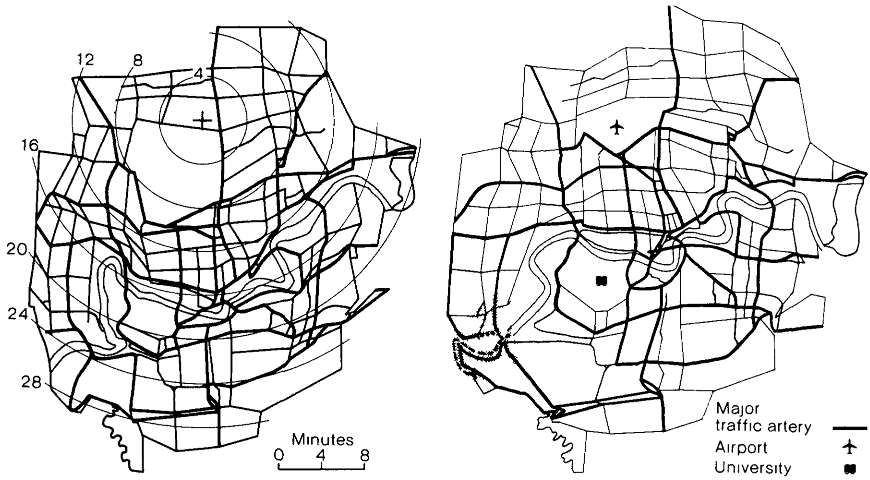


Figure 1 and Figure 2: A Polar Equidistant (left) and Multipolar (right) representations of Time-Distances in Edmonton, Alberta. Shading shows areas topologically fuzzy.

Non-Euclidean Space

Consider a square street map. The geographical distance $d_1(M,N)$ can be interpreted as the length of minimal path traversed by a car driver moving from M to N that is constrained to move only along line segments parallel to the directions North-South and East-West. Theoretically, there are an infinity of such minimal paths. Only limitations in the number of streets accessible reduce the choice of the driver who must decide only between two or three possible alternate routes. Geographic distances in a square street pattern can be calculated by using the definition of distance in the city-block metric:

$$d_1 \ (M,N) = | \ X_1^m - X_1^n \ | + | \ X_2^m - X_2^n \ | \qquad (1)$$

where $X_1^m$ designates the position of M on the first coordinate axis (say the East-West direction) and $X_2^m$ designates the position of the same point on the second coordinate axis (the North-South direction). Assume that the ground speed is the same all over the city. Hence travel time is equated to geographical distances. It is clearly impossible to map in two dimensions travel time according to the definition of Euclidean distances:

$$d_2\ (M,N) = [(x_1^m - x_1^n)^2 + (x_2^m - x_2^n)^2]^{\frac{1}{2}} \qquad (2)$$

since the sum of the lengths of the legs of a right triangle is always greater than the length of its hypotenuse. In order to represent travel time by Euclidean distances one is forced to add dimensions to the mapping structure, since a set of distances between n points can always be mapped isometrically in an n-1 dimensional Euclidean space.[5] This solution would lead to a self-defeating cartography, however, since one would need $n!/2(n-2)!$ maps to portray all facets of the manifold.

Non-Metric Space

A distance-space does not necessarily imply a metric space. Let d(M,N) be a function defined in the real plane $R^2$ for all pairs (M,N). The function is called metric on the plane if it satisfies the following conditions:[6]

$$\begin{aligned} &1)\quad M = N \rightarrow d(M,N) = 0 \\ &2)\quad M \neq N \rightarrow d(M,N) \neq 0 \\ &3)\quad d(M,N) = d(N,M) \\ &4)\quad d(M,N) + d(N,O) \geq d(M,O) \end{aligned} \qquad (3)$$

Spaces which do not comply with postulate 4) are semi-metric. An example of semi-metric time-distance space was previously mentioned. Furthermore, a time-distance matrix is usually non-symmetrical with $d(M,N) \neq d(N,M)$. One can think of other geographical examples where the metric function is "degenerated".[7] Consider a cost-space, where distances are equated to freight charges. Should terminal costs of merchandise loading and unloading be included in the total charge? If so the distance-space would violate the first postulate ($M = N \rightarrow d(M,N) \neq 0$). Imagine a transportation firm which would increase the charging cost up to a given distance, but would maintain a fixed total charge beyond that limit; for the remaining part of the trip the corresponding cost-space would contradict the second postulate ($M \neq N \rightarrow d(M,N) = 0$).

Semi-metric or "degenerated" metric spaces are not isometrically mappable. Any attempt in trying to map those distance-spaces will yield a cartographic product which remains intrinsically fuzzy. The uncertainty or fuzziness of the resulting positions leads to a problem in representing scales. As there is a range of angle as well as distance associated with each position computed, a conical symbol might be used to imply that the measuring-stick on the map is somewhat elastic in both angular orientation and length.

## Generalizations

It is possible to associate a set of elements E with more than one metric function. If $d_1$ and $d_2$ are two metric functions for E, then $(E,d_1)$, $(E,d_2)$ are two distinct metric spaces. In fact, there can be an infinite number of metric functions which can be associated with a given set E, as many invisible components of the geographical space are associated with many distinct metric functions.[8] The examples of metric and non-metric spaces given below can be found in most textbooks.

a) $$d(M,N) = |X^m - X^n| \tag{4}$$

This is the unbounded metric space called the real line.

b) $$d(M,N) = \left\{ \sum_{i=1}^{k} |X_i^m - X_i^n|^p \right\}^{1/p} \tag{5}$$

This is the well known family of Minkowski spaces (see Appendix). These are also called lp spaces.[9] If p = 1, we have the city block metric; if p = 2 we return to the k-dimensional Euclidean space. If k = 1, the metric is not distinguishable from a). Note the limit:

c) $$\lim_{p \to +\infty} d(M,N) = \underset{1 \leq i \leq k}{\mathrm{MAX}} \; | X_i^m - X_i^n | \tag{6}$$

named the maximum p-metric (see Appendix). However, d(M,N) is non-metric for $p < 1$ (the triangular inequality is not satisfied). Most researchers hesitate using the Minkowski model outside of the domain $1 < p < \infty$, although there is no particular reason to believe that a social space must be metric.[10]

d) $$d(M,N) = 0 \text{ if } M = N \text{ and } 1 \text{ if } M \neq N \text{ for all elements of } E \tag{7}$$

d is the standard discrete metric for E, often nicknamed the "pathological" example of a metric.

e) $$d(M,N) = \sum_{i=1}^{k} |X_i^m - X_i^n|^2 \tag{8}$$

This function is semi-metric since the triangular inequality is not satisfied. Statisticians avoid it, and prefer the use of the root-distance-squared, which is Euclidean, although this attitude implies an interpretation about nature which is not always warranted.

Despite differences among the names of distance measurement, all meanings taken on by the functions above have in common a measure of "how far apart" places M and N are. Earth scientists do not usually feel comfortable with metric functions which do not permit the classical analytical formulation. A connection between metric space and differential geometry would be advantageous to the cartographer, since it would unify the problems raised by the representation of non-Euclidean spaces.

A generalization, for instance, may be attempted from the Riemannian distance and the Minkowskian metric. In Riemann geometry, distance is defined as (Einstein's conventional summation over repeated i,j):

$$ds^2 = g_{ij}\, dx^i dx^j, \quad i, j = 1, \ldots k \tag{9}$$

where $g_{ij}$ is a tensor of transformation.

A special case is the Euclidean distance which can be written, in tensor notation

$$ds^2 = \delta_{ij} dx^i dx^j; \quad i,j = 1, \ldots . k \tag{10}$$

with $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise

Recall the geometric notation for Euclidean distance

$$d(M,N) = \left\{ \sum_{i=1}^{k} |x_i^m - x_i^n|^2 \right\}^{1/2} \tag{11}$$

which is a special case of the Minkowskian generalization

$$d(M,N) = \left\{ \sum_{i=1}^{k} |x_i^m - x_i^n|^p \right\}^{1/p} \tag{12}$$

or, in tensor notation

$$ds^p = \delta_{ij} |dx^i|^{p/2} |dx^j|^{p/2}; \quad i, j = 1, \ldots k \tag{13}$$

with $p \geq 1$

Combining (10) and (13), substituting $q = p/2$ as parameter, we get the expression:

$$ds^{2q} = g_{ij}\, (dx^i)^q \ (dx^j)^q; \quad i, j = 1, \ldots k \tag{14}$$

with $q \geq 1/2$.

We have the following cases:

0) If $q < 1/2$, the space is non-metric.

1) If $q = 1$, the space is Riemannian with $g_{ij}$ any metric tensor

2) If $g_{ij} = \delta_{ij}$ and $dx^i$, $dx^j$ are positive definite, the space is Minkowskian for $q \geq 1/2$.

3) Both $q = 1$ and $g_{ij} = \delta_{ij}$ give the Euclidean space

4) If $q \neq 1$ and $g_{ij} \neq \delta_{ij}$, we have a more general class of space where Euclidean, Reimanian and Minkowskian spaces are subsets (Fig. 3).

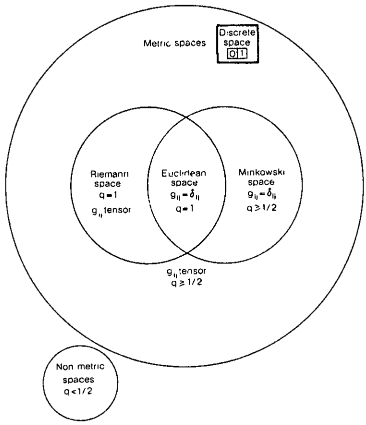


Figure 3: Venn Diagram of Various Classes of Spaces.

A concern of this study is the mapping in two dimensions of various non-Euclidean and non-metric spaces. Consider a distance space (described by a set of observed separations) whose metric function is unknown. The problem is to find a mapping which approximates the distance space. To our knowledge there has been no attempt to free the map from its age-old Euclidean scale servitude and to create a representation which incorporates a non-Euclidean scale for distance measurement. Multidimensional scaling algorithms, for instance, usually proceed on the idea of crow's flight distances. The limitations of the Euclidean framework may be shown through examples.

## Cartographic Experiments on Odd Geometries and Odd Metrics

The algorithm adopted in this study is based on a "trilateration" procedure, an iterative method suggested by Tobler and essentially identical to the Young-Togerson's method.[11] An APL program was written to perform two tasks: 1) find a mapping which best approximates a set of given distances, and 2) which has most similarity to an arbitrary starting configuration. The latter transformation ensures unity and comparability between various solutions of the experiment. The algorithm is implemented with an Euclidean distance function. Results are evaluated using three parameters: 1) the number of iterations necessary to arrive at a solution, 2) the stress value of the solution configuration when the solution is not identical to the target (the distance space), and 3) the variance and distribution of the residuals (the residual is based on the average of half the difference between the solution and the distance of all links involved).[12] Four types of situations may arise: 1) the target is two-dimensional Euclidean, 2) the target is metric but not Euclidean in two dimensions, 3) the target is semi-metric, and 4) the target is metrically degenerated.[13]

1) Two-Space Euclidean

The distance space is easily recovered and solutions converge iteratively toward a perfect degree of fit with the target. The rate of convergence, however, may be affected by large deviations between the position of homologous points in the starting configuration and the target. Discrepancies in neighbourliness relationships are particularly critical and may considerably increase the number of iterations.

2) Not Euclidean in two dimensions

Consider the discrete space where all elements $d_{ij}$ are equal and non-zero if $i \neq j$ and zero if $i = j$. The space is always mappable into an n-1 Euclidean manifold, where n is the number of points. The Euclidean mapping of five points in two dimensions was attempted for analysis of the conditions of approximation and stress (Fig. 4a). As expected the solution configuration shows a high stress and a poor correlation with the target. Notice the general symmetry of the solution configuration and even distribution of the residuals symbolized by the size of the circles. A larger residual lies at the center point of the configuration, since the distance of this point to all other points of the configuration is the least fitted to the target.

3) Semi-Metric

In a semi-metric space the distance between one pair of points does not depend on the distance assigned to other pairs. Hence the triangular inequality $d(M,N) + d(N,O) \geq d(M,O)$ is generally not true. An accurate representation of semi-metric spaces is impossible. A fuzzy map can be constructed, however, which best approximates the target space (Fig. 4b). The mapping of various semi-metric spaces was attempted showing increasing stress and fuzziness as the number of violations of the triangular rule increases. Stress and fuzziness also vary according to the strength of the inequality $d(M,N) + d(N,O) < d(M,O)$. A single strong inequality may increase the stress considerably. Several strong inequalities of this type quickly increase the size of the residuals to a point where the location of the points in the solution is so fuzzy that the map is useless. Triangular equalization may improve the map considerably. For instance, we found that the stress value on a map showing flying times between eleven major cities in Canada dropped from 16% to 5%, after "metrizing" the space by triangular equalization (in this procedure, the right side of the inequality is set equal to the left).

4) Degenerated

The solution configuration (the map) is never identical to the target (the distance-space). A few mapping experiments were attempted considering the successive violations of postulates 1), 2), and 3) in equation (3) (Fig. 4c and 4d). In all cases the stress is relatively high although only few entries of the corresponding distance-matrices do not fulfill the metric rule. Notice that the fuzziest locations in the solution configurations correspond to points which are the "trouble-makers" in the distance-matrices (for instance c and e whose distance d(c,e) = 0 and c whose distance d(e,e) ≠ 0).

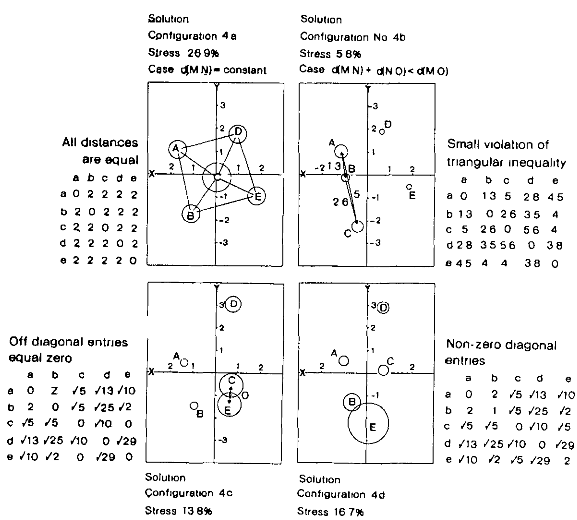


Figure 4: Mapping of Various Odd Metric and Non-Metric Spaces.

## Discussion

The failure of mapping a distance space without considerable errors may be related to two classes of reasons: 1) the algorithm is not operating properly, and 2) the distance function implemented in the algorithm is inadequate. The first problem is "mechanical". For instance, we have observed that when there is a large discrepancy between the distance space and the a priori starting configuration, the iterative procedure of trilateration oscillates between two rather unsatisfactory solutions, instead of converging to a "best" approximation. There are a number of ways to deal with this. We used the simple expedient of reducing the calculated displacement of each point at each step in the iteration. It should be emphasized that this damping procedure does not affect the converged solution in any way.

The second class of problems refers to a lack of knowledge about nature. If the metric function is unknown, a trial and error procedure by implementing various types of distance functions into the algorithm may be necessary. Further, if the configuration solution is scaled in non-Euclidean terms (a configuration of points which approximates a given set of city-block distances for instance), the goodness of the solution depends on the orientation of the axis since distances are rotationally invariant in the Euclidean case only (Fig. 5). In the trilateration method, as well as in most multidimensional scaling algorithms, the solution is arrived at by a series of directional corrections (as well as distance) which must assume rotational invariance. Kruskal[14] briefly mentioned this problem, but it has been mostly ignored.

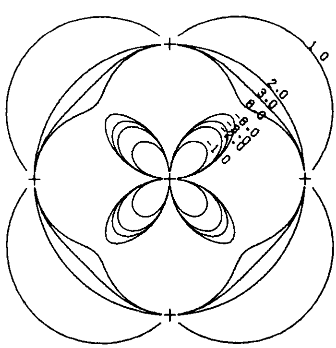

Figure 5: Variance of distances with respect to a center point for various values of the Minkowski distance exponent p. The axis are being rotated from 0 to 360°.

## Conclusions and Perspectives

The mapping of conceptual spaces either not Euclidean or non-metric raises specific cartographic issues which should merge into a new body of theories. Traditional notions of scale and locational accuracy become meaningless when the spatial structure of a space cannot be represented adequately on a plane.

From a communication view point, one may wonder if a representation which is partly erroneous and which involves geographic distortions is really useful. For instance, the discontinuous view of space of the walker or automobilist suggests that a discontinuous map may be more appropriate. Time-distance may be plotted at "scale" along highways and streets, instead of trying to infer from observations along highways the metric space around them and create a continuous representation. On the other hand, the underlying geometry of a continuous map reflects a spatial process which may be desirable to represent. In this case the vagueness of the representation becomes a precious source of information since it was shown that fuzziness concentrates on points whose locations do not conform with the distance function stipulated in the map scale. One may also question the use of maps whose scales are specified in non-Euclidean terms. The visual map of a non-Euclidean space indicates positions only, and people intuitively might still apply the concept of crow's flight distances when looking at the map pattern. This would lead to a misinterpretation of place separations. Indeed, the whole concept of map pattern based on geographical variance of Euclidean closeness and remoteness might need revision. The extent to which this new kind of maps will challenge the Euclidean framework of thought which presently governs map reading and interpretation is not foreseeable yet, but surely an assessment of their practicality will be needed.

---

Acknowledgement: Help from Dr. Werner Israel, dept. of Physics and Geoff Lester, Cartographic Laboratory, University of Alberta, are gratefully acknowledged.

Footnotes

1 Warntz, William, "Global Science and the Tyranny of Space," Proceedings of the Regional Science Association, Vol. 19, (1967) 7 - 19.

2 This example was given by J.W. Clark, "Time-Distance Transformations and Transportation Networks," Geographical Analysis, Vol. 9, (1977), 195 - 205.

3 Bertin has argued that three dimensional maps destroy the geographical content and is not suitable for visual analysis. See Jacques Bertin, "The Constants of Cartography," paper presented at the fifth conference of the International Cartographic Association, Stresa, (1970).

4 Boggs, S.W., "Mapping the Changing World: Suggested Developments in Maps," *Annals*, The American Association of Geographers, Vol. 31, (1941), 119 - 128.

5 Isometric means length preserving.

6 Pitts, C.G.C., *Introduction to Metric Spaces* (Edinburgh: Oliver and Boyd, Inc., 1972).

7 The term "degeneracy" is used by Yu. A. Shreider, *What is Distance?* (Chicago: The University of Chicago Press, 1974).

8 There is one space for each definition of distance. Warntz mentions the concept of "elastic mile". See Warntz op. cit. footnote 1.

9 Shreider, op. cit. footnote 7.

10 Lundberg, Ulf and Gosta Ekman, "Subjective Geographic Distance: A Multidimensional Comparison," *Psychometrika*, Vol. 38, (1973), 113 - 121.

11 Tobler's trilateration procedure is cited by R.G. Golledge and G. Rushton,"Multidimensional scaling: Review and Geographical Applications," *Geographical Technical Paper Series*, No. 10 (1972). See W.S. Torgenson, *Theory and Methods of Scaling* (New York: Wiley, 1958).

12 See Kruskal for the definition of stress. Kruskal, J.B., "Nonmetric Multidimensional Scaling: A Numerical Method," *Psychometrica*, Vol. 29 (1964), 115 - 129.

13 We avoid the expression 'non-Euclidean', which is often used with a much more restrictive meaning (elliptical or hyperbolical spaces).

14 Kruskal, op. cit., footnote 12.

Appendix: Limiting Distance in Minkowskian Spaces

The term "Minkowski space" here does not refer to the four dimensional space used in the theory of relativity.

$$s^m = \sum_{i=1}^{n} | x_i |^m \qquad (1)$$

where n is the number of coordinates in the space, and $| x_i | = | a_i - b_i |$ is the displacement between two points A and B along the i'th coordinate axis. Taking the logarithm

$$m \log s = \log ( \Sigma | x_i |^m )$$

If the longest displacement along any of the n coordinates is $| x |_{max}$, then define $p_i = | x_i | / | x |_{max} < 1$ for all i but max.

$$\text{Thus} \quad \Sigma | x_i |^m = | x |^m_{max} (1 + \Sigma p_i^m) \qquad (2)$$

Since $\lim_{m \to \infty} \Sigma p_i^m = 0$, when $m \to \infty$, (2) can be rewritten.

$$m \log s = m \log | x |_{max} + \log (1) = \log | x |_{max}$$

$$\text{or} \quad \lim_{m \to \infty} s = | x |_{max} \qquad (3)$$

If the smallest displacement along any of the n coordinates is $| x |_{min}$ then define $q_i = | x_i | / | x |_{min} > 1$ for all i but min.

$$\text{Thus} \quad \Sigma | x_i |^m = | x |^m_{min} (1 + \Sigma q_i^m) \qquad (4)$$

Since $\lim_{m \to -\infty} \Sigma q_i^m = 0$, when $m \to -\infty$, (4) becomes

$$m \log s = m \log | x |_{min} + \log (1) = m \log | x |_{min}$$

$$\text{or} \quad \lim_{m \to -\infty} s = | x |_{min} \qquad (5)$$

We know by definition that

$$| x_i |^0 = 1$$

$$\text{Thus} \quad \lim_{m \to 0} s^m = \sum_{i=1} | x_i |^m = n \qquad (6)$$

Taking the logarithm of this limit

$$\lim_{m \to 0} (m \log s) = \log n = \text{Positive constant } k$$

By the definition of the logarithm (e.g. base e):

$$\lim_{m \to 0} s \equiv \lim_{m \to 0} e^{\log s} = \lim_{m \to 0} e^{k/m} \qquad (7)$$

Giving two values, depending on whether m is approaching 0 from positive or negative values:

$$\lim_{m \to 0\pm} s = \lim_{m \to 0\pm} e^{k/m} = e^{\pm\infty} = \begin{cases} \infty & (+) \\ 0 & (-) \end{cases} \qquad (8)$$

Thus we have the limiting values for the distance in four cases:

$$s = \begin{cases} |X|_{max} & \text{for } m \to +\infty \\ \infty & \text{for } m \to 0^{+} \\ 0 & \text{for } m = 0^{-} \\ |X|_{min} & \text{for } m \to -\infty \end{cases}$$