



RETURNING MATERIALS:

Place in book drop to
remove this checkout from
your record. FINES will
be charged if book is
returned after the date
stamped below.

JUL 1 1980

THE GEOMETRY OF MULTIDIMENSIONAL SCALING

By

Daryl W. Tingley

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

1980

ABSTRACT

THE GEOMETRY OF MULTIDIMENSIONAL SCALING

By

Daryl W. Tingley

Multidimensional Scaling (M.D.S.) is a process of finding representations of a given class of objects as points in a metric space. The metric represents the similarity (or dissimilarity) between objects of the given class. Examples of M.D.S. are the representations of colors as points in a metric space, and of pure tones as points on a helix.

Once such a representation has been found, it is natural to ask if the representation is, in some sense, unique. The bulk of this thesis is devoted to the study of uniqueness questions.

An order transformation f between two metric spaces M_1 and M_2 is a function such that

$$d_1(x,y) \leq d_1(z,w) \iff d_2(f(x),f(y)) \leq d_2(f(z),f(w)) .$$

That is, order transformations preserve "the order of the distances". The uniqueness question can be stated as follows: If $f: M_1 \rightarrow M_2$ is an order transformation, is f

necessarily a similarity? In general, the answer is no, and examples are given in Chapter 1. However, by considering only certain types of spaces, such as subsets of E^n , we demonstrate many situations where f is necessarily a similarity.

Many of the results in this thesis are valid for the more general metric transformation. A function between two metric spaces is called a metric transformation if it preserves equality of distances.

Typical theorems in this thesis are:

Theorem: If M_1 and M_2 are convex metric spaces, and f is a bicontinuous metric transformation from M_1 onto M_2 , then f is a similarity.

Theorem: Let $S \subseteq E^m$, $m \geq 2$, be a connected set with non-empty interior. Then any metric transformation of S into E^n is either a similarity, or maps S onto a single point.

In Chapter 6 some results are obtained about the existence of order transformations. Typically, these results discuss whether or not a given metric space can be order embedded into another metric space.

ACKNOWLEDGMENTS

This author wishes to express his gratitude to Professor L. M. Kelly for his helpful suggestions and guidance during the research. In addition, he would like to thank the guidance committee, Professor Dubes, Professor Ludden, Professor Moran, and Professor Sonneborn, for their many helpful suggestions and comments.

TABLE OF CONTENTS

	PAGE
LIST OF FIGURES	v
LIST OF SYMBOLS	vi
LIST OF DEFINITIONS	x
INTRODUCTION	1
CHAPTER 1. PRELIMINARIES AND EXAMPLES	9
§1. Definitions	9
§2. Definitions of Order and Metric Transformations	12
§3. Formal Statement of the Problem	15
§4. Examples	17
CHAPTER 2. RESULTS IN GENERAL METRIC SPACES	23
§1. Elementary Lemmas	24
§2. On the Continuity of Metric Transformations	29
§3. Metric Transformations Between Convex Sets	40
§4. A Relationship Between the Curvatures of an Arc and of its Image Under a Metric Transformation	48
CHAPTER 3. RESULTS IN NORMED LINEAR SPACES	55
§1. Definitions and Elementary Lemmas	57
§2. On the Continuity of Metric Transformations Between Subsets of Normed Linear Spaces	66
§3. Metric Determinacy in Normed Linear Spaces	72

	PAGE
CHAPTER 4. ORDER AND METRIC TRANSFORMATIONS IN E^n .	84
§1. A Discussion of Metric Bases	86
§2. On Metric Transformations	89
Between Subsets of Euclidean Spaces	
§3. Screw Lines in E^n	101
§4. Order Transformations of the	122
Unit Sphere in E^n	
§5. Appendix to Chapter 4	131
CHAPTER 5. FURTHER RESULTS IN EUCLIDEAN SPACE . . .	133
§1. Notation, Definitions, and	136
Elementary Lemmas	
§2. Metric Transformations Between . . .	143
Hypersurfaces in E^n	
CHAPTER 6. EXISTENCE QUESTIONS	152
§1. On the Existence of Order	152
Embeddings of Finite Sets Into E^n	
§2. On the Existence of Order	162
Embeddings of Spheres Into Normed Linear Spaces	
CONCLUSION	171
BIBLIOGRAPHY	

LIST OF FIGURES

FIGURE	PAGE
1.1	19
1.2	19
1.3	21
2.1	34
3.1	64
3.2	65
3.3	80
4.1	91
4.2	94
4.3	98
4.4	125
4.5	127
4.6	128
4.7	129
5.1	140
6.1	156
6.2	157
6.3	159
6.4	163
6.5	167
6.6	168

LIST OF SYMBOLS

$\text{aff } A$		See page 165
a_s, b_s		Real numbers. See page 104
$B(x, r)$		The open ball with center x and radius r
$\mathcal{C}$		A class of sets
$\mathbb{C}$		The complex numbers
$\mathbb{C}^n$		n -dimensional complex space
$C \setminus T$		Theoretic set difference. See page 28
$\overline{C}$		The closure of the set C
$D(N)$		The distance set of N . See page 14
$d(x, y), d_1(x, y), \dots$		The distance between x and y
dN_p		The differential of N at the point p , called the Weingarten Map
E^2		The Euclidean plane
E^n		n -dimensional Euclidean space
$f, \tilde{f}, \dots$		Metric transformations
$f _C$		The function f restricted to the set C
$\mathcal{H}$		Hilbert space. See page 11
H^n		n -dimensional hyperbolic space
I		The identity transformation

$\kappa_H(\gamma, p)$	The Haantjes-Finsler curvature of the arc γ at the point p . See page 49
$\kappa_\gamma(p)$	The classical curvature of the arc γ at the point p . See pages 49, 137
$\kappa_N(p, v)$	The normal curvature of a given hypersurface at the point p and in the direction v
$l(\gamma)$	The length of the arc γ
$l(P)$	See page 31
$l_\gamma(q, r)$	The length of the arc γ from the point q to the point r
l_p^n	See page 12
l_∞^n	See page 12
$(M, d), (M_1, d_1), \dots$	Metric spaces
M.D.S.	Multidimensional Scaling
mesh(P)	See page 43
N.L.S.	Normed linear space (or spaces)
$(N, d), (N_1, d_1), \dots$	Distance spaces
$N(p)$	The unit normal to the given hypersurface, at the point p , and in the given orientation
$n(p, \alpha)$	The unit normal to the arc α at the point p
$p, p_1, p', q, \dots$	Points in a metric or distance space
pq	The line through the points p and q
(p, q)	The open segment from the point p to the point q
$[p, q]$	The closed segment from the point p to the point q

ρ	A scale function. See page 14
$\mathbb{R}$	The real line
S^n	n-dimensional spherical space
S_p	The space tangent to the hypersurface S at the point p
$S(p,r)$	The sphere about the point p with radius r
$S \times S$	The cartesian product of the set S with itself
s, s_j	Real numbers. See page 104
T	A set or a transformation
$T^s, T_j^s, \dots$	Affine transformations
θ, ϕ	Usually angles
U	A set or a transformation
$U^s, U_j^s, \dots$	Unitary transformations
$U_1, U_2, \dots, U_7$	See page 105
$V, V_j, \dots$	Vector spaces
$V^\perp$	The subspace perpendicular to the subspace V
$v, v_i, v_{ij}, \dots$	Usually vectors. See page 104 or 138
$\ v\ $	The norm of the vector v
$(v_1, v_2), \langle v_1, v_2 \rangle$. .	The inner product of the vectors v_1 and v_2
$[v_1, \dots, v_m]$	The space spanned by the vectors $v_1, \dots, v_m$
$W, W_j, \dots$	Vector spaces
w	Usually a vector. See page 104
$x, x_j, \dots$	Usually vectors

y, y_j	Usually vectors. See page 104
z, z_j	Usually vectors. See page 104
$I_p(u, v)$	See page 137
$\Pi_p(u, v)$	See page 137
(**)	An equation. See page 85 or 89

LIST OF DEFINITIONS

Affine, similarity, 57
 subspace, 57
 transformation, 57
Arc, 31
 geodesic, 32
 length of, 31
 rectifiable, 31

Ball, unit, 59
Basis, metric, 86
Bicontinuous, 10

C-metrized space, 3
Continuous, 10
Converge, 10
Convex, 40
 algebraically, 58
 pseudo, 41
 strictly algebraically, 58
Curvature, classical, 49
 Haantjes-Finsler, 49
 normal, 137

Discrete space, 24
Distance, set, 14
 space, 9

Equilong map, 138
Extreme point, 165

Face, 165
Facet, 165
Flat, m-flat, 11
Fundamental form, 137

Geodesic arc, 32

Isometric, 10
 order, 12
Isometry, 10
Isomorphism, order 13

Join, 31

Line, 31
 in a N.L.S., 58
Long legged local isosceles property, 34

- m-flat, 11
- Mesh, 43
- Motion, 87
- Metric, basis, 86
 - space, 9
 - transform, 13
 - transformation, 13
 - transformation, trivial, 13
- Normally ordered set, 31
- One parameter subgroup of motions, 108
- Open function, 73
- Order embeddable, 12
 - embedded, 13
 - isometric, 12
 - isomorphism, 13
 - transformation, 12
- Rectifiable arc, 31
- Relative interior point, 166
- Scale function, 14
- Scalene configuration, 155
- Screw curve or line, 20
- Segment, 31
 - algebraic, 58
- Similar, 10
- Similarity, 10
 - affine, 57
- Space, C-metrized, 3
 - distance, 9
 - metric, 9
- Span, 86
- Spread, 14
- Stress, 14
- Support, 165
- Supporting set, 165
- Transform, metric, 13
- Transformation, affine, 57
 - metric, 13
 - order, 12
 - unitary, 105
- Triangle equality, 62
- Trivial metric transformation, 13
- Ultra metric inequality, 154
- Unit ball, 59
- Unitary transformation, 105
- Weingarten map, 137

INTRODUCTION

The various psycho-physical theories associated with color, sound and other sense perceptions have given rise to many efforts to "represent" perceptual dissimilarities as distances in a conventional geometric structure. Newton [30], Helmholtz [16], Schrodinger [33], and Henning [17], are but a few of the scientific figures who have contributed to the geometrization of psycho-physical perception theories.

More recently a school of perception theorists led by the contemporary psychologists Torgerson and Shephard have advanced more purely psychological theories explicitly distinguishing between the physical properties of light and the subjective sensations which it produces. These theories have spawned a technique known as multidimensional scaling with applications far beyond the boundaries of perception theory. It is with this notion of multidimensional scaling (MDS) that this thesis is concerned.

Some sense of the explicitly geometrical thinking of the Shephard-Torgerson school can be gathered from the following introductory remarks taken from a recent paper by J.P. Cunningham and R.N. Shephard [7] entitled "Monotone Mappings of Similarities onto a General Metric Space".

STATEMENT OF THE PROBLEM

"In ordinary, nontechnical discussions of the perceived similarities and differences among things (whether faces, voices, tastes, odors, colors, etc.), we readily make use of a spatial metaphor. Thus we may say that one shade of color is very close to or, alternatively, far from another; or even that one shade seems to be somewhere between two others. This automatic use of clearly spatial terms such as "near", "far", or "between" to characterize subjective similarities and differences suggests that there is an implicit connection, in human cognition, between the concepts of similarity and dissimilarity, on the one hand, and the concepts of spatial proximity and distance on the other. Specifically, it suggests that similarities among objects are related by some sort of monotone decreasing function to distances among points corresponding to those objects in some sort of metric space."

As a first step in trying to determine the nature of the presumed real number distance function of an individual's color perception space, much experimenting has taken the form of studying the responses when the individual is asked to decide which of two pairs of presented colors "is the better match". If all such comparisons are made for a given finite set of colors presumably some evidence is adduced about the presumed underlying metric. In the early days of such activity an effort was made to "realize" the resulting

structure "on the blackboard" in such a way that the "blackboard distances" (presumably Euclidean 2-dimensional) reflected the order properties of the observed comparisons.

This leads in the direction of the following abstractions. If S is a set and C a chain (linearly ordered set) with minimal element 0 , and if e is a mapping from $S \times S$ into C satisfying $e(x,y) = e(y,x)$, $e(x,y) = 0$ if and only if $x = y$, then (S,C,e) is a C-metrized space. If C is a subset of the nonnegative real numbers, (S,C,e) is called a semi-metric, or distance space. The blackboard embedding problem is an attempt to "order embed" a given C-metrized set S into E^2 , that is to find a map $f: S \rightarrow E^2$ such that $e(x,y) \leq e(u,v)$ if and only if $d(f(x),f(y)) \leq d(f(u),f(v))$.

There is, of course, no reason to believe that such embeddings are always possible. In fact, if S contained four points x_1, x_2, x_3, x_4 such that $e(x_i, x_j) = e(x_k, x_l)$ for all permutations i, j, k, l of $1, 2, 3, 4$ then S could not be order embedded in E^2 . While the space (S,C,e) may not order embed in E^2 it may be possible to "nearly" embed it. That is, it may be possible to find a mapping of S into E^2 such that there are relatively few reversals of the order of distances. This idea is at the root of a scheme devised by J.B. Kruskal of Bell Telephone Labs to produce the "best" approximate order embedding possible into a Euclidean space of specified dimension [20]. The measure of approximation is called the stress of the

embedding and essentially measures the departure from strict monotonicity of the realized distances. Kruskal and others have concocted computer programs which make this embedding process quite simple and efficient. This thesis does not consider approximate order embeddings.

There are a number of natural questions about order embeddings, the first of course being, is such an embedding possible? Then, is it unique, in some sense? What about the possibility of embedding in spaces other than E^n ? In practice the psychologist is trying to adduce the nature of the "underlying" metric by examining a large number of finite subsets. Is there any reason to suppose that, even if every finite subset of an "unknown" metric space is order embeddable in, say, E^2 , then the space itself must be order embeddable in E^2 ?

Broadly speaking, three questions suggest themselves. These questions will also be asked about metric embeddings (see Chapter 1).

- I. Given two distance spaces N_1 and N_2 , is N_1 order embeddable into N_2 ? We call this the existence question.
- II. Is the embedding unique, in some sense? We call this the uniqueness question.
- III. If each finite subset of N_1 is order embeddable into N_2 , what is the relation of N_1 to N_2 ?

Question I would seem to be the most important. Note that there is no loss of generality in beginning with a distance space, rather than some C-metrized space (S, C, e) . For if there is an order embedding f of (S, C, e) into a distance space (N_2, d) there is induced by this embedding a one-to-one map from C to $\mathbb{R}^+$ given by $e(x, y) \rightarrow d(f(x), f(y))$. Thus, for there to be any hope of (S, C, e) being order embedded into N_2 , there must be a one-to-one mapping of C into $\mathbb{R}^+$, and we may assume the minimal element of C goes to 0. Using this map, we may take C to be $\mathbb{R}^+$.

Normally, an M.D.S. theorist would want N_2 to be a familiar space, or a subset of a familiar space, such as E^n or a normed linear space. We will present some answers to questions of type I, although many of them are negative in that they show, for certain N_1 and N_2 , there is no order embedding of N_1 into N_2 .

Question II is very important in the applications of MDS to problems of data analysis. It is widely assumed that two different order embeddings of a distance space in E^n are "nearly" equivalent "up to a scale factor". This presumably means that they are "approximately" similar. But a moment's reflection shows that this is not literally true. Imagine a set in E^2 with large (finite) cardinality and one additional point, very far removed from this set. Certainly this latter point can be rather freely perturbed without affecting the order of the distances, but the

resulting set will not be even approximately similar to the original set.

MDS theorists would characterize this set as exceptional and maintain their continued faith in the "principle of metric determinancy" which asserts that in a "highly structured" space such as E^n the order of distances of a configuration "essentially" determines the configuration, if the cardinality of the configuration is "large" and the configuration is not "exceptional".

There is strong intuitive reason to feel that there is an element of validity to this "principle" but its precise formulation let alone its proof is very elusive. One of our main purposes in this thesis is to examine the proposition that the order of distances of a configuration in a highly structured space determines the configuration.

Unfortunately we can make no contribution to the third question. Results of this type would be extremely useful, for in practice the MDS user deals with finite subsets of infinite sets - such as a finite set of colors, and tries to determine the underlying distance function from this.

This thesis is divided into 6 Chapters. To help orient the reader, and to indicate what we are trying to achieve we briefly outline these Chapters.

Chapter 1, called Preliminaries, is used to introduce definitions and to make precise some of the ideas discussed in the Introduction. In particular, §2 formally defines the

concepts of order transformation and metric transformation, while §3 further discusses the problems to be considered in this thesis. Chapter 1 concludes with a number of examples, designed to help the reader become familiar with order and metric transformations, and with the types of problems we consider.

Chapter 2 investigates order and metric transformations in very general metric spaces. The results obtained form the foundations for much of the work in later chapters. Sections 1 and 2 primarily investigate the continuity of order and metric transformations. Section 3 investigates metric transformations between convex spaces. Much of the work on convex spaces is a reorganization of previously known results. Finally, §4 investigates the effect of a metric transformation on the curvature of a curve. This is done using a metric definition of curvature.

Chapter 3 investigates metric transformations between subsets of normed linear spaces (N.L.S.). Section 1 consists largely of preliminaries such as definitions and known results. In the second section, we apply results from Chapter 2 and investigate the continuity of metric transformations between subsets of N.L.S. The main results of the chapter are presented in §3, and the following corollary is typical.

Corollary 3.19: Let M_1 be a N.L.S. of finite dimension at least 2, with a strictly convex unit ball. If

$U \subseteq M_1$ is a set with a non-empty interior, then any metric transformation from U into a N.L.S. with the same dimension as M_1 , is a similarity.

Chapters four and five investigate metric transformations between subsets of Euclidean spaces. Section 1 of Chapter four contains background material. In section 2, a theorem of Schoenberg is extended. Schoenberg showed that any metric (or order) transformation of E^m , $m \geq 2$, into E^n is necessarily a similarity. We extend this to subsets of E^m . In section 3, all metric (and order) transformations of the real line into E^n are characterized. This is an elaboration of a known result (von-Neumann and Schoenberg). Finally, section four of Chapter four investigates metric and order transformations from the unit sphere of E^n into arbitrary N.L.S.

Chapter five considers metric transformations between hypersurfaces in E^n . Techniques of differential geometry are used, forcing us to assume that the metric transformation is differentiable.

Finally, Chapter 6 briefly discusses "existence questions". That is, when do order transformations exist between two sets? There are two sections, the first of which deals primarily with finite sets, the second with infinite sets.

CHAPTER 1 PRELIMINARIES AND EXAMPLES

As indicated in the introduction, the principle concern of this thesis is with the problems of order and metric embeddings of one distance space into another, with emphasis on determining when such embeddings are unique. We now wish to make these concepts precise.

§1. By a distance space is meant what is commonly called a semi-metric space, that is a set N together with a mapping $d: N \times N \rightarrow \mathbb{R}^+$ (non-negative) such that for $(x,y) \in N \times N$, $d(x,y) = d(y,x)$ and $d(x,y) = 0$ if and only if $x = y$. For the most part, although not always, our distance spaces will also satisfy the triangle inequality, in which case the space is called a metric space. We denote a distance space by listing the set and the mapping symbol thus: (N,d) , or simply by writing the set symbol N . Symbols used for distance spaces in this thesis are (N,d) , (N',d') , (N_1,d_1) , and (N_2,d_2) . If the space is to also satisfy the triangle inequality the symbols (M,d) , (M',d') , (M_1,d_1) , (M_2,d_2) are used, and we also state explicitly that the space is to be metric.

As a metric space is a topological space, the concept of continuity is very familiar. We briefly discuss continuity in general distance spaces.

A sequence $\{p_i\}$ in a distance space N is said to converge to p if and only if $\lim d(p, p_i) = 0$. This is denoted by $\lim p_i = p$. A function f from N_1 into N_2 is said to be continuous at p if and only if for any sequence $\{p_i\}$ with $\lim p_i = p$ it is true that $f(p) = \lim f(p_i)$. Considering f as a function onto its range, f is said to be bicontinuous if and only if f is invertible and both f and f^{-1} are continuous. The distance function d of a distance space N is said to be continuous if and only if for any two sequences $\{p_i\}$ and $\{q_i\}$ with $\lim p_i = p$ and $\lim q_i = q$ it is true that $\lim d(p_i, q_i) = d(p, q)$. It is easily seen that a metric is continuous. Distance spaces with continuous distance functions are important for, unlike distance spaces in general, the sets of the form $\{x | d(p, x) < \delta, \delta > 0\}$ form a basis for a topology on the set N . We refer the reader to Blumenthal [4] for more details.

Two distance spaces, N_1 and N_2 , are said to be isometric if and only if there is a function g from N_1 onto N_2 such that for $x, y \in N_1$, $d_1(x, y) = d_2(g(x), g(y))$. The function g is called an isometry. If g is such that for $x, y \in N_1$, $d_1(x, y) = kd_2(g(x), g(y))$, $k > 0$ then N_1 and N_2 are said to be similar and g is called a similarity.

Two isometric distance spaces are, from a geometer's viewpoint, "the same", while two similar distance spaces are "the same except for the unit of measurement".

We assume a familiarity with the elementary topology of metric spaces, including commonly used terms. Strictly geometric concepts such as arclength, geodesic, curvature, etc. are defined when they are first used. For the most part the terminology of Blumenthal [4] is used.

In addition to some topology, the reader should be familiar with the concept of a normed linear space, and a linear transformation for Chapter 3, some linear algebra, including some knowledge of the standard theorems on simultaneous diagonalization of linear transformations for Chapter 4, and a knowledge of the differential geometry of hypersurfaces in E^n for Chapter 5.

We often look at very specific metric spaces. These include

E^n - n -dimensional Euclidean space

S^n - n -dimensional spherical space

H^n - n -dimensional hyperbolic space.

By an m -flat of E^n, S^n , or H^n we mean an isometric image of E^m, S^m , or H^m in E^n, S^n , or H^n respectively.

$\mathcal{H}$ - Hilbert space. By this we mean all sequences

$x_1, \dots, x_n, \dots$ of real numbers such that $\sum_{i=1}^{\infty} x_i^2 < \infty$,
with the usual inner product.

We also consider normed linear spaces (N.L.S.). In particular, the spaces ℓ_p^n which consist of all n-tuples of real numbers with the norm given by $\|(x_1, \dots, x_n)\|_p = (\sum |x_i|^p)^{1/p}$ and ℓ_∞^n whose norm is given by $\|(x_1, \dots, x_n)\|_\infty = \max_i \{|x_i|\}$, will be used.

We refer the reader to Blumenthal [4] for a more detailed discussion of these spaces.

§2. We now make precise the notion of order preserving transformation.

Definition: If N_1 and N_2 are distance spaces and f is a mapping from N_1 into N_2 such that for all $x, y, u, v \in N_1$

$$(1) \quad d_1(x, y) \leq d_1(u, v) \Rightarrow d_2(f(x), f(y)) \leq d_2(f(u), f(v))$$

then f is called an order preserving transformation or simply order transformation of N_1 into N_2 . The image $f(N_1)$ is said to be order isometric to N_1 and N_1 is said to be order embeddable into N_2 .

Some properties of order transformations follow easily from the definition. It is easy to see that if (1) holds for all $x, y, u, v \in N_1$ then

$$(2) \quad d_1(x, y) = d_1(u, v) \Rightarrow d_2(f(x), f(y)) = d_2(f(u), f(v)) .$$

From (2) it follows that f is one-to-one (to see this let $x \neq y$), and then because of the symmetry of (1) it

is easily seen that f^{-1} is an order transformation from $f(N_1)$ onto N_1 . We often refer to an onto order transformation as an order isomorphism, and when we ask if N_1 can be order embedded into N_2 we simply mean: Is there an order transformation from N_1 into N_2 ?

Much of what we do is valid under the weaker assumption that the transformation be a metric transformation.

Definition: If N_1 and N_2 are distance spaces and f is a mapping from N_1 onto N_2 such that

$$(3) \quad d_1(x, y) = d_1(u, v) \Rightarrow d_2(f(x), f(y)) = d_2(f(u), f(v))$$

then f is called a metric transformation and N_2 is said to be a metric transform of N_1 .

The term "metric transformation" was first used by Wilson in [40]. Some of the results from this paper will be mentioned later.

A metric transformation can be described as equality preserving, rather than order preserving. Note that (3) differs from (2) in that the implication is only in one direction. A metric transformation need not be invertible, and even if it is, the inverse need not be a metric transformation. The term metric transformation is used rather than, say, "distance transformation" to stay close to the terminology established in the literature. It is clear that if f maps a distance space to a single point, then f is a metric transformation. This is called the trivial metric transformation.

The distance set $D(N)$ of a distance space (N, d) is merely the set of non-negative real numbers which occur as distances in (N, d) , in other words the range of d . Associated with a metric or order preserving transformation f from N_1 into N_2 is a mapping ρ_f from $D(N_1)$ into $D(N_2)$ defined by $\rho_f(d_1(x, y)) = d_2(f(x), f(y))$. The function ρ_f is called the scale function of the transformation f . If no confusion arises it is denoted simply by ρ . Of course $\rho(0) = 0$. If $\lim_{d \rightarrow 0} \frac{\rho(d)}{d}$ exists, it is called the spread of f . This limit is used extensively in Chapter 3, and consequences of its existence are used throughout the thesis.

It is easily seen that a metric transformation is an order transformation if and only if its scale function ρ is strictly monotone. In this case, ρ is one-to-one, and ρ^{-1} is the scale function of f^{-1} .

It is the monotone characteristic of the scale function that the M.D.S. theorists are anxious to attain in their embeddings. If their scale function fails to be monotone, then a least squares measure of the departure from monotonicity is what Kruskal calls the stress of the mapping. Zero stress corresponds to strict monotonicity while low stress satisfies the M.D.S. theorist that he has a "good" representation. Our concern will always be with zero stress. It is easy to see that a metric or order transformation is an isometry or similarity if and only if $\rho(d) = d$ or

$\rho(d) = kd$, $k > 0$ respectively. Note that isometries and similarities are order preserving.

§3. In the introduction we said the purpose of this thesis is primarily to work on the "uniqueness" question, with some work on the "existence" question. We now try to be more explicit as to the meaning of these terms.

The "existence" question asks: When is one distance space order embeddable into another? This can only mean the following: Given two distance spaces N_1 and N_2 , is there an order transformation f of N_1 into N_2 ? The scope of the question may be widened by asking: Given a distance space N_1 and a class of distance spaces $\mathcal{C}$, is there an order transformation f of N_1 onto some member N_2 of $\mathcal{C}$? Note the assumption of onto here, rather than into. If desired, $\mathcal{C}$ could be such that any subspace of any member of $\mathcal{C}$ is also in $\mathcal{C}$, hence the above question includes the same question with into, rather than onto, used. On the other hand, it may not be desirable for $f(N_1)$ to be certain subspaces of N_2 . Thus phrasing the question as we do allows more flexibility. The same types of questions can also be asked of metric transformations.

The uniqueness question is harder to state, for it is not clear what "unique" means. If f and g are order transformations of N_1 onto members of a class $\mathcal{C}$ of distance spaces we could ask if necessarily $f \equiv g$.

However, this is too restrictive for our purpose. For example if $\mathcal{C}$ is all subspaces of E^n , $f: N \rightarrow f(N) \subseteq E^n$ is an order embedding, and h is a similarity of E^n onto itself, then $h \circ f$ is also an order embedding of N into E^n . As $f(N)$ and $h(f(N))$ are similar, so essentially "the same" from a geometric view point, it would be more desirable to ask if f were unique, "up to a similarity". This could be phrased as follows: If f and g are order transformations of N onto members of a class of distance spaces $\mathcal{C}$, is there a similarity h such that $h \circ f \equiv g$?

Order isometries are invertible, so in the above question we have $h \equiv g \circ f^{-1}$. Now because g and f are order isomorphisms, $g \circ f^{-1}$ is also an order isomorphism and $g \circ f^{-1}: f(N_1) \rightarrow g(N_1)$. Thus equivalent to the above question is the following: If f is an order transformation from N_1 onto N_2 , N_1 and N_2 members of a class $\mathcal{C}$ of distance spaces, is f a similarity? This phrasing is nice as we stay within the class $\mathcal{C}$ of spaces, yet the seemingly more general question (above) is answered.

Either of the above two questions could be asked of metric transformations as well, however they would no longer be equivalent. (A metric transformation need not be invertible.) Most of our work will consider the second question, with metric rather than order transformations used. That is, we ask: If f is a metric transformation from N_1 onto N_2 , N_1 and N_2 members of a class $\mathcal{C}$ of

distance spaces, is f a similarity? Occasionally we will insist that f have further properties such as continuity, or differentiability, but we try to stay away from this, putting restrictions on the class $\mathcal{C}$, rather than on the function f .

§4. At this point we wish to present some examples of order and metric transformations. Further examples are presented in appropriate Chapters - usually as counter-examples. The purpose of the following examples is to show that, even in "highly structured" spaces, the order of the distances of a set does not determine the set, even up to a similarity. In addition, in none of these examples is there any hope of formulating a notion of "approximate" similarity.

Example 1 shows that, given any metric or distance space, there are many metric or distance spaces order isomorphic to it.

Example 1 (Wilson) [39]: Let ρ be any real valued function with domain the non-negative real numbers, and with the following properties

- 1) $\rho(0) = 0$
- 2) ρ is strictly increasing
- 3) $\rho(\lambda t_1 + (1 - \lambda)t_2) \geq \lambda \rho(t_1) + (1 - \lambda)\rho(t_2)$ for $t_1, t_2 \geq 0, 0 \leq \lambda \leq 1$.

It is shown in [39] that if (M, d) is any metric space,

then $(M, \rho \circ d)$ is also a metric space. That (M, d) is order isomorphic to $(M, \rho \circ d)$ is easy to see. The scale function of the order transformation (the function which assigns each point p of M to itself) is ρ . Example 3 is of this form. Condition (3) is used only in proving $(M, \rho \circ d)$ satisfies the triangle inequality. Thus if one were interested only in distance spaces, Condition (3) could be omitted.

Example 2: Let S be a set of points on the line segment $[qr]$, on the y -axis, and let p and p' be any two points on the x -axis, further than $d(q, r)$ from the origin (see Figure 1.1). Then $S \cup \{p\}$ is order isomorphic to $S \cup \{p'\}$. The function

$$f(x) = \begin{cases} x & x \in S \\ p' & x = p \end{cases}$$

establishes the order isomorphism. See Figure 1.1.

Example 3: A semicircle of radius r and a line segment of length l are order isomorphic (see Figure 1.2). The order isomorphism is $f(x) = (r, \theta)$ where $\theta = \frac{\pi x}{l}$, and x and (r, θ) are as in Figure 1.2.

Example 4: The mapping from $\mathbb{R}$ to a helix given by $t \rightarrow (a \cos t, a \sin t, bt)$ where a and b are constants, is easily seen to be a metric transformation, whose scale function is given by $\rho^2(d) = 4a^2 \sin^2(\frac{d}{2}) + b^2 d^2$. If $b^2 \geq 2a^2$, by checking the derivative of ρ , it is seen that f is an order transformation.

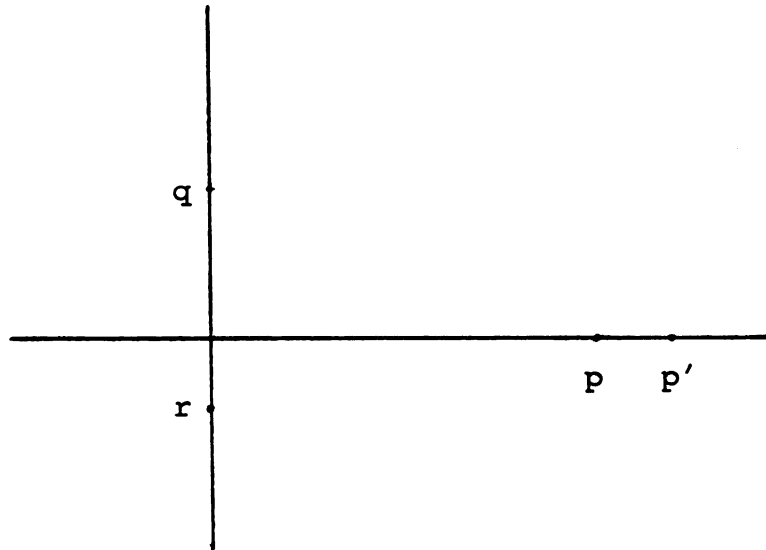


Figure 1.1

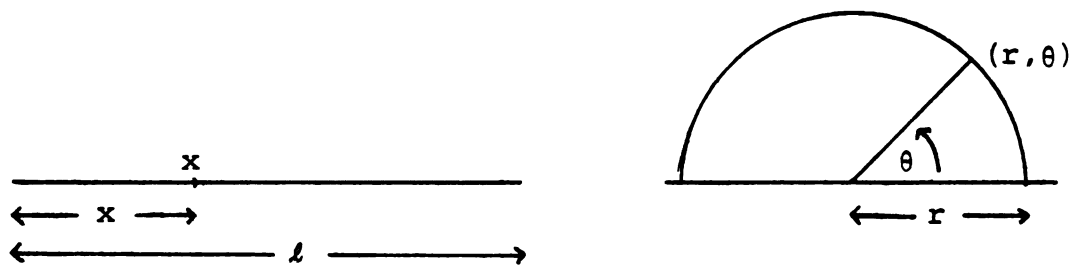


Figure 1.2

Example 5: Let C be the unit circle

$\{(\cos \theta, \sin \theta) \mid 0 \leq \theta < 2\pi\}$. We map C into E^4 by

$$f((\cos \theta, \sin \theta)) = (\cos \theta, \sin \theta, \frac{1}{2} \cos 2\theta, \frac{1}{2} \sin 2\theta) .$$

If $p = (\cos \theta_1, \sin \theta_1)$ and $q = (\cos \theta_2, \sin \theta_2)$ then

$$d(p, q) = \sqrt{4 \sin^2 \left(\frac{\theta_2 - \theta_1}{2} \right)}$$

$$d(f(p), f(q)) = \sqrt{4 \sin^2 \left(\frac{\theta_2 - \theta_1}{2} \right) + \sin^2(\theta_2 - \theta_1)} .$$

Because both $d(p, q)$ and $d(f(p), f(q))$ depend only on the quantity $|\theta_2 - \theta_1|$, f defines a metric transformation of C into E^4 . Furthermore, by investigating the functions

$$g(t) = \sqrt{4 \sin^2 t / 2} \quad \text{and} \quad h(t) = \sqrt{4 \sin^2 t / 2 + \sin^2 t} , \quad \text{it}$$

can be shown that f is an order transformation.

The helix, the circle, and the curve in E^4 of Example 5 are known as screw curves. A screw curve in a metric space M is a metric transformation of $\mathbb{R}$ into M . von-Neumann and Schoenberg [29] characterized all continuous screw curves in E^n as those curves of the form

$$(A_1 \cos k_1 t, A_1 \sin k_1 t, A_2 \cos k_2 t, A_2 \sin k_2 t, \dots, ct)$$

where A_i, k_i , and c are all constants. We shall present a proof and extension of this in Chapter 4.

Example 6: Consider the space $(\mathbb{R}, d)$ where $\mathbb{R}$ is the set of real numbers and $d(x, y) = \sqrt{|x - y|}$. It is easily seen that $(\mathbb{R}, d)$ is a metric space and is order isomorphic

to $\mathbb{R}$. It is shown in Blumenthal ([4], Sec. 54) that this space is not isometric to any subset of E^n , for any n , but it is isometrically embeddable in Hilbert space. Note that the scale function of the order isomorphism, $\rho(d) = \sqrt{d}$, has no spread, for $\lim_{d \rightarrow 0} \frac{\sqrt{d}}{d}$ does not exist. This space, considered as a subset of Hilbert space, is also interesting because it is a curve which has a tangent at no point.

Example 7. Let M be either hyperbolic or spherical n -space. Let $S \subseteq M$ be an m -flat in M , $m < n$, and let α (α less than $\frac{\pi}{2} \times$ radius, for spherical space) be given. For $s \in S$, let $f(s)$ be that point of M which is a distance α above S . (i.e. $d(s, f(s)) = \alpha$ and the segment joining s to $f(s)$ is perpendicular to S). Then f is a metric transformation for, if $d(p, q) = d(p', q')$, it is easy to show the quadrilaterals $pqf(q)f(p)$ and $p'q'f(q')f(p')$ are congruent.

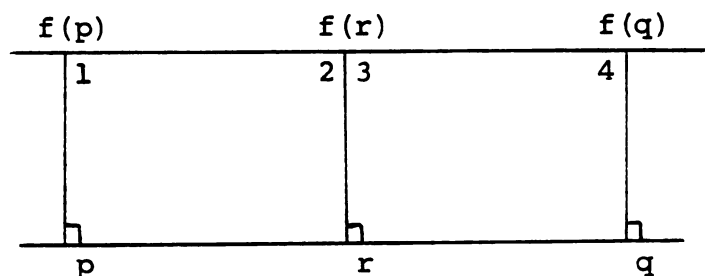


Figure 1.3

To see that f is not a similarity, consider three distinct, collinear points p, q , and r in S , with $d(p, r) + d(r, q) = d(p, q)$. If f is a similarity, then $d(f(p), f(r)) + d(f(r), f(q)) = d(f(p), f(q))$. Angles 1 and 2 (in the above diagram) are equal, as well as angles 3 and 4. (Using congruent triangles.) Thus $\angle 2 + \angle 3 = \angle 1 + \angle 4$. In both spherical and hyperbolic space, the sum of the angles of a quadrilateral is not 360° . Hence $\angle 1 + \angle 2 + \angle 3 + \angle 4 \neq 360^\circ$. So $\angle 2 + \angle 3 \neq 180^\circ$. Thus $f(r)$ does not lie on the line through $f(p)$ and $f(q)$ and

$$d(f(p), f(r)) + d(f(r), f(q)) \neq d(f(p), f(q)) . \quad \square$$

These examples show that in general the order of the distances of a distance space does not determine (even up to a similarity) the space. Even rather nice subspaces of E^n , such as Examples 3, 4, and 5, are not determined by the order of their distances. In none of these examples could one hope to say the two spaces are "approximately" similar.

One of our goals is to show that, at least in common well structured spaces, there may be some truth in the statement that these examples are exceptional. On the other hand, we hope any user of M.D.S. would recognize the limitations of M.D.S. shown by these examples.

CHAPTER 2

RESULTS IN GENERAL METRIC SPACES

In this Chapter results are obtained about order and metric transformations in general metric spaces. The results obtained are nice applications of metric geometry.

The Chapter is divided into four sections. The first two deal, for the most part, with the continuity and bicontinuity of metric and order transformations. The first section contains a series of lemmas, all of which are easy to prove. The major result shows that any order transformation between two distance spaces is necessarily bicontinuous, unless one of the two spaces is discrete. Some properties of the scale function are also developed.

The second section considers the continuity of metric transformations. It gives sufficient conditions for a metric transformation to be continuous. The conditions are mostly on the domain and range of the transformations but, unlike the same problem for order transformations, some (quite mild) conditions are imposed on f .

The third section shows that in the class of convex metric spaces, the order of the distances determines the space, up to a similarity. There is also a metric transformation version of this and several results which relate

the arc length of a curve, and the image of that curve by a metric or order transformation.

Finally, the fourth section includes an interesting result relating the curvature of a curve and the curvature of the image of that curve by a metric, or order, transformation.

§1. We begin Section 1 with a discussion of what is to be done. Let f be a metric or order transformation, from a distance space (N_1, d_1) to a distance space (N_2, d_2) , with scale function ρ . Lemmas 2.1 and 2.2 show that if f is an order transformation, and neither N_1 nor N_2 is discrete, then f is bicontinuous. This is done by showing that necessarily $\lim_{d \rightarrow 0} \rho(d) = 0$, from which continuity follows easily. (A distance space N is said to be discrete if for each $p \in N$, there is an $\epsilon > 0$ such that $\{x | d(p, x) < \epsilon\} = \{p\}$.)

Lemma 2.3 shows a similar result for f a metric transformation. Here we assume $\lim_{d \rightarrow 0} \rho(d) = 0$, hence the condition for continuity depends very much on the particular transformation involved.

Lemmas 2.4 and 2.5 show a converse of Lemma 2.3. They present conditions under which $\lim_{d \rightarrow 0} \rho(d) = 0$. We use this limit later in this Chapter, and in Chapter 3.

Lemma 2.6 does not really concern continuity. However it belongs with this series of lemmas, so we include it here. It shows that a metric transformation is trivial (i.e. maps the entire space to a single point) if the domain is connected, and for some $\epsilon > 0$, $\rho(d) = 0$ for all $d < \epsilon$. This result is used in Chapters 3 and 4.

Before proceeding to Lemma 2.1, consider the following example.

Example: Let (N, d) be any non-discrete distance space. Define a new distance r on N by

$$r(x, y) = \begin{cases} 0, & x = y \\ d(x, y) + 1, & x \neq y \end{cases}.$$

Define $f: (N, d) \rightarrow (N, r)$ by $f(x) = x$. Then f is clearly an order isomorphism. However, f is not continuous, for (N, r) is discrete while (N, d) is not. $\square$

Thus, not all order transformations need be continuous. Lemma 2.1 shows that many are.

Lemma 2.1: If f is an order isomorphism from N_1 onto N_2 , with scale function ρ , and N_2 is not discrete, then f is uniformly continuous, and $\lim_{d \rightarrow 0} \rho(d) = 0$.

Proof: Let $\epsilon > 0$ be given. Since N_2 is not discrete, there are $f(p), f(q)$ in N_2 with $0 < d_2(f(p), f(q)) < \epsilon$. Let $\delta = d_1(p, q)$ and notice that $\delta > 0$. For any $x, y \in N_1$ with $d_1(x, y) < \delta = d_1(p, q)$ we have

$$d_2(f(x), f(y)) = \rho(d_1(x, y)) \leq \rho(d_1(p, q)) = d_2(f(p), f(q)) < \epsilon.$$

Hence f is uniformly continuous and $\lim_{d \rightarrow 0} \rho(d) = 0$. $\square$

Lemma 2.2: If f is an order isomorphism from N_1 onto N_2 , then f is bicontinuous if and only if

- (a) both N_1 and N_2 are discrete
- (b) neither N_1 nor N_2 is discrete.

Proof: It is clear that if f is bicontinuous either (a) or (b) holds.

If (a) holds then, since an order isomorphism is one-to-one, and every set in each space is open, f is bicontinuous.

If (b) holds, the bicontinuity of f follows from the previous Lemma, and the fact that f^{-1} is also an order isomorphism. $\square$

Thus we have established that if N_1 and N_2 are non-discrete distance spaces, any order isomorphism between them is continuous.

Lemma 2.3 appears to be similar to Lemma 2.1, however the condition is on the scale function ρ (which depends on the metric transformation f). Thus Lemma 2.3 is not nearly as strong as Lemma 2.1.

Lemma 2.3: If f is a metric transformation from N_1 into N_2 with scale function ρ , and $\lim_{d \rightarrow 0} \rho(d) = 0$, then f is uniformly continuous.

Proof: Let $\epsilon > 0$ be given. Let $\delta > 0$ be such that $d < \delta \Rightarrow \rho(d) < \epsilon$. Let $p \in N_1$. Then if q is any point of N_1 with $d_1(p, q) < \delta$ we have $d_2(f(p), f(q)) = \rho(d_1(p, q)) < \epsilon$. The lemma follows. $\square$

The limit $\lim_{d \rightarrow 0} \rho(d) = 0$ which appears in the above Lemmas is important and we wish to study it further. Lemma 2.4, and a consequence of it, Lemma 2.5, will be used in future chapters. These may be considered as a converse to Lemma 2.3.

Lemma 2.4: Let f be a continuous metric transformation from N_1 into N_2 with scale function ρ . Suppose there is a $p \in N_1$, and a number $a > 0$ such that $0 \leq b \leq a$ implies there is a $q \in N_1$ with $d_1(p, q) = b$. Then $\lim_{d \rightarrow 0} \rho(d) = 0$.

Proof: Let $\epsilon > 0$ be given. Since f is continuous, there is a $\delta > 0$ such that $d_1(p, q) < \delta$ implies $d_2(f(p), f(q)) < \epsilon$. For every $b > 0$, $b < \max\{a, \delta\}$, there is a $q \in N_1$ with $d_1(p, q) = b$. Hence

$$\rho(b) = \rho(d_1(p, q)) = d_2(f(p), f(q)) < \epsilon.$$

That is, $b < \max\{a, \delta\} \Rightarrow \rho(b) < \epsilon$. Hence $\lim_{d \rightarrow 0} \rho(d) = 0$. $\square$

Lemma 2.5 states Lemma 2.4 in terms of familiar properties of distance spaces. A continuous distance function is needed here, as well as in Lemma 2.6, to consider the topological property of connectedness.

Lemma 2.5: If (N_1, d_1) is a distance space with a continuous distance function, containing a connected subset which is neither empty nor a singleton, then any continuous metric transformation f of N_1 onto N_2 , with scale function ρ , satisfies $\lim_{d \rightarrow 0} \rho(d) = 0$.

Lemma 2.5 follows easily from Lemma 2.4 and the Intermediate Value Theorem.

Lemma 2.6 is of a slightly different flavor, but we include it here as it also examines the scale function near 0. This lemma will be used several times in Chapter 4.

Notation: For any two sets C and T , $C \setminus T$ is defined to be $\{x | x \in C \text{ and } x \notin T\}$.

Lemma 2.6: Let f be a metric transformation from N_1 into N_2 with scale function ρ . If for some $\epsilon > 0$, $\rho(d) = 0$ for all $d, 0 < d < \epsilon$, then $f|_C$ is trivial for any connected component C of N_1 .

Proof: Let C be a connected component of N_1 and let $p \in C$. Let $T = \{x \in C | \rho(d_1(p, x)) = 0\}$.

Note that $f(T) = \{f(p)\}$. Let $x \in T$ and let y be such that $d_1(x, y) < \epsilon$. Then

$$\begin{aligned} d_2(f(p), f(y)) &= d_2(f(x), f(y)) \\ &= \rho(d_1(x, y)) \\ &= 0 \end{aligned}$$

Thus $y \in T$, and T is open.

Let $x \in C \setminus T$, and let y be such that $d_1(x, y) < \epsilon$.
 Now as above, if $y \in T$, it would follow that $x \in T$.
 Hence $y \in C \setminus T$, and $C \setminus T$ is open.

Thus $C = T \cup (C \setminus T)$ so is the union of two disjoint open sets, hence one must be empty (as C is connected).
 As $p \in T$, we have $C = T$, proving the lemma. $\square$

This concludes Section 1. We have seen that order transformations tend to be continuous, and even bicontinuous. Much of the work in Sections 3 and 4 of this Chapter uses the assumption that f is a bicontinuous metric transformation. Lemma 2.2 then shows that this assumption is very mild if one studies order, rather than metric transformations.

§2. Lemma 2.3 gave a condition sufficient to force the continuity of a metric transformation f , however it is not useful in answering a question such as this:

"If a given distance space N_2 is a metric transform of the given distance space N_1 , is the transformation involved continuous?"

Example 2, Chapter 3, shows a metric transformation $f: \mathbb{R} \xrightarrow{\text{onto}} \mathbb{R}$ which is not continuous. On the other hand there are certainly continuous metric transformations of $\mathbb{R}$ onto $\mathbb{R}$ (the identity for example). Thus for $N_1 = N_2 = \mathbb{R}$, the answer to the above question depends on the transformation.

If the above question were asked of order, rather than metric transformations, Lemma 2.1 would give sufficient conditions on N_2 so that the answer is yes, regardless of the transformation involved. The theorem we are about to present, Theorem 2.9, is used to answer the above question for metric transformations. In itself it does not answer the question, for some of the conditions involve the particular transformation. However, further information about the domain and range can often be used to show that the assumptions are satisfied, regardless of the particular transformation.

Before proceeding with this theorem, we introduce some definitions and two lemmas. The definitions are all standard in the study of curve theory and arclength. Not all will be needed here, but it seems appropriate to keep them together. We define arcs, arclength, etc. in distance spaces, however in general it is necessary to have a metric space to obtain satisfactory results involving them. Lemma 2.11, which is only true in metric spaces, illustrates why arcs are usually studied in metric spaces.

Definition: A subset of a distance space N is a segment if and only if it is isometric to an interval of the real line. A subset of N is a line if and only if it is isometric to the entire real line.

Definition: A subset γ of a distance space N is an arc if and only if it is a homeomorphic image of a closed segment of the real line. An arc $\gamma = \gamma([a, b])$ is said to join p and q if and only if $\gamma(a) = p$ and $\gamma(b) = q$.

Definition: Let $\gamma = \gamma([a, b])$ be an arc in a distance space N . If P is a finite subset of γ with $P = \{p_0 = \gamma(a), p_1, p_2, \dots, p_n = \gamma(b)\}$ where $p_i = \gamma(\alpha_i)$ and $\alpha_{i-1} \leq \alpha_i$, $i = 1, \dots, n$ then P is said to be normally ordered with respect to the arc γ .

Definition: Let P be normally ordered with respect to an arc γ . We define $l(P)$ by

$$l(P) = \sum_{i=1}^n d(p_{i-1}, p_i) .$$

The length of the arc γ , called $l(\gamma)$, is defined by

$$l(\gamma) = \sup_P l(P)$$

where the $\sup$ is taken over all sets P normally ordered with respect to γ . If $l(\gamma) < \infty$, then γ is called a rectifiable arc.

Definition: An arc $\gamma = \gamma([a,b])$ in a distance space N is said to be a geodesic arc if and only if its length is no greater than the length of any other arc joining $\gamma(a)$ to $\gamma(b)$. That is, if and only if $l(\gamma) = \inf_{\theta} l(\theta)$ where the $\inf$ is taken over all arcs θ joining $\gamma(a)$ to $\gamma(b)$.

Remark: This usage of the term "geodesic", although consistent with that of Blumenthal, contrasts sharply to its usage by Buseman and differential geometers. They define a geodesic as a curve which is locally a segment. Note that a closed segment in a metric space is a geodesic arc.

Lemmas 2.7 and 2.8 will be used in the proof of Theorem 2.9, and later in this Chapter.

Lemma 2.7: Let $\gamma = \gamma([a,b])$ be an arc in a distance space N with a continuous metric d , and let $r > 0$ be a given distance. Let $t \in [a,b]$ and assume $d(\gamma(t), \gamma(b)) \geq r$. Then there is an $s \in (t,b]$ with $d(\gamma(t), \gamma(s)) = r$.

Proof: This is a simple application of the intermediate value theorem.

Consider the function $g: [t,b] \rightarrow \mathbb{R}$ defined by $g(r) = d(\gamma(t), \gamma(r))$. Then g is continuous, since both the distance function and γ are continuous.

$$g(t) = d(\gamma(t), \gamma(t)) = 0$$

$$g(b) = d(\gamma(t), \gamma(b)) \geq r.$$

By the intermediate value theorem, there is an $s \in [t, b]$ with $g(s) = r$. Also $s \neq t$ for otherwise $0 = g(t) = g(s) = r$, contradicting $r > 0$. Thus, the lemma is proved. $\square$

Lemma 2.8: Let $\gamma = \gamma([a, b])$ be a rectifiable arc in a distance space N with continuous metric d . Let $r > 0$ be a given number. Then there exists a set $P_r = \{p_0, \dots, p_{n_r}\}$ in γ with $\gamma(a) = p_0$, $\gamma(b) = p_{n_r}$,

$$d(p_{i-1}, p_i) = r, \quad 1 \leq i < n_r$$

$$d(p_{n_r-1}, p_{n_r}) \leq r$$

and the set $\{p_i\}$ is normally ordered with respect to γ .

Proof: Let $p_0 = \gamma(a)$. Assume $p_0, \dots, p_m$ have been defined. If $d(p_m, \gamma(b)) \leq r$, let $p_{m+1} = \gamma(b)$, and let $n_r = m + 1$. Otherwise, we obtain p_{m+1} as in Lemma 2.7 and proceed, as above, to find p_{m+2} .

We need now to show that eventually the process terminates, that is, that for some m , $p_{m+1} = \gamma(b)$.

For any m , consider $S = \{p_0, \dots, p_m, \gamma(b)\}$. S is normally ordered with respect to γ , hence

$$\begin{aligned} l(\gamma) &\geq l(S) \\ &= \sum_{i=1}^m d(p_{i-1}, p_i) + d(p_m, \gamma(b)) \\ &= \sum_{i=1}^m r + d(p_m, \gamma(b)) \\ &\geq mr. \end{aligned}$$

Thus $m \leq \frac{l(\gamma)}{r}$, so m cannot be arbitrarily large. The lemma is now complete. $\square$

One more definition is needed before Theorem 2.9 can be stated. This definition is discussed further after the statement of Theorem 2.9.

Definition: The distance space N is said to satisfy the long legged local isosceles property at a point $p \in N$ if and only if there is a number $\lambda(p) > 0$ such that for any $\delta < \lambda(p)$, and for any q with $d(p, q) < \delta$ there exists $s \in N$ with $d(p, s) = d(q, s) = \delta$. The number $\lambda(p)$ may be ∞ .

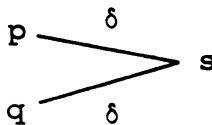


Figure 2.1

Theorem 2.9: Let f be a metric transformation with scale function ρ from a metric space (M_1, d_1) to a metric space (M_2, d_2) . Assume there is a point $p \in M_1$ and a number $\lambda > 0$ such that

- (a) M_1 satisfies the long legged local isosceles property at p with $\lambda(p) = \lambda$.
- (b) There is a point $q \in M_1$, with $f(p) \neq f(q)$, and a rectifiable arc $\gamma \subseteq M_1$, joining q to p .
- (c) The set $f(\{x \mid d_1(p, x) < \frac{\lambda}{2}\})$ is not uncountably discrete.

Then for every $t \in M_1$ the function $g_t = f|_{\{x | d_1(t, x) < \frac{\lambda}{2}\}}$ is bicontinuous. Furthermore f, g_t , and g_t^{-1} are all uniformly continuous.

The hypotheses of this theorem may seem awkward, but they are easy to apply. Many spaces satisfy the long legged local isosceles property. It will be shown in Chapter 3 that any open subset of a normed linear space (of dimension at least 2) satisfies it at each point. In Chapter 6, it is shown that a smooth hypersurface in E^n ($n \geq 3$) satisfies it at each of its points. Also, it can be shown that any point of an open subset of either hyperbolic or spherical n -space ($n \geq 2$) satisfies the long legged local isosceles property.

Assumptions (b) and (c) depend on the particular metric transformation, although they seem to be "mild" conditions. Fortunately, in practice, assumptions on the domain and range can often be substituted for these. For example, in Chapter 3 it will be shown that if $f: U \xrightarrow{\text{onto}} V$ is a metric transformation and U and V are both open subsets of a finite dimensional normed linear space, then (b) and (c) are satisfied.

Note that we require both the domain and range of f to be metric spaces. This is unfortunate, and may not be necessary. However we have been unable to prove or to disprove such a theorem without the triangle inequality.

The proof of Theorem 2.9 is in 5 parts. The first two show the continuity of g_t , from which the continuity

of f follows easily, the last three that g_t^{-1} is continuous.

In part (1) we show that some arbitrarily "small" distances are transformed to "small" distances. It is here that we use (c). In part (2) the long legged local isosceles property and the triangle inequality are used to show that if some arbitrarily small distances are transformed to "small" distances, then all "small" distances are transformed to "small" distances. Thus f is continuous.

To show g_t^{-1} is continuous it is necessary to show "small" distances come from "small" distances. Using the long legged local isosceles property and the rectifiability of γ , for each r , $0 < r < \lambda$, a set of points

$$p_0 = q, p_1, p_2, \dots, p_{n_r-1}, s, p_{n_r} = p$$

is constructed with the distance between any two adjacent points being r . Then $d_2(f(p), f(q)) \leq (n_r + 1)\rho(r)$. If $\lim_{r \rightarrow 0} (n_r + 1)\rho(r) = 0$ then $f(q) = f(p)$, contradicting (b). Thus as $\rho(r)$ becomes "small", n_r must become correspondingly large, which, it is seen, only occurs if r becomes "small". From this, the continuity of g_t^{-1} follows.

In part (3), the p_i 's and s are defined, and the relationship between n_r and r is studied. In part (4), g_t is shown to be one-to-one by verifying that $\rho(r) \neq 0$, $0 < r < \lambda$. Finally, in part (5), the continuity of g_t^{-1} is established.

Notation: Let $B(x, r) = \{y \in M_1 \mid d_1(x, y) < r\}$.

Proof of Theorem 2.9:

(1) Let $\epsilon > 0$ be given. Then there is a $\delta < \lambda$ such that $\rho(\delta) < \epsilon$.

Proof of (1): Assume the contrary. That is, assume there is an $\epsilon > 0$ such that for all $0 < \delta < \lambda$, $\rho(\delta) \geq \epsilon$. Then for any $x \neq y \in B(p, \frac{\lambda}{2})$ it must be that $d_2(f(x), f(y)) = \rho(d_1(x, y)) \geq \epsilon$. Hence $f|_{B(p, \lambda/2)}$ is one-to-one, and $f(B(p, \frac{\lambda}{2}))$ is discrete. On the other hand, $B(p, \frac{\lambda}{2})$ contains an open subset of the arc γ , so is uncountable. Thus $f(B(p, \frac{\lambda}{2}))$ is uncountably discrete, contradicting (c), and (1) has been proven.

(2) The function f is uniformly continuous.

Proof of (2): Let $\epsilon > 0$ be given. By (1), choose $\delta > 0$ with $\delta < \lambda$ such that $\rho(\delta) < \frac{\epsilon}{2}$. Let $d < \min\{\delta, d_1(p, q)\}$ be given. By Lemma 2.7, there is a $t \in \gamma$ such that $d_1(p, t) = d$, and by the long legged local isosceles property of M_1 at p , there is an s with $d_1(p, s) = d_1(s, t) = \delta$. Then

$$\rho(d) = d_2(f(p), f(t)) \leq d_2(f(p), f(s)) + d_2(f(s), f(t)) \leq 2\rho(\delta) < \epsilon.$$

Thus we have shown that $\lim_{d \rightarrow 0} \rho(d) = 0$, so (2) follows by Lemma 2.3.

(3) For each $r < \lambda$, $d_2(f(p), f(q)) \leq (\frac{l(\gamma)}{r} + 2) \rho(r)$.

Proof of (3): By Lemma 2.8 there is a normally ordered subset of γ $P_r = \{p_0 = q, p_1, \dots, p_{n_r-1}, p_{n_r} = p\}$ with $d_1(p_k, p_{k-1}) = r$, $k < n_r$, and $d_1(p_{n_r}, p_{n_r-1}) \leq r < 2r$. By the long legged isosceles property there is an s with $d_1(p_{n_r}, s) = d_1(s, p_{n_r-1}) = r$. Now $l(\gamma) \geq l(P_r) \geq (n_r - 1)r$. Thus $n_r \leq \frac{l(\gamma)}{r} + 1$. Also,

$$\begin{aligned} d_2(f(p), f(q)) &\leq \left[\sum_{k=1}^{n_r-1} d_2(f(p_k), f(p_{k-1})) \right] + \\ &\quad d_2(f(p_{n_r-1}), f(s)) + d_2(f(p_{n_r}), f(s)) \\ &= (n_r + 1) \rho(r) \\ &\leq (\frac{l(\gamma)}{r} + 2) \rho(r). \end{aligned}$$

Thus (3) has been shown.

(4) For each $t \in M_1$, the function f is one-to-one on $B(t, \frac{\lambda}{2})$.

Proof of (4): Assume that for some r with $0 < r < \lambda$ $\rho(r) = 0$. By (3) we have $d_2(f(p), f(q)) \leq (\frac{l(\gamma)}{r} + 2) \rho(r) = 0$. Since $f(p) \neq f(q)$, (assumption (b) of Theorem 2.9) this is a contradiction. Thus, $\rho(r) > 0$ for $r < \lambda$. If $x, y \in B(t, \frac{\lambda}{2})$ then $d_1(x, y) < \lambda$ so that $d_2(f(x), f(y)) = \rho(d_1(x, y)) \neq 0$, and hence $f(x) \neq f(y)$. This completes the proof of (4).

(5) For each $t \in M_1$, $f|_{B(t, \lambda/2)}$ is bicontinuous.

Proof of (5): Let ϵ be given, $0 < \epsilon < \lambda$. Assume that for each $\delta > 0$, there is an r , $\epsilon \leq r < \lambda$ such that $\rho(r) < \delta$. Then by (3),

$$\begin{aligned} d_2(f(p), f(q)) &\leq \left(\frac{\lambda(y)}{r} + 2\right) \rho(r) \\ &< \left(\frac{\lambda(y)}{\epsilon} + 2\right) \delta. \end{aligned}$$

Since δ can be chosen arbitrarily small, $d_2(f(p), f(q)) = 0$, which is a contradiction. Thus for ϵ , $0 < \epsilon < \lambda$, there is a $\delta > 0$ such that

$$\text{if } r < \lambda \text{ and } \rho(r) < \delta \text{ then } r < \epsilon.$$

Let $t \in M_1$ and let $g_t \equiv f|_{B(t, \lambda/2)}$. Then by Step 2, g_t is one-to-one and uniformly continuous. Let $\epsilon > 0$ be given (assume $\epsilon < \lambda$), and choose $\delta > 0$ such that

$$\text{if } r < \lambda \text{ and } \rho(r) < \delta \text{ then } r < \epsilon.$$

If $x, y \in B(t, \frac{\lambda}{2})$ and $d_2(f(x), f(y)) < \delta$ then $d_1(x, y) < \epsilon$. This shows that g_t^{-1} is uniformly continuous, and hence g_t is bicontinuous with both g_t and g_t^{-1} uniformly continuous. This completes the proof of Theorem 2.9. $\square$

Corollary 2.10: If $\lambda(p) = \infty$, then f is bicontinuous.

In Sections 3 and 4 of this chapter bicontinuity is usually assumed. Using the above theorem, or Theorem 2.2, this assumption can often be avoided. In Chapters 3 and 4, where metric transformations in normed linear spaces and Euclidean spaces are investigated, the above theorem is applied.

§3. The major result of this section is to show that in the class of convex metric spaces, the order of the distances determines the space, up to a similarity. This is the content of Theorem 2.11. Theorem 2.17 is a metric transformation version of this.

Before discussing Theorem 2.11 the following definition is needed.

Definition: A distance space (N,d) is said to be convex if and only if for any two points p and q of N , there is a segment joining p to q .

It should be noted that although this definition fits one's intuitive notion of convexity, and is similar to that normally used in linear spaces, it is not the definition commonly found in metric geometry. The common definition of metric geometry is that for each p and q in N there is an r , $r \neq p$ or q , such that $d(p,r) + d(r,q) = d(p,q)$. There are several additional assumptions that can be made on N so that these two definitions are equivalent. Perhaps the best known is that of completeness. For further discussion on this, see [4] sec 14.

Theorem 2.11: In the class of convex metric spaces, the order of the distances determines the metric up to a similarity.

In other words, if (M_1, d_1) and (M_2, d_2) are both convex metric spaces, and f is an order isomorphism from M_1 onto M_2 , then there is a constant D such that $d_1(p, q) = D \cdot d_2(f(p), f(q))$.

A metric transformation version of this was probably originally due to Wilson, for he states in [39], without proof, that a bicontinuous metric transformation with non-zero finite spread, between convex metric spaces, is a similarity. This is the content of Theorem 2.17, although the assumption that the spread is non-zero and finite is not made there. Theorem 2.11 is a consequence of Theorem 2.17.

In [2], Beals, Krantz, and Tversky show Theorem 2.11 with the added assumption of completeness. In [23] Lew proves a stronger version of Theorem 2.11. He insists on neither completeness nor as strong a version of convexity as we use here. He uses "pseudo-convexity".

Definition: A metric space (M, d) is said to be pseudo-convex if for any two points p and q , and λ in $[0, 1]$, and any $\epsilon > 0$ there exists u such that

$$d(p, u) \leq \epsilon + \lambda d(p, q) \quad \text{and} \quad d(q, u) \leq \epsilon + (1 - \lambda) d(p, q) .$$

It seems not to be known whether the completion of a pseudo-convex space is convex.

In the work of Beals, Krantz and Tversky, and that of Lew, the primary interest is that of existence. They show that, for a certain class of distance spaces, each space is order isomorphic to one and only one convex metric space.

Here we are considering only the uniqueness question and, like Wilson, use metric rather than order transformations. Our methods could be adapted to pseudo-convex spaces, giving Lew's result, although the proofs would become more involved.

Leading up to Theorem 2.17 we first show that if $f : M_1 \rightarrow M_2$ is a bicontinuous metric transformation, and γ is an arc in M_1 , then $\ell(f(\gamma)) = D \cdot \ell(\gamma)$, where D is the spread of the transformation (see Theorem 2.14). From this it follows that the image of a geodesic arc is a geodesic arc (Corollary 2.15) and then, in convex spaces, that the image of a segment is a segment, (theorem 2.17). Theorem 2.16 shows that, under very general conditions, the spread of a metric transformation is a non-zero finite number.

We begin with Lemma 2.12, which shows that arc length can be calculated by considering only "small" distances along the arc. Note that the triangle inequality is necessary.

Definition: For any normally ordered subset $P = \{p_0, p_1, \dots, p_m\}$ of an arc γ , the number $\text{mesh}(P)$ is defined to be $\max_i \{d(p_{i-1}, p_i) \mid 1 \leq i \leq m\}$.

Lemma 2.12: Let γ be any rectifiable arc of a metric space. Then for each $\eta > 0$ there is a $\delta > 0$ such that any normally ordered subset P of γ with $\text{mesh}(P) < \delta$ satisfies $\ell(P) > \ell(\gamma) - \eta$.

Proof: See Blumenthal [4], page 61, Lemma 24.1. $\square$

Remark: If γ is a non-rectifiable arc, then Lemma 2.12 may be stated as follows: For each $N > 0$, there is a $\delta > 0$ such that any normally ordered subset P of γ with $\text{mesh}(P) < \delta$ satisfies $\ell(P) > N$. Blumenthal does not prove this, but the proof is essentially the same as the proof of Lemma 2.12.

Corollary 2.13: If $r_n \rightarrow 0$ and P_n is a normally ordered subset of an arc γ with $\text{mesh}(P_n) \leq r_n$ then $\ell(\gamma) = \lim_{n \rightarrow \infty} \ell(P_n)$.

It is not assumed in this corollary that γ is rectifiable.

Theorem 2.14 (Wilson [40]): Let M_1 and M_2 be metric spaces, $f: M_1 \rightarrow M_2$ a bicontinuous metric transformation with finite non-zero spread D , and γ a rectifiable arc in M_1 . Then $f(\gamma)$ is a rectifiable arc in M_2 and $\ell(f(\gamma)) = D \cdot \ell(\gamma)$. If γ is a non-rectifiable arc, then $\ell(f(\gamma)) = \infty$.

Proof: Note that $f(\gamma)$ is an arc in M_2 , the equation of the arc being $f \circ \gamma(t)$, $a \leq t \leq b$. We will prove the case $\ell(\gamma) < \infty$. The case $\ell(\gamma) = \infty$ is essentially the same, using the above remark, rather than Lemma 2.12.

By Corollary 2.13 $\ell(\gamma) = \lim_{n \rightarrow \infty} \ell(P_n)$, where P_n is any sequence of normally ordered subsets of γ with $\lim_{n \rightarrow \infty} \text{mesh}(P_n) = 0$.

Let $\epsilon > 0$ be given. As $D = \lim_{d \rightarrow 0} \frac{\rho(d)}{d}$, there is a δ such that for $d < \delta$, $D - \epsilon < \frac{\rho(d)}{d} < D + \epsilon$, or $(D - \epsilon)d < \rho(d) < (D + \epsilon)d$.

For any normally ordered subset $P = \{p_0, \dots, p_m\}$ of γ with $\text{mesh}(P) < \delta$, $f(P)$ is a normally ordered subset of $f(\gamma)$ and

$$l(f(P)) = \sum_{i=1}^m d_2(f(p_{i-1}), f(p_i))$$

hence

$$(D - \epsilon) \sum_{i=1}^m d_1(p_{i-1}, p_i) \leq l(f(P)) \leq (D + \epsilon) \sum_{i=1}^m d_1(p_{i-1}, p_i)$$

or

$$(D - \epsilon) l(P) \leq l(f(P)) \leq (D + \epsilon) l(P) .$$

Let P_n be any sequence of normally ordered subsets of γ with $\text{mesh}(P_n) \rightarrow 0$. Then $f(P_n)$ is a sequence of normally ordered subsets of $f(\gamma)$, and as f is uniformly continuous on the compact set γ , $\lim_{n \rightarrow \infty} \text{mesh}(f(P_n)) = 0$.

Hence by Corollary 2.12 $l(f(\gamma)) = \lim_{n \rightarrow \infty} l(f(P_n))$.

For n such that $\text{mesh } P_n < \delta$ we have

$$(D - \epsilon) l(P_n) \leq l(f(P_n)) \leq (D + \epsilon) l(P_n) .$$

Letting $n \rightarrow \infty$, we get

$$(D - \epsilon) l(\gamma) \leq l(f(\gamma)) \leq (D + \epsilon) l(\gamma) .$$

As ϵ was arbitrary, it must be that $l(f(\gamma)) = D \cdot l(\gamma)$.

Thus if $l(\gamma) < \infty$, then $l(f(\gamma)) = D \cdot l(\gamma)$. As has been said, the case $l(\gamma) = \infty$ is much the same. In this case $l(f(\gamma)) = \infty$. $\square$

Recall from Chapter 1 that a geodesic arc joining p to q has the shortest length of any arc joining p to q .

Corollary 2.15: If $\gamma = \gamma([a, b])$ is a geodesic arc in M_1 and $f: M_1 \xrightarrow{\text{onto}} M_2$ is a bicontinuous metric transformation with finite non-zero spread D , then $f(\gamma)$ is a geodesic arc in M_2 , and $l(f(\gamma)) = D \cdot l(\gamma)$.

Proof: That $l(f(\gamma)) = D \cdot l(\gamma)$ is the content of Theorem 2.14. Let $\mathcal{O}$ be an arbitrary arc in M_2 , joining $f(\gamma(a))$ to $f(\gamma(b))$. Then by Theorem 2.14, $l(\mathcal{O}) = D \cdot l(f^{-1}(\mathcal{O}))$. Hence $\inf_{\mathcal{O}} l(\mathcal{O}) = \inf_{\mathcal{O}} D l(f^{-1}(\mathcal{O})) \geq D l(\gamma)$. However $l(f(\gamma)) = D \cdot l(\gamma)$, showing $f(\gamma)$ is a geodesic arc in M_2 . $\square$

Before continuing the study of metric transformations of arcs we show (Theorem 2.16) that the assumption in Theorem 2.14, that f has finite non-zero spread, is in many cases unnecessary.

Theorem 2.16 may be considered as a converse of Theorem 2.14.

Theorem 2.16: If the metric spaces M_1 and M_2 each contain a rectifiable arc of length greater than zero, and $f: M_1 \xrightarrow{\text{onto}} M_2$ is a bicontinuous metric transformation, then the spread of f is a non-zero finite number.

Proof: Let $\gamma = \gamma([a, b]) \subseteq M_1$ and $\mathcal{O} = \mathcal{O}([c, d]) \subseteq M_2$ be rectifiable arcs, and let ρ be the scale function of f . Let $\{r_i\}$ be any sequence of numbers such that $\lim_{i \rightarrow \infty} r_i = 0$ and $\lim_{i \rightarrow \infty} \frac{\rho(r_i)}{r_i} = D \leq \infty$. Since γ is connected, Corollary 2.5 guarantees that $\lim_{i \rightarrow \infty} \rho(r_i) = 0$. With this

fact, using notation as in Lemma 2.8, we have by Corollary

$$2.13, \quad l(f(\gamma)) = \lim_{i \rightarrow \infty} n_{r_i} \rho(r_i) \quad \text{and} \quad l(\gamma) = \lim_{i \rightarrow \infty} n_{r_i} r_i.$$

Then

$$D = \lim_{i \rightarrow \infty} \frac{\rho(r_i)}{r_i} = \lim_{i \rightarrow \infty} \frac{n_{r_i} \rho(r_i)}{n_{r_i} r_i} = \frac{l(f(\gamma))}{l(\gamma)}.$$

Since $l(f(\gamma))$ is bounded from below by $d_2(f(\gamma(a)), f(\gamma(b)))$ and $l(\gamma)$ is a positive real number, we see that $D > 0$.

Similarly we show that

$$\lim_{i \rightarrow \infty} \frac{r_i}{\rho(r_i)} = \frac{l(f^{-1}(0))}{l(0)} = \frac{1}{D} > 0.$$

Thus $0 < D < \infty$. Since $D = \frac{l(f(\gamma))}{l(\gamma)}$ is independent of $\{r_i\}$, it must be the case that $\lim_{r \rightarrow 0} \frac{\rho(r)}{r} = D$, proving the theorem. $\square$

Remark: At this point it should be noted that the assumption in Theorem 2.14 that f be a bicontinuous metric transformation with finite non-zero spread is satisfied for f an order isomorphism between two "reasonable" metric spaces. Using Lemma 2.2 and Theorem 2.16 we see that if both the domain and range of an order isomorphism f contain a rectifiable arc, then f is a bicontinuous metric transformation with finite non-zero spread.

With the above remark in mind, Theorem 2.11 follows easily from Theorem 2.17.

Theorem 2.17: If M_1 and M_2 are convex metric spaces, and f is a bicontinuous metric transformation from M_1 onto M_2 , then f is a similarity. The constant of similarity is the spread of f .

Proof: From their respective definitions, it is immediately clear that a segment is a geodesic arc, and any two geodesic arcs, joining the same points, have the same length.

Let p and q be arbitrary points in M_1 . Since M_1 and M_2 are convex, there are segments γ and ϕ joining p to q and $f(p)$ to $f(q)$, with lengths $d_1(p, q)$ and $d_2(f(p), f(q))$ respectively. By Theorem 2.16, f has finite spread D . By Corollary 2.15, $f(\gamma)$ is a geodesic arc of length $D \cdot l(\gamma)$. Furthermore, since $f(\gamma)$ is geodesic, it has the same length as ϕ . Thus

$$d_2(f(p), f(q)) = l(f(\gamma)) = D \cdot l(\gamma) = D \cdot d_1(p, q) .$$

Since p and q were arbitrary, the theorem is proved. $\square$

§4. The major result of this section, Theorem 2.20, concerns the curvature of arcs in general metric spaces. We first define what is meant by curvature of arcs in general metric spaces, and relate this to the more common definition of curvature in Euclidean and Riemannian spaces.

Definitions: Let $\gamma = \gamma([a, b])$ be an arc in a metric space (M, d) . If $q = \gamma(s)$ and $r = \gamma(t)$ we define

$l_\gamma(q,r) = l(\gamma([s,t]))$. That is, $l_\gamma(q,r)$ is the arc length of γ from q to r .

For p,q,r on r , if the following (non-negative) limit exists, it is called the Haantjes-Finsler curvature of γ at the point p :

$$K_H^2(\gamma,p) \stackrel{\text{def}}{=} \lim_{q,r \rightarrow p} 4! \frac{l_\gamma(q,r) - d(q,r)}{l_\gamma^3(q,r)}$$

This definition of curvature was first introduced by Finsler in his thesis of 1918 [13] and was studied extensively by Haantjes [15]. We are going to study the relationship between $K_H(\gamma,p)$ and $K_H(f(\gamma),f(p))$, where f is a metric transformation. The curvature K_H lends itself to this study for we have a nice grasp on the relationships of $l_\gamma(q,r)$ to $l_{f(\gamma)}(f(q),f(r))$ and $d_1(q,r)$ to $d_2(f(q),f(r))$. In Euclidean space, curvature is usually defined as follows:

Definition: If $\gamma(s)$ is a curve in Euclidean n -space, where the parameter s represents arc length along the curve, the curvature of γ at a point $\gamma(s_0)$ is defined to be $|\gamma''(s_0)|$. This is referred to as the classical curvature of γ at the point $\gamma(s_0)$.

More generally, classical curvature is defined for arcs in Riemannian spaces by means of the Frenet formulae. We refer the reader to a book on Riemann geometry, such as Spivak [37], for such definitions.

We now quote some theorems which relate these concepts of curvature.

Theorem 2.18: (Haantjes) If an arc in a Riemannian space has defining equations with continuous differentials of order 3, and if the classical curvature exists at a point on the arc, then the curvature K_H exists at that point, and the two curvatures are equal.

Proof: See [15], Theorem 5.

For arcs in Euclidean space we can combine another theorem of Haantjes ([15], Theorem 8) and a theorem of Egervary and Alexits ([11], Theorem 4.2) to obtain the following result.

Theorem 2.19: If γ is a curve in a Euclidean space, and the classical curvature exists at p , then it is equal to

$$\lim_{q \rightarrow p} 4! \frac{\ell_\gamma(q, p) - d(q, p)}{\ell_\gamma^3(q, p)}.$$

The limit in the above theorem certainly exists if the Haantjes-Finsler curvature of γ at p exists, in which case the two are equal. This limit could be used as a definition of curvature; however we will stay with the more common K_H . Theorem 2.20 remains true if K_H is replaced by the limit of Theorem 2.19.

In Theorem 2.20 it is assumed that f is a metric transformation with spread 1. There is essentially no loss

of generality since, if $f: M_1 \rightarrow M_2$ has finite non-zero spread D , we define a metric r on M_2 by $r(\bar{p}, \bar{q}) = \frac{1}{D} d_2(\bar{p}, \bar{q})$, $\bar{p}, \bar{q} \in M_2$, and let $\tilde{f}: (M_1, d_1) \rightarrow (M_2, r)$ be defined by $\tilde{f}(p) = f(p)$. It is now easily seen that $\tilde{f}$ is a metric transform with spread 1 and that (M_2, d_2) and (M_2, r) are similar.

Theorem 2.20: Let M_1 and M_2 be two metric spaces, f a bicontinuous metric transformation of M_1 onto M_2 , with spread 1. Then

(a) If γ is any arc in M_1 , $p \in \gamma$ and $K_H(\gamma, p)$ and $K_H(f(\gamma), f(p))$ both exist, then

$\lim_{d \rightarrow 0} \frac{\rho(d) - d}{d^3}$ exists and

$$(1) \quad K_H^2(\gamma, p) - K_H^2(f(\gamma), f(p)) = \lim_{d \rightarrow 0} 4! \frac{\rho(d) - d}{d^3}$$

(b) If $\lim_{d \rightarrow 0} \frac{\rho(d) - d}{d^3}$ exists, then for any arc γ

and any $p \in \gamma$, $K_H(\gamma, p)$ exists if and only if $K_H(f(\gamma), f(p))$ exists.

Proof: (a) Assume $K_H(\gamma, p)$ and $K_H(f(\gamma), f(p))$ exist. Consider

$$K_H^2(\gamma, p) = \lim_{q, r \rightarrow p} 4! \left(\frac{l_\gamma(r, q) - d_1(r, q)}{l_\gamma^3(r, q)} \right)$$

$$K_H^2(f(\gamma), f(p)) = \lim_{f(q), f(r) \rightarrow f(p)} 4! \left(\frac{l_{f(\gamma)}(f(r), f(q)) - d_2(f(r), f(q))}{l_{f(\gamma)}^3(f(r), f(q))} \right).$$

Because the spread of f is 1 we have

$l_{f(\gamma)}(f(r), f(q)) = l_\gamma(r, q)$ (see Theorem 2.14). Using this,

and the fact that f is continuous, it follows that

$$K_H^2(f(\gamma), f(p)) = \lim_{q, r \rightarrow p} 4! \left(\frac{l_\gamma(r, q) - \rho(d_1(r, q))}{l_\gamma^3(r, q)} \right).$$

Thus

$$K_H^2(\gamma, p) - K_H^2(f(\gamma), f(p)) = 4! \lim_{q, r \rightarrow p} \frac{\rho(d_1(r, q)) - d_1(r, q)}{l_\gamma^3(r, q)}.$$

For $K_H^2(\gamma, p)$ to exist, it must be the case that

$$\lim_{q, r \rightarrow p} \frac{l_\gamma(r, q) - d_1(r, q)}{l_\gamma(r, q)} = 0.$$

Hence

$$\lim_{q, r \rightarrow p} \frac{d_1(r, q)}{l_\gamma(r, q)} = 1 \quad \text{and so} \quad \lim_{q, r \rightarrow p} \frac{l_\gamma^3(r, q)}{d_1^3(r, q)} = 1.$$

Thus

$$K_H^2(\gamma, p) - K_H^2(f(\gamma), f(p)) = 4! \lim_{q, r \rightarrow p} \frac{\rho(d_1(r, q)) - d_1(r, q)}{d_1^3(r, q)}.$$

Letting $d = d_1(r, q)$, and noting that any sufficiently small distance can be written as $d_1(r, q)$ for some q , it follows that

$$(1) \quad K_H^2(\gamma, p) - K_H^2(f(\gamma), f(p)) = 4! \lim_{d \rightarrow 0} \frac{\rho(d) - d}{d^3}.$$

(b) Assume $\lim_{d \rightarrow 0} \frac{\rho(d) - d}{d^3}$ exists. Let γ be any arc, p

in the interior of γ and suppose that $K_H(\gamma, p)$ exists

(and is finite). As shown earlier, $\lim_{r, q \rightarrow p} \frac{d_1^3(r, q)}{l_\gamma^3(r, q)} = 1$.

Then

$$\begin{aligned}
& \kappa_H^2(\gamma, p) = 4! \lim_{d \rightarrow 0} \frac{\rho(d) - d}{d^3} \\
& = 4! \lim_{q, r \rightarrow p} \frac{l_\gamma(r, q) - d_1(r, q)}{l_\gamma^3(r, q)} = \\
& 4! \lim_{q, r \rightarrow p} \frac{\rho(d_1(r, q)) - d_1(r, q)}{d_1^3(r, q)} \cdot \frac{d_1^3(r, q)}{l_\gamma^3(r, q)} \\
& = 4! \lim_{q, r \rightarrow p} \frac{l_\gamma(r, p) - d_1(r, q) - \rho(d_1(r, q)) + d_1(r, q)}{l_\gamma^3(r, q)} \\
& = 4! \lim_{q, r \rightarrow p} \frac{l_{f(\gamma)}(f(r), f(q)) - d_2(f(r), f(q))}{l_{f(\gamma)}^3(f(r), f(q))} \\
& = \kappa_H^2(f(\gamma), f(p)) .
\end{aligned}$$

The proof that if $\kappa_H^2(f(\gamma), f(p))$ exists then $\kappa_H^2(\gamma, p)$ exists is much the same. $\square$

We wish to emphasize that in the above theorem, it has been shown that for an arbitrary curve γ in M_1 , and an arbitrary point p on γ , the number $\kappa_H^2(\gamma, p) = \kappa_H^2(f(\gamma), f(p)) = \lim_{d \rightarrow 0} \frac{\rho(d) - d}{d^3}$ is constant, regardless of the choice of γ and p .

Remark: Theorem 2.19 can easily be modified to the case when the spread D is not 1. The formula (1) then becomes

$$\kappa_H^2(\gamma, p) - D^2 \kappa_H^2(f(\gamma), f(p)) = \lim_{q, r \rightarrow p} 4! \frac{\frac{1}{D} \cdot \rho(d_1(r, q)) - d_1(r, q)}{d_1^3(r, q)}$$

and the existence of any two of the above terms implies the existence of the third.

This completes Chapter 2. In future chapters concrete metric spaces will be considered, and the results of this chapter will help us to show stronger theorems about metric and order transformations of these spaces.

CHAPTER 3 RESULTS IN NORMED LINEAR SPACES

This chapter is devoted to the study of metric and order transformations whose range and domain are both subsets of real normed linear spaces (N.L.S.). The objective is to discover subsets of N.L.S. which are "metrically determined".

The M.D.S. theorists have had some interest in N.L.S., particularly in the finite dimensional ℓ_p^n spaces. (That is, spaces consisting of points $x = (x_1, \dots, x_n)$, $x_i \in \mathbb{R}$, with the metric $d(x, y) = \left[\sum_{i=1}^n |x_i - y_i|^p \right]^{1/p}$). Beals, Krantz, and Tversky [3] present a set of axioms which, when satisfied by a distance space, are sufficient to guarantee that a space be order embeddable into an ℓ_p^n space. These axioms are satisfied by certain subsets of ℓ_p^n spaces, including open convex subsets. It follows from the work in [3] that any order transformation of these subsets, into any ℓ_p^n space of the same dimension, is a similarity. We extend this result to metric transformations, considering N.L.S. other than the ℓ_p^n spaces, and subsets of them more general than those consider in [3].

Metric transformations of N.L.S. were investigated by Vogt [38]. He showed that any metric transformation of a N.L.S. onto another N.L.S. is a similarity - a result which follows from our work in this chapter.

The study of N.L.S., often involves linear transformations. Mankiewicz [26] showed that any similarity of an open connected subset of a N.L.S. onto an open subset of another N.L.S. can be extended to a similarity of the first space onto the second space, and that this similarity is a linear transformation. This result is used extensively.

This chapter is divided into 3 sections. The first consists of definitions, some known results, and some lemmas which, although necessary for our proofs, do not concern metric transformations. In the second section we investigate the continuity of metric transformations between subsets of N.L.S. Much of this consists of showing when Theorem 2.9 can be applied to the present type of problems.

In the third section the main results are presented. These are Theorems 3.14, 3.16, 3.18 and Corollary 3.19. They give conditions under which metric transformations between subsets of N.L.S. are necessarily similarities. The nicest results are obtained for finite dimensional N.L.S. Theorem 3.16 shows that a metric transformation between open subsets of finite dimensional N.L.S. is a similarity. Theorem 3.18 strengthens this for spaces with strictly convex unit balls.

Before proceeding we should consider order transformations briefly. Although we have not been able to obtain stronger results for them, parts of this chapter can be greatly simplified in this case. In particular much of Section 2, which is devoted to the continuity of metric transformations between subsets of N.L.S. could be simplified, for Lemma 2.2 could be used rather than the more complex Theorem 2.9.

§1. The definitions presented in this section are used throughout the chapter. This section also contains some known results and elementary lemmas.

Throughout this chapter M , M_1 , and M_2 denote normed linear spaces over the field $\mathbb{R}$ of real numbers. If $x \in M$, the norm of x is denoted by $\|x\|$ and the vector dimension of M by $\dim M$.

The following definitions are standard in the geometry of normed linear spaces.

Definitions: An affine transformation is the composition of a linear transformation and a translation. That is, $A : M_1 \rightarrow M_2$ is an affine transformation if and only if $A(x) = Tx + d$, where T is linear and $d \in M_2$. An affine similarity of M_1 onto M_2 is an onto affine transformation, which is also a similarity. An affine similarity of a subset $U \subseteq M_1$ into M_2 is a similarity of U into M_2 which has an extension to an affine similarity of M_1 onto M_2 .

In this chapter, segments, lines, collinearity etc. will be dealt with from an algebraic viewpoint, rather than a metric viewpoint as in Chapter 2. The following definitions are needed:

Definitions: For p, q in a N.L.S. the algebraic segment joining p to q is the set $\{\lambda p + (1 - \lambda)q \mid 0 \leq \lambda \leq 1\}$. This set is denoted by $[p, q]$. The set $\{\lambda p + (1 - \lambda)q \mid 0 < \lambda < 1\}$ is denoted by (p, q) and the set $\{\lambda p + (1 - \lambda)q \mid \lambda \text{ real}\}$ by pq . The set pq is called the line joining p to q . The points p, q , and r are said to be collinear if $r \in pq$.

Definition: A set U in a N.L.S. is said to be algebraically convex if and only if for all $p, q \in U$ the set $[p, q] \subseteq U$. The set U is said to be strictly algebraically convex if and only if for all $p, q \in U$, $p \neq q$, the set (p, q) lies in the interior of U .

Segments and convexity were defined in Chapter 2. It is easily seen that an algebraic segment is a segment and an algebraically convex set is convex. In general, convexity in a N.L.S. does not imply algebraic convexity and a segment is not an algebraic segment. In the case of a N.L.S. with strictly convex unit ball, however, these ideas are equivalent, and stronger results are obtained in these spaces.

Note that the results of Chapter 2 remain valid when algebraic segments and algebraic convexity are used.

Throughout this chapter, "segment" and "convex" will mean "algebraic segment" and "algebraically convex".

Notation: For $p \in M$, define

$$B(p,r) = \{x \in M \mid \|p-x\| < r\}$$

$$S(p,r) = \{x \in M \mid \|p-x\| = r\}.$$

For $U \subseteq M$, $\bar{U}$ will denote the closure of U in M .

Note that $\overline{B(p,r)} = B(p,r) \cup S(p,r)$, so the boundary of $\overline{B(p,r)}$ is $S(p,r)$. Also, for $q,s \in \overline{B(p,r)}$, $0 < \lambda < 1$,

$$\|p - (\lambda q + (1-\lambda)s)\| = \|\lambda(p-q) + (1-\lambda)(p-s)\| \leq \lambda\|p-q\| + (1-\lambda)\|p-s\| \leq r.$$

Thus, $\overline{B(p,r)}$ is algebraically convex.

Definition: The unit ball of a N.L.S. M is the set $\overline{B(0,1)}$, where 0 is the origin in M .

Since $B(p,r) = rB(0,1) + p$, it follows that if $\overline{B(0,1)}$ is strictly convex, then so is $\overline{B(p,r)}$ for any p and r .

The following theorem is probably the most important tool used in this chapter.

Theorem 3.1: (Mankiewicz) Let M_1 and M_2 be N.L.S. Let U be a non-empty open connected subset of M_1 , and let V be an open subset of M_2 . Then every isometry (similarity) of U onto V is an affine isometry (similarity).

Using this theorem it is often only necessary to show that a metric transformation is a similarity to conclude that it is also affine.

Lemma 3.2: Two affine transformations A_1 and A_2 from M_1 onto M_2 which agree on a non-empty open set U of M_1 are identical.

The proof of lemma 3.2 is a standard vector space argument so it will not be reproduced here. It, and Corollaries 3.3 and 3.4 are proved in [26].

Corollary 3.3: Let U and V be open subsets of M_1 , $f: U \cup V \rightarrow M_2$ any function. Assume $U \cap V \neq \emptyset$. If $f|_U$ and $f|_V$ are affine similarities, then so is f .

This idea can be extended to obtain the following.

Corollary 3.4: Let U be a connected open subset of M_1 and $g: U \rightarrow M_2$ be such that for some open covering $\mathcal{C}$ of U , $g|_C$ is an affine similarity (isometry) for each $C \in \mathcal{C}$. Then g is an affine similarity (isometry).

Our best results have been for finite dimensional N.L.S. We are able to obtain these results because of the following theorem from algebraic topology, and some of its consequences. These consequences (Corollary 3.6) are standard results of dimension theory.

Theorem 3.5: (Invariance of Domain) If $g: U \rightarrow E^n$ is bicontinuous and U is open in E^n , then $g(U)$ is open in E^n .

Proof: This, of course, is a basic theorem of topological dimension theory.

Corollary 3.6: Let M_1 and M_2 be N.L.S.,
 $\dim M_1 = n < \infty$, $\dim M_2 = m \leq \infty$. Assume $B(p,r) \subseteq U \subseteq M_1$,
 and let $g : U \xrightarrow{\text{onto}} V \subseteq M_2$ be bicontinuous. Then

(a) $m \geq n$

(b) If $m > n$, then for any subset B of U we
 have $g(\overline{B} \cap U) = \overline{g(B \cap U)} \cap V$ and this set
 contains no open subset of M_2 .

(c) If $m = n$ and U is open in M_1 then $g(U)$
 is open in M_2 .

Proof: If $m < \infty$, then because the topologies on any
 two finite dimensional N.L.S. of the same dimension are
 equivalent, it is necessary only to prove these for $M_1 = E^n$,
 $M_2 = E^m$ (Euclidean spaces). Thus (c) follows directly from
 Theorem 3.5.

(a) Assume $m < n$. Then $m < \infty$ and we need only
 consider $M_1 = E^n$, and $M_2 = E^m$. Consider E^m as a
 subspace of E^n . Then $g : B(p,r) \rightarrow E^m \subseteq E^n$, and by
 Theorem 3.5, $g(B(p,r))$ is open in E^n . However as
 $g(B(p,r)) \subseteq E^m$ this cannot be true. Thus $m \geq n$, and
 (a) has been proved.

(b) That $g(\overline{B} \cap U) = \overline{g(B \cap U)} \cap V$ for any subset B
 of U follows directly from the fact that g is bicontinuous.
 Assume $m > n$. If $g(\overline{B} \cap U)$ contains an open subset V' of
 M_2 , let $M \subseteq M_2$ be a linear subspace of M_2 , of finite
 dimension greater than n , such that $M \cap V' \neq \emptyset$. Then

$g^{-1}|_{M \cap V'}$ is bicontinuous from $M \cap V' \subseteq M$ into M_1 ,

hence by part (a) $\dim M \leq n$. This contradicts the choice of M . Thus $g(\bar{B} \cap U)$ contains no open subsets of M_2 . $\square$

To conclude this section three lemmas are presented. The second, Lemma 3.8 is a commonly used property of convex quadrilaterals. Lemma 3.7 is presented only to prove Lemma 3.8. Lemma 3.9 has surely been shown by others, although we have not seen a proof of it. It does not extend to 3 dimensions. Note that Lemmas 3.7 and 3.8, while stated in terms of N.L.S., are in more general metric spaces.

Definition: Three vectors a_1, a_2 , and a_3 are said to satisfy the triangle equality if and only if for some permutation (i, j, k) of $(1, 2, 3)$, $\|a_i - a_j\| + \|a_j - a_k\| = \|a_i - a_k\|$.

Lemma 3.7: Let M_1 be a N.L.S. with a strictly convex unit ball. Then any three vectors are collinear if and only if they satisfy the triangle equality.

Proof: If three vectors are collinear it is easy to check that they satisfy the triangle equality.

To prove the converse assume the vectors are p, q , and r and that $\|p - r\| + \|r - q\| = \|p - q\|$. We show that $r \in [p, q]$.

If any two of p, q , or r are equal it is easy to see that $r \in [p, q]$. Assume they are distinct. Then none of $\|p - q\|$, $\|p - r\|$, $\|r - q\|$ is zero, and

$$\frac{\|p - r\|}{\|p - q\|} \cdot \frac{p - r}{\|p - r\|} + \frac{\|r - q\|}{\|p - q\|} \cdot \frac{r - q}{\|r - q\|} = \frac{p - q}{\|p - q\|}.$$

Also

$$1 - \frac{\|p-r\|}{\|p-q\|} = \frac{\|p-q\| - \|p-r\|}{\|p-q\|} = \frac{\|r-q\|}{\|p-q\|} .$$

Thus, letting $\lambda = \frac{\|p-r\|}{\|p-q\|}$, we have

$$\lambda \frac{p-r}{\|p-r\|} + (1-\lambda) \frac{r-q}{\|r-q\|} = \frac{p-q}{\|p-q\|} .$$

Certainly $0 < \lambda < 1$, and the three vectors $\frac{p-r}{\|p-r\|}$, $\frac{r-q}{\|r-q\|}$, $\frac{p-q}{\|p-q\|}$ all lie on the boundary of the unit ball. By the definition of strict convexity, they must all be equal. In particular

$$\frac{p-r}{\|p-r\|} = \frac{p-q}{\|p-q\|} .$$

Rearranging this last equation we obtain $r = \lambda q + (1-\lambda)p$, concluding the proof of Lemma 3.7. $\square$

Lemma 3.8: In a 2-dimensional N.L.S. let a, b, c, d be the vertices of a convex quadrilateral, given in a cyclic order around the edges of the quadrilateral. Then

$$(a) \quad \|a-b\| + \|c-d\| \leq \|a-c\| + \|b-d\| .$$

That is, the lengths of two opposite sides add up to no more than the sum of the lengths of the two diagonals.

(b) If in addition a, b, c , and d are distinct, not all collinear, and the unit ball is strictly convex then the inequality above is strict.

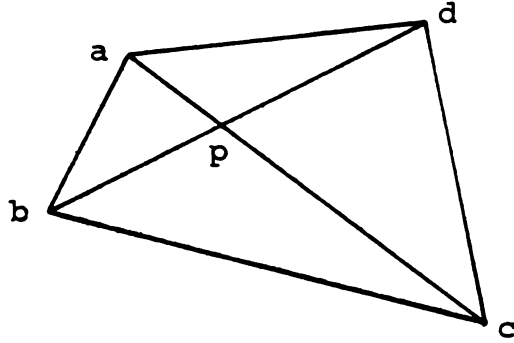


Figure 3.1

Proof: The proof rests on the fact that the diagonals $[a, c]$ and $[b, d]$ intersect at a point p which lies in the interior, or on the boundary of, the quadrilateral $abcd$.

$$(1) \quad \begin{aligned} \|a - c\| &= \|a - p\| + \|p - c\| \\ \|b - d\| &= \|b - p\| + \|p - d\| . \end{aligned}$$

However

$$(2) \quad \begin{aligned} \|a - b\| &\leq \|a - p\| + \|p - b\| \\ \|c - d\| &\leq \|d - p\| + \|p - c\| . \end{aligned}$$

Adding the two inequalities of (2) and combining with (1) we obtain

$$\|a - b\| + \|c - d\| \leq \|a - c\| + \|b - d\| .$$

To show part (b) of the lemma it is only necessary to show that one of the inequalities of (2) is strict. If not, Lemma 3.7 shows that a, b , and p are collinear and that c, d , and p are collinear. Since $p \in ac \cap bd$ we can readily conclude that either not all the points are distinct, or all are

collinear, contradicting the hypothesis of (b). Thus the inequality must be strict. $\square$

Lemma 3.9: In any two dimensional N.L.S. with strictly convex unit ball, there is at most one point equidistant from three distinct points.

Proof: Let p, q, r, c_1 , and c_2 be such that $\|p - c_i\| = \|q - c_i\| = \|r - c_i\|$, $i = 1, 2$. Also assume $c_1 \neq c_2$, and p, q, r are distinct. If p, q , and r are collinear, the strict convexity of the unit ball is violated. Consider Figure 2. If either c_1 or c_2 lies in one of the closed regions 1, 2, or 3 the convexity of the unit sphere is violated. The line c_1c_2 can intersect at most two of the segments (p, q) , (q, r) , or (p, r) . Assume it does not intersect (p, q) . Then either pqc_1c_2 or pqc_2c_1 forms a convex quadrilateral. Assume the latter. The points p, q, c_1, c_2 are distinct and not collinear, so by

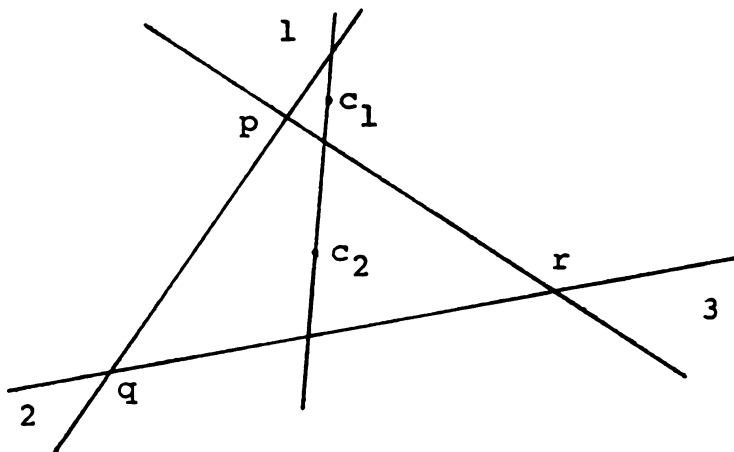


Figure 3.2

Lemma 3.8(b)

$$\|p - c_2\| + \|q - c_1\| > \|p - c_1\| + \|q - c_2\| .$$

However $\|p - c_2\| = \|q - c_2\|$ and $\|p - c_1\| = \|q - c_1\|$, giving a contradiction. Thus $c_1 = c_2$. $\square$

§2. We are now prepared to investigate the continuity of metric transformations between subsets of N.L.S. Theorem 2.9 gave sufficient conditions for a metric transformation to be continuous. This section proceeds by determining when these conditions are satisfied. We recall the definition of the long legged local isosceles property.

Definition: The metric space M has the long legged local isosceles property at P if and only if if there is a number $\lambda(p) > 0$ such that for any $\delta < \lambda(p)$ and for any q with $d(p,q) < \delta$ there exists $s \in M$ with $d(p,s) = d(s,q) = \delta$. The number $\lambda(p)$ may be taken to be ∞ .

There are four results in this section. These are Lemmas 3.10, 3.11, 3.13 and Theorem 3.12. Lemma 3.10 shows that any point p of an open subset of a N.L.S. has the long legged local isosceles property. Lemmas 3.11 and 3.13 then show sufficient conditions for the continuity of a metric transformation between subsets of N.L.S. Lemma 3.13 differs from Lemma 3.11 in that it considers only finite dimensional N.L.S., obtaining stronger results in this case. Theorem 3.12 is a theorem of Vogt, which follows easily from our work.

Lemma 3.10: Let U be a subset of a N.L.S. M , with $\dim M \geq 2$. Assume that p and $\lambda > 0$ are such that $B(p, \lambda) \subseteq U$. Then U has the long legged local isosceles property at p , and $\lambda(p)$ can be taken to be λ .

Proof: Let $\delta < \lambda$ be given, and let q be given with $\|p - q\| < 2\delta$. If $S(p, \delta) \cap S(q, \delta) = \emptyset$ then

$$S(q, \delta) = (B(p, \delta) \cap S(q, \delta)) \cup (\{x \mid d(p, x) > \delta\} \cap S(q, \delta)) .$$

Since both members of the above union are non-empty (for example $\frac{\delta(p-q)}{\|p-q\|} + q$ is in the first member, while $-\frac{\delta(p-q)}{\|p-q\|} + q$ is in the second), both members are open subsets of $S(q, \delta)$, and $S(q, \delta)$ is connected, we have a contradiction. Thus, $S(p, \delta) \cap S(q, \delta) \neq \emptyset$. Let $s \in S(p, \delta) \cap S(q, \delta)$. Then $d(p, s) = d(q, s) = \delta$. Since $d(p, s) = \delta < \lambda$ then $s \in B(p, \lambda)$ and hence $s \in U$. $\square$

Remark: In the above lemma, one could replace $B(p, \lambda)$ by any subset B , $p \in B \subseteq U$, such that B is an isometric image of an open ball of radius λ in a N.L.S. of dimension 2. Then it would be necessary to verify only the case $\dim M = 2$ and the result would follow immediately. Lemma 3.10 is used in proving Lemma 3.11, for it allows Theorem 2.9, on the continuity of metric transformation to be used. Example 2 at the end of this chapter shows the necessity of the assumption that $\dim M_1 \geq 2$.

There are undoubtedly many conditions which force a metric transformation from a subset $U \subseteq M_1$ to M_2 to be continuous. Lemma 3.11 presents two such conditions, the proofs of which use Theorem 2.9.

Lemma 3.11: Let $f: U \rightarrow M_2$ be a metric transformation, $U \subseteq M_1$, $\dim M_1 \geq 2$. Let p and $\lambda > 0$ be such that $B(p, \lambda) \subseteq U$.

- (a) If $f(B(p, \lambda/2))$ is not discrete then $f|_{B(t, \lambda/2) \cap U}$ is bicontinuous for each $t \in U$, and f is uniformly continuous.
- (b) If M_2 is separable, then $f|_{B(t, \lambda/2) \cap U}$ is either bicontinuous for each $t \in U$, or is trivial for each $t \in U$. If $f|_{B(t, \lambda/2) \cap U}$ is trivial for each $t \in U$, then $f|_C$ is trivial for each connected component C of U .

Proof: As $B(p, \lambda) \subseteq U$, Lemma 3.10 shows that $p \in U$ has the long-legged local isocles property (as a point of the metric space U) with $\lambda(p) = \lambda$.

- (a) If $f(p) = f(q)$ for all $q \in B(p, \lambda/2)$ then $f(B(p, \lambda/2))$ is a singleton, hence discrete. Thus there is a $q \in B(p, \lambda/2)$ with $f(p) \neq f(q)$ and letting $\gamma = [p, q] \subseteq U$, Theorem 2.9 shows that $f|_{B(t, \lambda/2)}$ is bicontinuous for all $t \in U$, and that f is uniformly continuous.

(b) Assume that there is a $q \in B(p, \lambda/2)$ with $f(p) \neq f(q)$. If $f([p, q])$ is uncountable then, because M_2 is separable, $f([p, q])$ contains a limit point of $f([p, q])$, (in fact a condensation element). Hence $f(B(p, \lambda/2))$ is not discrete. Letting $\gamma = [p, q]$ Theorem 2.9 now shows that $f|_{B(t, \lambda/2)}$ is bicontinuous for all $t \in U$.

If $f(p) = f(q)$ for all $q \in B(p, \lambda/2)$ then certainly $\rho(d) = 0$ for $0 < d < \lambda$. The proof follows by Lemma 2.6. $\square$

At this point the following theorem, due to Vogt [38], is easy to show.

Theorem 3.12: (Vogt) Let M_1 and M_2 be N.L.S., and $\dim M_1 \geq 2$. Let $f: M_1 \xrightarrow{\text{onto}} M_2$ be a metric transformation. Then f is an affine similarity.

Proof: By Lemma 3.10, the origin has the long legged local isosceles property, with $\lambda = \infty$. Since $M_2 = f(M_1) = f(B(0, \frac{\lambda}{2}))$ is not discrete, Theorem 2.9 shows that f is bicontinuous. Theorem 2.17 then shows that f is a similarity, and by the theorem of Mankiewicz (Theorem 3.1) we conclude that f is an affine similarity. $\square$

Remark: Corollary 3.15 generalizes this to the case $f: M_1 \xrightarrow{\text{onto}} V \subseteq M_2$, V open in M_2 . It is shown there that necessarily $V = M_2$.

Although our proof of Theorem 3.12 is somewhat different than Vogt's, our work on the continuity of metric transformations

(Theorem 2.9) was motivated by his work in [38] where he proves the above theorem. Mankiewicz makes use of techniques from this same paper in proving Theorem 3.1.

Lemma 3.13 strengthens Lemma 3.11 when separable and finite dimensional N.L.S. are considered. It is here that the Invariance of Domain Theorem (Theorem 3.5) is used.

Lemma 3.13: Let M_1 and M_2 be separable N.L.S. and $\dim M_1 \geq 2$. Let $f: U \subseteq M_1 \xrightarrow{\text{onto}} V \subseteq M_2$ be a metric transformation. Assume that there exists $p, q \in M_1$ and $r_1, r_2 > 0$ such that $B(p, r_1) \subseteq U$ and $B(f(q), r_2) \subseteq V$. Then $f|_{B(t, r_1/2) \cap U}$ is bicontinuous for all $t \in U$ and f is uniformly continuous. In particular $f|_{B(p, r_1/2)}$ is bicontinuous and nontrivial. If in addition, $\dim M_1 = n < \infty$, then $\dim M_2 = n$.

Proof: As M_2 is separable, and $B(p, r_1) \subseteq U$, Lemma 3.11 shows $f|_{B(t, r_1/2) \cap U}$ is either trivial for each $t \in U$ or bicontinuous for each $t \in U$. Since M_1 is separable, and a subspace of a separable metric space is separable, then U is separable. Let T be a countable dense subset of U . Let $\mathcal{C}$ be the countable collection of subsets of U of the form $\overline{B(t, r_1/4)} \cap U$, $t \in T$. Note $\mathcal{C}$ covers U .

If $f|_{B(t, r_1/2) \cap U}$ is trivial for each $t \in T$, then $V = \bigcup_{t \in T} f(\overline{B(t, r_1/4)} \cap U)$ would consist of a countable set of points. Since $B(f(q), r_2) \subseteq V$, this is impossible. Thus it

follows that $f|_{B(t, r_1/2) \cap U}$ is bicontinuous for each $t \in U$. In particular, $f|_{B(p, r_1/2)}$ is bicontinuous, and hence non-trivial.

If $\dim M_1 = n < \infty$, then Corollary 3.6(a) shows $\dim M_2 \geq n$. Assume $\dim M_2 > n$. Then by Corollary 3.6(b)

$$\overline{f(B(t, r_1/4) \cap U)} = \overline{f(B(t, r_1/4) \cap U) \cap V}$$

and this set contains no open subsets of M_2 . Hence

$$\begin{aligned} B(f(q), r_2) &= \left[\bigcup_{t \in T} f(B(t, r_1/4) \cap U) \right] \cap B(f(q), r_2) \\ &= \bigcup_{t \in T} [f(B(t, r_1/4) \cap U) \cap B(f(q), r_2)] \end{aligned}$$

Thus $B(f(q), r_2)$ is the countable union of closed subsets of the subspace $B(f(q), r_2)$ of M_2 , none of which contains an open subset of $B(f(q), r_2)$, (an open subset of $B(f(q), r_2)$ would also be open in M_2). This contradicts the Baire category theorem. Thus $\dim M_2 = n$, completing the proof of Lemma 3.13. $\square$

This concludes our study of the continuity of metric transformations between subsets of N.L.S. These results are used in Section 3 to establish theorems showing when such transformations are similarities.

§3. The main results of the chapter are in this section. These are Theorems 3.14, 3.16, 3.18, and Corollary 3.19. As has been stated before, the main goal of this thesis is to consider the "uniqueness question". That is, if U and V are spaces in a class $\mathcal{C}$ of distance spaces, and $f: U \xrightarrow{\text{onto}} V$ is an order transformation, is f a similarity? In this chapter, U and V are subsets of normed linear spaces, called M_1 and M_2 respectively, and transformations, f , are invariably metric rather than order transformations.

The class $\mathcal{C}$ just referred to, may be taken to be some suitable collection of subsets of M_1 . For example, one of the hypotheses of Theorem 3.14 is that the domain and range be open, hence a suitable choice for $\mathcal{C}$ would be all open subsets of M_1 .

Each of Theorems 3.14, 3.16, and 3.18 and Corollary 3.19 involve some type of openness hypothesis on the domain and/or range of f . This is done in order to use methods developed in Chapter 2 for convex sets. In addition, some further hypotheses is needed. For example, in Corollary 3.19, M_1 and M_2 are assumed to have the same finite dimension.

Note that these results could be considered as generalizations of Mankiewicz's Theorem (Theorem 3.1) to metric transformations.

Definition: A function is said to be open if and only if it maps open sets onto open sets.

Theorem 3.14: Let M_1 and M_2 be N.L.S., $\dim M_1 \geq 2$. Let f be an open metric transformation, from an open connected set U of M_1 onto an open set V of M_2 . Then f is an affine similarity.

Proof: Let p , and $\lambda > 0$ be such that $B(p, \lambda) \subseteq U$. Then by hypothesis, $f(B(p, \lambda))$ is open, hence not discrete. So, by Lemma 3.11(a), $f|_{B(t, \lambda/2) \cap U}$ is bicontinuous for each $t \in U$, and f is uniformly continuous.

Let t be an arbitrary point of U . By hypothesis, $f(B(t, \lambda/2) \cap U)$ is open in V , and hence in M_2 . Let r_2 be such that $B(f(t), r_2) \subseteq f(B(t, \lambda/2) \cap U)$. As f is continuous $f^{-1}(B(f(t), r_2))$ is open in U , and hence in M_1 . Let r_1 be such that

$$B(t, r_1) \subseteq f^{-1}(B(f(t), r_2)) \cap B(t, \lambda/2) .$$

Then $f(B(t, r_1)) \subseteq B(f(t), r_2) \subseteq f(B(t, \lambda/2) \cap U)$. Since $B(t, \lambda/2) \cap U$ and $f(B(t, \lambda/2) \cap U)$ contain segments (which are rectifiable arcs), and $f|_{B(t, \lambda/2) \cap U}$ is bicontinuous, Theorem 2.16 shows that f has finite spread D . For q_1 and q_2 in $B(t, r_1) \subseteq U$, the algebraic segment $[q_1, q_2] \subseteq B(t, r_1)$ and hence by Corollary 2.14 $f([q_1, q_2])$ is a geodesic arc in $f(B(t, \lambda/2) \cap U)$ of length $D \cdot \|q_2 - q_1\|$.

On the other hand,

$$[f(q_1), f(q_2)] \subseteq B(f(t), r_2) \subseteq f(B(t, \lambda/2) \cap U)$$

and has length $\|f(q_2) - f(q_1)\|$. As $[f(q_1), f(q_2)]$ is an arc, then $D \cdot \|q_2 - q_1\| = \|f(q_1) - f(q_2)\|$. (This argument is used in the proof of Theorem 2.16.) Since q_1 and q_2 are arbitrary points of $B(t, r_1)$, we conclude that $f|_{B(t, r_1)}$ is a similarity. Since t is an arbitrary point of U , Corollary 3.4 shows that f is an affine similarity. $\square$

The following generalizes Vogt's Theorem 3.12.

Corollary 3.15: If $f : M_1 \xrightarrow{\text{onto}} V \subseteq M_2$, $\dim M_1 \geq 2$, V a non-empty open subset of M_2 , and f is a metric transformation, then f is an affine similarity and $V = M_2$.

Proof: As in the proof of Theorem 3.12, f is necessarily bicontinuous, hence is open. So by Theorem 3.14, f is an affine similarity. Thus, for each $x \in M_1$, $f(x) = T(x) + d$ where T is a linear transformation and $d \in M_2$.

To show $V = M_2$, let $y \in M_2$ and $p \in M_1$. Then because V is open there is a $\lambda \neq 0$ such that $(1 - \lambda)f(p) + \lambda y \in V$. Hence there is a $q \in M_1$ such that $f(q) = (1 - \lambda)f(p) + \lambda y$. Then

$$\begin{aligned} y &= \frac{f(q) - (1 - \lambda)f(p)}{\lambda} \\ &= \frac{T(q) - (1 - \lambda)T(p)}{\lambda} + d \\ &= T\left(\frac{q - (1 - \lambda)p}{\lambda}\right) + d \\ &= f\left(\frac{q - (1 - \lambda)p}{\lambda}\right) \end{aligned}$$

and hence $y \in f(M_1) = V$, so $V = M_2$. $\square$

Theorem 3.16: Let M_1 and M_2 be N.L.S., $2 \leq \dim M_1 < \infty$. Let f be a metric transformation, $f: U \xrightarrow{\text{onto}} V$, U an open connected subset of M_1 , and $V \subseteq M_2$.

(a) If V contains an open subset of M_2 and M_2 is separable, then f is an affine similarity and $\dim M_1 = \dim M_2$.

(b) If $\dim M_1 = \dim M_2$, then f is either trivial or an affine similarity.

Proof: If V contains an open subset of M_2 , then Lemma 3.13 shows that $\dim M_1 = \dim M_2$, and f is non-trivial. Thus if we assume $\dim M_1 = \dim M_2$, and f is non-trivial, and then show that f is an affine similarity both (a) and (b) will have been proven. This is what is now done.

Assume $\dim M_1 = \dim M_2$ and f non-trivial. Let p and λ be such that $B(p, \lambda) \subseteq U$. Then by Lemma 3.11(b) (because M_2 is a finite dimensional vector space, it is separable) $f|_{B(t, \lambda/2) \cap U}$ is either bicontinuous for each $t \in U$ or trivial for each $t \in U$. Furthermore, Lemma 3.11(b) shows that if $f|_{B(t, \lambda/2) \cap U}$ is trivial for each $t \in U$, then $f|_C$ is trivial for any connected component C of U . Since U is assumed to be connected, this would imply that f is trivial, contradicting the assumption that f is non-trivial. Thus we may assume that $f|_{B(t, \lambda/2) \cap U}$ is bicontinuous for each $t \in U$.

As $B(t, \lambda/2) \cap U$ is an open subset of M_1 , and by the assumption $\dim M_1 = \dim M_2$, Corollary 3.6(c) shows that $f(B(t, \lambda/2) \cap W)$ is open in M_2 , for any $t \in U$, and any open subset $W \subseteq U$. Thus, given an arbitrary open set $W \subseteq U$, $f(W) = \bigcup_{t \in U} f(B(t, \lambda/2) \cap W)$ is an open subset of M_2 , so f is an open mapping. Theorem 3.14 now shows that f is an affine similarity. $\square$

Note that Theorem 3.16 (a) generalizes Mankiewicz's theorem to metric transformations between open subsets of finite dimensional N.L.S.

We have tried, without success, to establish the results of Theorem 3.16 for the case in which both M_1 and M_2 are separable. Theorem 3.17 summarizes the progress we have made.

Theorem 3.17: Let M_1 and M_2 be separable N.L.S., $\dim M_1 \geq 2$. Let $U \subseteq M_1$, $V \subseteq M_2$ be open, and let $f: U \xrightarrow{\text{onto}} V$ be a metric transformation with scale function ρ . Then there is a $\lambda > 0$ such that $f|_{B(t, \lambda/2) \cap U}$ is bicontinuous for all $t \in U$, and an $r > 0$ and $q \in U$ such that $f|_{B(q, r)}$ is an affine similarity, and $f|_{B(t, r) \cap U}$ is a similarity for each $t \in U$.

Remark: If it could be shown that $f(B(t, r))$ is open in M_2 for all t , it would follow that $f|_{B(t, r)}$ is an affine similarity and Theorem 3.1 would show that f is an affine isometry.

Proof of Theorem 3.17: Let p and λ be such that $B(p, \lambda) \subseteq U$. Then Lemma 3.13 shows that $f|_{B(t, \lambda/2) \cap U}$ is bicontinuous for all t . As in the proof of Lemma 3.13, for some $b \in U$, $f(B(b, \lambda/2) \cap U)$ must contain an open subset of M_2 . Let $f(a)$, and r_1 be such that $B(f(a), r_1) \subseteq f(B(b, \lambda/2) \cap U)$. Now $f|_{B(b, \lambda/2) \cap U}$ is bicontinuous, hence $W = (f|_{B(b, \lambda/2) \cap U})^{-1}(B(f(a), r_1))$ is open in $B(b, \lambda/2)$, and hence open in M_1 . Let $B(q, r) \subseteq W$, and $g \equiv f|_{B(q, r)}$. Then g is a bicontinuous metric transformation of $B(q, r)$. The set $g(B(q, r))$ is open in M_2 , for $g(B(q, r)) = f(B(q, r)) \subseteq f(W) = B(f(a), r_1)$ and $f|_{B(b, \lambda/2) \cap U}$ is bicontinuous. It then follows from Theorem 3.14 that

$g \equiv f|_{B(q,r)}$ is an affine similarity, hence $\rho(d) = D \cdot d$ for some $D > 0$, and for all $d < 2r$ so $f|_{B(t,r) \cap U}$ is a similarity for all t . $\square$

When working in a N.L.S. with a strictly algebraically convex unit ball (see the definition at the beginning of this chapter) we are able to obtain results which are stronger than those of Theorem 3.14 and 3.16. These results, contained in Theorem 3.18 and Corollary 3.19, might be suprising since the set $U \subseteq M_1$ is not assumed to be connected.

Theorem 3.18: Let M_1 , $\dim M_1 \geq 2$ be a N.L.S. with a strictly convex unit ball. Let $U \subseteq M_1$ be any set with non-empty interior. Let f be a metric transformation, $f: U \rightarrow f(U) \subseteq M_2$. Assume for some $p \in U$, $\lambda > 0$ that $B(p, \lambda) \subseteq U$, and $f(B(p, \lambda/2))$ is open in M_2 . Then f is an affine similarity.

More important than this theorem may be the following corollary.

Corollary 3.19: Let M_1 , $2 \leq \dim M_1 < \infty$, be a N.L.S. with a strictly convex unit ball. Let $U \subseteq M_1$ be any non-empty set which contains an open subset of M_1 . Let $f: U \rightarrow M_2$ be a metric transformation,

(a) If $\dim M_1 = \dim M_2$ then either f is an affine similarity or $f|_C$ is trivial for each connected component C of U .

(b) If $f(U)$ contains an open subset of M_2 , and M_2 is separable, then f is an affine similarity.

Proof of Corollary 3.19:

(a) Assume $\dim M_1 = \dim M_2$. Note that $\dim M_2$ is finite, hence M_2 is separable. Let $p \in U$ and $\lambda > 0$ be such that $B(p, \lambda) \subseteq U$. Then by Lemma 3.11(b), either $f|_C$ is trivial for each connected component of C of U or $f|_{B(t, \lambda/2) \cap U}$ is bicontinuous for each $t \in U$. In the latter case, $f|_{B(p, \lambda/2)}$ is bicontinuous, hence by Corollary 3.6(c), $f(B(p, \lambda/2))$ is open in M_2 . Part (a) now follows from Theorem 3.18.

(b) If $f(U)$ contains an open subset of M_2 , and M_2 is separable, Lemma 3.13 shows that $n = m$ and $f|_{B(p, \lambda/2)}$ is bicontinuous. Hence $f(B(p, \lambda/2))$ is open in M_2 (Corollary 3.6(c)). Then (b) follows from Theorem 3.18. $\square$

Proof of Theorem 3.18: Let $g = f|_{B(p, \lambda/2)}$. Then g is a metric transformation, $g : B(p, \lambda/2) \rightarrow f(B(p, \lambda/2))$. Since $f(B(p, \lambda/2))$ is open, it cannot be discrete. Hence Lemma 3.11(a) shows that g is bicontinuous, and so is an open function. From this it follows by Theorem 3.14 that g is an affine similarity. Thus $g = A|_{B(p, \lambda/2)}$ where A is an affine transformation and also a similarity of M_1 onto M_2 . Let $\alpha > 0$ be such that $\|A(x) - A(y)\| = \alpha\|x - y\|$, for all $x, y \in M_1$.

Let $q \in U$ and consider $f(q)$ and $A(q)$. Let $d > 0$ be such that $S(q, d) \cap B(p, \lambda/2) \neq \emptyset$. Let π be any (two

dimensional) plane in M_2 containing $f(q)$ and $A(q)$, and intersecting $g(S(q,d) \cap B(p,\lambda/2))$. Since A is an affine similarity from $M_1 \xrightarrow{\text{onto}} M_2$ and $g \equiv A|_{B(p,\lambda/2)}$ we have

$$\begin{aligned} g(S(q,d) \cap B(p,\frac{\lambda}{2})) &= A(S(q,d) \cap B(p,\frac{\lambda}{2})) \\ &= S(A(q),\alpha d) \cap B(A(p),\frac{\alpha\lambda}{2}). \end{aligned}$$

The set $C = \pi \cap S(A(q),\alpha d) \cap B(A(p),\frac{\alpha\lambda}{2})$ is the intersection of the "circle" in π , (with center $A(q)$ and radius αd) with the set $B(A(p),\frac{\alpha\lambda}{2})$, hence is an open subset of the circle.

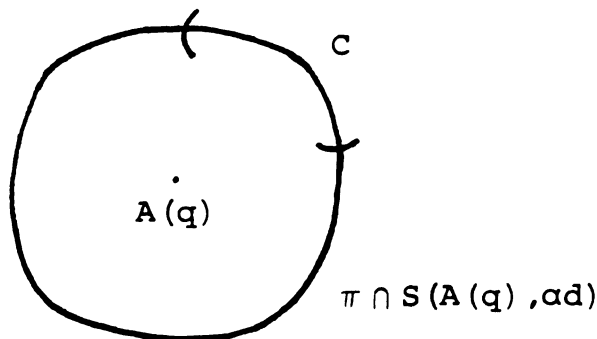


Figure 3.3

Since C is non empty, it contains an infinite number of points of $M_2 = A(M_1)$. Let $A(x_1), A(x_2), A(x_3) \in C$. Then

$$(1) \quad \|A(x_1) - A(q)\| = \|A(x_2) - A(q)\| = \|A(x_3) - A(q)\|.$$

On the other hand,

$$\|A(x_i) - A(q)\| = \alpha \|x_i - q\|, \quad i = 1, 2, 3.$$

Hence,

$$\|x_1 - q\| = \|x_2 - q\| = \|x_3 - q\| .$$

As f is a metric transformation,

$$(2) \quad \|f(x_1) - f(q)\| = \|f(x_2) - f(q)\| = \|f(x_3) - f(q)\| .$$

Since π is the translation of a two dimensional N.L.S. with strictly convex unit ball and $A(x_i) = f(x_i)$ for each i we can apply Lemma 3.9 to the equations (1) and (2) to see $A(q) = f(q)$. Now q being an arbitrary point of U , we have $f \equiv A|_U$, and f is thus an affine similarity. $\square$

Corollary 3.19 is our final result showing metric transformations between subsets of N.L.S. are similarities. Theorems 3.14, 3.16, 3.18 and Corollary 3.19 would seem to give some validity to the statement "the order of the distances determines the set up to a similarity". On the other hand, we have looked at specialized situations and undoubtedly there are many order transformations between subsets of N.L.S. which are not similarities. Unfortunately the only type of example we have been able to find has a disconnected domain - Example 1.

Example 1 shows that the strict convexity of the unit ball is necessary for Theorem 3.18, and that the connectedness condition of Theorem 3.14 is also necessary.

Example 1: Let $M_1 = l_\infty^2$. That is, M_1 is the set of all ordered pairs (x, y) with $\|(x, y)\| = \max\{x, y\}$. Let

$$S_1 = \{ (x,y) \mid \| (x,y) \| < 1 \}$$

$$S_2 = \{ (x,y) \mid \| (x,y) - (4,0) \| < 1 \} .$$

For any real number $a > 0$, define f by

$$f(x,y) = (x,y), \quad (x,y) \in S_1$$

$$f(x,y) = (x+a,y), \quad (x,y) \in S_2 .$$

Then f is a metric transformation, with scale function

$$\rho(d) = \begin{cases} d, & d < 2 \\ d+a, & d > 2 \end{cases}, \text{ but } f \text{ is certainly not a similarity. } \square$$

Throughout this chapter it has been assumed $\dim M_1 \geq 2$. This was used to obtain the continuity of the metric transformation. In the case of order transformations, as has been mentioned, Lemma 2.2 rather than Theorem 2.9 can be used to show continuity. In this case all of the theorems could be changed to include $\dim M_1 = 1$.

To show the assumption $\dim M_1 \geq 2$ is necessary for metric transformations, consider the following, due to Vogt [38].

Example 2: (Vogt) Consider $\mathbb{R}$, the real numbers, as a vector space over the field $\mathbb{Q}$, the rational numbers. Let A be a basis for this vector space, and assume $1 \in A$. For $a \in A$, define f by $f(a) = \begin{cases} 2a, & a \neq 1 \\ a, & a = 1 \end{cases}$ and extend f to $\mathbb{R}$ by linearity.

Now considering $\mathbb{R}$ as a N.L.S., (of dimension 1) the function $f : \mathbb{R} \rightarrow \mathbb{R}$ is a metric transformation because

$$|f(x) - f(y)| = |f(x - y)| = |\pm f(|x - y|)| = |f(|x - y|)|$$

and the scale function of f is $\rho(d) = |f(d)|$.

On the other hand f is not a similarity for $\rho(1) = 1$, while $\rho(|a|) = 2|a|$, $a \in A$, $a \neq 1$.

More generally, define a metric transformation $f: \mathbb{R} \rightarrow M$, M any N.L.S., by defining $f(a)$, $a \in A$ arbitrary and then extending f by linearity to $\mathbb{R}$ (treating $\mathbb{R}$ as a vector space over Q). As above

$$\|f(x) - f(y)\| = \|f(x - y)\| = \|\pm f(|x - y|)\| = \|f(|x - y|)\|.$$

This may be stated more simply by noting that f is a group homomorphism from $\mathbb{R}$ into M , although possibly the above description may yield more insight. Thus many non-continuous metric transformations of $\mathbb{R}$ into any N.L.S. can easily be constructed. Even metric transformations onto finite dimensional N.L.S. can be constructed as the cardinality of a basis of a finite dimensional N.L.S., treated as a vector space over the rationals, is the same as the cardinality of $\mathbb{R}$, treated as a vector space over the rationals. In Chapter 4 all metric transformations, both continuous and non-continuous from $\mathbb{R}$ into E^n are characterized. $\square$

This concludes our work on metric and order transformations of general normed linear spaces. In Chapter 4 we shall consider the problem further when working with subsets of E^n .

CHAPTER 4

ORDER AND METRIC TRANSFORMATIONS IN E^n

The original interest of M.D.S. Theorists was in order transformations into E^n . Typically their concern was with a finite space, which we regard as a distance space, and with the possibility of finding a subset of a low dimensional Euclidean space order isomorphic to it. Thus, order embeddings into E^n have assumed a prominent role in M.D.S. theory.

A number of interesting questions on the existence of order embeddings into E^n have been examined by Holman, Kelly and others. We have made some efforts to extend these without much success. For completeness, we summarize these results in Chapter 6.

Some of our earlier results have immediate implications for embeddings into E^n . The most interesting of these may be Corollary 3.19 which shows that if S is any subset of E^n , containing a ball, then any order or metric transformation of S into E^n (in fact, into any n -dimensional normed linear space) is a similarity. This result is extended in Section 2 of this chapter.

It was mentioned in Chapter 1 that von-Neumann and Schoenberg [29] have characterized all continuous metric transforms of $\mathbb{R}$ into Hilbert space. These are the so called screw curves, whose associated scale functions are given by

$$(**) \quad \rho^2(d) = c^2 d^2 + \sum A_i \sin^2 \kappa_i d, \quad c_i, A_i, \kappa_i \text{ being constant.}$$

Much of this chapter depends on this characterization.

The chapter is divided into four sections. Section 1 is background material for the rest of the chapter. Here we briefly discuss the concept of a metric basis. These are often implicitly used in metric geometry, although rarely mentioned by name.

In Section 2, a theorem of Schoenberg is extended. In [32] Schoenberg showed that if $f: E^m \rightarrow E^n$, $m \geq 2$ is a continuous metric transformation, then f is a similarity. Vogt [38] showed that the assumption of continuity could be dropped. We show the domain need not be all of E^m , giving two different types of subsets of E^m for which the theorem is still valid.

Section 3 contains the characterization by von-Neumann and Schoenberg of the continuous metric transformations of $\mathbb{R}$ into E^n . This characterization is extended to the non-continuous case. We also show in this section that any metric transformation of an interval of $\mathbb{R}$ into E^n can be uniquely extended to a metric transformation of $\mathbb{R}$ into E^n .

This is important, in Section 2, where it is used to show that (**) is a characterization of the scale function of a metric transformation of an interval, as well as of the real line.

Although parts of Section 3 are used in Section 2 it seems appropriate to present the results of Section 2 first. The main result of Section 3 (the characterization (**)) is not new, and our contributions are not surprising. In addition, we suspect that all of these results are contained somewhere in the literature.

Section 4 of this chapter is quite separate from the other sections. It considers metric and order transformations of the unit ball in E^n (and in other inner product spaces) into normed linear spaces. This will be discussed further when we come to it.

§1. In this section we define and state some theorems concerning metric bases. This material is used in subsequent sections of this chapter.

Definition: A subset B of a distance space N is said to be a metric basis of N if and only if each point p of N is uniquely determined by the set $\{d(p,b) \mid b \in B\}$.

Metric bases abound in the classical spaces. A subset of E^n , S^n , or H^n is a metric basis if and only if it does not lie in an $(n-1)$ -flat. Such a set is said to span E^n , S^n , or H^n respectively.

A metric basis of E^n , S^n or H^n must contain at least $(n+1)$ points, and any metric basis (spanning set) of E^n , S^n or H^n contains a subset of exactly $(n+1)$ points which is itself a metric basis. We refer the reader to Blumenthal [4] for the details of these statements.

Metric bases have appeared earlier in this thesis, although they were not called such. The proof of Theorem 3.18 partly consisted of showing that any open subset of a normed linear space with strictly convex unit ball is a metric basis for that space. Metric bases also appear briefly in Chapter 5.

The following two lemmas are important properties of metric bases. These are stated for E^n , but they are also valid for S^n and H^n . Lemma 4.1 is a commonly used property of E^n . It is often referred to as the property of "free mobility".

Definition: An isometry of E^n onto itself is called a motion of E^n .

Lemma 4.1: If $f : S \subseteq E^n \rightarrow E^n$ is an isometry, then f can be extended to a motion $\bar{f}$ of E^n . If S contains a metric basis B , then the extension is unique.

Proof: See Blumenthal [4, Sec 38].

Lemma 4.2 shows that a metric basis is not always needed to use the same type of arguments, provided more information is available.

Lemma 4.2: Let E^m be an m -flat in E^n and let $B \subseteq E^m$ be a metric basis of E^m . If $p \in E^m$ and $q \in E^n$, and $d(p, b) = d(q, b)$ for all $b \in B$, then $p = q$.

Proof: See Blumenthal [4, Sec 40].

This essentially says that points of E^m are determined by a metric basis of E^m , even within the larger space E^n .

Corollary 4.3: If f is a motion of E^n which maps a metric basis of E^m into E^m , then $f(E^m) = E^m$.

The following lemma is used in Section 2.

Lemma 4.4: Let $f: S \subseteq E^V \rightarrow E^V$. Let E^u be a u -flat in E^V , and $U \subseteq S \cap E^u$ be such that $f|_U$ is an isometry. Let $B \subseteq U$ be a metric basis of E^u and let $p \in S$ be such that $f|_{B \cup \{p\}}$ is an isometry. Then $f|_{U \cup \{p\}}$ is an isometry. In particular, $d(p, q) = d(f(p), f(q))$ for all $q \in U$.

Proof: By Lemma 4.1, $f|_{B \cup \{p\}}$ and $f|_U$ can be extended to motions $\bar{f}$ and $\bar{\bar{f}}$ respectively of E^V . The set $f(B)$ is a metric basis of both the u -flats $\bar{f}(E^u)$ and $\bar{\bar{f}}(E^u)$, from which it follows that $\bar{f}(E^u) = \bar{\bar{f}}(E^u)$.

Both $\bar{f}|_{E^u}$ and $\bar{\bar{f}}|_{E^u}$ extend $f|_B$ to an isometry from E^u to $\bar{f}(E^u)$. By Lemma 4.1, there is only one such extension. Hence $\bar{f}|_{E^u} = \bar{\bar{f}}|_{E^u}$.

For $q \in U \subseteq E^u$, $f(q) = \bar{f}(q) = \bar{f}(q)$. Also, $f(p) = \bar{f}(p)$.
Hence, $f|_{U \cup \{p\}} = \bar{f}|_{U \cup \{p\}}$ is an isometry. $\square$

§2. In this section we extend the result of Schoenberg [32] which shows that any continuous metric transformation of E^m into E^n , is necessarily a similarity. The result depends on the following characterization, by von-Neumann and Schoenberg [29], of the scale functions associated with screw curves in E^n .

$$(**) \quad \rho^2(d) = c^2 d^2 + \sum_{i=1}^v A_i^2 \sin^2 \kappa_i d$$

where $v = \frac{n-1}{2}$, n odd, $v = \frac{n}{2}$ for n even and $c = 0$, and $v = \frac{n-2}{2}$ for n even and $c \neq 0$.

An examination of Schoenberg's proof reveals that in fact he proves the following. (Recall that a trivial metric transformation is one which maps the entire set to a single point.)

Theorem 4.5: Let $S \subseteq E^m$ contain a circle C and a line l . Then any continuous metric transformation of S into E^n is either a similarity or trivial.

Theorems 4.6 and 4.7, which follow, extend this result even further.

Theorem 4.6: If $S \subseteq E^m$ contains a circle C , an unbounded connected subset T , and a line segment L with $L \subseteq T$, then any continuous metric transformation of S into E^n is either trivial or a similarity.

Theorem 4.7: Let $S \subseteq E^m$, $m \geq 2$ be a connected set with non-empty interior. Then any metric transformation of S into E^n is either trivial or is a similarity.

The subject of continuity needs comment. Theorem 3.11 (b) shows that any metric transformation from an open subset of a normed linear space (of dimension greater than one) into a separable normed linear space is either continuous or trivial. Thus the assumption of continuity is not necessary in Schoenberg's version of Theorem 4.5. This is proved by Vogt in [39]. Theorem 4.7 includes the assumption of an open subset of E^m , so Lemma 3.11 (b) shows continuity there. On the other hand we are not guaranteed an open subset of E^m in the set S of Theorem 4.6, so Lemma 3.11 cannot be applied. We conjecture that continuity need not be assumed in this case, but follows from the remaining hypotheses. However we have not been able to show this.

The proofs of Theorem 4.6 and 4.7 are adaptations of Schoenberg's proof of Theorem 4.5. For this reason we present Schoenberg's proof of Theorem 4.5.

Proof of Theorem 4.5: Let f be a continuous metric transformation of S into E^n . We proceed by examining the scale function, ρ , associated with f . First consider f restricted to ℓ . Since ℓ is isometric to $\mathbb{R}$, $f(\ell)$ is a screw line, thus ρ must satisfy (**). That is,

$$(**) \quad \rho^2(d) = c^2 d^2 + \sum_{i=1}^s A_i^2 \sin^2 \kappa_i d,$$

for all d and some non-negative values of c , A_i , and κ_i .

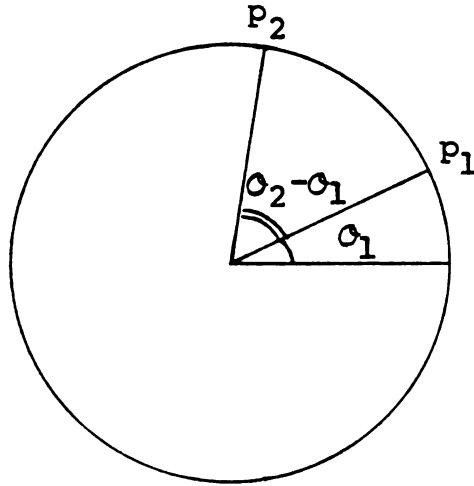


Figure 4.1

Next consider f restricted to C . For any $p \in C$, let $\mathcal{O}(p)$ be the central angle, measured between the radius with endpoint p , and some fixed radius. For any two points p_1 and p_2 of C , with $\mathcal{O}_1 = \mathcal{O}(p_1)$, $\mathcal{O}_2 = \mathcal{O}(p_2)$ we have

$$d(p_1, p_2) = \left| 2r \sin \frac{|\mathcal{O}_2 - \mathcal{O}_1|}{2} \right| \text{ and}$$

$$\rho \left(\left| 2r \sin \frac{|\mathcal{O}_2 - \mathcal{O}_1|}{2} \right| \right) = \rho(d(p_1, p_2)) = d(f(p_1), f(p_2)) .$$

Since $\rho\left(\left|2r \sin \frac{|\phi_2 - \phi_1|}{2}\right|\right)$ depends only on $|\phi_2 - \phi_1|$, the mapping $\varphi : \mathbb{R} \rightarrow E^m$ defined by $\varphi(\phi(p) + 2k\pi) = f(p)$ for each k , is a continuous metric transformation of $\mathbb{R}$ into E^n with scale function $\tilde{\rho}(\alpha) = \rho(|2r \sin \frac{\alpha}{2}|)$. Hence by (**)

$$(1) \quad \rho^2(|2r \sin \frac{\alpha}{2}|) = \tilde{\rho}^2(\alpha) = \tilde{c}^2 \alpha^2 + \sum_{i=1}^v B_i^2 \sin^2 h_i \alpha.$$

Since C is bounded, $f(C)$ and hence $\varphi(\mathbb{R})$ are also bounded, so it must be the case that $\tilde{c} = 0$. Thus $\rho(|2r \sin \frac{\alpha}{2}|) = \sum_{i=1}^v B_i^2 \sin^2 h_i \alpha$. Let $d = |2r \sin \frac{\alpha}{2}|$, and setting the expressions (**) and (1) for ρ equal, we get

$$\sum_{i=1}^v B_i^2 \sin^2 h_i \alpha = 4r^2 c^2 \sin^2 \frac{\alpha}{2} + \sum_{i=1}^s A_i^2 \sin^2(2k_i r \sin \frac{\alpha}{2})$$

It follows from this (see the appendix at the end of this chapter) that either A_i or k_i is 0 for each i . Thus (**) reduces to $\rho^2(d) = c^2 d^2$, or simply $\rho(d) = cd$, $c \geq 0$. Since (**) is true for any distance d , f is either a similarity, or is trivial. $\square$

To adapt this idea to Theorems 4.6 and 4.7, several problems arise. The most serious of these would seem to be the substitution of a line segment L for a line ℓ in the hypotheses. It is not clear that the scale function associated with a continuous metric transformation of a line segment, rather than a line, has the form (**). It is shown in Section 3 of this chapter (Theorem 4.12) that any metric transformation of an open interval can be extended, uniquely, to a metric

transformation of the entire real line. Assuming the validity of this result, it follows that the scale function associated with a continuous metric transformation of a line segment has the form (**).

It can then be shown in the same manner as in the proof of Theorem 4.5 that if $S \subseteq E^n$ contains a circle and a line segment, then any metric transformation of S into E^n satisfies $\rho(d) = cd$, for any distance d which occurs as the distance between two points on the circle, or two points on the line segment. In Schoenberg's case (Theorem 4.5) this shows $\rho(d) = cd$ for all d (hence that f is a similarity) however it does not do so for us. To overcome this problem, extra hypotheses are needed, (Example 1 at the end of this section shows that some extra conditions are indeed needed), such as the connectedness condition in each theorem.

Before proceeding to the proofs of Theorems 4.6 and 4.7, we show the following lemma.

Lemma 4.8: Let S be a connected set and ℓ a line in E^m such that $S \cap \ell$ contains an interval U of ℓ . If $f: S \rightarrow E^n$ is a metric transformation, with scale function ρ , and $f|_U$ is an isometry, then

$$d(f(p), f(q)) = d(p, q) \quad \text{for all } p \in S, q \in U.$$

Proof: Let q_0 be an arbitrary point of U . Since U is open in ℓ , there are points q_1 and q_2 of U such that the segment joining q_1 to q_2 is in U , and

$d(q_0, q_1) = d(q_0, q_2)$. Let

$$r = d(q_0, q_1) = d(q_0, q_2)$$

$$d_0 = \sup\{d \mid d = d(p, q_0), p \in S\}$$

$$d_1 = \sup\{d' \mid \rho(d) = d, \text{ for all } d < d', d = d(p, q), p, q \in S\}.$$

In other words, d_1 is the largest distance such that $\rho(d) = d$ for all d in $[0, d_1)$.

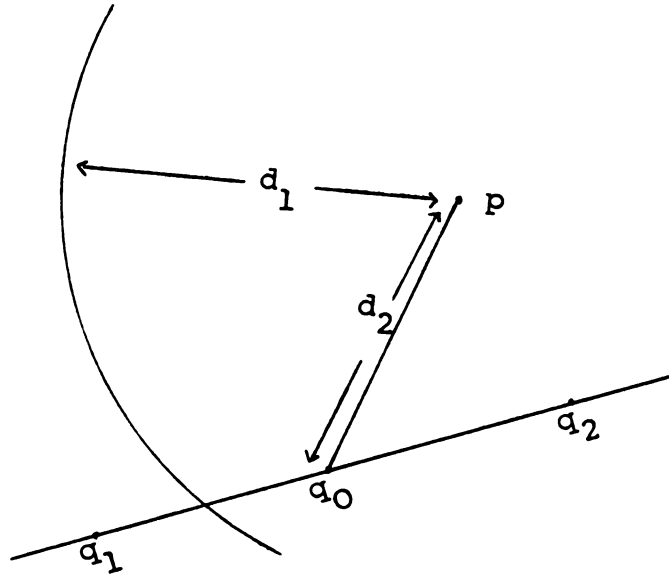


Figure 4.2

Since $f|_U$ is an isometry we have $d_1 \geq 2r > 0$ and $d_0 \geq r > 0$, hence $\min\{d_0, d_1\} > 0$. Let d_2 be an arbitrary distance less than $\min\{d_0, d_1\}$. Let p be an arbitrary point such that $d(p, q_0) = d_2$. There is at least one such point because S is connected, and $d_2 < d_0$. Now since $d_2 < d_1$,

there are two points b_1 and b_2 of U with $d(p, b_i) < d_1$ and hence with $\rho(d(p, b_i)) = d(p, b_i)$.

Let M be an Euclidean space with $E^m \subseteq M$, $E^n \subseteq M$. Then we can apply Lemma 4.4, with f, p, S and U as above, $E^u = \ell$, $E^v = M$, and $B = \{b_1, b_2\}$ to show that

$$\rho(d(p, q)) = d(f(p), f(q)) = d(p, q) \quad \text{for all } q \in U.$$

In particular, $\rho(d) = d$ for all d in $[0, \sqrt{d_2^2 + r^2})$. (See Figure 4.2. Apply the law of cosines, noting that one of the angles $p q_0 q_2$ or $p q_0 q_1$ is between $\pi/2$ and π .) As d_2 is an arbitrary distance less than $\min\{d_0, d_1\}$, it follows that $\rho(d) = d$ for all d in $[0, \sqrt{(\min\{d_0, d_1\})^2 + r^2})$.

If $d_1 = \min\{d_0, d_1\}$ this is a contradiction. Therefore $d_0 = \min\{d_0, d_1\}$ and hence $\rho(d(p, q_0)) = d(p, q_0)$ for all $p \in S$. Since q_0 is an arbitrary point of U , the lemma is proved. $\square$

Corollary 4.9: Let S be a connected set in E^m , and assume U is an open subset of a line ℓ in E^m with $U \subseteq S$.

Let $f : S \rightarrow E^n$ be a metric transformation such that $f|_U$ is a similarity with

$$\rho(d(q_1, q_2)) = cd(q_1, q_2), \quad c > 0, \quad q_1, q_2 \in U.$$

Then

$$\rho(d(p, q)) = cd(p, q) \quad \text{for all } p \in S, q \in U.$$

Proof: Define $\tilde{f}$ by $\tilde{f}(p) = \frac{1}{c} f(p)$. Then it is easily seen that $\tilde{f}$ is a metric transformation and $\tilde{f}|_U$ is an isometry. Let $\tilde{\rho}$ be the scale function associated with $\tilde{f}$. Then certainly $\tilde{\rho}(d) = \frac{1}{c} \rho(d)$. It follows from Lemma 4.8 that $\tilde{\rho}(d(p,q)) = d(p,q)$ for any $p \in S, q \in U$. Hence $\rho(d(p,q)) = cd(p,q)$ for any $p \in S, q \in U$. $\square$

We are now prepared to prove Theorems 4.6 and 4.7.

Theorem 4.6: If $S \subseteq E^m$ contains a circle C , an unbounded connected subset T , and a line segment L with $L \subseteq T$, then any continuous metric transformation of S into E^n is either trivial or is a similarity.

Proof: Let f be a continuous metric transformation of S into E^n , with scale function ρ . Using Theorem 4.12 $f|_L$ can be extended to a screw curve, hence ρ must have the form

$$(**) \quad \rho^2(d) = c^2 d^2 + \sum_{i=1}^r A_i^2 \sin^2 \kappa_i d$$

for any distance d that can be written as $d = d(q_1, q_2)$, where $q_1, q_2 \in L$. Using C , the proof of Theorem 4.5 can be imitated to show that $\rho(d) = cd$, $c \geq 0$, for any distance d that can be written as $d = d(q_1, q_2)$, $q_1, q_2 \in L$. (Recall that the proof of Theorem 4.5 depended upon (**), the existence of a circle C in S , and the continuity of f .)

If $c = 0$, then $\rho(d) = 0$ for d less than the length of L and, by Lemma 2.6, f is trivial. If $c \neq 0$ Corollary 4.9 shows that $\rho(d(p,q)) = cd(p,q)$ for any $p \in T, q \in L$, which defines $\rho(d)$ for any distance d . Hence f is a similarity. $\square$

Remark: When applying the above theorem, Lemma 3.11(b) could well guarantee that the continuity assumption is satisfied.

Theorem 4.7: Let $S \subseteq E^m$, $m \geq 2$ be a connected set with non-empty interior. Then any metric transformation of S into E^n is either a similarity or is trivial.

Proof: Let f be a metric transformation of S into E^n , with scale function ρ . That f is continuous follows from Lemma 3.11 (b). Since $m \geq 2$, and S contains an interior point of E^m , S contains both a line segment and a circle. Therefore, as in the proof of Theorem 4.6, $\rho(d) = cd$ for small d . If $c = 0$ Lemma 2.6 shows that f is trivial.

Now consider the case $c > 0$. Let $\tilde{f}(p) = \frac{1}{c} f(p)$, and let $\tilde{\rho}$ be the scale function corresponding to $\tilde{f}$. Clearly, $\tilde{\rho} = \frac{1}{c} \rho$, and hence $\tilde{\rho}(d) = d$ for small d . Thus there is an open ball B of S such that $\tilde{f}|_B$ is an isometry. Noting that for $q \in B$, there is an open segment L with $q \in L \subseteq B$, Lemma 4.8 shows that

$$d(\tilde{f}(p), \tilde{f}(q)) = \tilde{\rho}(d(p,q)) = d(p,q) \quad \text{for every } p \in S, q \in B.$$

Now $\tilde{f}(B)$ is an isometric image in E^n of an open set of E^m , so it must be that $m \leq n$. Let M_1 be any n -dimensional Euclidean space containing E^m , and let M_2 be the m -flat of E^n which contains $\tilde{f}(B)$. See Figure 4.3.

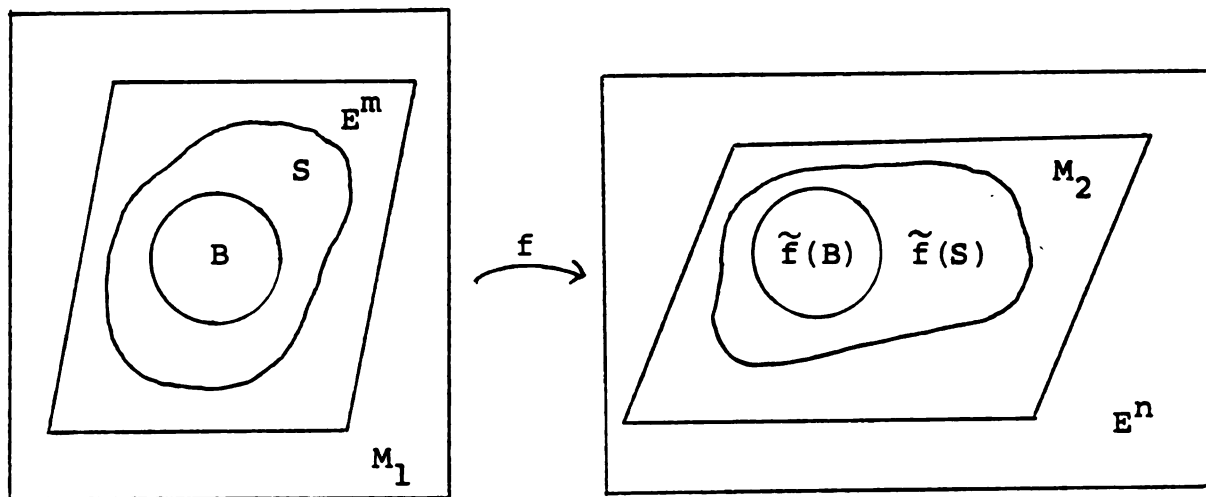


Figure 4.3

Recall that an open subset of E^m is a metric basis of E^m , and hence B is a metric basis of E^m . The isometry $\tilde{f}|_B$ can be extended to an isometry $\bar{f}: E^m \rightarrow M_2$. Let $p \in S$. For any $q \in B$, it has been shown that $d(p, q) = d(\tilde{f}(p), \tilde{f}(q))$. Also

$$d(p, q) = d(\bar{f}(p), \bar{f}(q)) = d(\bar{f}(p), \tilde{f}(q)) .$$

Hence

$$d(\tilde{f}(p), \tilde{f}(q)) = d(\bar{f}(p), \tilde{f}(q)) \quad \text{for all } q \in B .$$

Lemma 4.2 now shows $\tilde{f}(p) = \bar{f}(p)$.

In particular, $\tilde{f}(p) \in M_2$. Thus $\tilde{f}: S \subseteq E^m \rightarrow M_2$. The spaces E^n and M_2 are normed linear spaces of the same finite dimension, so Corollary 3.19 (a) shows that $\tilde{f}$ is a similarity. (In fact $\tilde{f}$ is an isometry.) From this it follows that f is a similarity. $\square$

Example 4.1: To show the connectedness conditions of Theorems 4.6 and 4.7 cannot be eliminated (although they can surely be modified) consider the following.

Let D_n , n any integer, be an open disc of radius one in the xy -plane, centered at $(4n, 0)$. Let $f: \bigcup_{n=-\infty}^{\infty} D_n \rightarrow E^3$ be defined by

$$f(x, y) = (x, y, n) \quad \text{for } (x, y) \in D_n.$$

If $d(p, q) = d(p', q')$ where $p \in D_n$, $q \in D_m$, $p' \in D_n$, $q' \in D_m$, then $|n - m| = |n' - m'|$ and

$$d^2(f(p), f(q)) = d^2(p, q) + (n - m)^2$$

$$d^2(f(p'), f(q')) = d^2(p', q') + (n' - m')^2$$

Thus it follows that f is a metric transformation. However f is clearly not a similarity. $\square$

To conclude this section we present a curious example. The proof of Theorem 4.5 proceeded as follows. First, distances of the form $d(p, q)$, where p and q lie on a line l were considered. Then distances of the form $d(p, q)$, where p and q lie on a circle C were considered. In each case a

characterization of the scale function ρ was obtained. Using the two characterizations, f was then shown to be a similarity. Our first thought for proving a theorem such as Theorem 4.7, where S contains an open subset of E^m , was to do the following. Take two circles of different radii. For each circle we have a characterization of the scale function ρ . From this it was hoped that it would follow that f was a similarity. This approach would eliminate the necessity of Theorem 4.12. In the following example we show a function f which maps the unit disc of the xy -plane into E^4 and which is clearly not a similarity. However f restricted to any circle centered at the origin is a metric transformation and if p_1 and q_1 lie on the circle of radius r_1 centered at $(0,0)$, p_2 and q_2 lie on the circle of radius r_2 centered at $(0,0)$, and $d(p_1, q_1) = d(p_2, q_2)$ then $d(f(p_1), f(q_1)) = d(f(p_2), f(q_2))$. Thus the above approach would not work without considering something further about the circles.

Example 4.2: For any $r > 0$, let C_r be the circle of radius r , centered at the origin. Let (r, θ) be the polar co-ordinates of a point p on the circle. Define $f: C_r \rightarrow E^4$ by

$$f(p) = (\sqrt{3r^2 - 2r^4} \cos \theta, \sqrt{3r^2 - 2r^4} \sin \theta, \sqrt{\frac{1}{2}} r^2 \cos 2\theta, \sqrt{\frac{1}{2}} r^2 \sin 2\theta) .$$

It is easy to see that $f(C_r)$ is a metric transformation of C_r . The scale function associated with $f|_{C_r}$, called ρ_r , is given by

$$\begin{aligned}\rho_r^2(d) &= (3r^2 - 2r^4)\sin^2(\sin^{-1} \frac{d}{2r}) + \frac{1}{2} r^4 \sin^2(2 \sin^{-1} \frac{d}{2r}) \\ &= \frac{3}{4} d^2 - \frac{1}{8} d^4 .\end{aligned}$$

Since this last expression is independent of r , it has been shown that for $p_1, q_1 \in C_{r_1}$ and $p_2, q_2 \in C_{r_2}$ with $d(p_1, q_1) = d(p_2, q_2) = d$, then $\rho_{r_1}(d) = \rho_{r_2}(d)$.

We have now shown that to prove Theorem 4.7, it is not enough to consider a set of circles and distances which occur between two points of the same circle. $\square$

A more interesting example would be a set in E^m containing two circles of different radii and a metric transformation of that set into E^n which is not a similarity. We conjecture that no such example exists.

§3. The work in Section 2 of this chapter is based on the characterization (**) of the scale function of a continuous metric transformation of $\mathbb{R}$ into E^n . The transformation, relative to a suitable co-ordinate system, is that which takes $t \in \mathbb{R}$ to

$$(A_1 \cos k_1 t, A_1 \sin k_1 t, \dots, A_v \cos k_v t, A_v \sin k_v t, ct)$$

where A_i, k_i , and c are constants. If $f(\mathbb{R})$ spans E^n , then $v = \frac{n}{2}$, $c = 0$ for n even, and $v = \frac{n-1}{2}$, $c \neq 0$ for n odd.

In E^2 this says the only continuous metric transforms of $\mathbb{R}$, (screw curves) that span the plane, are circles. Since an order transformation is necessarily bicontinuous, then there are no order transforms of $\mathbb{R}$ spanning E^2 . In E^3 the screw curves are necessarily circular helices. As in Example 4 Chapter 1, these may be order transforms of $\mathbb{R}$. This behaviour is typical of the even and odd dimensional cases. In even dimensions the screw curves spanning the space are bounded simple closed curves lying on a sphere. Hence there are no order transforms of $\mathbb{R}$ spanning even dimensional spaces. The simple closed curves referred to above are sometimes called "trigonometric moment" curves.

In odd dimensions the screw curves spanning the space are unbounded and may be thought of as generalizations of the circular helix in E^3 . Such a curve lies on a "spherical cylinder" and its projection onto the base of the cylinder is one of the trigonometric moment curves referred to above. The tangents to the curve make a fixed angle with its axis as in E^3 .

The von-Neuman-Schoenberg study shows much more, being set in general Hilbert spaces. For example it is shown that bounded screw curves in $\mathcal{H}$ must lie on a sphere. In E^n screw curves are, of course, rectifiable, but such is not the case in $\mathcal{H}$. For example the order transform of $\mathbb{R}$ given by the square root function is an example of a homeomorphic image of the line no sub-arc of which is rectifiable and it was shown in [32] that this is a screw curve in $\mathcal{H}$. Von-Neuman

and Schoenberg characterize those screw curves in $\mathcal{K}$ which are rectifiable.

Their proofs involve rather sophisticated analytic techniques applicable to the Hilbert space setting. Since M.D.S. theorists are more concerned with finite dimensions and these characterizations are finding their way into the M.D.S. literature we thought a strictly finite-dimensional derivation of the characterization in E^n might be both simpler and more revealing.

We have another reason for presenting a proof in finite dimensional space. Vogt [38] gives an intriguing example, presented and generalized in Chapter 3 (Section 3), showing non-continuous metric transformations of $\mathbb{R}$. It is not hard to show that this example is typical of all metric transformations of $\mathbb{R}$ into $\mathbb{R}$. That is, any metric transformation of $\mathbb{R}$ into $\mathbb{R}$ is a group homomorphism, followed by a translation. A metric transformation of $\mathbb{R}$ into E^2 that is not of this type is the following. Let $\theta : \mathbb{R} \rightarrow \mathbb{R}$ be a group homomorphism. Define $f : \mathbb{R} \rightarrow E^2$ by $f(t) = (\cos \theta(t), \sin \theta(t))$. Certainly this need not be continuous, and it is not hard to show f is a metric transformation. In Theorem 4.10 all metric transformations of $\mathbb{R}$ into E^n (be they continuous or not) are characterized. We see that they are "made up" from the above types of metric transformations.

To assist the reader, we list the various symbols used in the remainder of this section. To be consistent with the literature, norm notation shall be used for distance here.

Notation for Theorems 4.10, 4.11, 4.12:

s, s_j, a_s, b_s - real numbers

$x, y, x_j, y_j, v_{1j}, v_{2j}, v_{3j}, w$ - real vectors (The definition of a real vector is introduced in the proof.)

z, z_j - complex vectors

V_j, W_j, W - subspaces of $\mathbb{C}^n$

U^s, U_j^s - unitary transformations of $\mathbb{C}^n$

T^s, T_j^s - affine isometries

I - the identity transformation

Note: U^s (or T^s) does not represent the transformation U (or T) raised to the power s , but is one symbol.

We write it in this way because it is shown that $U^s \circ U^r = U^{s+r}$ ($T^s \circ T^r = T^{s+r}$).

Let $v_1, \dots, v_m$ be vectors in $\mathbb{C}^n$

(v_1, v_2) - the standard inner product in $\mathbb{C}^n$

$[v_1, \dots, v_m]$ - the subspace of $\mathbb{C}^n$ spanned by $v_1, \dots, v_m$

If $V \subseteq \mathbb{C}^n$ is a subspace, we denote by $V^\perp$ the set

$\{z \in \mathbb{C}^n \mid (z, v) = 0 \text{ for all } v \in V\}$.

$\|v_1\| = \sqrt{(v_1, v_1)}$

Definition: A unitary transformation $\mathbb{C}^n \rightarrow \mathbb{C}^n$ (or $E^n \rightarrow E^n$) is a function U such that $U(0) = 0$ and $\|z_1 - z_2\| = \|U(z_1) - U(z_2)\|$ for all $z_1, z_2 \in \mathbb{C}^n$ (or E^n).

The following is a list of elementary properties of unitary transformations. These are all easy to show, and can be found in linear algebra or functional analysis reference books.

(See for example [1], [18]).

Let U be a unitary transformation.

U1. U is a linear transformation of $\mathbb{C}^n$ (or E^n). Thus U is an affine isometry.

U2. $(z_1, z_2) = (U(z_1), U(z_2))$ for all $z_1, z_2 \in \mathbb{C}^n$ (or E^n).

U3. The absolute value of any eigenvalue of U is 1.

U4. U^{-1} exists.

U5. $(z_1, U(z_2)) = (U^{-1}(z_1), z_2)$.

U6. If V is a subspace of $\mathbb{C}^n$ (or E^n) and $U(V) = V$, then $U(V^\perp) = V^\perp$.

U7. If $a + bi$ is an eigenvalue of U , then $\frac{1}{a + bi} = a - bi$ is an eigenvalue of U^{-1} .

Theorem 4.10: Let $f: \mathbb{R} \rightarrow E^n$ be a metric transformation with scale function ρ , such that $f(\mathbb{R})$ spans E^n . Then there are mutually orthogonal two dimensional subspaces $V_1, \dots, V_m$ of E^n , an $(n - 2m)$ -dimensional subspace W of E^n with $W \perp V_j$ for all j , vectors

v_{1j}, v_{2j} , and v_{3j} in V_j , with $v_{1j} \perp v_{2j}$ and $\|v_{1j}\| = \|v_{2j}\| = 1$, a vector $w \in W$, group homomorphisms

$$\theta_j : \mathbb{R} \rightarrow \mathbb{R}/2\pi, \quad 1 \leq j \leq m$$

a group homomorphism $g : \mathbb{R} \rightarrow W$, and constants A_j such that

$$(*) \quad \begin{cases} f_j(t) = A_j((\cos \theta_j(t))v_{1j} + (\sin \theta_j(t))v_{2j}) + v_{3j} \\ f_W(t) = g(t) + w \end{cases}$$

where $f_j(t)$ is the projection of $f(t)$ into V_j , and $f_W(t)$ is the projection of $f(t)$ into W .

Conversely, any function $f : \mathbb{R} \rightarrow E^n$ of the form $(*)$ is a metric transformation of $\mathbb{R}$ into E^n .

Proof of Converse: We show that if f satisfies $(*)$ for some set of normal subspaces $V_1, \dots, V_n$ and W of E^n , then f is a metric transformation. In this case we have

$$\begin{aligned} \|f(t_1) - f(t_2)\|^2 &= \sum_{j=1}^m 4A_j^2 \sin^2\left(\frac{\theta_j(t_1) - \theta_j(t_2)}{2}\right) + \|f_W(t_1) - f_W(t_2)\|^2 \\ &= \sum_{j=1}^m 4A_j^2 \sin^2\left(\frac{\theta_j(t_1 - t_2)}{2}\right) + \|g(t_1 - t_2)\|^2 \\ &= \sum_{j=1}^m 4A_j^2 \sin^2\left(\pm \frac{\theta_j(|t_1 - t_2|)}{2}\right) + \|\pm g(|t_1 - t_2|)\|^2 \\ &= \sum_{j=1}^m 4A_j^2 \sin^2\left(\frac{\theta_j(|t_1 - t_2|)}{2}\right) + \|g(|t_1 - t_2|)\|^2. \end{aligned}$$

This shows f is a metric transformation with scale function $\rho(d)$ satisfying

$$\rho^2(d) = \sum_{j=1}^m 4A_j^2 \sin^2 \theta_j\left(\frac{d}{2}\right) + \|g(d)\|^2. \quad \square$$

We now provide some motivation for the proof of this theorem by informally discussing the situation in E^3 .

Any motion of E^3 can be described (see [6], Chapter 7) as either

(A) A rotation about a line, and a translation along that line. This line is called the axis of rotation.

(B) A rotation about a line, and a reflection through a plane perpendicular to that line.

(C) A reflection through a plane, and a translation along a line in that plane.

We are given a metric transformation f of the real line, with $\{f(t), t \in \mathbb{R}\}$ spanning E^3 . We can define an isometry T^s of the set $\{f(t) \mid t \in \mathbb{R}\}$ by $T^s(f(t)) = f(t+s)$. It is not hard to verify that this is an isometry. Extend this to a motion of E^3 , also called T^s . As $\{f(t) \mid t \in \mathbb{R}\}$ spans E^3 , this extension will be unique (Theorem 4.1).

Using these properties of T^s , it can be shown that

$$(1) \quad T^{s+r} \equiv T^s \circ T^r \equiv T^r \circ T^s.$$

In particular $T^s \equiv T^{s/2} \circ T^{s/2}$ for all s , from which we can conclude that T^s is not of type (B) or type (C), hence is of type (A).

Next, using (1), it can be shown that the axis of rotation of each T^s must be the same. Thus, if T^s , T^r , and T^{s+r} have rotation angles $\theta(s)$, $\theta(r)$, and $\theta(s+r)$ respectively, then $\theta(s) + \theta(r) = \theta(s+r)$ (modulo 2π). It also follows

that if T^s has a translation along this axis of a distance $g(s)$, T^r of a distance $g(r)$, and T^{s+r} a distance $g(s+r)$ then $g(s) + g(r) = g(s+r)$ (where $g(s)$, $g(r)$, $g(s+r)$, can be either positive or negative, depending on the direction of the translation).

Definition: A set of motions $\{T^s \mid s \in \mathbb{R}\}$ of E^n satisfying $T^s \circ T^r = T^{s+r}$ is called a one parameter subgroup of motions.

It follows immediately that any two members of a one parameter subgroup of motions commute.

Comment on Notation: Note that T^s is a motion, and does not represent "T raised to the power s". When we wish to use exponentiation, and we shall, we write $(T^s)^n$. Thus

$$(T^s)^n = \underbrace{T^s \circ T^s \circ \dots \circ T^s}_{n \text{ times}}. \text{ It is easy to see that if}$$

$\{T^s \mid s \in \mathbb{R}\}$ is a one parameter subgroup of motions, then $(T^s)^n = T^{n \cdot s}$. As T^s and T^r commute, it follows that $(T^{s+r})^n = (T^s \circ T^r)^n = (T^s)^n (T^r)^n$.

Proof of Theorem 4.10: The proof consists of eight steps.

Step 1. For each $s \in \mathbb{R}$, define a motion T^s of E^n as follows. Let $T^s(f(t)) = f(t+s)$. As f is a metric transformation it follows that

$$\|T^s(f(t_1)) - T^s(f(t_2))\| = \|f(t_1+s) - f(t_2+s)\| = \rho(|t_2 - t_1|).$$

Hence T^s defines an isometry of $f(\mathbb{R})$ into E^n . By hypothesis $f(\mathbb{R})$ spans E^n , hence T^s can be uniquely extended to a motion of E^n (Theorem 4.1). Because of this uniqueness, and the fact that

$$T^s \circ T^r(f(t)) = f(t+s+r) = T^{s+r}(f(t)),$$

it follows that $T^s \circ T^r = T^{s+r}$. Thus $\{T^s \mid s \in \mathbb{R}\}$ forms a one parameter subgroup of motions of E^n .

The function $T^s(x) - T^s(0)$ is a norm preserving map from E^n onto E^n , which takes 0 to 0, hence is a unitary transformation (by definition). Thus $T^s(x) = U^s(x) + T^s(0)$ where U^s is a unitary transformation. By U1, U^s is a linear transformation, hence T^s is an affine isometry.

Remark: Step 1 shows that every metric transformation f of $\mathbb{R}$ into E^n can be written as $f(s) = T^s(f(0))$, where T^s is a one parameter subgroup of motions of E^n .

Step 2. Consider $\mathbb{C}^n$, the complexification of E^n . Extend U^s to a unitary transformation on $\mathbb{C}^n$, also called U^s , by

$$U^s(x + iy) = U^s(x) + iU^s(y)$$

where x and y are real vectors. (That is to say, x and y lie in the natural embedding of E^n into $\mathbb{C}^n$. See [1] for a discussion of complexification.) Note that this is the unique extension of U^s to a linear map of $\mathbb{C}^n$. As $(U^s)^{-1}$ exists (by U4) it follows that $U^s(y) = 0$ if and only if $y = 0$.

Hence if $z = x + iy$ (x, y real) then $U^S(z) = U^S(x + iy) = U^S(x) + iU^S(y)$ is real if and only if z is real, that is $y = 0$. This property of U^S will be used frequently

Extend T^S to $\mathbb{C}^n$ by

$$T^S(z) = U^S(z) + T^S(0) .$$

Note that this is the unique extension of T^S to $\mathbb{C}^n$ such that $T^S(z)$ is an affine isometry of $\mathbb{C}^n$. Also, because $T^S(0)$ is real, and $U^S(z)$ is real if and only if z is real, then $T^S(z)$ is real if and only if z is real.

Step 3. Because the extension of T^S to $\mathbb{C}^n$ is unique it follows that

$$T^S \circ T^R = T^{S+R} = T^R \circ T^S$$

showing T^S is a one parameter subgroup of motions of $\mathbb{C}^n$.

Now

$$\begin{aligned} T^S \circ T^R(z) &= T^S(U^R(z) + T^R(0)) \\ &= U^S(U^R(z) + U^S(T^R(0))) + T^S(0) \end{aligned}$$

and

$$T^{S+R}(z) = U^{S+R}(z) + T^{S+R}(0) .$$

Hence, combining the above equations

$$\begin{aligned} (U^{S+R} - U^S U^R)z &= U^S(T^R(0)) + T^S(0) - T^{S+R}(0) \\ &= \text{constant} . \end{aligned}$$

As this is true for all z , the constant is 0, and

$$(1) \quad U^S U^R = U^{S+R} \quad \text{and} \quad T^{S+R}(0) = U^S(T^R(0)) + T^S(0) .$$

Similarly

$$U^R U^S = U^{S+R} \quad \text{and} \quad T^{S+R}(0) = U^R(T^S(0)) + T^R(0) .$$

Thus

$$(2) \quad U^S U^R = U^R U^S$$

and

$$U^S(T^R(0)) + T^S(0) = U^R(T^S(0)) + T^R(0)$$

or

$$(3) \quad (I - U^R)T^S(0) + U^S(T^R(0)) = T^R(0) ,$$

where I is the identity transformation.

Step 4. We now wish to "simultaneously diagonalize" the transformations $U^S, s \in \mathbb{R}$. That is, we wish to find a basis B of $\mathbb{C}^n$ such that the matrix of U^S in B is diagonal, for all s .

We have enough information to apply standard theorems of linear algebra, however we need to be quite specific in the choosing of the basis vectors so we present some details of the usual arguments. Equation (2) and a standard theorem of linear algebra ([2], P 206) show that the transformations $U^S, s \in \mathbb{R}$ possess a common eigenvector which we call $x + iy$, where x and y are real. Assume $\|x + iy\| = \sqrt{2}$. Let the eigenvalue of U^S corresponding to $x + iy$ be $a_s + ib_s$. For any $s \in \mathbb{R}$,

$$\begin{aligned} U^S(x) + iU^S(y) &= U^S(x + iy) \\ &= (a_s + ib_s)(x + iy) \\ &= a_s x - b_s y + i(b_s x + a_s y) . \end{aligned}$$

As observed earlier, $U^S(x)$ is real if and only if x is real, hence it follows that

$$U^s(x) = a_s x - b_s y$$

$$U^s(y) = b_s x + a_s y .$$

Now

$$U^s(x - iy) = U^s(x) - iU^s(y) = (a_s - b_s i)(x - iy),$$

showing that $x - iy$ is an eigenvector for all U^s , the corresponding eigenvalue being $a_s - b_s i$.

Case I. If $b_s = 0$ for all s , let $W_1 = [x + iy, x - iy]$. It is easy to verify that $W_1 = [x, y]$. Note that W_1 may be either one or two dimensional, but in any case, it has a real basis. Clearly $U^s(W_1) = U^s([x + iy, x - iy]) = W_1$ hence by U6, $U^s(W_1^\perp) = W_1^\perp$.

Case II. Assume that $b_{s_1} \neq 0$, for some $s_1 \in \mathbb{R}$. We now show $(x + iy) \perp (x - iy)$ and that $x \perp y$. We have

$$(4) \quad (x + iy, U^{s_1}(x - iy)) = (a_{s_1} + b_{s_1} i)(x + iy, x - iy) .$$

By U7,

$$(5) \quad ((U^{s_1})^{-1}(x + iy), x - iy) = (a_{s_1} - b_{s_1} i)(x + iy, x - iy) .$$

By U5, the left hand sides of (4) and (5) are equal, hence subtracting (5) from (4) yields

$$0 = 2b_{s_1} i(x + iy, x - iy) .$$

As $b_{s_1} \neq 0$, we have $(x + iy, x - iy) = 0$. Then

$$0 = (x + iy, x - iy) = \|x\|^2 - \|y\|^2 + 2i(x, y) .$$

From which it follows that $(x, y) = 0$ and $\|x\| = \|y\|$. Also

$$2 = \|x + iy\|^2 = \|x\|^2 + \|y\|^2 + 2i(x, y) .$$

Combining these, we see that

$$\|x\| = \|y\| = 1 .$$

Let $x_1 = x$, $y_1 = y$ and let V_1 be the space spanned by the orthonormal vectors x_1 and y_1 . For any s , $U^s(V_1) = U^s([x_1 + iy_1, x_1 - iy_1]) = V_1$, hence by U6, $U^s(V_1^\perp) = V_1^\perp$.

For any s , the restriction of U^s to $W_1^\perp$ in Case I, or $V_1^\perp$ in Case II is a unitary transformation hence we can continue the decomposition of $\mathbb{C}^n$, in the same manner, by looking at this restriction. Thus it follows that

$$\mathbb{C}^n = V_1 \oplus V_2 \oplus \dots \oplus V_m \oplus W_1 \oplus W_2 \oplus \dots \oplus W_k .$$

The subspace V_j has the orthogonal basis $\{x_j + iy_j, x_j - iy_j\}$ and the real orthonormal basis $\{x_j, y_j\}$. For some s , say s_j , the eigenvalue of U^{s_j} associated with $x_j + iy_j$, call it $a_{s_j} + ib_{s_j}$, is not real, ie $b_{s_j} \neq 0$. The subspace W_j is spanned by two real vectors (which may be dependent) so it too has a real basis. If $W = W_1 \oplus \dots \oplus W_k$, then W has a real basis. Note, it may be that $\mathbb{C}^n = V_1 \oplus V_2 \oplus \dots \oplus V_m$ or $\mathbb{C}^n = W$.

Step 5. Let $U_j^s = U^s|_{V_j}$, $U_W^s = U^s|_W$. For each s let

$$T^s(0) = T_1^s(0) + \dots + T_m^s(0) + T_W^s(0)$$

where $T_j^s(0) \in V_j$, $T_W^s(0) \in W$. As V_j , $j = 1, \dots, m$ and W have real bases, and $T^s(0)$ is real, it is not hard to conclude that $T_j^s(0)$, $j = 1, \dots, m$ and $T_W^s(0)$ are themselves real.

Define

$$T_j^s(z) = U_j^s(z) + T_j^s(0) \quad \text{and} \quad T_W^s(z) = U_W^s(z) + T_W^s(0) .$$

If $z = z_1 + \dots + z_m + z_W$, $z_i \in V_i$, $z_W \in W$, it follows that

$$\begin{aligned} T^s(z) &= \sum (U^s(z_j) + T_j^s(0)) + U^s(z_W) + T_W^s(0) \\ &= \sum (U_j^s(z_j) + T_j^s(0)) + U_W^s(z_W) + T_W^s(0) \\ &= \sum T_j^s(z_j) + T_W^s(z_W) . \end{aligned}$$

Step 6. In this step, and in step 7 the situation in V_j is examined. As noted in step 4, the eigenvalue $a_{s_j} + b_{s_j} i$, associated with $x_j + iy_j$, is not real. The eigenvalue associated with $x_j - iy_j$ is $a_{s_j} - b_{s_j} i$.

We wish to find a vector v_{3j} such that

$$(6) \quad v_{3j} = T_j^{s_j}(v_{3j}) = U_j^{s_j}(v_{3j}) + T_j^{s_j}(0) .$$

Since 1 is not an eigenvalue of $U_j^{s_j}$, $(I - U_j^{s_j})^{-1}$ exists.

Solving for v_{3j} shows that

$$v_{3j} = (I - U_j^{s_j})^{-1} T_j^{s_j}(0) .$$

The vector v_{3j} is real, for if $v_{3j} = u + iv$ then

$$T_j^{s_j}(0) = (I - U_j^{s_j})(u + iv) = (I - U_j^{s_j})u + i(I - U_j^{s_j})v .$$

As $(I - U_j^{s_j})^{-1}$ exists and $T_j^{s_j}(0)$ is real, this shows $v = 0$.

For $z_j \in V_j$ (6) shows

$$\begin{aligned} T_j^{s_j}(z_j) &= (U_j^{s_j}(z_j) + (T_j^{s_j}(0) - v_{3j})) + v_{3j} \\ &= U_j^{s_j}(z_j) - U_j^{s_j}(v_{3j}) + v_{3j} \\ &= U_j^{s_j}(z_j - v_{3j}) + v_{3j} . \end{aligned}$$

Consider Equation (3). Projecting both sides of this onto V_j , it follows that for arbitrary s and r ,

$$(7) \quad (I - U_j^r) T_j^s(0) + U_j^s(T_j^r(0)) = T_j^r(0) .$$

For arbitrary s ,

$$T_j^s(z_j) = U_j^s(z_j - v_{3j}) + U_j^s(v_{3j}) + T_j^s(0) .$$

Now $U_j^s(v_{3j}) + T_j^s(0)$ is equal to

$$(I - U_j^{s_j})^{-1} (I - U_j^{s_j}) (U_j^s (I - U_j^{s_j})^{-1} (T_j^{s_j}(0)) + T_j^s(0)) .$$

As $U_j^{s_j}$ and U_j^s commute (see (2)) this becomes

$$(I - U_j^{s_j})^{-1} (U_j^s T_j^{s_j}(0) + (I - U_j^{s_j}) T_j^s(0)) .$$

From (7) it follows that this equals

$$(I - U_j^{s_j})^{-1} T_j^{s_j}(0) = v_{3j} \quad (\text{definition of } v_{3j}).$$

Hence $T_j^s(z_j) = U_j^s(z_j - v_{3j}) + v_{3j}$. Note this is true for all s .

Step 7. We continue examining the situation in V_j .

Call the projection of $f(s)$ onto V_j , $f_j(s)$. Let

$$f_j(0) - v_{3j} = A_j(\alpha_j x_j + \beta_j y_j) ,$$

where $|\alpha|^2 + |\beta|^2 = 1$, $A_j \geq 0$. As $f_j(0) - v_{3j}$ is real, it must be that α_j and β_j are real.

$$\begin{aligned} \text{Let} \quad v_{1j} &= \alpha_j x_j + \beta_j y_j \\ v_{2j} &= \beta_j x_j - \alpha_j y_j . \end{aligned}$$

These form a real orthonormal basis of V_j . In the basis $\{v_{1j}, v_{2j}\}$ the matrix M_j^s of U_j^s is

$$M_j^s = \begin{pmatrix} a_s & -b_s \\ b_s & a_s \end{pmatrix} .$$

(This is easily checked by applying it to $x_j = \alpha_j v_{1j} + \beta_j v_{2j}$, and $y_j = \beta_j v_{1j} - \alpha_j v_{2j}$.) As $a_s^2 + b_s^2 = 1$, it follows that

$$M_j^s = \begin{pmatrix} \cos \theta_j(s) & -\sin \theta_j(s) \\ \sin \theta_j(s) & \cos \theta_j(s) \end{pmatrix}$$

for some angle $\theta_j(s)$.

We have

$$\begin{aligned} f_j(s) &= T_j^s(f(0)) \\ &= U_j^s(f(0) - v_{3j}) + v_{3j} \\ &= A_j U_j^s(v_{1j}) + v_{3j} \\ &= A_j((\cos \theta_j(s))v_{1j} + (\sin \theta_j(s))v_{2j}) + v_{3j} . \end{aligned}$$

We need to show that θ_j is a group homomorphism from $\mathbb{R}$ to $\mathbb{R}/2\pi$.

As $U_j^{s+r} = U_j^s \cdot U_j^r$, elementary trigonometric identities show that

$$M_j^{s+r} = \begin{pmatrix} \cos(\theta_j(s) + \theta_j(r)) & -\sin(\theta_j(s) + \theta_j(r)) \\ \sin(\theta_j(s) + \theta_j(r)) & \cos(\theta_j(s) + \theta_j(r)) \end{pmatrix}$$

and hence

$$\theta_j(s+r) = \theta_j(s) + \theta_j(r) \pmod{2\pi},$$

showing that θ_j is group homomorphism. Thus the required form of $f_j(s)$, is achieved.

Step 8. Here the situation in W is examined. In choosing the basis vectors for C^n , it was seen that U_W^s has real eigenvalues for all s , and hence by $U3$, these eigenvalues are ± 1 . Now $U^{s/2} \cdot U^{s/2} = U^s$, so $U_W^{s/2} \cdot U_W^{s/2} = U_W^s$, and hence the eigenvalues of U_W^s are all 1. This shows U_W^s is the identity.

Now

$$T_W^{s+r}(z) = U_W^{s+r}(z) + T_W^{s+r}(0) = z + T_W^{s+r}(0)$$

$$T_W^s T_W^r(z) = T_W^s(z + T_W^r(0)) = z + T_W^r(0) + T_W^s(0).$$

As $T_W^{s+r} \equiv T_W^s T_W^r$, it follows that

$$T_W^{s+r}(0) = T_W^r(0) + T_W^s(0).$$

Letting $g(s) = T_W^s(0)$, it follows that $g(s+r) = g(s) + g(r)$,

hence g is a group homomorphism from $\mathbb{R}$ to W , and

$$f_W(s) = T_W^s(f(0)) = g(s) + f_W(0) .$$

Letting $w = f_W(0)$ the required form in W is achieved.

The proof of the theorem is now complete. $\square$

Remark: The von-Neumann-Schoenberg result, where $f(t)$ is continuous, follows easily from this. For if f is continuous then θ_j , $j = 1, \dots, m$ and g must be continuous, in which case it is not difficult to conclude that

$$\theta_j(s) = ks \pmod{2\pi}$$

and

$$g(s) = s \cdot u, \quad u \text{ a fixed vector in } W.$$

This then gives the characterization of a metric transformation of $\mathbb{R}$ into E^n given in the von-Neumann-Schoenberg paper.

In Section 2 it was asserted that a metric transform of a segment can be extended to a metric transform of $\mathbb{R}$. This is proven in Theorem 4.12. We first need a lemma about one parameter subgroups of motions of E^n .

Lemma 4.11: Let $\{T^s \mid |s| < \delta\}$ be a set of motions of E^n satisfying $T^s \circ T^r = T^{s+r}$ whenever s, r and $s+r$ are in $(-\delta, \delta)$. Then there is a unique one parameter subgroup of motions, called $\{\bar{T}^s \mid s \in \mathbb{R}\}$ such that $\bar{T}^s \equiv T^s$, $|s| < \delta$.

Proof: Let s be given. Define $\overline{T}^s$ as follows.

Let m be an integer such that $s/m \in (-\delta, \delta)$ and define $\overline{T}^s$ by $\overline{T}^s = (T^{s/m})^m$. If $s/l \in (-\delta, \delta)$ then

$$(T^{s/m})^m = ((T^{\frac{s}{l \cdot m}})^l)^m = (T^{s/l})^l.$$

Thus the definition of $\overline{T}^s$ is independent of the choice of m .

Let s and r be given. Let m be such that $\frac{s}{m}, \frac{r}{m}$ and $\frac{s+r}{m}$ are in $(-\delta, \delta)$. Then

$$\overline{T}^s \circ \overline{T}^r = (T^{s/m})^m (T^{r/m})^m = (T^{\frac{s}{m} + \frac{r}{m}})^m = \overline{T}^{s+r}$$

Thus $\{T^s\}$ is a one parameter sub-group of motions. Clearly this extension of $\{T^s \mid |s| < \delta\}$ to a one parameter sub-group of motions is the only such extension. $\square$

Theorem 4.12: Let f be a metric transformation of $(-a, a)$ into E^n . Then f can be uniquely extended to a metric transformation $\overline{f}$ of $\mathbb{R}$ into E^n . If $E^m \subseteq E^n$, and $f((-a, a)) \subseteq E^m$, then $\overline{f}(\mathbb{R}) \subseteq E^m$.

Proof: Case I: Assume $f((-a, a))$ spans E^n . The case that it does not will be covered in II. Let $-a < t_0 \leq t_1 \leq \dots \leq t_n < a$ be such that $\{f(t_i)\}$ spans E^n , and let $\delta = \min \{a - t_n, t_0 + a\}$.

For each s , $|s| < \delta$, the function given by $f(t) \rightarrow f(t+s)$ is an isometry of $f([-a+\delta, a-\delta])$ into E^n , hence can be uniquely extended to a motion T^s of E^n (Theorem 4.1).

For s and r such that s, r , and $s+r$ are in $(-\delta, \delta)$

$$T^s \circ T^r (f(t)) = f(t+s+r) = T^{s+r} (f(t)).$$

Because T^{s+r} is the unique motion such that $T^{s+r}(f(t)) = f(t+s+r)$, then

$$T^s \circ T^r = T^{s+r}.$$

Thus $\{T^s \mid |s| < \delta\}$ satisfies the hypotheses of Lemma 4.11, so there is a unique one parameter subgroup of motions $\{\overline{T}^s\}$ which extends $\{T^s \mid |s| < \delta\}$.

Define $\overline{f}(s)$ by $\overline{f}(s) = \overline{T}^s(f(0))$. For $s \in (-a, a)$, and m such that $\frac{s}{m} \in (-\delta, \delta)$.

$$\overline{f}(s) = \overline{T}^s(f(0)) = (T^{s/m})^m(f(0)) = f(m \cdot \frac{s}{m}) = f(s).$$

Thus $\overline{f}$ extends f . The uniqueness follows from the uniqueness in Lemma 4.11, and the fact that every metric transformation $\overline{f}$ of $\mathbb{R}$ can be written as

$$\overline{f}(s) = \overline{T}^s(f(0))$$

for some one parameter subgroup of motion $\{\overline{T}^s\}$ of E^n (Step 1, Theorem 4.10).

Case II. Consider now the case $f((-a, a))$ does not span E^n . Let E^m be the m -flat of E^n which contains, and is spanned by $f((-a, a))$. Let $\overline{f}$ be any extension of f to $\mathbb{R}$ and assume $\overline{f}(\mathbb{R})$ spans E^l . As above, let δ be such that $f([-a+\delta, a-\delta])$ spans E^l . As in Theorem 4.10, let $\{T^s\}$ be a one parameter subgroup of motions of E^l such that $T^s(f(t)) = f(t+s)$. For $|s| < \delta$,

$$T^s(f([-a+\delta, a-\delta])) = f([-a+\delta+s, a-\delta+s]) \subseteq E^m.$$

Thus, by Corollary 4.3, $T^S(E^m) \subseteq E^m$.

For arbitrary s , let p be an integer such that $\frac{s}{p} \in (-\delta, \delta)$. Then $T^S(E^m) = (T^{s/p})^p(E^m) \subseteq E^m$. Thus $\bar{f}(s) = T^S(f(0)) \in E^m$ for all s , so $E^\ell = E^m$, and the uniqueness of the extension follows from Case I. $\square$

Differential geometers cannot help but feel that this work is closely related to the several studies of generalized helices in the literature and those interested in the study of topological groups, transformation groups, and topological dynamics will observe that this characterization of metric transforms of $\mathbb{R}$ in E^n can be viewed as characterizing the orbits of a one parameter subgroup of the group of motions of E^n . No doubt the extensive literature in these fields contains similar characterizations of screw curves, but probably under stronger hypotheses than we impose here. In any event, we hope that this derivation under these weak hypotheses and resorting to relatively elementary tools will prove useful.

Finally we should note the close connection between these ideas and some recent work of Grunbaum and his students concerning regular polygons in E^n . In that work, polygons of varying degrees of regularity are studied. A polygon $P_1, P_2, \dots, P_m$ is called k -regular if all $\overline{P_i P_{i+j}} = c_j$ for all i and for $j = 1, 2, \dots, k$. We could describe a regular polygon in E^n as one whose vertex space is an order isomorph of a regular planar polygon. It is easy to see that there are regular polygons with $2k$ vertices spanning E^n for any n and $k > \frac{n}{2}$, but J. Lawrence [22] has shown that odd sided

regular polygons span only even dimensional spaces. His reasoning employs transformation ideas similar to those in our proof.

Our early reflection on the metric determinancy problem led us to conjecture that regular polygons with a large number of vertices would probably be planar but, as the above discussion shows, this was rather naive. Even the ultimate regular convex planar figure, the circle, has order isomorphs spanning arbitrarily high-dimensional Euclidean spaces.

§4. Most of the results of this thesis, up to this point, have been of the following type. The function f is a metric transformation, whose domain contains a convex subset. Then f is shown to be a similarity by first showing that the range also contains a convex subset (compare Theorems 2.16, 3.16, 3.18, 4.7). In this section we shall consider order transformations of the unit sphere of an inner product space. The unit sphere of such a space is certainly not convex, so it seems we have managed to move away from the convexity conditions. However, as the surface metric of the sphere of an inner product space, and the norm metric, are related by an order transformation (see Lemma 4.14), convexity again plays a role.

Notice that we have been talking in terms of order transformations here. The main theorem of this section, Theorem 4.15, is stated, and proved only for order transformations. This theorem shows that an order transformation from the unit sphere of an inner product space onto the unit sphere of a

normed linear space must be a similarity, and the N.L.S. must be an inner product space. The theorem is, at least for finite dimensions greater than two, true for metric, rather than order transformations. The proof of this generalization is outlined after the proof of Theorem 4.15.

The proof of Theorem 4.15 relies on the following characterization of inner product spaces, due to Senechalle [34].

Theorem 4.13 (Senechalle): Let M be a normed linear space with unit sphere S . Then M is an inner product space if and only if there is a function F such that for each p and q in S

$$F(\|p - q\|) = \|p + q\|.$$

Proof: See [34].

The following lemma is an easy consequence of some of our work in Chapter 2. It is an interesting result about metric transformations, although it certainly is not surprising. The assumption of bicontinuity is not necessary for $\dim E_1 \geq 3$, as Theorem 2.9 can be applied. However, for our purposes, the following suffices.

Lemma 4.14: Let S_1 and S_2 be the unit spheres of inner product spaces E_1 and E_2 respectively. Let $f: S_1 \xrightarrow{\text{onto}} S_2$ be a bicontinuous metric transformation. Then f is an isometry.

Proof: Let the metric given by the inner product on S_1 and S_2 be d_1 and d_2 respectively, and let the intrinsic metrics be d'_1 and d'_2 respectively. (The intrinsic distance between x and y being the length of the shortest arc of the great circle joining x to y .) Note that (S_1, d'_1) and (S_2, d'_2) are convex metric spaces.

Let $g_1 : (S_1, d_1) \rightarrow (S_1, d'_1)$ and $g_2 : (S_2, d_2) \rightarrow (S_2, d'_2)$ be given by $g_1(x) = x$, $g_2(x) = x$. Both g_1 and g_2 are order transformations with scale function ρ , where $\rho(r) = 2 \sin^{-1} \frac{r}{2}$. Thus the function $g_2 f g_1^{-1}$ is a bicontinuous metric transformation between the convex metric spaces (S_1, d'_1) and (S_2, d'_2) . Theorem 2.17 then shows that $g_2 f g_1^{-1}$ is a similarity, and it follows that indeed it is an isometry. Hence, f is an isometry. $\square$

Before proceeding to the proof of Theorem 4.15, recall from Chapter 1 that an order transformation has a one-to-one strictly increasing scale function. Hence the scale function and the order transformation are both invertible. See Chapter 1, §2.

Theorem 4.15: Let (E, d) be an inner product space with unit sphere S . Let M be a normed linear space, with unit sphere U . If $f : S \xrightarrow{\text{onto}} U$ is an order transformation, then f is an isometry, and M is an inner product space.

Proof: Let $\|\cdot\|$ be the norm in M and let ρ be the scale function for f . Then ρ is one-to-one. First it is shown that U is strictly convex, and then that antipodal points of

S are mapped by f to antipodal points of U .

Since 2 is the maximum distance between points of S , f preserves the order of the distances, and 2 is the maximum distance between points of U , it follows that $\rho(2) = 2$.

Let $x \in S$. The scale function ρ of f is one to one, and $-x$ is the only point a distance 2 from x , so it follows that $f(-x)$ is the only point a distance $\rho(2) = 2$ from $f(x)$. This shows that $f(-x) = -f(x)$.

Let $\bar{x}, \bar{y} \in U$. Then

$$\begin{aligned}
 \|\bar{x} + \bar{y}\| &= \|\bar{x} - (-\bar{y})\| \\
 &= \rho(d(f^{-1}(\bar{x}), f^{-1}(-\bar{y}))) \\
 &= \rho(d(f^{-1}(\bar{x}), -f^{-1}(\bar{y}))) \\
 &= \rho(\sqrt{4 - d^2(f^{-1}(\bar{x}), f^{-1}(\bar{y}))}) \\
 &= \rho(\sqrt{4 - (\rho^{-1}(\|\bar{x} - \bar{y}\|))^2})
 \end{aligned}$$

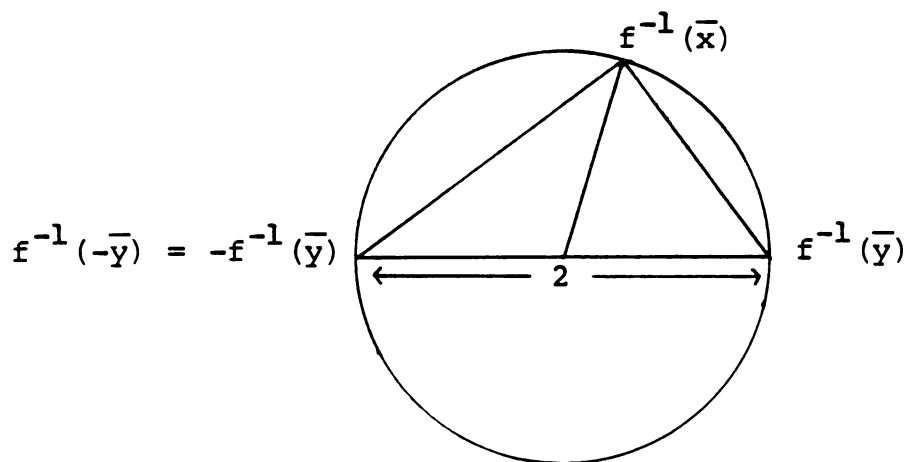


Figure 4.4

Senechalle's Theorem, Theorem 4.13, now shows that M is an inner product space. Lemma 4.14 now shows that f is an isometry. $\square$

Corollary 4.16: Let (E, d) be an inner product space with unit sphere S . Let M be a N.L.S. with unit sphere U and $\dim M = \dim E < \infty$. If $f: S \rightarrow U$ is an order transformation, then f is an isometry, and M is an inner product space.

Proof: Let $\dim M = \dim E = n$. Note that U is homeomorphic to S . Any proper open subset of S is homeomorphic to an open subset of E^{n-1} , and hence by the Invariance of Domain Theorem (Theorem 3.5), is mapped by f onto an open subset of U . Thus $f(S)$ is open in U . On the otherhand $f(S)$ is a compact subset of U , hence (as U is a Hausdorff space) $f(S)$ is closed in U . Since U is connected, it follows that $f(S) = U$. $\square$

Corollary 4.17: If $\dim E = \dim M = 2$, and $f: S \rightarrow M$ is an order transformation then M is an inner product space (ie the Euclidean plane) and $f(S)$ is similar to S .

Proof: Fix a point p of S . Then for $x \in S$ we have the first configuration shown in Figure 4.5.

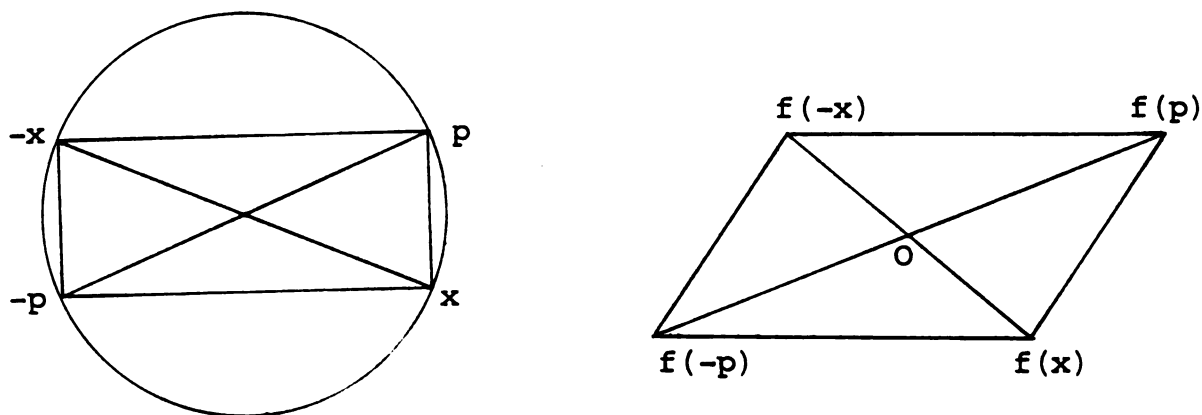


Figure 4.5

Since f is an order transformation, it is easy to see that the rectangle $(-x)px(-p)$ is transformed into a parallelogram, whose diagonals have equal length. These diagonals bisect each other at a point O , the midpoint of $[f(-p), f(p)]$. Thus $\|f(p) - O\| = \|f(x) - O\|$, and hence $f(S)$ lies on the circle $\{y \mid \|y - O\| = \frac{1}{2}\rho(2)\}$, so the corollary follows from Corollary 4.16. $\square$

As indicated earlier, Theorem 4.15 remains valid if the assumption of "order transformation" is replaced by that of "metric transformation", and $3 \leq \dim E < \infty$. The proof of this is only outlined here.

Theorem 4.15(a): Let M be a N.L.S. with unit sphere U . Let S be the unit sphere of E^n , $n \geq 3$. If $f: S \xrightarrow{\text{onto}} U$ is a metric transformation, then f is an isometry, and M is an n -dimensional inner product space.

Outline of Proof: That it is necessary for n to be at least 3 follows by noting that a circle is a screw curve, and can be mapped onto itself by a discontinuous metric transformation. For example, $e^{it} \rightarrow e^{i\phi(t)}$ where ϕ is any group homomorphism of $\mathbb{R}$. (Example 2 of Chapter 3 discussed this type of metric transformation.)

The proof that f is an isometry is the same as that of Theorem 4.15, once the following three facts are established. The function f is bicontinuous, $f(-x) = -f(x)$, and the scale function ρ is one-to-one. We shall now informally discuss the proofs of these in terms of the three dimensional case. The proofs remain virtually the same for higher finite dimensions.

To show that f is bicontinuous, Theorem 2.9 is used. It is not hard to see that the unit sphere S of E^3 has the long legged local isosceles property with $\lambda(p) = \sqrt{3}$. Theorem 2.9 then shows that f is locally bicontinuous and, in particular, $\rho(d) \neq 0$ for $d < \sqrt{3}$.

We now show that f is one-to-one. It is here, and only here, that we have not settled the theorem for the case that

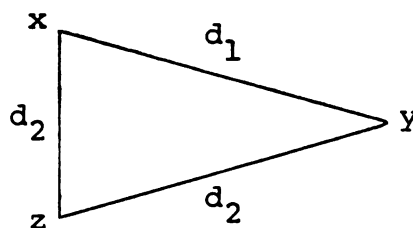


Figure 4.6

E and M have infinite dimension. Let $d_1, \sqrt{3} \leq d_1 \leq 2$ be given. Then there would be points x, y, z of S as in Figure 4.6, where d_2 is some distance less than $\sqrt{3}$.

If $\rho(d_1) = 0$, it follows that $\rho(d_2) = 0$, contradicting $\rho(d) \neq 0$ for all $d < \sqrt{3}$.

We have now shown $\rho(d) \neq 0$ for any d , hence f is one-to-one.

To show f is onto, an argument such as the one in Corollary 4.16 can be used. Hence, f is bicontinuous.

Next we show that $f(-p) = -f(p)$ for any $p \in S$. Let p be an arbitrary point of S and let $A = \{x \in S \mid \rho(d(p, x)) = 2\}$ and $B = \{y \in U \mid \|f(p) - y\| = 2\}$. Then $f(A) = B$, so the sets A and B , as well as their complements, are homeomorphic.

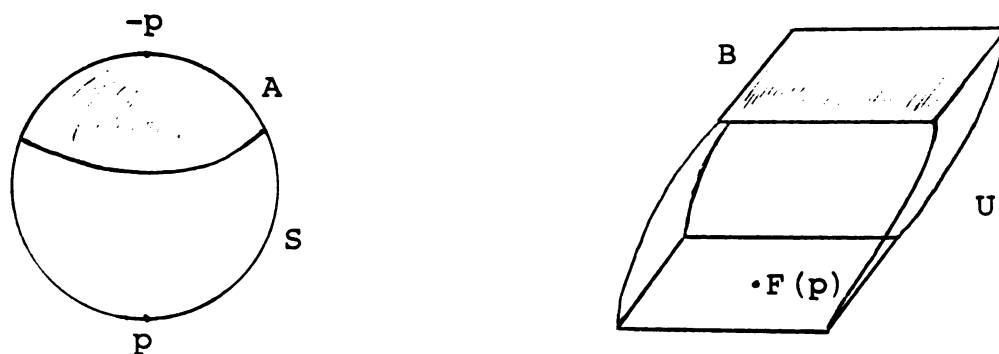


Figure 4.7

The set B can be shown to be a convex subset in U . The set A must consist of a family of parallel "lines of latitude" with p as the south pole). As A is homeomorphic to B , A must be a spherical cap and must contain $-p$.

If A is not equal to $\{-p\}$, then every spherical cap the same size as A is mapped onto a convex set in U . Clearly this leads to a contradiction. Thus $A = \{-p\}$, and $B = \{f(-p)\}$. Since $-f(p) \in B$, it follows that $f(-p) = -f(p)$.

To show ρ is one-to-one, a similar approach is taken. Let $A = \{x \in S \mid \rho(d(p, x)) = r\}$ and $B = \{y \in U \mid d(f(p), y) = r\}$. As before, A and B are homeomorphic and A contains a number of lines of latitude of S . However, B can be shown to be a simple closed curve in U (for $r < 2$). Thus A must consist of one line of latitude, so ρ is one-to-one.

The proof that f is an isometry can now proceed as in Theorem 4.15. $\square$

The questions raised in this section could be asked of order and metric transformations between the spheres of arbitrary N.L.S. We have made no progress on this type of question. While considering it, we asked the following: If U_1 and U_2 are the unit spheres of N.L.S. M_1 and M_2 respectively, and $f : U_1 \xrightarrow{\text{onto}} U_2$ is an isometry, is M_1 isometric to M_2 ? In other words, can f be extended to M_1 ? This problem seems to be quite difficult, and yet should be easier than questions about metric and order transformations between spheres. If such theorems are true, they would supercede theorems of the Mankiewicz type discussed in Chapter 3.

§5.

Appendix to Chapter 4

In proving Theorem 4.5, the following fact was used.

Let

$$h(\alpha) = \sum_{j=1}^v B_j^2 \sin^2 h_j \alpha$$

$$g(\alpha) = 4r^2 c^2 \sin^2 \frac{\alpha}{2} + \sum_{j=1}^s A_j^2 \sin^2 (2k_j r \sin \frac{\alpha}{2}) ,$$

where $B_j, h_j, c, A_j,$ and k_j are all non-negative constants. If $h(\alpha) = g(\alpha)$ for all $\alpha \geq 0$ on a non-empty interval I , then for each $j = 1, \dots, s$ either A_j or k_j is zero.

The following proof of this is due to Dr. L. Sonneborn.

Assume that $h(\alpha) = g(\alpha)$, for $\alpha \in I$. Consider h and g as functions of a complex variable z . Both are analytic functions on the entire complex plane.

By a well known theorem of complex analysis, if $h(z) = g(z)$ for all z in any set having an accumulation point, then $h(z) = g(z)$ for all z . Since we know that $h(\alpha) = g(\alpha)$ for every α in I , it must be the case that $h(z) = g(z)$ for all z .

Let $z = it$, t real. Then

$$\begin{aligned} h(it) &= \sum_{j=1}^v B_j^2 \sin^2 h_j it \\ &= \sum_{j=1}^v B_j^2 \left(\frac{e^{-h_j t} - e^{h_j t}}{2i} \right)^2 \\ &= - \sum_{j=1}^v \frac{B_j^2}{4} (e^{-2h_j t} - 2 + e^{2h_j t}) . \end{aligned}$$

Let $b = \max\{h_j\}$. Then

$$\lim_{t \rightarrow \infty} e^{-(2b+1)t} h(it) = 0.$$

Similarly, $\lim_{t \rightarrow \infty} e^{-(2b+1)t} \cdot 4r^2 c^2 \sin^2 \frac{it}{2} = 0$. Now

$$\begin{aligned} & e^{-(2b+1)t} \sum_{j=1}^s A_j^2 \sin^2(2k_j r \sin \frac{it}{2}) \\ &= e^{-(2b+1)t} \sum_{j=1}^s A_j^2 \left(\frac{e^{i2k_j r \sin \frac{it}{2}} - e^{-i2k_j r \sin \frac{it}{2}}}{2i} \right)^2 \\ &= e^{-(2b+1)t} \sum_{j=1}^s A_j^2 \left(\frac{e^{2k_j r (e^{-t/2} - e^{t/2})} - 2 + e^{-2k_j r (e^{-t/2} - e^{t/2})}}{-4} \right). \end{aligned}$$

Taking the limit as $t \rightarrow \infty$, we get $-\infty$ unless either A_j or k_j is zero for each $j, j = 1, \dots, s$. From the above we know this limit must be 0, so A_j or k_j is zero for each j . $\square$

CHAPTER 5 FURTHER RESULTS IN EUCLIDEAN SPACE

In this chapter we consider metric transformations between hypersurfaces immersed in E^{n+1} , $n \geq 2$. Since we make use of results from differential geometry, the hypersurfaces are required to be smooth, and the metric transformations to be diffeomorphisms. It is not our intention here to carry out a complete investigation of diffeomorphic metric transformations, but rather to show some applications of the theory and techniques outlined in chapters one through four. We suspect that there is much room for generalizing the results of this chapter.

Before proceeding, a word of caution. Normally, a differential geometer calls a function which preserves arc length an isometry. In our work we consider, for the most part, hypersurfaces in E^{n+1} , and continue to mean by an isometry a function which preserves the Euclidean metric. A map which preserves arc length is called equilong.

The links connecting preceeding chapters to differential geometry are Theorems 2.14 and 2.20. Let $f: S \xrightarrow{\text{onto}} \bar{S}$ be a diffeomorphic metric transformation with spread 1. Theorem 2.14 shows that, for any rectifiable arc γ in S , the length of

γ equals the length of $f(\gamma)$. That is, f is equilong. Thus, the first fundamental forms of S and $\bar{S}$ are the same. (See §1 for definitions, symbols and references). Theorem 2.20 shows that, for a differentiable curve γ in S and a point $p \in \gamma$, the difference of the squares of the curvatures of γ at p and of $f(\gamma)$ at $f(p)$ is constant. The constant is independent of both γ and p . This enables us to relate the second fundamental forms of S and $\bar{S}$ and, in the case of hypersurfaces in E^{n+1} , to show that they must be the same, up to sign. See Theorem 5.6.

Because f , a metric transformation with spread 1, preserves the first fundamental form (ie. arc length), a large body of knowledge is immediately applicable. This body of knowledge, known as "rigidity theory", investigates diffeomorphisms between submanifolds of a manifold. The diffeomorphisms are assumed to preserve the first fundamental form. The question asked is whether or not such a diffeomorphism can be extended to a diffeomorphism from the manifold to itself, with the extension also preserving the first fundamental form. One of the most famous rigidity theorems is the following.

Theorem 5.1: (Cohn-Vossen) If $S \subset E^3$ is a compact surface which is the boundary of a convex set, then every equilong diffeomorphism of S into E^3 is an isometry.

Proof: See [36] 5, p 280 Theorem 12.

It follows immediately from Theorem 5.1 that a diffeomorphic metric transformation, defined on the boundary of a compact convex set in E^3 , into E^3 , is a similarity.

Another famous rigidity theorem, used later in this chapter, is the following.

Theorem 5.2: Let S and $\bar{S}$ be differentiable hypersurfaces immersed in E^{n+1} , and let $f: S \xrightarrow{\text{onto}} \bar{S}$ be an equilong diffeomorphism. Assume also that the rank of the differential is at least 3 for all points of S . Then f is an isometry.

Proof: See [36] 5, p 244 Theorem 1.

Theorem 5.10, the main result of this chapter, eliminates the assumption of Theorem 5.2 that the rank of the differential be at least 3.

The rigidity theorems for hypersurfaces in E^{n+1} (such as 5.1 and 5.2) in themselves seem to support the principle of "metric determinacy". That is, they seem to confirm that an order transformation is "usually" an isometry. (See Chapter 1, §3). Theorem 5.10, and the results of Chapter 4, may make the reader feel very comfortable with the principle of metric determinacy, for subsets of E^{n+1} . However, questions about metric transformations between submanifolds of dimension less than n , in E^{n+1} , have not been addressed. For example, we have not determined whether or not a metric

transformation of an $n-1$ dimensional manifold immersed in E^{n+1} is necessarily an isometry.

The case of a submanifold of dimension one is particularly interesting. In Chapter 4, all the metric transformations of the real line into E^n are classified. It is easily seen that if a continuous metric transformation of the real line has an inverse which is also a metric transformation, then that metric transformation is necessarily an order transformation. If f_1 is an order transformation of $\mathbb{R}$ into E^n , f_2 a metric transformation of $\mathbb{R}$ into E^m . Then $f_2 \circ f_1^{-1}$ is a metric transformation from a curve in E^n into E^m . One may ask whether or not this is the only type of metric transformation from a curve in E^n into E^m . The answer to this is by no means clear.

The remainder of this chapter is divided into 2 sections. Section 1 contains a list of symbols, some definitions, and a few elementary lemmas. Section 2 contains the main results of this chapter, in particular Theorem 5.10.

§1. Notation, Definitions, and Elementary Lemmas.

Throughout this chapter S and $\bar{S}$ (note that $\bar{S}$ is not, as in Chapter 4, the closure of S , but a new symbol) are n -dimensional differentiable hypersurfaces immersed in E^{n+1} ($n \geq 2$). The function $f: S \xrightarrow{\text{onto}} \bar{S}$ is a diffeomorphic metric transformation and, as usual, ρ is the scale

function associated with f . For any $U \subseteq S$, $\bar{U} \equiv f(U)$ and in particular, for $x \in S$, $\bar{x} \equiv f(x)$. If v is any vector tangent to S at p , then $\bar{v}$ is the corresponding vector tangent to $\bar{S}$ at $\bar{p}$, under the map between the tangent spaces induced by f .

We have tried to keep the notation and terminology "standard" in this chapter. For the reader's convenience we now present a list of symbols used in this chapter, along with their meanings. We refer the reader to Spivak [37] for any terms he may wish defined in more detail.

Let $p \in S$, and $\gamma \subseteq S$ be a differentiable arc with $p \in \gamma$.

$\kappa_\gamma(p)$ The curvature of the curve γ at p , as defined in differential geometry. Called the "classical curvature" in Chapter 2.

S_p The tangent space to S at p .

The symbols $N(p)$, dN_p , $\kappa_n(p,v)$, and Π_p are only defined for S an orientable hypersurface.

$N(p)$ The unit normal at p in the given orientation of S .

dN_p The differential of N at the point p , called the Weingarten Map.

Let $u, v \in S_p$.

$\kappa_N(p,v)$ The normal curvature of S at p in the direction v .

$I_p(u,v)$ The first fundamental form of S at p , evaluated at u and v .

$\Pi_p(u,v)$ The second fundamental form of S at p , evaluated at u and v .

$l(\gamma)$ The length of the arc γ .

$n(p, \alpha)$ The normal to the curve α at p .

Let $v_1, \dots, v_k$ be vectors in E^{n+1} .

$\|v_1\|$ The standard norm in Euclidean space.

$\langle v_1, v_2 \rangle$ The inner product of v_1 and v_2 .

$[v_1, \dots, v_k]$ The subspace of E^n spanned by $v_1, \dots, v_k$.

$[v_1, \dots, v_k]^\perp = \{u \mid \langle u, v \rangle = 0, v \in [v_1, \dots, v_k]\}.$

Note that a symbol with "bars" is the corresponding symbol in $\bar{S}$. For example $\kappa_{\bar{N}}(\bar{p}, \bar{v})$ is the normal curvature of $\bar{S}$ at $\bar{p}$ in the direction $\bar{v}$.

Equilong maps were referred to earlier in this chapter. Their formal definition now follows.

Definition: A diffeomorphism $f: S \xrightarrow{\text{onto}} \bar{S}$ is said to be equilong if and only if $I_p \equiv \bar{I}_p$ for all $p \in S$.

A standard result of differential geometry shows that an equilong function preserves arc length. As noted earlier, differential geometers use the term isometry, rather than equilong. In this chapter, f is an isometry will continue to mean $d(x, y) = d(f(x), f(y))$ for all $x, y \in S$.

Throughout this thesis, we have considered a metric transformation, or order transformation, f . We have tried to assume as little as possible about f , preferring to make hypotheses about the domain and range, rather than the

mapping itself. For example, although we often assume f to be continuous and to have finite spread, Theorem 2.9 and 2.16 show that this is true if the domain and range contain rectifiable arcs. In this chapter, we are not so fortunate.

For Theorem 5.10, our main result, we ask that f have spread 1, and be a diffeomorphism. The assumption of spread 1, rather than finite spread, allows us to prove f is an isometry, rather than a similarity, and is included for convenience.

Unfortunately, we are unable to show anything about the differentiability of a metric transformation between two differentiable hypersurfaces. However, Theorem 5.3 shows that a metric transformation from a differentiable Euclidean hypersurface is necessarily continuous and locally bicontinuous.

Theorem 5.3: If $f : \Sigma \xrightarrow{\text{onto}} \bar{\Sigma} \subset E^n$ is a metric transformation, where Σ is a connected regular hypersurface, then f is continuous and there is a $\lambda > 0$ such that $f|_{B(t, \lambda/2) \cap \Sigma}$ is bicontinuous for any $t \in \Sigma$.

Since none of our future work depends on this, we only outline the proof.

Proof: We need the following standard facts about a hypersurface Σ regular at a point p .

1. Σ has a unique tangent hyperplane at p and a unique normal line p_n at p . For any sequence of points $\{p_i\}$ on Σ approaching p the measures of angles $\{\angle n p p_i\}$ approach 90° .
2. There exists a cylindrical neighborhood σ of p , with axis p_n , such that $\sigma \cap \Sigma$ is a topological $(n-1)$ -ball D and $\sigma - D$ consists of exactly two components.

Based on these observations we can conclude that there is a cylindrical neighborhood, σ , of p , axis p_n , radius λ such that for any $x \in D = \Sigma \cap \sigma$ the measure of angle $\angle n p x$ is between 89° and 91° . This means that no point of D is in the double napped cone with vertex p , axis p_n and vertex angle 89° .

Now for $x \in D$, consider the plane $n p x$ and note that there are two points in this plane, q and q^* such that $pq = qx = pq^* = q^*x = \lambda$ and further that q and q^* are in σ with q in one component of $\sigma - D$ and q^* in the other.

Finally the locus of points y in E^n such that $py = yx = \lambda$ is an $(n-2)$ sphere which is a subset of σ

This connected set must intersect D and thus there is a point m of Σ such that $pm = mx = \lambda$. This implies that Σ has the $\mathcal{L}_1$ property at p .

If there are two points p and q of Σ such that $f(p) \neq f(q)$ then there is a rectifiable arc joint p to q . Hence by Theorem 2.9 $f|_{B(t, \lambda/2)}$ is bicontinuous for any $t \in \Sigma$. $\square$

Let $f : S \xrightarrow{\text{onto}} \bar{S}$ be a metric transformation, and let $S, \bar{S}$ be smooth hypersurfaces in E^{n+1} , $n \geq 2$. Combining Theorems 5.3, 2.9 and 2.16, it is seen that f is necessarily continuous, locally bicontinuous, and has finite spread. Unfortunately, since we require f to be a diffeomorphism, this information is superfluous for our present purposes.

The following two lemmas are used to prove Theorem 5.6. They are standard results of linear algebra.

Notation: Let T be a linear transformation. We write Tx for $T(x)$.

Lemma 5.4: Let $T : E^n \rightarrow E^n$ be a symmetric linear transformation. Let e_1 be such that

$$|\langle Te_1, e_1 \rangle| = \max_{\|x\|=1} |\langle Tx, x \rangle|.$$

Then e_1 is an eigenvector of T , with $\langle Te_1, e_1 \rangle$ the corresponding eigenvalue.

Proof: Since T is symmetric there is a set of orthonormal eigenvectors of T spanning E^n , call them $\{b_i\}_{i=1}^n$, with $Tb_i = \kappa_i b_i$. Let A be the matrix for T in the basis $\{b_i\}$,

$$A = \begin{bmatrix} \kappa_1 & & 0 \\ & \ddots & \\ 0 & & \kappa_n \end{bmatrix}.$$

Let $x = \sum \alpha_i b_i$, with $\sum \alpha_i^2 = 1$, (i.e. $\|x\| = 1$). Then $\langle Tx, x \rangle = \sum \kappa_i \alpha_i^2$. From this we can see that both $\max_{\|x\|=1} \langle Tx, x \rangle$ and $\min_{\|x\|=1} \langle Tx, x \rangle$ are eigenvalues, and the values of x where these are obtained are eigenvectors. Thus it follows that if

$$|\langle Te_1, e_1 \rangle| = \max_{\|x\|=1} |\langle Tx, x \rangle|,$$

then e_1 is an eigenvector, and $\langle Te_1, e_1 \rangle$ is the corresponding eigenvalue. $\square$

Lemma 5.5: We can define a complete set of eigenvectors for the symmetric linear transformation T as follows. Let e_1 be such that

$$|\langle Te_1, e_1 \rangle| = \max_{\|x\|=1} |\langle Tx, x \rangle|.$$

Assume $e_1, \dots, e_{k-1}$ have been defined, $k \leq n$. Let e_k be defined by

$$|\langle Te_k, e_k \rangle| = \max_{\substack{\|x\|=1 \\ x \in [e_1, \dots, e_{k-1}]^\perp}} |\langle Tx, x \rangle|.$$

Proof: This follows immediately from Lemma 5.4 by noting that for $k \leq n$, $T_k = T|_{[e_1, \dots, e_k]^\perp}$ is a symmetric linear transformation on $[e_1, \dots, e_k]^\perp$, and that any eigenvalue or eigenvector of T_k is an eigenvalue or eigenvector respectively of T . $\square$

§2. This section contains the main result of this chapter, Theorem 5.10. As has been mentioned, Theorem 5.10 shows that a diffeomorphic metric transformation f , with spread 1, between connected hypersurfaces S and $\bar{S}$ in E^{n+1} , is an isometry. Theorem 5.6 contains the crucial argument: that, up to sign, f preserves the second fundamental form. If $p \in S$ is such that $\Pi_p \neq 0$, standard theorems of differential geometry then show that f is, in a neighborhood of p , an isometry. Theorem 5.9 then shows, by purely metric means, that f itself must be an isometry.

Theorem 5.6: Let S be an orientable differentiable hypersurface immersed in E^{n+1} . Let $f: S \xrightarrow{\text{onto}} \bar{S} \subseteq E^{n+1}$ be a diffeomorphic metric transformation with spread 1. Then,

for any $p \in S$ and $u, v \in S_p$,

$$\bar{\Pi}_p(\bar{u}, \bar{v}) = \Pi_p(u, v) \quad \text{or} \quad \bar{\Pi}_p(\bar{u}, \bar{v}) = -\Pi_p(u, v) .$$

Proof: We abbreviate this by saying $\bar{\Pi}_p \equiv \pm \Pi_p$. As f has spread 1, by Theorem 2.14, f is equilong. Also, by Theorem 2.20,

$$(1) \quad \kappa_{\bar{\gamma}}^2(\bar{p}) - \kappa_{\gamma}^2(p) = -4 : \lim_{d \rightarrow 0} \frac{\rho(d) - d}{d^3} \stackrel{\text{def}}{=} c .$$

Assume $c \geq 0$. If not, interchange f and f^{-1} . (Although f^{-1} is not necessarily a metric transformation, (1) still holds and f^{-1} is equilong. As we do not use the fact that f is a metric transformation further in this proof, we may now interchange, if necessary, f and f^{-1} .)

Let $p \in S$, $v \in S_p$, $\|v\| = 1$. Let $\alpha(s)$ be a regular curve, parametrized by arc length, with $\alpha'(0) = v$, and $\alpha(0) = p$. Then,

$$\begin{aligned} \Pi_p(\alpha'(0), \alpha'(0)) &= \langle \alpha''(0), N(p) \rangle \\ &= \kappa_{\alpha}(p) \langle n(p, \alpha), N(p) \rangle \\ (2) \quad &= -\langle dN_p(\alpha'(0)), \alpha'(0) \rangle \\ &= \kappa_N(p, \alpha'(0)) . \end{aligned}$$

More generally, for $u, v \in S_p$,

$$\Pi_p(u, v) = -\langle dN_p(u), v \rangle .$$

To prove Theorem 5.6, it suffices to show that $d\bar{N}_p \equiv \pm dN_p$.

First we show $\kappa_N(\bar{p}, \bar{v}) = \pm \sqrt{\kappa_N^2(p, v) + c}$ for each $p \in S$, $v \in S_p$, $\|v\| = 1$. In the following, $\alpha(s)$ is a curve with $\alpha(0) = p$ and $\alpha'(0) = v$. Now

$$|\kappa_N(p, v)| = |\kappa_\alpha(p)| |\langle n(p, \alpha), N(p) \rangle|.$$

As there are curves $\alpha(s)$, (for example, the normal section), with $|\langle n(p, \alpha), N_p \rangle| = 1$, we can write

$$(3) \quad |\kappa_N(p, v)| = \min_{\alpha(s)} |\kappa_\alpha(p)|.$$

From (1) it follows that

$$(4) \quad |\kappa_N(\bar{p}, \bar{v})| = \min_{\alpha(s)} |\kappa_\alpha(\bar{p})| = \min_{\alpha(s)} \sqrt{\kappa_\alpha^2(p) + c}.$$

Clearly the minima (3) and (4) occur for the same curve $\alpha(s)$. Hence,

$$|\kappa_N(\bar{p}, \bar{v})| = \sqrt{\kappa_N^2(p, v) + c}.$$

Next, apply Lemmas 5.4 and 5.5 to the symmetric transformation $dN_p: \mathbb{R}^n \rightarrow \mathbb{R}^n$ and define the eigenvector e_1 of dN_p by

$$\begin{aligned} |\kappa_N(p, e_1)| &= |\langle dN_p e_1, e_1 \rangle| \\ &= \max_{\|x\|=1} |\langle dN_p x, x \rangle| \\ &= \max_{\|x\|=1} |\kappa_N(p, x)|. \end{aligned}$$

Let $|\kappa_N(p, e_1)| = \kappa_1$. Thus e_1 is an eigenvector of dN_p and κ_1 the corresponding eigenvalue. Since f preserves the first fundamental form, $\|x\| = \|\bar{x}\|$ and

$$\max_{\|\bar{x}\|=1} |\langle d\bar{N}_p \bar{x}, \bar{x} \rangle| = \max_{\|x\|=1} |\kappa_N(\bar{p}, \bar{x})| = \max_{\|x\|=1} \sqrt{\kappa_N^2(p, x) + c}.$$

This maximum occurs for the same value of x as the preceding maximum, that is for $x = e_1$. Hence

$$\max_{\|\bar{x}\|=1} |\langle d\bar{N}_p \bar{x}, \bar{x} \rangle| = |\langle d\bar{N}_p \bar{e}_1, \bar{e}_1 \rangle| = \sqrt{\kappa_N^2(p, e_1) + c}.$$

Thus, using Lemma 5.4, $\bar{e}_1$ is an eigenvector, corresponding to the eigenvalue $\kappa_1 \stackrel{(\text{defn})}{=} \langle d\bar{N}_p \bar{e}_1, \bar{e}_1 \rangle$. From the above,

$$\bar{\kappa}_1 = \pm \sqrt{\kappa_N^2(p, e_1) + c} = \pm \sqrt{\kappa_1^2 + c}.$$

Using Lemma 5.5, and an argument as above, choose $e_2, \dots, e_n$ eigenvectors of dN_p , corresponding to the eigenvalues $\kappa_i = \kappa_N(p, e_i)$ while $\bar{e}_2, \dots, \bar{e}_n$ are eigenvectors of $d\bar{N}_p$ corresponding to the eigenvalues $\bar{\kappa}_i = \pm \sqrt{\kappa_i^2 + c}$. In the bases $e_1, \dots, e_n$ and $\bar{e}_1, \dots, \bar{e}_n$ we have

$$dN_p = \begin{pmatrix} \kappa_1 & & 0 \\ & \ddots & \\ 0 & & \kappa_n \end{pmatrix} \quad d\bar{N}_p = \begin{pmatrix} \bar{\kappa}_1 & & 0 \\ & \ddots & \\ 0 & & \bar{\kappa}_n \end{pmatrix}.$$

Assume that dN_p has rank at least 3. Since dN_p is a continuous function in p , and a differentiable hypersurface is locally connected, there is a neighborhood $U \subseteq S$ of p such that dN_q has rank at least 3 for all $q \in U$. Remembering that f is equilong, and applying Theorem 5.2,

we see that $f|_U$ is an isometry (i.e. an Euclidean motion) and hence $\Pi_p \equiv \pm \bar{\Pi}_p$.

Now assume that dN_p has rank at most 2 and that $\kappa_3 = \dots = \kappa_n = 0$. By the Theorema Egregium ([36] 4, p 98, Corollary 23),

$$|\kappa_1 \kappa_2 \dots \kappa_n| = |\bar{\kappa}_1 \bar{\kappa}_2 \dots \bar{\kappa}_n| = \sqrt{\kappa_1^2 + c} \sqrt{\kappa_2^2 + c} \dots \sqrt{\kappa_n^2 + c}.$$

Since $c \geq 0$, it is immediately clear that $c = 0$. Hence

$$d\bar{N}_p = \begin{pmatrix} \pm\kappa_1 & & & 0 \\ & \pm\kappa_2 & & \\ & & 0 & \ddots \\ 0 & & & 0 \end{pmatrix}.$$

On the other hand, the set $\{\kappa_i \kappa_j \mid i < j\}$ is identical, including multiplicities, to the set $\{\bar{\kappa}_i \bar{\kappa}_j \mid i < j\}$.

([36] 4, p 97, Proposition 22). From this it follows that $\bar{\kappa}_1 \bar{\kappa}_2 = \kappa_1 \kappa_2$ and hence either

$$d\bar{N}_p = \begin{pmatrix} \kappa_1 & & & 0 \\ & \kappa_2 & & \\ & & 0 & \ddots \\ 0 & & & 0 \end{pmatrix} \text{ or } d\bar{N}_p = \begin{pmatrix} -\kappa_1 & & & 0 \\ & -\kappa_2 & & \\ & & 0 & \ddots \\ 0 & & & 0 \end{pmatrix}$$

The proof of Theorem 5.6 is now complete. $\square$

Corollary 5.7: If, in addition to the hypotheses of Theorem 5.6, S is connected and Π_p does not vanish identically for any $p \in S$, then f is an isometry.

Proof: As Π is continuous in p , $\Pi_p \neq 0$ for any p , and S is connected, we must have either $\Pi_p \equiv \overline{\Pi_p}$ for all $p \in S$ or $\Pi_p \equiv -\overline{\Pi_p}$ for all $p \in S$. Remembering that f is equi-long, the corollary now follows from [36], 4, p 23, Theorem 21. $\square$

Lemma 5.8: Let V be a connected subset of a metric space M . Let $p \in M$, $q_1, q_2 \in V$ and assume $d(p, q_1) < d(p, q_2)$. Then for any d such that $d(p, q_1) \leq d \leq d(p, q_2)$, there is a $v \in V$ such that

$$d(p, v) = d.$$

Proof: Since $d(p, x)$ is a real valued continuous function defined on a connected set, the lemma is a consequence of the Intermediate Value Theorem. $\square$

Remark: Let $B_1 \subseteq S_1 \cap S_2$ be a metric basis of E^{n+1} . If $f|_{S_1}$ and $f|_{S_2}$ are isometries, Lemma 4.1 shows that there are unique extensions $\bar{f}$ and $\overline{\bar{f}}$ of $f|_{S_1}$ and $f|_{S_2}$ to motions of E^{n+1} . Because $f|_{B_1}$ has a unique extension to a motion of E^{n+1} , and $\bar{f}$ and $\overline{\bar{f}}$ both extend $f|_{B_1}$ to a motion of E^{n+1} , it follows that $\bar{f} \equiv \overline{\bar{f}}$ and $f|_{S_1 \cup S_2}$ is an isometry. In particular, if $B \subseteq U \subseteq S$, B is a metric basis of E^{n+1} and $f|_{U \cup \{p\}}$ is an isometry for each $p \in S$, then $f|_S$ is an isometry.

Lemma 5.9. Let S be a connected set in E^{n+1} such that for any two points $x_1, x_2 \in S$ it is true that $S \setminus \{x_1, x_2\}$

is also connected. Let $U \subseteq S$ be a connected orientable differentiable hypersurface in E^{n+1} , with $\Pi_q \neq 0$ for all $q \in U$. Let $f: S \rightarrow E^{n+1}$ be a metric transformation with spread 1, such that $f|_U$ is a diffeomorphism. Then f is an isometry.

Proof: By Corollary 5.7, $f|_U$ is an isometry. Let $y_1, y_2 \in U$. By Lemma 5.8, $\rho(d) = d$ for all $d < d(y_1, y_2)$. It follows from Lemma 2.3 that f is continuous. Let

$$d_0 = \sup\{d' \mid \rho(d) = d \text{ for all } d < d'\}.$$

Note that $d_0 \geq d(y_1, y_2) > 0$.

Our first objective is to show that $d(p, q) = d(f(p), f(q))$ for all $p \in S, q \in U$. Let q_0 be a fixed but arbitrary point of U . On the line normal to U at q_0 there are only two points $x_1, x_2 \in E^{n+1}$ such that $d(x_1, q_0) = d(x_2, q_0) = d_0$. Let $S' = S \setminus \{x_1, x_2\}$. By hypothesis, S' is connected. Let $p \in S'$. Assume $d(p, q_0) \geq d_0$. Then by Lemma 5.8 there is a $p_0 \in S'$ with $d(p_0, q_0) = d_0$. The line through p_0 and q_0 is not normal to the hypersurface U at q_0 . Thus there are points q_1 and q_2 of U with $d(p_0, q_1) < d_0 = d(p_0, q_0) < d(p_0, q_2)$.

Consider $B = \{x \mid d(p_0, x) < d_0\} \cap U$. Since $\Pi_q \neq 0$ for $q \in B$, B does not lie in an n -flat of E^{n+1} , hence is a metric basis of E^{n+1} (see Chapter 4, §1). By the definitions of B and p_0 , $f|_{B \cup \{p_0\}}$ is an isometry. As $f|_U$ is also an isometry and $B \subseteq U$, the above remark shows that $f|_{U \cup B \cup \{p_0\}} = f|_{U \cup \{p_0\}}$ is an isometry.

Let d be such that $d(p_0, q_1) < d < d(p_0, q_2)$. By Lemma 5.8, there is a $q \in U$ with $d(p_0, q) = d$. As $f|_{U \cup \{p_0\}}$ is an isometry then

$$\rho(d) = \rho(d(p_0, q)) = d(f(p_0), f(q)) = d(p_0, q) = d.$$

It now follows that $\rho(d) = d$ for all $d < d(p_0, q_2)$, showing that $d_0 \geq d(p_0, q_2)$. This contradicts the choice of q_2 , hence the assumption that $d(p, q_0) \geq d_0$ is false. However if $d(p, q_0) < d_0$ it follows from the definition of d_0 that $d(p, q_0) = d(f(p), f(q_0))$ for all $p \in S'$. If either x_1 or x_2 is in S - say x_j , the continuity of f shows that $d(x_j, q_0) = d(f(x_j), f(q_0))$.

As q_0 is an arbitrary point of U , it follows that $\rho(d(p, q)) = d(f(p), f(q))$ for all $p \in S, q \in U$, and hence $f|_{U \cup \{p\}}$ is an isometry for each $p \in S$. The above remark now shows that f is an isometry. $\square$

Theorem 5.10: Let S be a connected differentiable hypersurface in E^{n+1} , $f: S \xrightarrow{\text{onto}} \bar{S} \subseteq E^{n+1}$ a diffeomorphic metric transformation with spread 1. Then f is an isometry.

Proof: Any differentiable hypersurface is locally orientable, hence locally one may consider Π_p .

If $\Pi_p \equiv 0$ for all $p \in S$, then by Theorem 5.6, $\bar{\Pi}_p \equiv \Pi_p \equiv 0$, for all $p \in S$. Then S and $\bar{S}$ both lie in n -flats of E^{n+1} . Hence Theorem 5.10 follows from either Theorem 4.7 or Corollary 3.19(a).

Otherwise there is a $q_0 \in S$ such that (in a local orientation) $\Pi_{q_0} \neq 0$. As Π is continuous in q_0 , and a differentiable hypersurface is both locally arcwise connected and locally orientable, there is an open arcwise connected subset U of S , containing q_0 , for which we can assign an orientation $N(p)$ such that $\Pi_q \neq 0$ for all $q \in U$. Noticing that any connected manifold S of dimension at least 2 is such that $S \setminus \{x_1, x_2\}$ is connected, for any x_1 and x_2 , we now apply Lemma 5.9 to complete the proof of Theorem 5.10. $\square$

Corollary 5.11: Let $f: S \rightarrow \bar{S} \subseteq E^{n+1}$ be a diffeomorphic metric transformation, S a differentiable hypersurface in E^{n+1} . Then f is a similarity.

We think that Theorem 5.10 has many possibilities for generalizations, the most promising may be to submanifolds, rather than hypersurfaces, immersed in E^{n+1} .

Two of the principal theorems from differential geometry that we use in the proofs of Theorem 5.6 and Corollary 5.7 ([36] 4, p 97, Proposition 22 and [36] 4, p 93, Theorem 21) remain valid for both hyperbolic and spherical space, and one might suspect that Theorem 5.10 could be generalized to hypersurfaces immersed in manifolds of constant curvature. However Example 7, Chapter 1, shows that this is not the case.

CHAPTER 6 EXISTENCE QUESTIONS

This chapter is an attempt to summarize and extend some of the work that has been done on the "Existence Question".

The "Existence Question" was discussed in Chapter 1. Essentially, it is this: Given a distance space (N,d) and a class $\mathcal{C}$ of distance spaces, is (N,d) order embeddable onto some member of $\mathcal{C}$? Frequently, $\mathcal{C}$ is taken to be all subsets of some particular distance space - say E^n . All of our results are of this type.

The chapter is divided into two sections. Section 1 deals exclusively with finite sets, while Section 2 deals primarily with infinite sets. In both sections, most of the results are of a "negative" type. That is, they show that a certain space is not order embeddable into a certain class $\mathcal{C}$.

§1. Of most interest to the M.D.S. Theorists are finite sets of points, with an ordering of $S \times S$. That is, they begin with a set S , a totally ordered set C and a mapping $e : S \times S \rightarrow C$. Then (S,C,e) is called a C -metrized space. Because all that is of interest is the ordering

induced on $S \times S$ by C and e , and as $e(S \times S)$ is a finite subset of C , C may always be taken to be a subset of the positive reals. In this case, (S, C, e) is a distance space and is denoted by (S, e) .

One type of Multidimensional Scaling order embeds (S, e) into a metric space (M, d) . One way of doing this is to find a metric d on S such that the function f taking (S, e) into (S, d) , given by $f(x) = x$, is an order transformation. This is easy to do. If d is defined by
$$d(x, y) = \begin{cases} e(x, y) + \kappa & x \neq y \\ 0 & x = y \end{cases} \quad \text{and } \kappa \text{ is chosen}$$
 suitably large, then (S, d) will be a metric space. For example, if $\kappa \geq \max_{x, y \in S} \{e(x, y)\}$, then for any distinct $x, y, z \in S$,

$$d(x, y) + d(y, z) = e(x, y) + \kappa + e(y, z) + \kappa \geq 2\kappa \geq e(x, z) + \kappa = d(x, z)$$

This shows d is a metric, but the metric space (M, d) is not very useful without knowing something more about κ . There has been some work done on the so called "additive constant" problem, which attempts to find "appropriate" values for κ (see [27], [38]). Deciding on the "appropriate" value of κ requires more detailed information. One possibility often considered is the value that minimizes κ .

In general the M.D.S. user would like to embed (S, e) into a metric space he knows - such as a Euclidean space or a normed linear space. Naturally, he would prefer a Euclidean space. We now look at a few results in this area.

If S contains n points, (S, e) can be order embedded into E^{n-1} as follows. Consider an equilateral simplex in E^{n-1} . Any edge can be either shortened or lengthened slightly by a rotation about the opposite $(n-2)$ -face, without changing the lengths of any of the other edges. Clearly, the lengths of these sides could be arranged to correspond to any ordering of $S \times S$. If the length of the longest side is attained for only one pair of points, it can be lengthened until the simplex lies in an $(n-2)$ -flat of E^{n-1} . These remarks apply also to H^{n-1} and S^{n-1} .

In [19] Holman characterizes all the distance spaces (S, e) of n points which can be order embedded into E^{n-2} . He shows that if for some x, y, z in S the inequality

$$(1) \quad e(x, z) \leq \max[e(x, y), e(y, z)]$$

is not satisfied then (S, e) is order embeddable into E^{n-2} . The inequality $e(x, z) \leq \max[e(x, y), e(y, z)]$ is often referred to as the ultra-metric inequality. It is easily seen that the ultra-metric inequality implies the triangle inequality, and that it is satisfied if and only if all the triangles in the space are "long legged isosceles"; that is if and only if all triangles are isosceles and the length of the two equal sides is greater than, or equal to, the length of the third side. Thus the only metric n -tuples not order embeddable are the ultrametric n -tuples.

The presence of equal distances is clearly a complicating factor in the geometric embedding problem, as well as in the various scaling interpretations. Thus we consider only scalene configurations for the time being. A distance space (S, e) is said to be scalene if and only if $e(x, y) \neq e(z, w)$ for any $x, y, z, w \in S$ with $\{x, y\} \neq \{z, w\}$, $x \neq y$. Holman's Theorem implies that any scalene distance space (S, e) of n points can be order embedded into E^{n-2} . The problem of order embedding an arbitrary scalene configuration of n -points into lower dimensional spaces now arises.

In general, we have not been able to make much headway against this question. We now present some results, first considered by Kelly and Erdos (unpublished), which answer the question for $n = 4, 5$, and 6 , and give a bound for the minimum dimension needed for larger n .

The case $n = 4$ is easy, so it will not be considered here. For $n = 5$ and 6 an example is presented which cannot be order embedded into E^2 or E^3 respectively. Thus an arbitrary 5 or 6 point scalene configuration need not be order embeddable into E^2 or E^3 respectively. Unfortunately, obvious generalizations of these examples to $n = 7$ are order embeddable into E^4 .

The result for $n = 5$ (Theorem 6.1) shows the stronger fact that there exists a scalene distance space of 5 points not order embeddable into any two dimensional normed linear space.

Theorem 6.1: Let (S, e) be any distance space,
 $S = \{p_1, p_2, p_3, q_1, q_2\}$, such that

$$(2) \quad e(p_i, p_j) > e(q_1, q_2) > e(p_m, q_n), \quad i \neq j.$$

Then (S, e) cannot be order embedded into any two dimensional
 normed linear space.

Proof: Let $f: S \rightarrow E^2$ be an order embedding of S .
 For simplicity's sake, call $f(x)$ simply x .

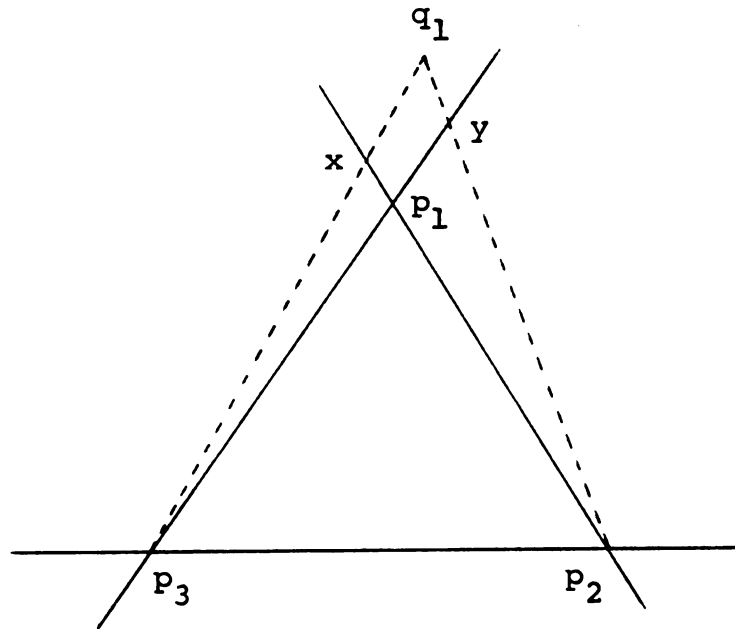


Figure 6.1

Consider the above diagram. We begin by assuming that
 one of the points q_1, q_2 is in one of the closed regions

labeled 1, 2, or 3. Say q_1 lies in 1. We have

$$q_1 p_3 + q_1 y \geq p_3 p_1 + p_1 y$$

$$p_2 y + p_1 y \geq p_1 p_2$$

Adding these, and simplifying, we obtain

$$q_1 p_3 + q_1 y + p_2 y \geq p_3 p_1 + p_1 p_2$$

or

$$q_1 p_3 + q_1 p_2 \geq p_3 p_1 + p_1 p_2 .$$

However, this contradicts (2), since f is to be an order embedding.

Assume neither q_1 nor q_2 lies in any of 1, 2, or 3.

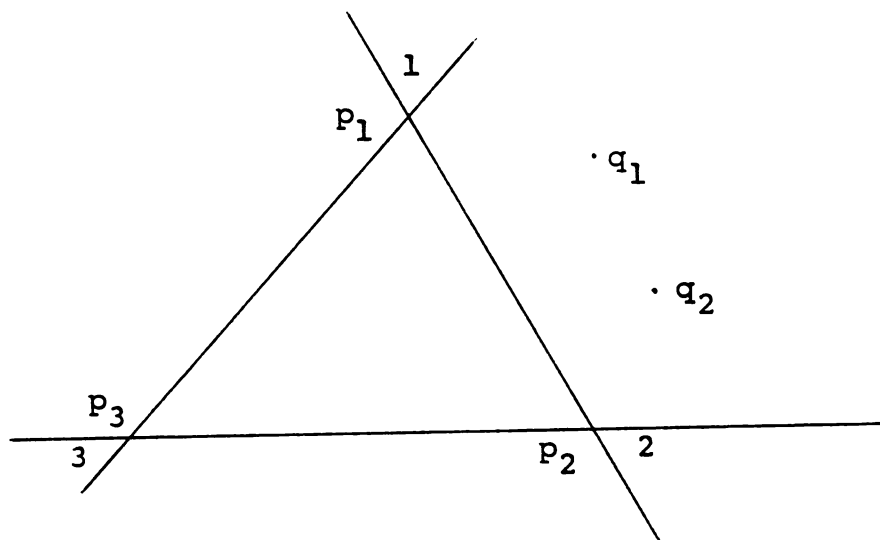


Figure 6.2

Then there is a pair of points p_i, p_j , such that the line $q_1 q_2$ does not intersect $[p_i, p_j]$ and the line $p_i p_j$ does not intersect $[q_1, q_2]$. The points p_i, p_j, q_2, q_1 form a convex quadrilateral with $[q_2, q_1]$ and $[p_i, p_j]$ being opposite sides. Say the quadrilateral is $p_1 p_3 q_2 q_1$. By Lemma 3.8,

$$d(p_1, p_3) + d(q_2, q_1) \leq d(p_1, q_2) + d(p_3, q_1) .$$

Hence, one of the terms on the left is less than or equal to one of the terms on the right, contradicting (2). A similar contradiction is obtained in all cases. This completes the proof. $\square$

We have been able to generalize this to 6 points only for order embeddings into E^3 .

Theorem 6.2: Let (S, e) , $S = \{p_1, p_2, p_3, q_1, q_2, q_3\}$, be any distance space such that

$$(3) \quad e(p_i, p_j) > e(q_k, q_l) > e(p_m, q_n), \quad i \neq j, \quad k \neq l .$$

Then (S, e) cannot be order embedded into E^3 .

Proof: As in Theorem 6.1, let $f: S \rightarrow E^3$ be any order embedding of S , and call $f(x)$ simply x .

Let π be any plane containing p_1, p_2 , and p_3 .

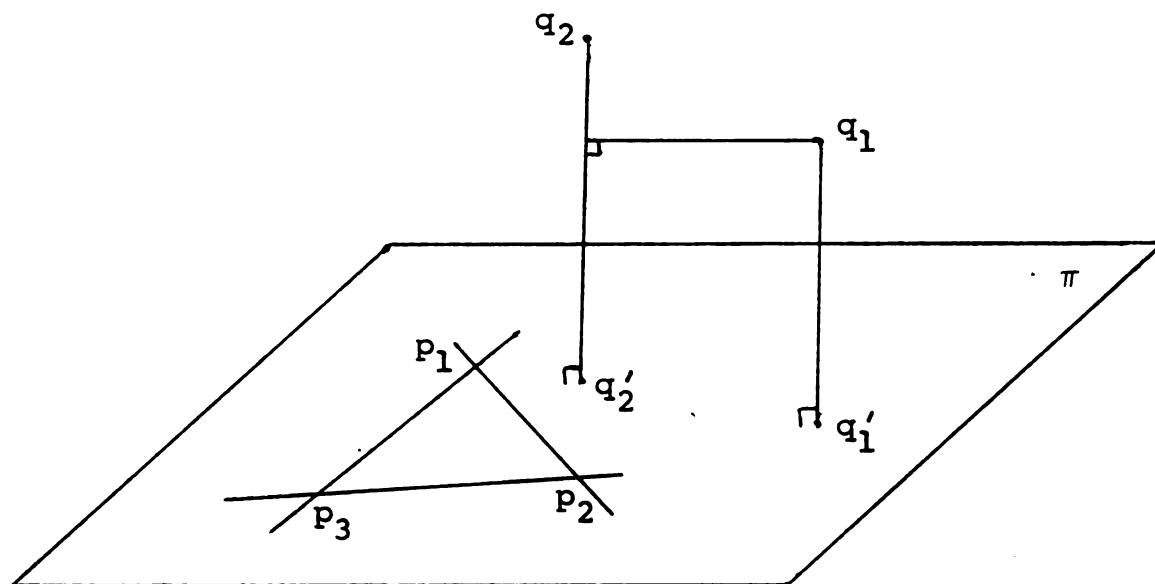


Figure 6.3

At least two of q_1, q_2, q_3 lie on the same side of π , say q_1 and q_2 . Let q'_1, q'_2 be the perpendicular projections of q_1 and q_2 onto π . The situation in π is very similar to that in the previous theorem, using q'_1, q'_2 instead of q_1 and q_2 . The only difference is that $d(q'_1, q'_2) \leq d(q_1, q_2)$, so $d(q'_1, q'_2)$ may be less than $d(p_i, q'_j)$ for some i, j .

However, as in Theorem 6.1 it can be shown that

$$d(p_m, p_n) + d(q'_2, q'_1) \leq d(p_m, q'_2) + d(p_n, q'_1)$$

for some $m, n = 1, 2, 3$, $m \neq n$. By (3), $d(p_m, p_n)$ is the largest of these distances (for $d(p_m, q'_2) \leq d(p_m, q_2)$ and

$d(p_n, q'_1) \leq d(p_n, q_1)$, so $d(q'_2, q'_1)$ must be the smallest. It follows that $d(p_m, q'_2) \geq d(q'_2, q'_1)$ and $d(p_n, q'_1) \geq d(q'_2, q'_1)$. Assume $d(q_2, q'_2) \geq d(q_1, q'_1)$, (otherwise change the roles of q_1 and q_2 , as well as p_n and p_m in the following).

Then

$$\begin{aligned} d(p_m, q_2)^2 &= d(p_m, q'_2)^2 + d(q_2, q'_2)^2 \\ &\geq d(q'_2, q'_1)^2 + (d(q_2, q'_2) - d(q_1, q'_1))^2 \\ &= d(q_2, q_1)^2. \end{aligned}$$

This contradicts (3) so the proof is complete. $\square$

As has been said, there seems to be no obvious generalization of Theorems 6.1 and 6.2, even to $n = 4$.

The next result shows that if every distance space of n points is order embeddable into some N.L.S. of dimension m , then it is necessary that $\frac{\log(n-2)}{\log 5} \leq m$. This shows that, as n tends to ∞ , so does the minimum dimension such that every distance space of n points is order embeddable into a N.L.S. of that dimension.

Theorem 6.3: In order for every scalene distance space of n points to be order embeddable into some N.L.S. M of dimension m , it is necessary that $\frac{\log(n-2)}{\log 5} \leq m$ and sufficient that $m \leq n-2$.

Proof: That it is sufficient that $m \leq n-2$ follows from Holman's Theorem.

Consider any distance space (S, e) consisting of the n points $\{p_1, p_2, q_1, \dots, q_{n-2}\}$ such that

$$(4) \quad e(p_1, p_2) > e(q_i, q_j) > e(q_k, p_\ell), \quad i \neq j.$$

As in Theorems 6.1 and 6.2, let $f: S \rightarrow M$ be an order isomorphism, and denote $f(x)$ by x .

Let $r = d(p_1, p_2)$. Because the triangle inequality is satisfied,

$$d(q_i, p_1) + d(q_i, p_2) \geq d(p_1, p_2) = r \quad \text{for all } i.$$

From this, and (4) it follows that

$$d(q_i, q_j) > \frac{r}{2} \quad \text{for all } i \text{ and } j.$$

Again because of (4), $q_i \in B(p_1, r)$ for all i . Thus

$B(q_i, r/4) \subseteq B(p_1, 5r/4)$ for all i . Also, because

$d(q_i, q_j) > r/2$ it follows that $B(q_i, r/4) \cap B(q_j, r/4) = \emptyset$ for all i, j .

Let V_t be the volume of a ball of radius t . Then $(n-2)V_{r/4} \leq V_{5r/4}$. Now $V_t = t^m V_1$ (see [5], page 158), so $(n-2)(r/4)^m \leq (5r/4)^m$. Thus, it follows that $m \geq \frac{\log(n-2)}{\log 5}$. $\square$

Clearly, the above theorem could be improved. However we have not been able to obtain any lower bound on m which is not a logarithmic function of n . At least for Euclidean space, it should be possible to do better than this.

§2. In Section 1 we presented a few results concerning the existence question for finite sets. We now turn to infinite sets.

In [24] Lew gives "Some Counterexamples in M.D.S.". He shows the following:

(1) The spaces ℓ_1^m and ℓ_∞^m have no order embedding into E^n for $m \geq 2$. However for each m , ℓ_1^m can be order embedded into $\mathcal{H}$, Hilbert space.

(2) The space C_0 of all real sequences with limit zero, and norm defined by $\|(x_n)\| = \sup_n |x_n|$ has no non-trivial metric transformation into Hilbert space.

Lew calls these counterexamples because they show that, in applying M.D.S., it is necessary to consider spaces other than E^n and $\mathcal{H}$. The proofs of (1) and (2), as presented by Lew, are quite difficult, relying heavily on the work of Schoenberg, von-Neumann, and Einhorn ([12], [29], [32]) on positive definite forms.

We now present an elementary proof and generalization of the first part of (1).

Theorem 6.4: A N.L.S. which has a segment on its unit sphere has no metric transform in E^n , S^n , or H^n .

Proof: Let M be a normed linear space with a segment on its unit sphere. Consider a plane π that contains the origin, and a segment of the unit sphere with endpoints

p_1 and q_1 . It is sufficient to show that this plane, with the induced norm, has no metric transform into E^n, S^n , or H^n . Let q_0 be the

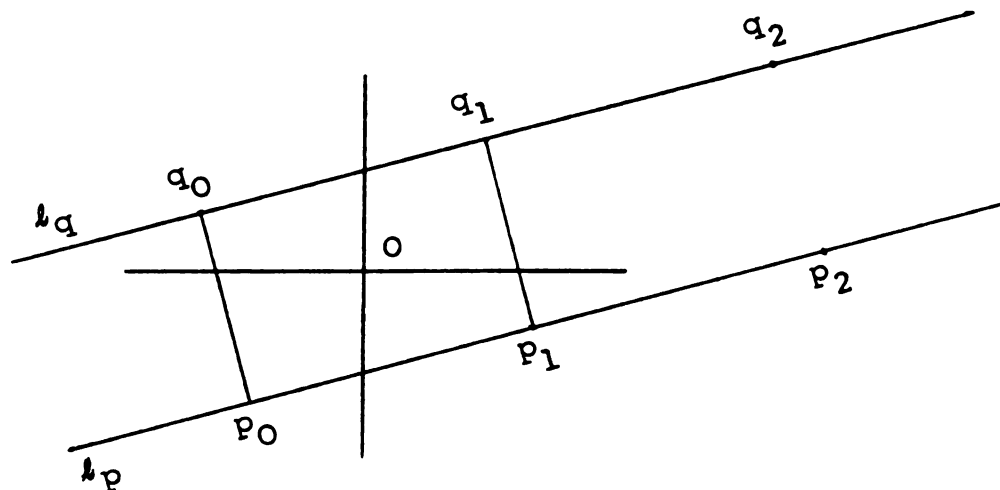


Figure 6.4

reflection of p_1 through the origin, and p_0 be the reflection of q_1 through the origin. Then the segment joining p_0 to q_0 also lies on the unit sphere. Let l_q and l_p be the lines through q_0 and q_1 , and p_0 and p_1 respectively. For $i > 1$, let q_i be the point on l_q which is a distance 2 from q_{i-1} , but is not equal to q_{i-2} . Similarly define $\{p_i\}_{i=0}^{\infty}$ on l_p . It is easily seen that

$$(3) \quad \|q_i - q_j\| = \|q_i - p_j\| = \|p_i - p_j\| = \|p_i - q_j\| \quad \text{for all } i \neq j.$$

Let M_1 be any of E^n, S^n , or H^n and let $f: \pi \rightarrow M_1$ be a metric transformation. For $a, b \in M_1$ let $B(a, b) = \{x \mid d(x, a) = d(x, b), x \in M_1\}$. Now $B(a, b)$ is an $(n-1)$ -flat of M_1 . (See Busemann [5] p. 309). Thus, if $a', b' \in B(a, b)$, then

$$\{x \in B(a, b) \mid d(x, a') = d(x, b')\} = B(a', b') \cap B(a, b)$$

is an $(n-2)$ -flat of M_1 . Continuing, if $a_i, b_i \in \bigcap_{k=1}^{i-1} B(a_k, b_k)$, then $\bigcap_{i=1}^m B(a_i, b_i)$ is an $(n-m)$ -flat of M_1 .

Now because of (3) and the fact that f is a metric transformation,

$$d(f(q_i), f(q_j)) = d(f(q_i), f(p_j)) = d(f(p_i), f(p_j)) = d(f(p_i), f(q_j)).$$

Thus $f(q_i)$ and $f(p_i)$ are in $B(f(q_j), f(p_j))$ for $i \neq j$.

Therefore $f(q_i), f(p_i) \in \bigcap_{j=1}^m B(f(q_j), f(p_j))$ for $i > m$.

Hence $\bigcap_{j=1}^m B(f(q_j), f(p_j))$ contains an infinite number of points for all m . However, $\bigcap_{j=1}^n B(f(q_j), f(p_j))$ is zero-dimensional, hence contains 1 or 2 points (1 if M_1 is E^n or H^n , 2 if M_1 is S^n). This is a contradiction, and the theorem is proved. $\square$

Corollary 6.5: If a N.L.S. has a segment on its unit sphere, then no open subset of that space has a metric transform in E^n, S^n or H^n .

The proof of this is essentially the same as the above.

The space ℓ_{∞}^n seems to lend itself nicely to problems involving order transformations. Theorem 6.10 show that many N.L.S., in particular E^n , cannot be order embedded into ℓ_{∞}^n . The proof of these theorems is based on the characterization of "flat spots" of the unit sphere of a N.L.S., given in Lemma 6.8.

Before proceeding to Lemma 6.8 several definitions and preliminary lemmas are needed. In these definitions, and throughout the rest of this chapter, convex always means algebraically convex.

Definitions: Let K be a closed algebraically convex set in a N.L.S., M . A hyperplane H in M is said to support K if K lies in one of the closed half spaces determined by H . A subset S of K is said to be a supporting set of K if $S = K \cap H$ for some supporting hyperplane H of K . Clearly a supporting set of K is convex. An extreme point of K is a supporting set of one point. Supporting sets of more than one point are called faces of K . A face of K which is properly contained in no other face of K is called a facet of K .

Definition: An affine subspace of M is a translate of a linear subspace of M . For any set A in a N.L.S. M , aff A is defined by

$$\text{aff } A = \{x \mid x = \sum_{i=1}^n \lambda_i a_i, \sum_{i=1}^n \lambda_i = 1, n < \infty, a_i \in A\}.$$

A point $p \in A$ is said to be a relative interior point of A if $p \in V \subseteq A$, and V is open in the relative topology of $\text{aff } A$. If K is algebraically convex it is not hard to show that the relative interior of K is non-empty (see [14], page 9).

For the following work (Lemmas 6.6, 6.7, 6.8, 6.9) let B be the closed unit ball, and U the unit sphere of a N.L.S. M .

The proofs of Lemmas 6.6 and 6.7 are omitted.

Lemma 6.6: If K is a convex subset of U , then there is a supporting hyperplane H of B with

$$K \subseteq (\text{aff } K) \cap B \subseteq H \cap B \subseteq U$$

Proof: See Day [8] p. 43.

In other words, K lies in a face of B .

Lemma 6.7: If p is a relative interior point of a facet F of B , then p lies in no other facet of B .

Notation:: For $p \in U$, define $E(p)$ to be $\{x \mid x \in U, \|p - x\| = 2\}$. Let $-p$ be the point of U diametrically opposite p . Certainly, $-p \in E(p)$, hence $E(p)$ is not empty.

As mentioned earlier, Lemma 6.8, which characterizes a facet of B , is the crucial result to obtain Theorem 6.10.

Lemma 6.8: If F is a facet of B , then $F = E(p)$ where $-p$ is any relative interior point of F .

Proof: Assume $x \in F$. Then, since F is convex, $[-p, x] \subseteq F$.

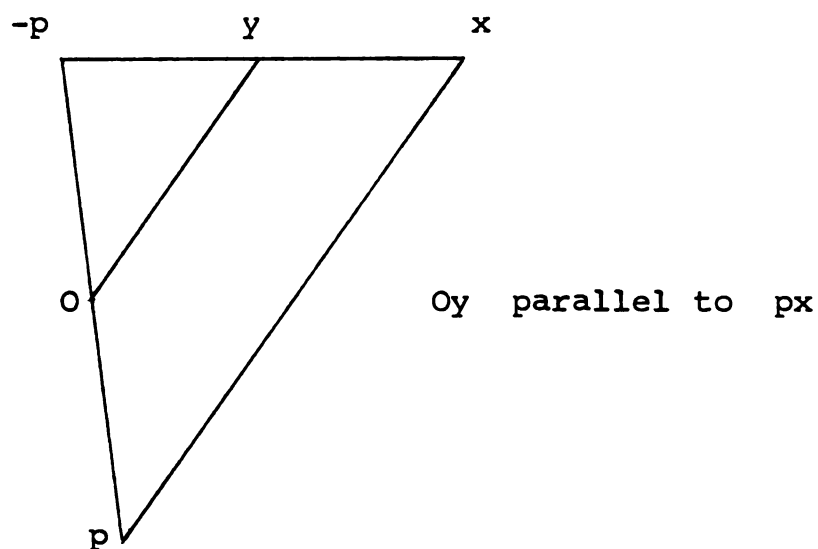


Figure 6.5

In Figure 6.5, let Oy be parallel to px , with $y \in [-p, x]$. Then $\|y\| = 1$, so it follows that $\|p - x\| = 2$, and hence $x \in E(p)$. Thus, $F \subseteq E(p)$.

Assume $x \in E(p)$, $x \notin F$. We will show that F could not be a facet of B , giving a contradiction.

Consider Figure 6.5, where $x \in E(p)$, and Oy is parallel to px . Then $\|y\| = 1$, so it follows from the convexity of B that $[-p, x] \subseteq U$. (Recall that $x \in U$, so $\|x\| = 1$).

Let $y = \lambda x + (1 - \lambda)q$, $q \in F$, $0 \leq \lambda \leq 1$

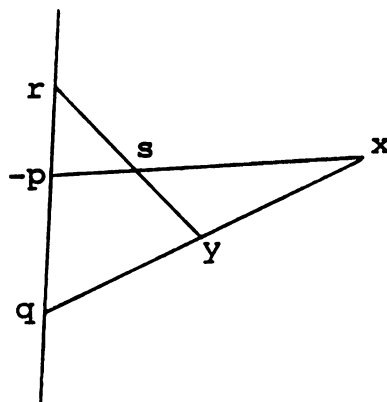


Figure 6.6

Consider Figure 6.6, which is the situation in the plane containing $x, -p$ and q . As $-p$ is in the relative interior of F , and $q \in F$, then $r \in F$ can be chosen with $-p, q$ and r collinear, and q and r on opposite sides of $-p$. Let $[r, y] \cap [-p, x] = s$. Then r, s, q , and x all are in U , so it follows that $y \in U$. Thus the convex set

$$\begin{aligned} & \text{(def)} \\ K &= \{ \lambda x + (1 - \lambda)q \mid q \in F, 0 \leq \lambda \leq 1 \} \subseteq U. \end{aligned}$$

Lemma 6.6 shows that K lies in some face of B . However $F \not\subseteq K \subseteq B$, contradicting the fact that F is a facet of U . Thus, $E(p) \subseteq F$ and the proof is complete. $\square$

Let f be an order transformation with scale function ρ , from U (the unit sphere of a N.L.S.) into ℓ_∞^n . Let K be the smallest closed ball of ℓ_∞^n containing $f(U)$,

and let S be the corresponding sphere (that is, S is the boundary of K). Then, clearly, two opposite facets of K touch $f(U)$. So K and $f(U)$ have the same diameter, which is $\rho(2)$. Note, we use the fact that f is an order transformation to determine that $\rho(2)$ is the maximum transformed distance.

For $p \in U$, both $f(p)$ and $f(-p)$ lie in K and

$$\|f(p) - f(-p)\| = \rho(\|p - (-p)\|) = \rho(2).$$

Hence, $f(p)$ and $f(-p)$ lie in opposite facets of K .

Thus, $f(U) \subseteq S$.

Lemma 6.9: Let f be an order transformation from U into $\mathbb{R}^n$. If $p_1, \dots, p_m$ are arbitrary points of U and $m \geq 2n+1$, then for some j , $\bigcup_{i=1}^m E(p_i) = \bigcup_{\substack{i=1 \\ i \neq j}}^m E(p_i)$.

Proof: Let F_i , $i = 1, \dots, r$ be the facets of S containing points of $f(\bigcup_{i=1}^m E(p_i))$. Let $-F_i$ be the facet of S opposite F_i . Note that $\{f(p_i); i = 1, \dots, m\} \subseteq \bigcup (-F_i)$ and for each k there is some members of $\{f(p_i)\}$ in $-F_k$. As $m \geq 2n+1 > r$, for some j it must be that there is some member of $\{f(p_i), i = 1, \dots, m, i \neq j\}$ in $-F_k$ for each k . Let $f(x) \in F_i$. Let $f(p_k) \in \{f(p_i), i = 1, \dots, m, i \neq j\}$ be in $-F_i$. Then $\|f(x) - f(p_k)\| = \rho(2)$, hence $\|x - p_k\| = 2$, hence $x \in E(p_k)$. The lemma is now complete. $\square$.

Theorem 6.10: Let B be the unit ball and U be the unit sphere of a N.L.S. M . If B has more than $2n$ facets, then U is not order embeddable into ℓ_∞^n .

Proof: Assume B has more than $2n$ facets. Let $-p_i$, $i = 1, \dots, 2n+1$ be relative interior points of distinct facets F_i . Then $\bigcup_{i=1}^{2n+1} E(p_i) \subseteq U$. By Lemma 6.9, $\bigcup_{i=1}^{2n+1} E(p_i) = \bigcup_{\substack{i=1 \\ i \neq j}}^{2n+1} E(p_i)$ for some j . However Lemma 6.7 shows that $-p_j \notin E(p_i) = F_i$ for any $i \neq j$, giving a contradiction. $\square$.

It should be noted that a facet F may be an extreme point p . In this case $\text{aff } F = \text{aff}\{p\} = \{p\}$, so p is a relative interior point, and Theorem 6.10 still holds. If B is strictly convex then every point of U is a facet, hence U cannot be order embedded into ℓ_∞^n .

It is annoying that Theorem 6.10 depends on the order embedding being into ℓ_∞^n . These proofs do not hold up in other spaces, even a space such as ℓ_1^n which often behaves in a very similar fashion. However, Professor L.M. Kelly claims that E^n is not order embeddable into ℓ_1^m , for any n, m .

CONCLUSION

This thesis represents the first effort that we know of to bring together and to expand on the geometric embedding problems underlying the data analysis technique known as multidimensional scaling. Organization of the material has been difficult and the results may seem discursive and confused. It may be well, then, to summarize and reflect on what has and has not been achieved.

The principal concern has been with the study of two seemingly very general classes of transformations from one distance space to another, namely metric transformations which merely preserve equality of distances and order transformations which preserve the order of the distances. It is the latter which are of prime interest to the MDS theorists. In fact it is their fervent hope that in many "highly structured" spaces the only order transformations are "essentially" similarities.

The Beals-Krantz result [2], which we reprove in Chapter 2, supports this idea in the class of convex metric spaces.

As pleasing as this result is it fails to answer such simple and relevant questions as: "does the order of distances determine a sphere in euclidean 3-space"? or

"what do order transforms of the real line look like in E^n "?
 The first of these questions is answered in Chapter 5 and
 the second in Chapter 2.

We will now attempt to highlight what we consider the
 major results and achievements of the thesis making clear
 in the process which are more or less original.

Among the more important essentially original contributions
 are the following:

1. The arc length, arc curvature and various
 continuity theorems in Chapter 2.
2. Theorem 3.16 generalizing the Mankiewicz
 theorem from isometric transformations to
 metric transformations.
3. Theorem 5.10 to the effect that two order
 isomorphic hypersurfaces in E^n are similar.
4. Proof (Theorem 4.15) that any Minkowski space
 whose unit sphere is order isomorphic to the
 sphere in E^n must be congruent to E^n .
5. There are k -point scalene metric spaces which
 are not order embeddable into M^n (where k
 is a function of n , Theorem 6.3).
6. Examples of non-similar order isomorphic
 hypersurfaces in hyperbolic n -space, showing that
 Theorem 5.10 and Theorem 4.5 of Schoenberg cannot
 be directly extended even to spaces of constant
 Gauss curvature.

7. The impossibility of order embedding a non-strictly convex NLS into E^n , H^n , or S^n (Theorem 6.4).
8. The impossibility of order embedding a strictly convex NLS into ℓ_∞^n . (Theorem 6.10).

Much of our effort has gone into summarizing, adapting, refining and extending the work of others, notably:

1. The Schoenberg-Von Neumann characterization of screw curves in E^n , the proof of which we have made more accessible in that setting. We have also extended this characterization to screw segments and shown that the situation in hyperbolic n -space is much more complicated.
2. The Schoenberg proof that a metric transformation from E^m into E^n is necessarily a similarity and our extension of this result to metric transformations defined on sets in E^m with a non-null interior.
3. Our simplification and extension of the work of Vogt [37] in Chapter 3, of Lew [24] in Chapter 6, and Holman [19] in Chapter 6.

A number of obvious questions remain unanswered.

Some of the more appealing are the following:

1. If the spheres in two normed linear spaces are order isomorphic must the spaces be congruent?

In this connection it seems to be even unknown whether two such spaces are congruent if their unit spheres are congruent.

2. Is there an order isomorph of E^2 in any normed linear space other than E^n ?

3. Are there two arcs in E^n which are order isomorphic which are not of constant curvature?

Such arcs exist in H^n .

4. The euclidean sphere in n-space is characterized in the class of n-dimensional Minkowski spaces by the order of its distances. Is this true of other hypersurfaces in E^n ? e.g. is it true of the boundaries of convex bodies?

5. Is there a finite set in a euclidean space characterized in the class of subsets of all euclidean spaces by the order of its distances?

6. What is the smallest scalene space not order embeddable in E^4 ?

7. The perception theorist seems to operate on the assumption that order analyzing large numbers of finite subsets of his postulated "underlying" space will give him a clue to a characterization of the space itself. To what extent is this true?

In deference to those MDS theorists and users who have persevered thus far we should concede that what we have done here possibly has little direct bearing on their concerns. The principle of metric determinacy is certainly central in scaling theory and practice but in the context of finite spaces it seldom can be claimed that the order of distances determines a set up to a similarity. Scaling theorists tend to feel that if a finite metric space is order embedded in the "proper dimensional" euclidean space, that embedding should be unique up to an "approximate" isometry if it is properly^{*} scaled".

Furthermore, in practice, embeddings with low "stress" are tolerated on the grounds that real data is subject to stochastic uncertainties and noise. So the order of distances of a configuration is claimed to "determine" the configuration subject to much hedging.

We claim that our work could be a valuable first step in trying to make the "practical principle" more precise. In the absence of that precision some of our results can be viewed as lending some support to the principle.

The literature on the applications of MDS is enormous and for the reader unfamiliar with it we can do no better than to refer to the extensive and authoritative writings of R.N. Shephard. His popular paper [35] in Science is particularly recommended both for its attractive survey and its extensive bibliography.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Bachman G., Narici L., Functional Analysis. Academic Press, 1972.
2. Beals, R., Krantz, D., Metrics and Geodesics Induced by Order Relations. Math Zeitschr. 1967, 101, 285-298.
3. Beals, R., Krantz, D., Tversky, A., Foundations of Multidimensional Scaling. Psychological Review 1968, 75, 127-142.
4. Blumenthal, L.M., Theory and Applications of Distance Geometry. London: Oxford University Press, 1953.
5. Busemann, H., The Foundations of Minkowskian Geometry. Comm. Math. Helv., 1950, 24, 156-186.
6. Coxeter, H.S.M., Introduction to Geometry, New York, Wiley.
7. Cunningham, J.P., Shephard, R.N., Monotone Mappings of Similarities into a General Metric Space. Journal of Math. Psych., 1974, 11, 335-363.
8. Day, M.M., Normed Linear Spaced. Springer Verlag, 1973.
9. Drobish, M.W., Abh. Math. Phys. Kl. Konigl. Sachs, Ges. Wiss., 1855, 4.
10. Dugundgi, J., Topology. Allyn and Bacon, Boston 1966.
11. Egervary, E., Alexits, G., Fondements d'une Theorie Générale de la Courbure Linéaire. Commentarii Math.
12. Einhorn, S.J., Functions Positive Definite in $C[0,1]$. Proceedings of the A.M.S., 1969, 22, 702-703.
13. Finsler, P., Uber Kurven und Flächen in Allgemeinen Raumen., Verlag Birkhauser, Basel, 1951.
14. Grunbaum, B., Convex Polytopes, Interscience Publishers, 1967.

15. Haantjes, J., Distance Geometry: Curvature in Abstract Metric Spaces. Proceedings Akademie van Wetenschappen, 1947, 50, 496-508.
16. Helmholtz, H., Handbuch der Physiologischen Optik. Voss, Hamburg, 1896.
17. Henning, H., Der Geruch. (Barth, Leipzig, 1916) Z. Psychology, 1916, 74, 203.
18. Hoffman, K., Kunze, K., Linear Algebra. Prentice Hall Inc., New Jersey, 1971.
19. Holman, E.W., The Relation between Hierarchical and Euclidean Models for Psychological Distances. Psychometrika, 1972, 37, 417-423.
20. Kruskal, J.B., Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. Psychometrika, 1964a, 29, 1-27.
21. Kruskal, J.B., Nonmetric Multidimensional Scaling: A Numerical Method. Psychometrika, 1964b, 29, 115-129.
22. Lawrence, J., k -Equilateral $(2k+1)$ -gons span only even dimensional spaces. The Geometry of Metric and Linear Spaces. Springer-Verlag, 1974.
23. Lew, J.S., Preorder Relations and Pseudoconvex Metrics. American Journal of Math, 1975, 97, 344-363.
24. Lew, J.S., Some Counterexamples in Multidimensional Scaling. Journal of Math. Psych., 1978, 17, 247-254.
25. Lindman, H., and Caelli, T., Constant Curvature Riemannian Scaling. Journal of Math. Psych. 1978, 17, 89-109.
26. Mankiewicz, P., On Extension of Isometries in Normed Linear Spaces. Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronomy Phys. 1972, 20, 367-371.
27. Messick, S., Abelson, R.P., The Additive Constant Problem in Multidimensional Scaling. Psychometrika, 1956, 21, 1-15.
28. Munkres, J.R., Topology: A First Course Prentice Hall Inc., New Jersey, 1975.
29. von-Neumann, J. and Schoenberg, I.J., Fourier Integrals and Metric Geometry. Transactions of the A.M.S., 1941, 50, 226-251.

30. Newton, I., Opticks. Smith and Walford, London, 1704.
31. Pieszko, H., Multidimensional Scaling in Riemannian Space. Journal of Math. Psych. 1975, 12, 449-477.
32. Schoenberg, I.J., Metric Spaces and Positive Definite Functions. Transactions of the A.M.S., 1938, 44, 522-536.
33. Schrodinger, E., Ann. Physik, 1920, 63.
34. Senechalle, D.A., A Characterization of Inner Product Spaces. Proc. Amer. Math. Soc., 1968, 19, 1306-1312.
35. Shephard, R.N., Multidimensional scaling, Tree-Fitting, and Clustering. Science, 1980, 210, 390- ?
36. Spivak, M., Differential Geometry. Publish or Perish Inc., Berkeley, 1970.
37. Torgerson, W.S., Multidimensional Scaling, Theory and Method. Psychometrika, 1952, 17, 401-419.
38. Vogt, A., Maps which Preserve Equality of Distance. Studia Mathematica, 1973, 45, 43-48.
39. Wilson, W.A., On Certain Types of Continuous Transformations of Metric Spaces. American Journal of Math, 1935, 57, 62-68.