

AN EXTENSION OF THE MAZUR-ULAM THEOREM

CONSTANTIN P. NICULESCU

ABSTRACT. One proves that the Mazur-Ulam theorem can be extended in the framework of metric spaces as long as a well behaved concept of midpoint is available. This leads to the new concept of Mazur-Ulam space. Besides the classical case of real normed spaces, other examples such as $Sym^{++}(n, \mathbb{R})$, the space of all $n \times n$ dimensional positive definite matrices, appear as cones attached to suitable Euclidean Jordan algebras. It turns out that the Mazur-Ulam spaces provide a framework for new generalizations of the concept of convex function.

1. INTRODUCTION

The Mazur-Ulam theorem asserts that every bijective isometry $T: E \rightarrow E$ acting on a real normed linear space is an affine map, that is,

$$(1.1) \quad T(\lambda x + (1 - \lambda)y) = \lambda T(x) + (1 - \lambda)T(y)$$

for all $x, y \in E$ and $\lambda \in \mathbb{R}$. See [8], [2], [12] and [13] for details and extensions within the framework of linear spaces. The essence of this result is the property of T to preserve midpoints of line segments, that is,

$$(1.2) \quad T\left(\frac{x + y}{2}\right) = \frac{T(x) + T(y)}{2}$$

for all $x, y \in E$. In fact, the condition (1.2) implies (1.1) for dyadic affine combinations, and thus for all convex combinations (since every isometry is a continuous map). Finally, it is routine to pass from convex combinations to general affine combinations in (1.1).

Surprisingly, the linear structure of E is needed only to support the notion of midpoint. In fact, a property like (1.2), of midpoint preservation, works in the framework of metric spaces as long as a well-behaved concept of midpoint is available. This is made clear by the following definition:

Definition 1. *Definition A Mazur-Ulam space is any metric space $M = (M, d)$ on which there is given a pairing $\sharp: M \times M \rightarrow M$ with the following four properties:*

(the idempotent property) $x \sharp x = x$ for all $x \in M$;

(the commutative property) $x \sharp y = y \sharp x$ for all $x, y \in M$;

(the midpoint property) $d(x, y) = 2d(x, x \sharp y) = 2d(y, x \sharp y)$ for all $x, y \in M$;

(the transformation property) $T(x \sharp y) = T(x) \sharp T(y)$, for all $x, y \in M$ and all bijective isometries $T: M \rightarrow M$.

2000 *Mathematics Subject Classification.* Primary 54E40; Secondary 46B20, 52A41.

Key words and phrases. Metric space, isometry, midpoint.

Published in vol. *Global Analysis and Applied Mathematics*, Intern. Workshop on Global Analysis, Ankara, Turkey, 15-17 April, 2004. American Institute of Physics Conference Proceedings, Vol. **729**, pp. 248-256, New York, 2004 (K. Tas, D. Krupka, O. Krupkova and D. Baleanu, Editors). ISBN 0-7354-0209-4. ISSN 0094-243X. Revised, March 30, 2008.

A Mazur–Ulam space should be viewed as a triplet (M, d, \sharp) . In this context, the point $x\sharp y$ is called a *midpoint* between x and y .

In a real normed space, the midpoint has the classical definition,

$$x\sharp y = \frac{x + y}{2},$$

and the Mazur–Ulam theorem is equivalent to the assertion that *every real normed space is a Mazur–Ulam space*. It is this way we want to extend the Mazur–Ulam theorem, by investigating other classes of Mazur–Ulam spaces. And they are many.

In the above example, \sharp coincides with the arithmetic mean, A . The simplest example of a Mazur–Ulam space where the midpoint is associated to the geometric mean is $M = (0, \infty)$, endowed with the metric

$$\delta(x, y) = \left| \log \frac{x}{y} \right|,$$

and the midpoint pairing $x\sharp y = G(x, y) = \sqrt{xy}$.

Notice that no discrete metric space M (of cardinality greater than 1) supports a midpoint pairing \sharp .

In Section 2 we give a new proof of the classical Mazur–Ulam theorem by taking into account the existence of sufficiently many reflections in a real normed space. This leads to new examples of Mazur–Ulam spaces such as $Sym^{++}(n, \mathbb{R})$, the space of all $n \times n$ dimensional positive definite matrices with real coefficients (which provides a higher dimensional generalization of (\mathbb{R}_+^*, δ)). In this case, the midpoint is precisely the geometric mean in the operator-theoretical sense.

In Section 3 we briefly discuss the class of Bruhat–Tits spaces, which also includes $Sym^{++}(n, \mathbb{R})$, (and other symmetric cones attached to suitable Euclidean Jordan algebras). Our exposition follows the approach by J. D. Lawson and Y. Lim [6]. See also [5].

In Section 4 we make an attempt to built up a theory of generalized convexity in the framework of Mazur–Ulam spaces. The idea is to replace the arithmetic mean by a midpoint combination. Precisely, if $M' = (M', d', \sharp')$ and $M'' = (M'', d'', \sharp'')$ are two Mazur–Ulam spaces, with M'' a subinterval of \mathbb{R} , it is natural to say that a continuous function $f: M' \rightarrow M''$ is *convex* (more precisely, (\sharp', \sharp'') -convex) if

$$f(x\sharp' y) \leq f(x)\sharp'' f(y) \quad \text{for all } x, y \in M'.$$

The theory encompasses a large variety of functions such as the usual convex functions, the log-convex functions, the multiplicatively convex functions etc. The case of multiplicatively convex functions corresponds to the choice $M' = M'' = (\mathbb{R}_+^*, \delta, G)$ and refers to all functions $f: \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ such that

$$f(\sqrt{xy}) \leq \sqrt{f(x)f(y)} \quad \text{for all } x, y > 0.$$

It turns out that this property is shared by a large variety of special functions like Γ (the gamma function), Si (the integral sine), Li (the integral logarithm) etc. See our paper [9] for details. While the entire theory of convex functions has a companion within multiplicative convexity, things become unclear in the case of $Sym^{++}(n, \mathbb{R})$, and we end this paper with a number of open problems.

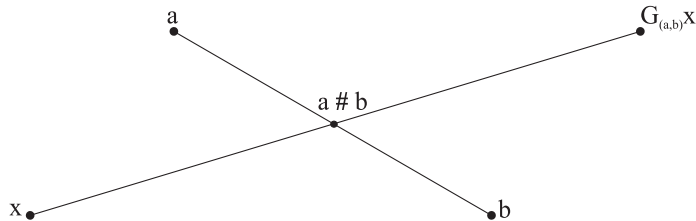


FIGURE 1. The geometrical framework of Theorem 1.

2. MIDPOINTS AS FIXED POINTS

The classical Mazur–Ulam theorem can be proved easily by noticing the presence of sufficiently many reflections on any normed vector space. This idea can be considerably extended.

Theorem 1. *Suppose that $M = (M, d)$ is a metric space such that for every pair (a, b) of points of M there exists a bijective isometry $G_{(a,b)}$, from M onto itself, having the following two properties:*

$$(MU1): G_{(a,b)}a = b \text{ and } G_{(a,b)}b = a;$$

$$(MU2): G_{(a,b)} \text{ has a unique fixed point } z \text{ (denoted } a\#b) \text{ and}$$

$$d(G_{(a,b)}x, x) = 2d(x, z) \quad \text{for all } x \in M.$$

Then M is a Mazur–Ulam space.

The geometrical framework of Theorem 1 is illustrated in Fig. 1, while its proof will constitute the objective of Lemma 2 below.

Every normed vector space verifies the hypotheses of Theorem 1. In fact, in that case the maps $G_{(a,b)}$ are precisely the reflections

$$G_{(a,b)}x = a + b - x.$$

The unique fixed point of $G_{(a,b)}$ is the midpoint of the line segment $[a, b]$, that is, $a\#b = (a + b)/2$.

In the case of $M = (\mathbb{R}_+^*, \delta, G)$, the hypotheses of Theorem 1 are fulfilled by the family of isometries

$$G_{(a,b)}x = \frac{ab}{x};$$

the fixed point of $G_{(a,b)}$ is precisely the geometric mean \sqrt{ab} , of a and b .

A higher dimensional generalization of this example is provided by the space $Sym^{++}(n, \mathbb{R})$, endowed with the *trace metric*,

$$(2.1) \quad d_{trace}(A, B) = \left(\sum_{k=1}^n \log^2 \lambda_k \right)^{1/2},$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of AB^{-1} . Since similarities preserve eigenvalues, this metric is invariant under similarities, that is,

$$d_{trace}(A, B) = d_{trace}(C^{-1}AC, C^{-1}BC) \quad \text{for all } C \in GL(n, \mathbb{R}).$$

Notice that AB^{-1} is similar with

$$A^{-1/2}(AB^{-1})A^{1/2} = A^{1/2}B^{-1/2}(A^{1/2}B^{-1/2})^* > 0$$

and this fact assures the positivity of the eigenvalues of AB^{-1} .

The proof that $Sym^{++}(n, \mathbb{R})$ admits a midpoint pairing follows from Theorem 1. We shall need the following technical result:

Lemma 1. *Given two matrices A and B in $Sym^{++}(n, \mathbb{R})$, their geometric mean*

$$A\sharp B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$$

is the unique matrix C in $Sym^{++}(n, \mathbb{R})$ such that

$$d_{trace}(A, C) = d_{trace}(B, C) = \frac{1}{2} d_{trace}(A, B).$$

The geometric mean $A\sharp B$ of two positive definite matrices A and B was introduced by Pusz and Woronowicz [11]. It is the unique solution of the equation

$$XA^{-1}X = B$$

and this fact has a number of useful consequences such as:

$$A\sharp B = (AB)^{1/2} \quad \text{if } A \text{ and } B \text{ commute}$$

$$A\sharp B = B\sharp A$$

$$(C^*AC)\sharp(C^*BC) = C^*(A\sharp B)C \quad \text{for all } C \in GL(n, \mathbb{R})$$

as well as the fact that the maps

$$G_{(A,B)}X = (A\sharp B)X^{-1}(A\sharp B)$$

verify the condition (MU1) in Theorem 1 above. As concerns the condition (MU2), let us check first the fixed points of $G_{(A,B)}$. Clearly, $A\sharp B$ is a fixed point. It is the only fixed point because any solution $X \in Sym^{++}(n, \mathbb{R})$ of the equation

$$CX^{-1}C = X,$$

with $C \in Sym^{++}(n, \mathbb{R})$, verifies the relation

$$(X^{-1/2}CX^{-1/2})(X^{-1/2}CX^{-1/2}) = I.$$

Since the square root is unique, we get $X^{-1/2}CX^{-1/2} = I$, that is, $X = C$. The second part of the condition (MU2) asks for

$$d_{trace}(G_{(A,B)}X, X) = 2d(X, A\sharp B),$$

that is,

$$d_{trace}((A\sharp B)X^{-1}(A\sharp B), X) = 2d_{trace}(X, A\sharp B),$$

for every $X \in Sym^{++}(n, \mathbb{R})$. This follows directly from the definition (2.1) of the trace metric. Notice that $\sigma(C^2) = \{\lambda^2 \mid \lambda \in \sigma(C)\}$ for all C in $Sym^{++}(n, \mathbb{R})$.

Lemma 2. *Suppose that $M_1 = (M_1, d_1)$ and $M_2 = (M_2, d_2)$ are two metric spaces which verify the conditions (MU1) and (MU2) of Theorem 1. Then*

$$T(x\sharp y) = Tx\sharp Ty$$

for all bijective isometries $T: M_1 \rightarrow M_2$.

Proof. For $x, y \in M_1$ arbitrarily fixed, consider the set $\mathcal{G}_{(x,y)}$ of all bijective isometries $G: M_1 \rightarrow M_1$ such that $Gx = x$ and $Gy = y$. Notice that the identity of M_1 belongs to $\mathcal{G}_{(x,y)}$. Put

$$\alpha = \sup_{G \in \mathcal{G}_{(x,y)}} d(Gz, z),$$

where $z = x\sharp y$. Since

$$d(Gz, z) \leq d(Gz, x) + d(x, z) = d(Gz, Gx) + d(x, z) = 2d(x, z),$$

we infer that $\alpha < \infty$. If $G \in \mathcal{G}_{(x,y)}$, so is $G' = G_{x,y}G^{-1}G_{x,y}G$, which yields

$$d(G_{(x,y)}G^{-1}G_{(x,y)}Gz, z) \leq \alpha.$$

Then

$$\begin{aligned} d(G'z, z) &= d(G_{(x,y)}G^{-1}G_{(x,y)}Gz, z) = d(G_{(x,y)}G^{-1}G_{(x,y)}Gz, G_{x,y}z) \\ &= d(G^{-1}G_{(x,y)}Gz, z) \\ &= d(G_{(x,y)}Gz, Gz) \\ &= 2d(Gz, z) \end{aligned}$$

and thus $d(Gz, z) \leq \alpha/2$ for all G . Consequently $\alpha = 0$ and this yields $G(z) = z$ for all $G \in \mathcal{G}_{(x,y)}$.

Now, for $T: M_1 \rightarrow M_2$ an arbitrary bijective isometry, we want to show that $Tz = z'$, where $z' = Tx\sharp Ty$. In fact, $G_{(x,y)}T^{-1}G_{(Tx,Ty)}T$ is a bijective isometry in $\mathcal{G}_{(x,y)}$, so

$$G_{(x,y)}T^{-1}G_{(Tx,Ty)}Tz = z.$$

This implies

$$G_{(Tx,Ty)}Tz = Tz.$$

Since z' is the only fixed point of $G_{(Tx,Ty)}$, we conclude that $Tz = z'$. \square \square

As observed by A. Vogt [13], the Mazur–Ulam theorem can be extended to all surjective maps $T: E \rightarrow F$ (acting on real normed spaces of dimension ≥ 2) which preserve equality of distances,

$$\|x - y\| = \|u - v\| \text{ implies } \|Tx - Ty\| = \|Tu - Tv\|.$$

It is open whether this result remains valid in the more general framework of Theorem 1.

3. MIDPOINTS WITHIN BRUHAT-TITS SPACES

The presence of $Sym^{++}(n, \mathbb{R})$ among the Mazur–Ulam spaces is just the peak of the iceberg. In fact, many other symmetric cones play the same property due to the presence of a special metric structure.

A *Bruhat–Tits space* is a complete metric space $M = (M, d)$ which verifies the semiparallelogram law, that is, for every pair of points x and y in M there is a point z such that

$$d(x, y)^2 + 4d(z, w)^2 \leq 2d(x, w)^2 + 2d(y, w)^2 \quad \text{for all } w \in M.$$

As in the particular case of Hilbert spaces (when equality occurs), the point z appearing in the semiparallelogram law is the unique point in M satisfying

$$d(x, z) = d(y, z) = \frac{1}{2}d(x, y).$$

We shall call z the midpoint between x and y .

A basic source of Bruhat–Tits spaces (which are also Mazur–Ulam spaces) is as follows:

Theorem 2. *Let $M = (M, d)$ be a metric space such that:*

i) For each pair of points x and y in M there is an isometry $G : M \rightarrow M$ with $Gx = y$;

ii) There exists a Hilbert space H and a continuous bijection $\exp : H \rightarrow M$ such that $\|a - b\| \leq d(\exp(a), \exp(b))$ for all $a, b \in H$, and \exp restricted to any line $\mathbb{R}c$ is an isometry;

iii) There is a pairing $\# : M \times M \rightarrow M$ such that any equality of the form

$$G(x\#y) = \exp(0)$$

for $G : M \rightarrow M$ an isometry is possible only when $\log Gx = -\log Gy$, where $\log : M \rightarrow H$ is an inverse of $\exp : H \rightarrow M$.

Then M is a Bruhat-Tits space and $x\#y$ is the midpoint of x and y for all $x, y \in M$. Furthermore, each isometry $G : M \rightarrow M$ preserves midpoints and $\#$ is the only pairing satisfying $\exp((a + b)/2) = \exp(a)\#\exp(b)$ whenever $a, b \in \mathbb{R}c$ and $c \in H$.

The details are covered by Proposition 4.2 in [6], p. 805.

One can easily prove that $M = \text{Sym}^{++}(n, \mathbb{R})$ verifies the hypothesis of Theorem 2 above. In fact, in this case we may choose as H the space $\text{Sym}(n, \mathbb{R})$, of all $n \times n$ dimensional symmetric matrices with real coefficients, endowed with the inner product

$$\langle A, B \rangle = \text{trace}(AB).$$

The corresponding exponential map $\exp : H \rightarrow M$ is given by the familiar formula

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

It is worth to notice that $\text{Sym}^{++}(n, \mathbb{R})$ is actually a Cartan-Hadamard manifold. These manifolds are complete simply connected Riemannian manifolds with seminegative curvature. See [5], Ch. XI. In their case the Riemannian metric provides a structure of Bruhat-Tits space. Examples are the symmetric cones $\Omega = \exp(V)$, attached to Euclidean Jordan algebras V . See [6] and references therein.

4. GENERALIZED CONVEXITY IN MAZUR-ULAM SPACES

As mentioned in the introduction, the Mazur-Ulam spaces constitute a natural framework for a generalized theory of convexity, where the role of the arithmetic mean is played by a midpoint pairing.

Suppose that $M' = (M', d', \#')$ and $M'' = (M'', d'', \#'')$ are two Mazur-Ulam spaces, with M'' a subinterval of \mathbb{R} . A continuous function $f : M' \rightarrow M''$ is called *convex* (more precisely, $(\#', \#'')$ -convex) if

$$(4.1) \quad f(x\#'y) \leq f(x)\#''f(y) \quad \text{for all } x, y \in M'$$

and *concave* if the opposite inequality holds. If

$$(4.2) \quad f(x\#'y) = f(x)\#''f(y) \quad \text{for all } x, y \in M'$$

then the function f is called *affine*.

Every subinterval of \mathbb{R} (endowed with the pairing associated to the arithmetic mean) is a Mazur-Ulam space and thus the above framework provides a generalization of the usual notion of convex function.

When $M' = M'' = (\mathbb{R}_+^*, \delta, G)$, the convex functions $f: \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ in the sense of (4.1) are precisely the multiplicatively convex functions. Their theory can be easily deduced from the general theory of usual convex functions by a change of variable and function. In fact, $f: \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ is *multiplicatively convex* if and only if $\log \circ f \circ \exp: \mathbb{R} \rightarrow \mathbb{R}$ is convex in the usual sense. See [9] or [10] for details.

When $M' = \mathbb{R}$ and $M'' = (\mathbb{R}_+^*, \delta, G)$, we recover the class of *log-convex functions*, that is, of those continuous functions $f: \mathbb{R} \rightarrow \mathbb{R}_+^*$ such that¹

$$f\left(\frac{x+y}{2}\right) \leq \sqrt{f(x)f(y)} \quad \text{for all } x, y \in \mathbb{R}.$$

Due to the arithmetic mean-geometric mean inequality, this class is contained in the class of convex functions and every nondecreasing log-convex function is also multiplicatively convex.

Things become considerably more technical in the case where M' is the cone $Sym^{++}(n, \mathbb{R})$. A notable example of an affine function $f: Sym^{++}(n, \mathbb{R}) \rightarrow \mathbb{R}$ is $f = \log \det$, but few is known on the corresponding class of convex functions.

In fact, an important feature of the usual theory of convex functions is the possibility to extend the basic inequality (4.1) to all convex combinations of finitely many points, and then to random variables attached to probability fields.

The analogue of $(1-\lambda)x + \lambda y$ in the context of $Sym^{++}(n, \mathbb{R})$ is

$$A \sharp_{\lambda} B = A^{1/2} (A^{-1/2} B A^{-1/2})^{\lambda} A^{1/2},$$

and this formula was investigated by F. Kubo and T. Ando [4] from the point of view of noncommutative means. As concerns the noncommutative analogue of $A \sharp_{\lambda} B$ for three (or finitely many) positive matrices, the theory is only at the beginning. An interesting approach was recently proposed by C.-K. Li and R. Mathias [7].

Besides the question on the generality of the Jensen inequality² (as well as of all other basic inequality) within the above theory of convexity associated to midpoints, many others important questions remain open. We shall mention here the problem of an analogue of the gamma function,

$$\Gamma: (0, \infty) \rightarrow \mathbb{R}, \quad \Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad \text{for } x > 0$$

in the case of $Sym^{++}(n, \mathbb{R})$, $n \geq 2$.

The famous Bohr-Mollerup theorem asserts that Γ is the unique function $f: (0, \infty) \rightarrow \mathbb{R}$ which verifies the following three conditions:

- G1) *the functional equation* $f(x+1) = xf(x)$;
- G2) *the normalization condition* $f(1) = 1$;
- G3) *the condition of log-convexity.*

¹Corrected, March 3, 2008. In the original text, the definition of log-convexity was not accurate.

²Added March 3, 2008. For a generalization of Jensen's inequality within the Bruhat-Tits spaces see: K. T. Sturm, *Probability measures on metric spaces of nonpositive curvature*. In vol.: *Heat kernels and analysis on manifolds, graphs, and metric spaces* (Pascal Auscher et al. editors). *Lecture notes from a quarter program on heat kernels, random walks, and analysis on manifolds and graphs*, April 16–July 13, 2002, Paris, France. *Contemp. Math.* **338** (2003), 357-390. This paper offers also the right framework for considering the convex combinations of arbitrary finite length in a Bruhat-Tits space; the Bruhat-Tits spaces are called there *spaces with a global nonpositive curvature*.

See [1], or [10], pp. 50-52. On the other hand, D. Gronau and J. Matkowski [3] proved a characterization of Γ within the class of multiplicatively convex functions, by replacing $\Gamma 3$) with the following condition:

$\Gamma 3')$ f is multiplicatively convex on an interval (a, ∞) , for some $a > 0$.

The problem whether these results have an analogue when $(0, \infty)$ is replaced by $Sym^{++}(n, \mathbb{R})$, with $n \geq 2$, is left open.

REFERENCES

- [1] Artin, E. , *The Gamma Function*, Holt, Rinehart and Winston, New York, 1964. English translation of German original, *Einführung in die Theorie der Gammafunktion*, Teubner, 1931.
- [2] Day M. M., *Normed linear spaces*, 3rd Edition, Springer-Verlag, Berlin, 1973.
- [3] Gronau, D. and Matkowski, J., *Geometrical convexity and generalizations of the Bohr-Mollerup theorem on the Gamma function*, *Mathematica Panonica*, **4**, 153-160 (1993).
- [4] Kubo, F., and Ando, T., *Means of Positive Linear Operators*, *Math. Ann.*, **246**, 205-224 (1980).
- [5] Lang, S., *Fundamentals of Differential Geometry*, Graduate Texts in Math., Springer-Verlag, Berlin, 1999.
- [6] Lawson, J. D., and Lim, Y., *The Geometric Mean, Matrices, Metrics and More*, *Amer. Math. Month.*, **108**, 797-812 (2001).
- [7] Li, C.-K., and Mathias, R., *Geometric Means*. Preprint, 2003.
- [8] Mazur, S., and Ulam, S., *Sur les transformations isométriques d'espaces vectoriels normés*, *C. R. Acad. Sci. Paris*, **194**, 946-948 (1932).
- [9] Niculescu, C. P., *Convexity according to the geometric mean*, *Math. Inequal. Appl.* **3**, 155-167 (2000).
- [10] Niculescu, C. P., and Persson, L.-E., *Convex Functions and Their Applications. An Introduction*. Universitaria Press, Craiova, 2003.
- [11] Pusz, W., and Woronowicz, S. L., *Functional calculus for sesquilinear forms and the purification map*, *Rep. Math. Phys.*, **8**, 159-170 (1975).
- [12] Väisälä, J., *A proof of the Mazur-Ulam theorem*, *Amer. Math. Month.*, **110**, 633-635 (2003).
- [13] Vogt, A., *Maps which preserve equality of distance*, *Studia Math.*, **XLV**, 43-48 (1973).

UNIVERSITY OF CRAIOVA, DEPARTMENT OF MATHEMATICS, STREET A.I. CUZA 13, CRAIOVA 200585, ROMANIA

E-mail address: cniculescu47@yahoo.com

URL: <http://www.inf.ucv.ro/~niculescu>