DE TTK

1949

Preserver problems and separation theorems

Doktori (PhD) értekezés

Szerző: Szokol Patrícia Ágnes

Témavezetők: Dr. Molnár Lajos és
Dr. Bessenyei Mihály

DEBRECENI EGYETEM
Természettudományi Doktori Tanács
Matematika- és Számítástudományok Doktori Iskola

Debrecen, 2015.

Ezen értekezést a Debreceni Egyetem Természettudományi Doktori Tanács Matematika- és Számítástudományok Doktori Iskola Funkcionálanalízis programja keretében készítettem a Debreceni Egyetem természettudományi doktori (PhD) fokozatának elnyerése céljából.

Debrecen, 2015. november 25.

. . . .
Szokol Patrícia Ágnes
jelölt

Tanúsítjuk, hogy Szokol Patrícia Ágnes doktorjelölt 2010-2013 között a fent megnevezett doktori program keretében irányításunkkal végezte munkáját. Az értekezésben foglalt eredményekhez a jelölt önálló alkotó tevékenységével meghatározóan hozzájárult. Az értekezés elfogadását javasoljuk.

Debrecen, 2015. november 25.

. . . .
Dr. Molnár Lajos
témavezető

. . . .
Dr. Bessenyei Mihály
témavezető

Preserver problems and separation theorems

Értekezés a doktori (PhD) fokozat megszerzése érdekében
a matematika tudományágban.

Írta: Szokol Patrícia Ágnes okleveles alkalmazott matematikus

Készült a Debreceni Egyetem Matematika- és Számítástudományok
Doktori Iskolája (Funkcionálanalízis programja) keretében

Témavezetők: Dr. Molnár Lajos és Dr. Bessenyei Mihály

A doktori szigorlati bizottság:

elnök:	Dr. Páles Zsolt	....
tagok:	Dr. Sebestyén Zoltán	....
	Dr. Szilasi József	....

A doktori szigorlat időpontja: 2014. szeptember 3.

Az értekezés bírálói:

Dr. ....
Dr. ....
Dr. ....

A bírálóbizottság:

elnök:	Dr.	....
tagok:	Dr.	....
	Dr.	....
	Dr.	....
	Dr.	....

Az értekezés védésének időpontja: 20 . . ..

ACKNOWLEDGEMENTS

First and foremost I would like to gratefully and sincerely thank to my supervisors, Professor Lajos Molnár and Professor Mihály Bessenyei for their continuous help during my PhD program. I am very grateful to Professor Lajos Molnár for encouraging my scientific work related to my PhD studies and for supporting my scientific carrier. His brilliant ideas, comments and suggestions helped and inspired me a lot to write the present dissertation. I wish to express my deep gratitude to Professor Mihály Bessenyei for his guidance, help and most importantly, his friendship during my PhD studies. I appreciate all his contributions of time and ideas to make my PhD experience productive.

I am also very grateful to all Members of Department of Analysis for their continuous encouragement. I would especially like to thank to Professor Zsolt Páles for the possibility to participate and present my results at several international conferences and for providing me the excellent condition to write my dissertation.

Last but not least, I would like to express my thankfulness to my family and to Gábor Sándor for their continuous love, help, sacrifices and support during my studies.

The author was supported by the "Lendület" Program (LP2012-46/2012) of the Hungarian Academy of Sciences.

Contents

Introduction

The present dissertation deals with two different and widely studied kinds of mathematical problems. In Chapters 1–3 preserver results on different structures are presented; and Chapters 4–5 are devoted to the investigations of separation problems. In the mathematical literature the investigations relating to the so-called “preserver” transformations on different kinds of mathematical structures are called preserver problems. The aim of such a problem is to characterize all mappings on a given structure X which preserve operations defining on the elements of X, quantities or relations among elements relevant for the structure X or other similar objects. Preserver problems show up in most parts of mathematics. In the territory of algebra, the description of homomorphisms, i.e. transformations that preserve a given operation defined on certain algebraic structure, plays an important role. Concerning the field of geometry the structure of isometries, i.e. transformations that preserve the distance on a given metric space is widely studied. Moreover, we note that preserver transformations appear also in physics and even in chemistry where they are usually called symmetries. Because of the distinguished role these maps play in those sciences, the study of preserver transformations is an important area of research. However, in the majority of results appearing in the present dissertation linearity is not assumed, it is important to emphasize that such kinds of investigations started by linear preserver problems (LPPs), which represent one of the most active research areas in matrix theory during the last one hundred years. In the case of LPPs the structure and the corresponding mappings under consideration are linear. For survey papers see e.g. [**68**], [**69**].

As a particular example of linear preserver problems we present the well-known Frobenius theorem from 1897 [**33**], which can be regarded as the first result on LPPs. This result describes the structure of all linear transformations on the algebra $\mathbb{M}_n$ of all $n \times n$ complex matrices that preserve the determinant. It reads as follows.

THEOREM 0.1. *Let $n \in \mathbb{N}$ and suppose that $\phi\colon \mathbb{M}_n \to \mathbb{M}_n$ is a linear map which satisfies*

$$\det \phi(A) = \det A, \quad A \in \mathbb{M}_n. \tag{0.1}$$

Then there exist matrices $M, N \in \mathbb{M}_n$ such that $\det MN = 1$ and ϕ is either of the form

$$\phi(A) = MAN, \quad A \in \mathbb{M}_n,$$

or of the form

$$\phi(A) = MA^{tr}N, \quad A \in M_n.$$

Here A^{tr} denotes the transpose of $A \in \mathbb{M}_n$. Observe that every transformation which appears on the right hand side of the two displayed formulas preserves the determinant of all matrices, hence the main content of this result is, in fact, that the reverse statement is also true: every linear determinant preserving transformations on $\mathbb{M}_n$ is necessarily of one of those two forms.

There are several important and well-known results on preserver problems in the field of quantum mechanics. One of the most fundamental theorems in that field is the famous Wigner theorem on the structure of quantum mechanical symmetry transformations. To present this nice result we recall that in the mathematical formulation of quantum mechanics, which is due to John von Neumann, to each quantum system there corresponds a complex Hilbert space H. Moreover, rank-one projections on H represent pure states of the quantum system. Throughout the present dissertation the set of all rank-one projections acting on H will be denoted by $P_1(H)$.

A bijective map $\phi\colon P_1(H) \to P_1(H)$ is called a quantum mechanical symmetry transformation if it preserves the quantity $\operatorname{tr} PQ$ called transition probability between pure states in the sense that

$$\operatorname{tr} \phi(P)\phi(Q) = \operatorname{tr} PQ \tag{0.2}$$

holds for arbitrary rank-one projections P, Q on H. Here tr denotes the usual trace functional. It is obvious that every transformation of the form $P \mapsto UPU^*$ on the space of all rank-one projections of H induced by the unitary or antiunitary operator U on H is a symmetry transformation. Wigner's celebrated theorem says that the converse statement is also true, i.e. every quantum symmetry transformation can be obtained in that way. The original theorem was presented in [**88**] and the first proof of that was given by Lomont and Mendelson in [**49**].

The main result of Chapter 1 describes the structure of all transformations on the set of density operators which preserve the quantum f-divergence for any strictly convex function f defined on the non-negative real line. With particular choices of the function f, the definition of quantum f-divergence leads to certain well-known and important kinds of relative entropies. Hence the main result appearing in Chapter 1 gives the form of transformations on the set of density operators that leave different types of entropies invariant. In the proof of that result Wigner's theorem plays a significant role. In fact, in our main theorem the surjectivity of preserver transformations is not assumed. Therefore we need the non-surjective version of Wigner's theorem which in the finite dimensional case has the same conclusion as that of the original case. For a version of non-surjective Wigner's theorem in Hilbert spaces of arbitrary (not necessarily finite) dimension we refer to Theorem 2.1.4 in the book [**55**] or [**5**]. It reads as follows.

THEOREM 0.2. *Let H be a complex Hilbert space and $\phi\colon P_1(H) \to P_1(H)$ be a transformation which preserves transition probability between pure states, i.e. satisfies the equality* (0.2) *for all pairs of rank-one projections $P, Q \in P_1(H)$. Then we have either a linear or a conjugate-linear isometry V of H such that*

$$\phi(P) = VPV^*, \quad P \in P_1(H).$$

We remark that in the finite dimensional case (conjugate-)linear isometries are (anti)unitary operators. Therefore, if $\dim H < \infty$, then the mapping ϕ in Theorem 0.2 is implemented by a unitary-antiunitary operator. We note that there are several other proofs for both the bijective and non-surjective versions, see for instance [**52**], [**37**] and [**36**]. We emphasize that there is an important generalization of Wigner theorem by Uhlhorn in [**83**]. It states that if $\dim H \geq 3$ then assuming merely the preservation of zero transition probability the conclusion is the same as that of Wigner's theorem. For more interesting and important results on preserver problems on quantum structure see e.g. Chapter 2 in [**55**] and the references appearing there.

In Chapter 2 we are going to describe the structure of all surjective transformations of the space of positive definite matrices that preserve so-called generalized distance measures which are parameterized by unitarily invariant norms and continuous real functions satisfying certain conditions. We will consider similar preserver transformations acting on the subset of all complex positive definite matrices with unit determinant. These kinds of investigations are motivated by results appearing in [**62**]. In that paper Molnár determined all surjective isometries of the set of positive definite matrices with respect to certain metrics which can be regarded as particular cases of generalized distance measures. Using our new theorem, we also describe the surjective maps of the set of positive definite matrices that preserve the Stein's loss or several other types of divergences. The key step of the proof of the main result is to show that on certain substructures of groups surjective transformations that preserve a given generalized distance measure d which is compatible with the group operation, necessarily preserve locally the so-called inverted Jordan triple product (i.e., they respect the operation $xy^{-1}x$). We point out that results of this kind, which first appeared in the paper [**38**], can be considered as noncommutative versions of the famous Mazur-Ulam theorem. Now we recall the original version of Mazur-Ulam theorem, which states that every surjective isometry (i.e., surjective distance preserving map) between normed real linear spaces is necessarily affine.

THEOREM 0.3. *Let X and Y be normed real linear spaces. Assume that $\phi\colon X \to Y$ is a surjective isometry such that $\phi(0) = 0$. Then ϕ is linear.*

We remark that Mazur and Ulam proved this result in response to a question raised by Banach and their proof is appeared in [**51**] and [**4**]. In [**84**] Väisälä proposed a simpler and more elegant proof for this result, which is based on the ideas of Vogt [**86**]. For a short history of the Mazur-Ulam theorem, see [**30**] (pp 6–9, 20–21).

The study of linear isometries of linear function spaces has also been an extensive research area in functional analysis over the past decades. The starting point of those investigations was the famous Banach-Stone theorem which describes the structure of all surjective linear isometries between the Banach spaces of complex-valued continuous functions on compact Hausdorff spaces equipped with the supremum norm. Denoting the space of all continuous functions f from X to $\mathbb{R}$ by $C(X,\mathbb{R})$ the result reads as follows.

THEOREM 0.4. *Let X and Y be compact Hausdorff spaces and assume that $T\colon C(X,\mathbb{R}) \to C(Y,\mathbb{R})$ is a surjective linear isometry. Then there exists a homeomorphism $\varphi : Y \to X$ and a function $\tau \in C(Y,\mathbb{R})$ with*

$$|\tau(y)| = 1, \quad y \in Y$$

such that

$$T(f)(y) = \tau(y) f(\varphi(y)), \quad y \in Y, f \in C(X,\mathbb{R}).$$

Considering the continuous complex-valued functions this result remains valid. One can find a comprehensive and very nice overview of the topic in the two volume monograph [**30**], [**31**]. There are some important metric spaces of functions which are not linear spaces. For example, the set of all probability distribution functions on $\mathbb{R}$ that plays so fundamental role in probability theory and statistics is not a linear space. In [**26**] Dolinar and Molnár described the structure of all surjective isometry of the space of all probability distribution functions with respect to so-called Kolmogorov-Smirnov metric. The importance of this metric lies in its applications in statistics (Kolmogorov-Smirnov test). It comes from the supremum norm of bounded functions, hence the result in [**26**] can be regarded as a Banach-Stone type result for the isometries of the non-linear function space of probability distribution functions. Motivated by that result Molnár continued the investigation concerning surjective isometries of certain subspaces of all probability distribution functions. Namely, in [**58**] Kolmogorov-Smirnov isometries of the spaces of all absolute continuous, or singular, or discrete probability distribution functions on $\mathbb{R}$ were investigated. In Chapter 3, instead of the space of all probability distribution functions we are going to consider a larger space, the sets of so-called (continuous) generalized distribution functions, which plays also an important role in probability theory. Motivated by the mentioned preserver results, in Chapter 3 we describe the structure of all surjective Kolmogorov-Smirnov isometries on these larger spaces.

In the second part of the present dissertation (Chapter 4 and Chapter 5) we focus to separation problems. Separation theorems play a crucial role especially in the field of convex analysis [**50**], [**78**]. One of the most fundamental separation theorems states that if a convex and a concave function is given such that the convex function is "above" the concave one, then there exists an affine function between them. Of course, the assumptions on convexity/concavity are sufficient but not necessary for the existence of an affine separator. However, there exists a characterization of those pairs of real functions that can be separated by an affine function.

THEOREM 0.5. *If $I \subset \mathbb{R}$ is an interval, $f, g\colon I \to \mathbb{R}$, then the following statements are equivalent:*

(i) there exists an affine function $h\colon I \to \mathbb{R}$ such that $f \le h \le g$;

(ii) the inequalities

$$f\big(\lambda x + (1-\lambda)y\big) \le \lambda g(x) + (1-\lambda)g(y),$$
$$g\big(\lambda x + (1-\lambda)y\big) \ge \lambda f(x) + (1-\lambda)f(y). \tag{0.3}$$

hold for all $\lambda \in [0,1]$ *and* $x, y \in I$.

Concerning the history of this result, it was motivated by the well-known theorem of Baron, Matkowski and Nikodem [**6**], which states that the existence of a convex separator between two given functions can be characterized via a simple inequality.

THEOREM 0.6. *Let* I *be an interval and* $f, g : I \to \mathbb{R}$. *Then, the following conditions are equivalent:*

(i) there exists a convex function $h : I \to \mathbb{R}$ *such that* $f \le h \le g$*;*
(ii) for all $\lambda \in [0,1]$ *and* $x, y \in I$,

$$f\big(\lambda x + (1-\lambda)y\big) \le \lambda g(x) + (1-\lambda)g(y). \tag{0.4}$$

We remark that the existence of concave separator can be also characterized with the help of second inequality appearing in (0.3). Consequently, the separated inequalities appearing in (0.3) are responsible for the existence of convex/concave separations and the simultaneous inequalities guarantees the existence of an affine separator.

In Chapter 4 and Chapter 5 we are going to extend Theorem 0.5 and Theorem 0.6, respectively. To do this we need to generalize the concept of the affine and convex functions in the following way. There are two geometrical properties of affine functions which are the most determinative. Namely, each affine function is continuous and for every two points of the plain (with distinct first coordinates) there exists exactly one affine function interpolating them. Using this geometric idea the notion of Beckenbach family can be introduced in the following way. Let $n \in \mathbb{N}$ and $I \subset \mathbb{R}$ be an interval. A set of continuous functions defined on I is called an n-parameter Beckenbach family, if each n points of $I \times \mathbb{R}$ (with pairwise distinct first coordinates) can be interpolated by a unique element of the set. The members of a Beckenbach family are termed briefly generalized lines. In Chapter 4, we characterize such pairs of real valued functions that can be separated by a member of a given Beckenbach family of order n, which is closed under convex combinations.

Moreover, each Beckenbach family induces a convexity notion in the following way: a function is generalized convex with respect to the family in question if each generalized line that interpolates the points of the function's graph intersects the graph alternately. In the case of classical convexity, the properties that the chords are above the graph and that the interpolating affine functions intersect alternately are equivalent. This fact motivates the definition of generalized convexity with respect to Beckenbach families. Not claiming completeness, we quote here the works of Beckenbach [**7**], Hopf [**42**], Popoviciu [**77**] and Tornheim [**82**]. For further details, consult the introduction of [**10**]. Chebyshev systems are particular cases of Beckenbach families, therefore the notion of convexity with respect to Beckenbach families leads to that of Chebyshev systems. In Chapter 5 our aim is to give an extension of Theorem 0.6 to the setting of two dimensional Chebyshev systems (regular pairs).

1

Quantum f-divergences preserving maps on density operators

1.1. Introduction and statement of the results

In this chapter we present some results on preserver problems appearing in the field of quantum mechanics. In [**56**] L. Molnár has described the structure of all surjective transformations on the space of density operators which preserve the Umegaki relative entropy. We extend his result by removing the surjectivity condition, i.e. we prove that this result remains valid even if we omit the strongly restrictive condition that the transformation is surjective. In fact, our aim is to prove an even more general result. It is known, that the notion of Umegaki relative entropy is a particular case of the so-called quantum f-divergence, which is the quantum version of Csiszár's f-divergence. In the main theorem of the present chapter for an arbitrary strictly convex function f defined on the non-negative real line we determine the structure of all transformations on the set of density operators which preserve the quantum f-divergence. The results of the present chapter appeared in [**60**] and [**57**].

We begin with some necessary notation which will be used in this chapter. Let H be a finite dimensional complex Hilbert space. We denote by $B(H)$ the algebra of all linear operators on H and by $B(H)^+$ the cone of all positive semi-definite operators on H. Next, $S(H)$ stands for the set of all density operators which are the elements of $B(H)^+$ having unit trace. Finally, $P_1(H)$ denotes the set of all rank-one projections acting on H.

Relative entropy is one of the most important numerical quantities appearing in quantum information theory. It is used as a measure of distinguishability between quantum states, or their mathematical representatives, the density operators. In fact, there are several concepts of relative entropy, among which the most common one is due to Umegaki. That kind of relative entropy between the operators $A, B \in S(H)$ is defined by

$$S(A\|B) = \operatorname{tr} A(\log A - \log B). \tag{1.1}$$

Here tr stands for the usual trace functional and $\log$ denotes logarithm with base 2. It is well-known that the quantity $S(A\|B)$ is always nonnegative, it is finite if and only if $\operatorname{supp} A \subset \operatorname{supp} B$ (supp stands for the orthogonal complement of the kernel of density operators) and $S(A\|B)$ is zero if and only if $A = B$ ($A, B \in S(H)$) (for details and more information see Section 11.3 in [**65**]).

In [**56**] Molnár described the general form of all surjective transformations on the set of density operators which preserve the Umegaki relative entropy. The motivation to explore the structure of those transformations came from the fundamental theorem of Wigner on quantum mechanical symmetry transformations. As we have seen in the Introduction, Wigner's theorem states that any such transformation is implemented by either a unitary or an antiunitary operator on the underlying Hilbert space. In [**56**] the author showed that the same conclusion holds for those surjective transformations on the set of density operators which preserve the Umegaki relative entropy.

THEOREM 1.1. (Molnár [**56**])
Let $\phi\colon S(H) \to S(H)$ be a surjective transformation which preserves the Umegaki relative entropies, i.e. which satisfies

$$S(\phi(A)\|\phi(B)) = S(A\|B) \tag{1.2}$$

for all $A, B \in S(H)$. Then there exists either a unitary or an antiunitary operator U on H such that ϕ is of the form

$$\phi(A) = UAU^*, \quad A \in S(H).$$

We recall, that although originally Wigner's theorem was formulated for bijective transformations, it turned out that in the finite-dimensional case it holds true also for "a priori" non-surjective transformations, see Introduction Theorem 0.2. Non-surjective versions of classical theorems like Wigner's theorem or the fundamental theorem of projective geometry (for a recent proof see [**28**]), etc. are far more useful and applicable compared to their original bijective versions. This is what has motivated us to study transformations that preserve the Umegaki relative entropy without assuming the condition of surjectivity. Our first theorem shows that in the non-surjective case the conclusion is formally the same as that of the original result of Molnár.

THEOREM 1.2. (Molnár, Szokol [**57**])
Let $\phi\colon S(H) \to S(H)$ *be a transformation such that* (1.2) *holds for every* $A, B \in S(H)$*. Then there exists either a unitary or an antiunitary operator* U *on* H *such that* ϕ *is of the form*

$$\phi(A) = UAU^*, \quad A \in S(H).$$

We note that in [**59**] the authors determined all maps on $S(H)$ that leave further kinds of relative entropies invariant.

The main goal of this chapter is to give a far-reaching generalization of Theorem 1.2 and hence of Theorem 1.1. Namely, we describe all transformations on the set of density operators which preserve the quantum f-divergence with respect to an arbitrary strictly convex function f defined on the non-negative real line.

Classical f-divergences between probability distributions were introduced by Csiszár [**25**], and by Ali and Silvey [**1**] independently. They are widely used in information theory and statistics as distinguishability measures among probability distributions (see, e.g., [**48**]). Their quantum theoretical analogues, quantum f-divergences play a similar role in quantum information theory and quantum statistics (see, e.g., [**74**]) and were defined by Petz [**71**], [**72**], [**73**]. This concept is an essential common generalization and extension of several notions of quantum relative entropy including Umegaki's and Tsallis' relative entropies. We note, that quantum f-divergences are particular cases of the so-called quasi-entropies (for details see the introduction in [**40**]).

We define that concept following the approach given in [**40**]. We recall that $B(H)$ is a complex Hilbert space with the Hilbert-Schmidt inner product $\langle ., . \rangle_{\mathrm{HS}}\colon B(H) \times B(H) \to \mathbb{C}$ defined by

$$\langle A, B \rangle_{\mathrm{HS}} = \operatorname{tr} AB^* \quad A, B \in B(H).$$

For any $A \in B(H)$, let $L_A, R_A\colon B(H) \to B(H)$ be the left and the right multiplication operators defined as

$$L_A T = AT, \; R_A T = TA, \quad T \in B(H).$$

We remark that $L_A R_B = R_B L_A$ holds for every $A, B \in B(H)$. If $A, B \in B(H)^+$, then L_A and R_B are positive Hilbert space operators, hence so is $L_A R_B$.

Let now $f\colon [0, \infty[\to \mathbb{R}$ be a function which is continuous on $]0, \infty[$ and the limit

$$\alpha := \lim_{x \to \infty} \frac{f(x)}{x}$$

exists in $[-\infty, \infty]$. Following Definition 2.1 in [**40**], for $A, B \in B(H)^+$ with $\operatorname{supp} A \subset \operatorname{supp} B$ the f-divergence $S_f(A\|B)$ of A with respect to B is defined by

$$S_f(A\|B) = \left\langle \sqrt{B}, f(L_A R_{B^{-1}})\sqrt{B} \right\rangle_{\mathrm{HS}}.$$

In the general case we set

$$S_f(A\|B) = \lim_{\varepsilon \to 0^+} S_f(A\|B + \varepsilon I),$$

where I is the identity operator on H. By Proposition 2.2 [**40**] the limit above exists in $[-\infty, \infty]$. We next recall a useful formula which will play an important role in our arguments. Let $A, B \in B(H)^+$ and for any $\lambda \in \mathbb{R}$ denote by P_λ, respectively by Q_λ the projection on H projecting onto the kernel of $A - \lambda I$, respectively onto the kernel of $B - \lambda I$. According to Corollary 2.3 [**40**] we have

$$S_f(A||B) = \sum_{a \in \sigma(A)} \left(\sum_{b \in \sigma(B)\setminus\{0\}} bf\left(\frac{a}{b}\right) \operatorname{tr} P_a Q_b + \alpha a \operatorname{tr} P_a Q_0 \right), \tag{1.3}$$

where $\sigma(.)$ stands for the spectrum of elements in $B(H)$ and the convention $0 \cdot (-\infty) = 0 \cdot \infty = 0$ is used.

We can now formulate our result which describes the structure of all transformations on $S(H)$ leaving the quantum f-divergence invariant with respect to a given real valued and strictly convex function f on $[0, \infty[$. First observe that for any unitary or antiunitary operator U on H the transformation $A \mapsto UAU^*$ preserves the f-divergence on $S(H)$, i.e., we have $S_f(UAU^*\|UBU^*) = S_f(A\|B)$ for any $A, B \in S(H)$ (here the function f is as above). The theorem below states that for a strictly convex function f the reverse statement is also true: All transformations on $S(H)$ which leave the f-divergence invariant are of the preceding form, i.e., they are all implemented by unitary or antiunitary operators. Let us point out the fact that any convex function $f\colon [0, \infty[\to \mathbb{R}$ satisfies the requirements given in the definition of f-divergence: it is continuous on the open interval $]0, \infty[$ and since the difference quotient $(f(x) - f(0))/(x - 0)$ is increasing, the limit $\lim_{x \to \infty} f(x)/x$ exists and is finite or equal to ∞. The precise formulation of our result is as follows.

THEOREM 1.3. (Molnár, Nagy, Szokol [**60**])
Assume that $f\colon [0, \infty[\to \mathbb{R}$ *is a strictly convex function and* $\phi\colon S(H) \to$

$S(H)$ is a transformation satisfying

$$S_f(\phi(A)||\phi(B)) = S_f(A||B), \quad A, B \in S(H).$$

Then there is either a unitary or an antiunitary operator U on H such that ϕ is of the form

$$\phi(A) = UAU^*, \quad A \in S(H).$$

We emphasize that the bijectivity or the surjectivity of the transformation ϕ is not assumed in the theorem and we do not require any sort of linearity either. Let us make a remark also on the convexity assumption above. When they consider f-divergence in the classical setting, it is practically always assumed that the function f is convex. The main reason is that this condition guarantees the joint convexity and information monotonicity of the f-divergence which are significant properties. As for quantum f-divergence, to obtain similar important properties one needs to assume that f is operator convex (see, e.g., [**40**]). Therefore, our condition that f is a convex function is very natural and not restrictive. As for strict convexity, it is easy to see that if f is affine then $S_f(.\|.)$ is constant. Hence in that case every selfmap of $S(H)$ preserves the f-divergence which is obviously out of interest. A few important examples of quantum f-divergences between density operators follow. Let $A, B \in S(H)$.

(i) If

$$f(x) = \begin{cases} x \log_2 x, & x > 0 \\ 0, & x = 0, \end{cases} \tag{1.4}$$

then the definition of quantum f-divergence leads to that of Umegaki relative entropy of A with respect to B (which was defined in (1.1)).

(ii) Let $q \in]0, 1[$ and define the function $f_q \colon [0, \infty[\to \mathbb{R}$ by $f_q(x) = (1 - x^q)/(1 - q)$ $(x \geq 0)$. Then

$$S_{f_q}(A\|B) = \frac{1 - \operatorname{tr} A^q B^{1-q}}{1 - q}$$

which is the quantum Tsallis relative entropy of A with respect to B (see, e.g., [**35**]).

(iii) If $f(x) = (\sqrt{x} - 1)^2$ $(x \geq 0)$, then $S_f(A\|B) = \left\|\sqrt{A} - \sqrt{B}\right\|_{\mathrm{HS}}^2$, where $\|.\|_{\mathrm{HS}}$ stands for the Hilbert-Schmidt norm.

Observe that the functions appearing in (i)-(iii) are all strictly convex. Therefore, Theorem 1.3 can also be applied to determine all transformations of

$S(H)$ preserving Umegaki relative entropy which shows that it is a real generalization of Theorem 1.2.

1.2. Proofs

In this section we are going to prove Theorem 1.3. As we have already observed the function f which corresponds to the Umegaki relative entropy is the one appeared in (1.4). Since it is a strictly convex function we get that Theorem 1.2 is a particular case of Theorem 1.3. Therefore we skip the proof of Theorem 1.2 which can be found in the paper [**57**].

Before the proof of Theorem 1.3, we recall that the self-adjoint operators $A, B \in B(H)$ are said to be orthogonal if and only if $AB = 0$, which is equivalent to the fact that A and B have mutually orthogonal ranges.

PROOF OF THEOREM 1.3. Observe that for any real number a and operators $A, B \in S(H)$ we have $S_{f+a}(A\|B) = S_f(A\|B) + a$. Therefore without any loss of generality we may and do assume that $f(0) = 0$. As we have mentioned the quantum f-divergence is defined for any function $f\colon [0, \infty[\to \mathbb{R}$ which is continuous and the limit $\alpha = \lim_{x\to\infty} \frac{f(x)}{x}$ exists. According to the value of the limit α, we divide the proof into two cases.

CASE I. We assume that α is finite. First we show that ϕ preserves the orthogonality in both directions, i.e. it satisfies

$$\phi(A)\phi(B) = 0 \iff AB = 0$$

for any $A, B \in S(H)$. To see this we need the following characterization of orthogonality. For any $A, B \in S(H)$ we have

$$AB = 0 \iff S_f(A||B) = \alpha. \tag{1.5}$$

Indeed, if $AB = 0$, then a straightforward calculation using the formula (1.3) shows that $S_f(A||B) = \alpha$. Suppose now that the right-hand side of (1.5) holds. On the one hand, we have

$$S_f(A||B) = \sum_{a\in\sigma(A)\setminus\{0\}} \sum_{b\in\sigma(B)\setminus\{0\}} bf\left(\frac{a}{b}\right) \operatorname{tr} P_a Q_b + \sum_{a\in\sigma(A)\setminus\{0\}} \alpha a \operatorname{tr} P_a Q_0.$$

On the other hand, we have

$$\alpha = \alpha \operatorname{tr} A = \alpha \sum_{a\in\sigma(A)\setminus\{0\}} a \operatorname{tr} P_a.$$

Since the left-hand sides of the previous two equalities are equal, using the fact that

$$\sum_{b\in\sigma(B)} Q_b = I$$

we easily infer that

$$\sum_{a\in\sigma(A)\setminus\{0\}} \sum_{b\in\sigma(B)\setminus\{0\}} bf\left(\frac{a}{b}\right) \operatorname{tr} P_a Q_b = \alpha \sum_{a\in\sigma(A)\setminus\{0\}} \sum_{b\in\sigma(B)\setminus\{0\}} a \operatorname{tr} P_a Q_b.$$

This yields that

$$\sum_{a\in\sigma(A)\setminus\{0\}} \sum_{b\in\sigma(B)\setminus\{0\}} \left(\alpha a - bf\left(\frac{a}{b}\right)\right) \operatorname{tr} P_a Q_b = 0. \tag{1.6}$$

Let $a \in \sigma(A)\setminus\{0\}$ and $b \in \sigma(B)\setminus\{0\}$ and consider the quantity

$$\alpha a - bf\left(\frac{a}{b}\right) = b\left(\alpha\frac{a}{b} - f\left(\frac{a}{b}\right)\right). \tag{1.7}$$

It follows from the strict convexity of f that the function $f_1\colon]0,\infty[\to \mathbb{R}$ defined by

$$f_1(x) = \frac{f(x)-f(0)}{x-0} = \frac{f(x)}{x}, \quad x > 0 \tag{1.8}$$

is strictly increasing. Therefore, for any $x > 0$ we have

$$f(x)/x < \lim_{s\to\infty} f_1(s) = \alpha \tag{1.9}$$

and hence

$$f(x) < \alpha x \tag{1.10}$$

which implies that the quantity in (1.7) is positive. On the other hand, since $P_a, Q_b \in B(H)^+$ we have $\operatorname{tr} P_a Q_b \geq 0$. It follows that the terms of the sum on the left-hand side of (1.6) are nonnegative. We conclude that $\operatorname{tr} P_a Q_b = 0$ holds for all $a \in \sigma(A)\setminus\{0\}$ and $b \in \sigma(B)\setminus\{0\}$ which implies that $AB = 0$. This completes the proof of the equivalence in (1.5). Since ϕ preserves the f-divergence, it then follows that ϕ preserves the orthogonality in both directions.

Apparently, we can characterize the elements of $P_1(H)$ as those operators in $S(H)$ which belong to a set of n pairwise orthogonal density operators on H. By the orthogonality preserving property of ϕ, we infer that it maps $P_1(H)$ into itself. We claim that ϕ preserves the transition probability (the trace of products) on $P_1(H)$. To prove this, let $P, Q \in P_1(H)$ be arbitrary. A straightforward calculation gives that

$$S_f(P||Q) = (f(1) - \alpha)\,\mathrm{tr}\,PQ + \alpha$$

and similarly

$$S_f(\phi(P)||\phi(Q)) = (f(1) - \alpha)\,\mathrm{tr}\,\phi(P)\phi(Q) + \alpha.$$

By (1.10) one has $f(1) - \alpha \neq 0$ and it follows that

$$\mathrm{tr}\,\phi(P)\phi(Q) = \mathrm{tr}\,PQ.$$

This means that the restriction of ϕ to $P_1(H)$ preserves the transition probability. The non-surjective version of Wigner's theorem (Theorem 0.2) describes the structure of all such maps. Since H is finite dimensional, we obtain that there exists either a unitary or an antiunitary operator U on H such that

$$\phi(P) = UPU^*, \quad P \in P_1(H).$$

Consider the transformation $\psi\colon S(H) \to S(H)$ defined by

$$\psi(A) = U^*\phi(A)U, \quad A \in S(H).$$

It is clear that this map preserves the quantum f-divergence and has the additional property that it acts as the identity on $P_1(H)$. Define the function $f_2\colon [0,\infty[\to \mathbb{R}$ by

$$f_2(x) = \begin{cases} xf\left(\frac{1}{x}\right), & x > 0 \\ \alpha, & x = 0. \end{cases}$$

Let $A \in S(H)$ be fixed and $Q \in P_1(H)$ be arbitrary. Using (1.3), we easily have

$$S_f(Q||A) = \mathrm{tr}\,Qf_2(A)$$

and similarly

$$S_f(Q||\psi(A)) = \mathrm{tr}\,Qf_2(\psi(A)).$$

By the properties of ψ, the left-hand sides of the above equalities coincide, therefore

$$\mathrm{tr}\,f_2(\psi(A))Q = \mathrm{tr}\,f_2(A)Q$$

holds for every rank-one projection Q on H, which implies that $f_2(\psi(A)) = f_2(A)$. Observe that $f_2(x) = f_1(1/x)$ $(x > 0)$. Since f_1 is clearly injective, so is f_2 on $]0,\infty[$. Moreover, by (1.9) we have $f_1(x) < \alpha$ $(x > 0)$ and then we obtain that f_2 is injective on the whole interval $[0,\infty[$. It then follows that

$$A = \psi(A) = U^*\phi(A)U, \quad A \in S(H)$$

and this completes the proof in CASE I.

CASE II. We now assume that α is infinite. The basic strategy of the argument below is close to that of the proof of [**57**, Theorem]. However, due to the fact that here we consider general divergences, we necessarily face many problems which are of different levels of difficulties. Although at some parts in our argument we may directly refer to parts of the proof of [**57**, Theorem], for the sake of understandability, readability and completeness we present practically all necessary details. As mentioned before the formulation of Theorem, the possibility $\alpha = -\infty$ is ruled out by the convexity of the function f. Therefore, $\alpha = \infty$. We show that ϕ preserves the rank, i.e. for any $A \in S(H)$ the rank of $\phi(A)$ equals the rank of A. In order to see it, let $A, B \in S(H)$ be arbitrary. Using (1.3) it is easy to check that $S_f(A||B) < \infty$ holds if and only if $\operatorname{supp} A \subset \operatorname{supp} B$. It follows that

$$\operatorname{supp}\phi(A) \subset \operatorname{supp}\phi(B) \iff \operatorname{supp} A \subset \operatorname{supp} B$$

and next that

$$\operatorname{supp}\phi(A) \subsetneq \operatorname{supp}\phi(B) \iff \operatorname{supp} A \subsetneq \operatorname{supp} B. \tag{1.11}$$

Observe that the rank of A is k if and only if there is a strictly increasing chain (with respect to inclusion) of supports of n density operators on H such that its kth element is $\operatorname{supp} A$. Using this characterization and (1.11) we see that ϕ leaves the rank of operators invariant. In particular

$$\phi(P_1(H)) \subset P_1(H). \tag{1.12}$$

We next verify that ϕ is injective. Indeed, it is an immediate consequence of the following assertion. For any $A, B \in S(H)$ we have $f(1) \leq S_f(A\|B)$ and equality appears if and only if $A = B$. For the proof, it is clear that if the support of A is not contained in that of B, then this inequality holds and it is strict. Otherwise we have

$$S_f(A\|B) = \sum_{a\in\sigma(A)} \sum_{b\in\sigma(B)\setminus\{0\}} (b \operatorname{tr} P_a Q_b) f\left(\frac{a}{b}\right).$$

Observe that the numbers $b \operatorname{tr} P_a Q_b$ are nonnegative for all $a \in \sigma(A),\ b \in \sigma(B)\setminus\{0\}$ and their sum is 1. Thus, by the convexity of f it follows easily that

$$f(1) = f\left(\sum_{a\in\sigma(A)} \sum_{b\in\sigma(B)\setminus\{0\}} b\frac{a}{b} \operatorname{tr} P_a Q_b\right) \leq S_f(A\|B) \tag{1.13}$$

and this yields the desired inequality. Moreover, since f is strictly convex, in the above inequality we have equality exactly when for any $a \in \sigma(A)$ and

$b \in \sigma(B)\backslash\{0\}$ satisfying $b \operatorname{tr} P_a Q_b > 0$, the value a/b is constant. Since the sum of the numbers $b(a/b) \operatorname{tr} P_a Q_b$ over such values of a and b equals 1, we get that this constant is 1. By the previous observations we easily obtain that for any $a \in \sigma(A)$ and $b \in \sigma(B)\backslash\{0\}$ at least one of the relations $a = b$, $P_a Q_b = 0$ must hold. One can simply check that under the condition $\operatorname{supp} A \subset \operatorname{supp} B$ which we have supposed above, the latter property is equivalent to the equality $A = B$. We conclude that ϕ is injective.

We derive a formula which will be used several times in the rest of the proof. Define the function $f_3\colon]0,\infty[\to \mathbb{R}$ by

$$f_3(x) = xf\left(\frac{1}{x}\right) = f_1\left(\frac{1}{x}\right), \quad x > 0,$$

where f_1 is the function that has appeared in (1.8). Easy computation shows that for any $A \in S(H)$ and $P \in P_1(H)$ with $\operatorname{supp} P \subset \operatorname{supp} A$ we have

$$S_f(P||A) = \operatorname{tr} P|_{\operatorname{supp} A} f_3(A|_{\operatorname{supp} A}). \tag{1.14}$$

In the next part of our argument H is assumed to be 2-dimensional. We claim that for any $A \in S(H)$ we have

$$[\min \sigma(A), \max \sigma(A)] \subset [\min \sigma(\phi(A)), \max \sigma(\phi(A))]$$

meaning that ϕ can only enlarge the convex hull of the spectrum of the elements of $S(H)$. To verify this property, first observe that by (1.12) the inclusion above holds for all $A \in P_1(H)$. Now pick a rank-two operator $A \in S(H)$ and set $\lambda = \max \sigma(A) \in [1/2, 1[$. Then there are mutually orthogonal projections $P, Q \in P_1(H)$ such that $A = \lambda P + (1-\lambda)Q$. Applying (1.14) we easily get that for any $R \in P_1(H)$

$$S_f(R||A) = f_3(\lambda) \operatorname{tr} RP + f_3(1-\lambda) \operatorname{tr} RQ. \tag{1.15}$$

We have seen that f_1 is strictly increasing, so f_3 is strictly decreasing and thus $f_3(\lambda) \leq f_3(1-\lambda)$. It follows that as R runs through the set $P_1(H)$, the quantity $S_f(R||A)$ runs through $[f_3(\lambda), f_3(1-\lambda)]$. Similarly, we infer that for any $R \in P_1(H)$ the number $S_f(\phi(R)\|\phi(A))$ belongs to $[f_3(\mu), f_3(1-\mu)]$, where $\mu = \max \sigma(\phi(A))$. Since ϕ preserves f-divergence, we obtain that

$$f_3(\mu) \leq f_3(\lambda) \leq f_3(1-\lambda) \leq f_3(1-\mu).$$

Due to the fact that f_3 is strictly decreasing this implies

$$\min \sigma(\phi(A)) \leq \min \sigma(A) \leq \max \sigma(A) \leq \max \sigma(\phi(A))$$

which verifies our claim.

In the most crucial part of the proof that follows we show that $\phi\left(\frac{1}{2}I\right) = \frac{1}{2}I$. Assume on the contrary that there is a number $\lambda_1 \in]1/2, 1[$ and mutually orthogonal projections $P_1, Q_1 \in P_1(H)$ such that

$$\phi\left(\frac{1}{2}I\right) = \lambda_1 P_1 + (1-\lambda_1)Q_1. \tag{1.16}$$

By (1.15) for any $R \in P_1(H)$ one has $S_f\left(R \,\middle\|\, \frac{1}{2}I\right) = f_3\left(\frac{1}{2}\right)$ and then we deduce that

$$\begin{aligned} f_3\left(\frac{1}{2}\right) = S_f\left(\phi(R) \,\middle\|\, \phi\left(\frac{1}{2}I\right)\right) &= f_3(\lambda_1)\operatorname{tr}\phi(R)P_1 \\ &+ f_3(1-\lambda_1)\operatorname{tr}\phi(R)Q_1. \end{aligned} \tag{1.17}$$

Because $1 = \operatorname{tr}\phi(R) = \operatorname{tr}\phi(R)P_1 + \operatorname{tr}\phi(R)Q_1$, this gives us that $f_3\left(\frac{1}{2}\right)$ is a convex combination of $f_3(\lambda_1)$ and $f_3(1-\lambda_1)$. Since these two latter numbers are different (f_3 is strictly decreasing), we infer that $\operatorname{tr}\phi(R)P_1$ has the same value for any $R \in P_1(H)$ and the same holds for $\operatorname{tr}\phi(R)Q_1$, too. We next prove that

$$\operatorname{tr}\phi(R)P_1 > \operatorname{tr}\phi(R)Q_1. \tag{1.18}$$

To this end, we first show that f_3 is strictly convex. According to [**64**, Lemma 1.3.2], a real-valued function g defined on an interval J is strictly convex if and only if for any elements $x_1 < x_2 < x_3$ in J we have

$$\det\begin{pmatrix} 1 & x_1 & g(x_1) \\ 1 & x_2 & g(x_2) \\ 1 & x_3 & g(x_3) \end{pmatrix} > 0.$$

It is easy to check that for any positive reals $x_1 < x_2 < x_3$ we have

$$\frac{1}{x_1x_2x_3}\det\begin{pmatrix} 1 & x_1 & f_3(x_1) \\ 1 & x_2 & f_3(x_2) \\ 1 & x_3 & f_3(x_3) \end{pmatrix} = \det\begin{pmatrix} 1 & 1/x_3 & f(1/x_3) \\ 1 & 1/x_2 & f(1/x_2) \\ 1 & 1/x_1 & f(1/x_1) \end{pmatrix}$$

and the latter number is positive due to the strict convexity of f. This proves that f_3 is also strictly convex. Using that property and the fact that f_3 is strictly decreasing, referring to (1.17) one can verify in turn that $\operatorname{tr}\phi(R)P_1 \neq \frac{1}{2}$ and then that $\operatorname{tr}\phi(R)P_1 > \frac{1}{2}$. Therefore, we obtain that $\operatorname{tr}\phi(R)P_1 > \operatorname{tr}\phi(R)Q_1$. In fact, in any representation of $f_3\left(\frac{1}{2}\right)$ as a convex combination of $f_3(t)$ and $f_3(1-t)$ ($t \in]1/2, 1[$), the coefficient of the former term is greater than the coefficient of the latter one.

Now choose unit vectors u and v from the ranges of P_1 and Q_1. It is easy to check that the matrix of an element of $P_1(H)$ with respect to the basis $\{u, v\}$ is of the form

$$\begin{pmatrix} a & \varepsilon\sqrt{a(1-a)} \\ \overline{\varepsilon}\sqrt{a(1-a)} & 1-a \end{pmatrix},$$

where $a \in [0,1]$ and $\varepsilon \in \mathbb{C}$ with $|\varepsilon| = 1$. It follows from what we have observed above that when R runs through the set $P_1(H)$, the number $a = \operatorname{tr}\phi(R)P_1$ in the matrix representation of $\phi(R)$ remains constant, and since f_3 is clearly injective, a is different from the numbers $0, 1$. Now we can rewrite (1.17) in the form

$$af_3(\lambda_1) + (1-a)f_3(1-\lambda_1) = f_3\left(\frac{1}{2}\right). \tag{1.19}$$

Next let us consider $\phi\left(\phi\left(\frac{1}{2}I\right)\right)$. We have

$$\phi\left(\phi\left(\frac{1}{2}I\right)\right) = \lambda_2 P_2 + (1-\lambda_2)Q_2,$$

for some $\frac{1}{2} \leq \lambda_2 < 1$ and mutually orthogonal elements P_2, Q_2 of $P_1(H)$. In fact, as ϕ can only enlarge the convex hull of the spectrum and $\lambda_1 > \frac{1}{2}$, it follows that $\lambda_2 > \frac{1}{2}$. Pick an arbitrary rank-one projection R on H and set $R_2 = \phi(\phi(R))$. Since ϕ preserves $S_f(.\|.)$, by (1.17) we have

$$\begin{aligned} f_3\left(\frac{1}{2}\right) &= S_f\left(\phi(\phi(R))\,\middle\|\,\phi\left(\phi\left(\frac{1}{2}I\right)\right)\right) \\ &= S_f(R_2\|\lambda_2 P_2 + (1-\lambda_2)Q_2) \\ &= f_3(\lambda_2)\operatorname{tr} R_2P_2 + f_3(1-\lambda_2)\operatorname{tr} R_2Q_2. \end{aligned} \tag{1.20}$$

Here $\lambda_2 > \frac{1}{2}$ is fixed. Since we have $\operatorname{tr} R_2P_2 + \operatorname{tr} R_2Q_2 = 1$, it follows just as above that the numbers $\operatorname{tr} R_2P_2$ and $\operatorname{tr} R_2Q_2$ are also fixed, they do not change when R varies. Moreover, we necessarily have

$$\operatorname{tr} R_2P_2 > \operatorname{tr} R_2Q_2. \tag{1.21}$$

Consider a unit vector from the range of P_2. Let x, y be its coordinates with respect to the basis $\{u, v\}$ appearing in the previous paragraph. It is easy to see that the representing matrix of P_2 is

$$\begin{pmatrix} x \\ y \end{pmatrix}\begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix}^t,$$

where ^{t} denotes the transposition. Moreover, since R_2 is a rank-one projection which is the image (under ϕ) of a rank-one projection, its matrix representation is of the form

$$\begin{pmatrix} a & \varepsilon\sqrt{a(1-a)} \\ \overline{\varepsilon}\sqrt{a(1-a)} & 1-a \end{pmatrix},$$

where a is the same as in (1.19), and $\varepsilon \in \mathbb{C}$ with $|\varepsilon| = 1$ may vary. We have

$$\operatorname{tr} R_2 P_2 = \operatorname{tr}\left[\begin{pmatrix} a & \sqrt{a(1-a)}\varepsilon \\ \sqrt{a(1-a)}\overline{\varepsilon} & 1-a \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix}^t\right].$$

Elementary computations show that the latter quantity equals

$$ax\overline{x} + \sqrt{a(1-a)}\varepsilon\overline{x}y + \sqrt{a(1-a)}\overline{\varepsilon}x\overline{y} + (1-a)y\overline{y} =$$
$$a|x|^2 + (1-a)|y|^2 + 2\sqrt{a(1-a)}\Re(\varepsilon\overline{x}y).$$

As we have already noted, the value of $\operatorname{tr} R_2P_2$ does not change when R varies and a is also constant. Therefore, we obtain that the value of

$$a|x|^2 + (1-a)|y|^2 + 2\sqrt{a(1-a)}\Re(\varepsilon\overline{x}y)$$

is constant for infinitely many values of ε (by the injectivity of ϕ we see that R_2 runs through a set of continuum cardinality, so there is such a large set for the values of ε, too). It follows that $\Re(\varepsilon\overline{x}y)$ is constant for infinitely many values of ε which clearly implies that $\overline{x}y = 0$. Therefore, the column vector

$$\begin{pmatrix} x \\ y \end{pmatrix}$$

is a scalar multiple of

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ or } \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Obviously, this can happen only when $P_2 = P_1$ or $P_2 = Q_1$. Using the fact that R_2 is the image of a rank-one projection under ϕ, it follows from (1.18) that

$$\operatorname{tr} R_2P_1 > \operatorname{tr} R_2Q_1. \tag{1.22}$$

Therefore the equality $P_2 = Q_1$ is excluded due to (1.21). Consequently, $P_2 = P_1$ and $Q_2 = Q_1$ and hence we obtain

$$\phi\left(\phi\left(\frac{1}{2}I\right)\right) = \lambda_2P_1 + (1-\lambda_2)Q_1. \tag{1.23}$$

From (1.20) we have

$$f_3(\lambda_2)\,\mathrm{tr}\,R_2P_1 + f_3(1-\lambda_2)\,\mathrm{tr}\,R_2Q_1 = f_3\left(\frac{1}{2}\right).$$

On the other hand, referring to the preceding paragraph we see that

$$\mathrm{tr}\,R_2P_1 = a \quad \text{and} \quad \mathrm{tr}\,R_2Q_1 = 1-a,$$

thus it follows that

$$af_3(\lambda_2) + (1-a)f_3(1-\lambda_2) = f_3\left(\frac{1}{2}\right). \tag{1.24}$$

We assert that the equation

$$af_3(\lambda) + (1-a)f_3(1-\lambda) = f_3\left(\frac{1}{2}\right) \tag{1.25}$$

has at most two solutions in $]0,1[$. Indeed, consider the function

$$\lambda \mapsto af_3(\lambda) + (1-a)f_3(1-\lambda), \quad \lambda \in]0,1[.$$

Since f_3 is strictly convex, the same holds for this function, too. Therefore it is obvious that it cannot take the same values at three different places. Hence (1.25) does not have three different solutions in $]0,1[$. But by (1.19) and (1.24) λ_1, λ_2 and clearly $\frac{1}{2}$ too are solutions. Since $\lambda_2 \geq \lambda_1 > \frac{1}{2}$, it then follows that $\lambda_2 = \lambda_1$ and referring to (1.16) and (1.23) we see that $\phi\left(\phi\left(\frac{1}{2}I\right)\right) = \phi\left(\frac{1}{2}I\right)$. Since ϕ is injective, this gives us that $\phi\left(\frac{1}{2}I\right) = \frac{1}{2}I$. Therefore, ϕ sends $\frac{1}{2}I$ to itself.
Now let $\frac{1}{2}I \neq A \in S(H)$ be a rank-two operator and denote by $\lambda \in]1/2,1[$ its maximal eigenvalue. We assert that $\sigma(\phi(A)) = \sigma(A)$. Let $f_4\colon]0,1[\to \mathbb{R}$ be the function defined by

$$f_4(x) = \frac{f(2x) + f(2(1-x))}{2}, \quad x \in]0,1[.$$

Using the formula (1.3) we obtain $S_f\left(A\,\middle\|\,\frac{1}{2}I\right) = f_4(\lambda)$ and, similarly,

$$S_f\left(\phi(A)\,\middle\|\,\frac{1}{2}I\right) = f_4(\lambda'),$$

where $\lambda' = \max\sigma(\phi(A)) > \frac{1}{2}$. Since ϕ preserves the f-divergence and sends $\frac{1}{2}I$ to itself, it follows that $S_f\left(\phi(A)\,\middle\|\,\frac{1}{2}I\right) = S_f\left(A\,\middle\|\,\frac{1}{2}I\right)$, and hence that $f_4(\lambda) = f_4(\lambda')$. We have that f_4 is strictly convex and symmetric with respect to the middle point $\frac{1}{2}$ of its domain. By elementary properties of convex functions this implies that the restriction of f_4 to $]1/2,1[$ is strictly increasing. We necessarily obtain that $\lambda = \lambda'$ and this yields that the spectrum of A coincides with that of $\phi(A)$. Therefore, ϕ is spectrum preserving.

Select mutually orthogonal projections $P, Q \in P_1(H)$ and pick a number $\lambda \in]1/2, 1[$. Consider the operator $B = \lambda P + (1-\lambda)Q$. By the spectrum preserving property of ϕ we can choose another pair $P', Q' \in P_1(H)$ of mutually orthogonal projections such that $\phi(B) = \lambda P' + (1-\lambda)Q'$. We have learnt before that when R runs through the set of all rank-one projections, the quantity $S_f(R||B)$ runs through the interval $[f_3(\lambda), f_3(1-\lambda)]$. Using the equation (1.15) we easily see that $S_f(R\|B) = f_3(\lambda)$ if and only if $\operatorname{tr} RP = 1$ which holds exactly when $R = P$. Therefore, we obtain

$$R = P \iff S_f(R\|B) = f_3(\lambda) \iff S_f(\phi(R)\|\phi(B)) = f_3(\lambda)$$
$$\iff S_f(\phi(R)\|\lambda P' + (1-\lambda)Q') = f_3(\lambda) \iff \phi(R) = P'.$$

This gives us that $\phi(P) = P'$ and then we also obtain $\phi(Q) = Q'$. Consequently, ϕ preserves the orthogonality between rank-one projections. Moreover, we have

$$\phi(B) = \phi(\lambda P + (1-\lambda)Q) = \lambda\phi(P) + (1-\lambda)\phi(Q). \tag{1.26}$$

Next, we show that ϕ preserves also the nonzero transition probability between rank-one projections. Let P and R be different rank-one projections which are not orthogonal to each other. Choose a rank-one projection Q which is orthogonal to P. Pick $\lambda \in]1/2, 1[$. On the one hand, we have

$$S_f(R\|\lambda P + (1-\lambda)Q) = f_3(\lambda) \operatorname{tr} RP + f_3(1-\lambda) \cdot \operatorname{tr} RQ$$

and on the other hand, by (1.26), we compute

$$S_f(R\|\lambda P + (1-\lambda)Q) = S_f(\phi(R)\|\lambda\phi(P) + (1-\lambda)\phi(Q))$$
$$= f_3(\lambda) \operatorname{tr} \phi(R)\phi(P) + f_3(1-\lambda) \operatorname{tr} \phi(R)\phi(Q).$$

Comparing the right-hand sides, we infer

$$\operatorname{tr} RP = \operatorname{tr} \phi(R)\phi(P).$$

Consequently, ϕ preserves the transition probability between rank-one projections.

Above we have supposed that H is two-dimensional. Assume now that H is an arbitrary finite dimensional Hilbert space and $\phi\colon S(H) \to S(H)$ is a transformation which preserves the f-divergence. We show that ϕ preserves the transition probability between rank-one projections in this case too. In fact, we can reduce the general case to the previous one. To see this, first let H_2 be a two-dimensional subspace of H and $A_0 \in S(H)$ be such that $\operatorname{supp} A_0 = H_2$. Set $H_2' = \operatorname{supp} \phi(A_0)$. Since ϕ preserves the rank,

H_2' is also two-dimensional. By what we have learnt at the beginning of the proof in CASE II, ϕ maps any element of $S(H)$ whose support is included in H_2 to an element of $S(H)$ whose support is included in H_2'. In that way ϕ gives rise to a transformation $\phi_0 \colon S(H_2) \to S(H_2')$ which preserves the f-divergence. Consider a unitary operator $V : H_2' \to H_2$. The transformation $V\phi_0(.)V^*$ maps $S(H_2)$ into itself and preserves the f-divergence. We have already seen that such a transformation necessarily preserves the transition probability between rank-one projections which implies that the same holds for ϕ_0 as well. Since for any two rank-one projections P, Q there exists a rank-two element $A_0 \in S(H)$ such that $\operatorname{supp} P, \operatorname{supp} Q \subset \operatorname{supp} A_0$, it follows that we have

$$\operatorname{tr} PQ = \operatorname{tr} \phi(P)\phi(Q).$$

By the non-surjective version of Wigner's theorem we infer that there is either a unitary or an antiunitary operator U on H such that

$$\phi(P) = UPU^*, \quad P \in P_1(H).$$

Define the map $\psi \colon S(H) \to S(H)$ by $\psi(A) = U^*\phi(A)U$ $(A \in S(H))$. It is clear that ψ preserves $S_f(.||.)$ and it acts as the identity on $P_1(H)$. Let $A \in S(H)$. Since ψ leaves the quantum f-divergence invariant, it preserves the inclusion between the supports of elements of $S(H)$ (see the first part of the proof in CASE II). This implies that for every rank-one projection P on H we have

$$\operatorname{supp} P \subset \operatorname{supp} A \iff \operatorname{supp} P \subset \operatorname{supp} \psi(A).$$

We easily obtain that $\operatorname{supp} A = \operatorname{supp} \psi(A)$. Let P be an arbitrary rank-one projection which satisfies $\operatorname{supp} P \subset \operatorname{supp} A = \operatorname{supp} \psi(A)$. Using (1.14) and the equality $S_f(P||\psi(A)) = S_f(P||A)$ we deduce that

$$\operatorname{tr} Pf_3(\psi(A)) = \operatorname{tr} Pf_3(A).$$

It follows that $f_3(\psi(A))$ equals A on $\operatorname{supp} A$. Using the injectivity of f_3 we can infer that $\psi(A) = A$ and next that $\phi(A) = UAU^*$. This completes the proof of the theorem. □

2

Maps on positive definite matrices preserving generalized distance measures

2.1. Introduction and statement of the results

Motivated by former results on the structure of surjective isometries of spaces of positive definite matrices obtained in the paper [**62**], in the present chapter we study so-called generalized distance measures which are parameterized by unitarily invariant norms and continuous real functions satisfying certain conditions. In the present chapter we determine the structure of all transformations on the space of all positive definite matrices that preserve not only a true metric, but a given generalized distance measure. Among the many possible applications, we emphasize that using our new result it is easy to describe the surjective maps of the set of positive definite matrices that preserve the Stein's loss or several other types of divergences. We also present results concerning similar preserver transformations defined on the subset of all complex positive definite matrices with unit determinant. The results of this chapter appeared in [**63**].

We begin with a short history of the problem we are considering in this chapter. First of all we mention that in [**62**], L. Molnár has described the structure of all surjective isometries of the space $\mathbb{P}_n$ of all $n \times n$ complex positive definite matrices with respect to any element of a large family of metrics. Those distances can be regarded as generalizations of the geodesic distance in the natural Riemannian structure on $\mathbb{P}_n$. To explain this, a few details follow. The set $\mathbb{P}_n$ is an open subset of the normed linear space $\mathbb{H}_n$ of all $n \times n$ Hermitian matrices, hence it is a differentiable manifold which can naturally be equipped with a Riemannian structure as follows. For any $A \in \mathbb{P}_n$, the tangent space at A is identified with $\mathbb{H}_n$ on which we define an inner product by

$$\langle X, Y\rangle_A = \mathrm{tr}(A^{-1/2}XA^{-1}YA^{-1/2}), \quad X, Y \in \mathbb{H}_n.$$

Clearly, the corresponding norm is

$$||X||_A = ||A^{-1/2}XA^{-1/2}||_{HS}, \quad X \in \mathbb{H}_n,$$

where $||.||_{HS}$ stands for the Hilbert-Schmidt norm (Frobenius norm) defined by $||T||_{HS}^2 = \mathrm{tr}(T^*T)$ for every element T of $\mathbb{M}_n$ the linear space of all $n \times n$ complex matrices. In that way we obtain a Riemannian space whose geometry has been investigated deeply in the literature for many reasons. It is well known that in this space the geodesic distance $\delta_R(A, B)$ between $A, B \in \mathbb{P}_n$ is

$$\delta_R(A, B) = ||\log A^{-1/2}BA^{-1/2}||_{HS}. \tag{2.1}$$

That sort of distance measure appears in a more general setting, too. In fact, in a series of papers from the 1990's Corach and his collaborators studied the cone of invertible positive elements in general C^*-algebras equipped with a Finsler-type structure, see, e.g., [**22**], [**23**], [**24**]. They explored interesting and important connections among geodesics, operator means and operator inequalities. In the particular case of matrices (i.e., when the underlying C^*-algebra is just $\mathbb{M}_n$) the structure they studied is the following. At any point $A \in \mathbb{P}_n$, on the tangent space $\mathbb{H}_n$ a Finsler-type norm is given by

$$||X||_A = ||A^{-1/2}XA^{-1/2}||, \quad X \in \mathbb{H}_n,$$

where $||.||$ stands for the usual operator norm (spectral norm). The corresponding shortest path distance on $\mathbb{P}_n$ can be computed in a way similar to (2.1) but the Hilbert-Schmidt norm is replaced by the operator norm.

Proceeding further, we mention that in the paper [**34**], Fujii presented a common extension of the above two approaches in the setting of finite dimensional C^*-algebras. For the algebra $\mathbb{M}_n$ of all $n \times n$ complex matrices this means the following. Let N be a unitarily invariant norm on $\mathbb{M}_n$. For each point $A \in \mathbb{P}_n$ and every vector $X \in \mathbb{H}_n$ define

$$N(X)_A = N(A^{-1/2}XA^{-1/2})$$

which gives a Finsler-type metric on the tangent space at A. Theorem 5 in [**34**] states that in the corresponding structure on $\mathbb{P}_n$ the shortest path distance $d_N(A, B)$ between any pair $A, B \in \mathbb{P}_n$ of points is

$$d_N(A, B) = N(\log A^{-1/2}BA^{-1/2}). \tag{2.2}$$

In [**62**] L. Molnár has described the structure of all surjective isometries of $\mathbb{P}_n$ with respect to any such metric d_N. In the same paper another structural result has also been presented concerning the isometries of $\mathbb{P}_n$ with respect to a recently defined interesting metric originating from the so-called symmetric Stein divergence. The details in short are the following. For any pair $A, B \in \mathbb{P}_n$ of positive definite matrices the Stein's loss $l(A, B)$ is defined by

$$l(A, B) = \operatorname{tr} AB^{-1} - \log\det AB^{-1} - n.$$

The Jensen-Shannon symmetrization of $l(A, B)$ is the quantity

$$S_{JS}(A, B) = \frac{1}{2}\left(l\left(A, \frac{A+B}{2}\right) + l\left(B, \frac{A+B}{2}\right)\right)$$

which is called symmetric Stein divergence. It is easy to see that we have

$$S_{JS}(A, B) = \log\det\left(\frac{A+B}{2}\right) - \frac{1}{2}\log\det AB, \quad A, B \in \mathbb{P}_n.$$

In [**79**] Sra has proven that the square root of S_{JS}, i.e.,

$$\delta_S(A, B) = \sqrt{S_{JS}(A, B)}, \quad A, B \in \mathbb{P}_n,$$

gives a true metric on $\mathbb{P}_n$. (As a matter of curiosity we mention that in [**19**] it was conjectured that δ_S not a metric, shortly after that in [**18**] the opposite was claimed, and finally, Sra has shown that δ_S is indeed a true metric on $\mathbb{P}_n$.) In [**79**] he has pointed out the importance of this new distance function. Among others, he has emphasized that δ_S is a useful substitute of the widely applied geodesic distance δ_R, it respects a non-Euclidean geometry of a rather similar kind, but, compared to the case of δ_R, the calculation of δ_S is easier, it is much less time and capacity demanding which is a really considerable advantage from the computational point of view. In [**62**] the structure of all surjective isometries of the metric space $(\mathbb{P}_n, \delta_S)$ has also been determined.

This was the short history of the former results in [**62**] and now a few sentences about the new results we are going to exhibit. First of all, our main aim here is to give a far reaching and common generalization of the above mentioned results in [**62**]. Our idea comes from the following observation. The metrics d_N, δ_S can be regarded as particular distance measures of the form

$$d_{N,f}(A, B) = N(f(A^{-1/2}BA^{-1/2})), \quad A, B \in \mathbb{P}_n, \tag{2.3}$$

where N is a unitarily invariant on $\mathbb{M}_n$ and $f\colon]0, \infty[\to \mathbb{R}$ is an appropriate real function. We emphasize that $d_{N,f}$ is not a true metric in general only a so-called generalized distance measure. By this concept in this chapter we mean a function $d : X \times X \to [0,\infty[$ (X is any set) which has the definiteness property (for arbitrary $x, y \in X$ we have $d(x,y) = 0$ if and only if $x = y$), but neither the symmetry of d nor the triangle inequality for d is assumed.

In Theorem 2.1 below we determine the structure of all surjective maps on $\mathbb{P}_n$ that leave $d_{N,f}(.,.)$ invariant. To demonstrate that our new result really extends the ones we have obtained in [**62**], observe that the metric d_N considered in [**62**] (and also defined in (2.2)) coincides with $d_{N,\log}$ defined in (2.3). As for δ_S, for any $A, B \in \mathbb{P}_n$ we have

$$\delta_S(A,B)^2 = S_{JS}(A,B) = \operatorname{tr}\left(\log\frac{Y+I}{2} - \frac{1}{2}\log Y\right) = \left\|\log\frac{Y+I}{2} - \frac{1}{2}\log Y\right\|_1$$

with $Y = A^{-1/2}BA^{-1/2}$, where $\|.\|_1$ denotes the trace-norm on $\mathbb{M}_n$. Indeed, on the one hand, observe that $(\log(Y+I)/2 - 1/2\log Y)$ is a positive semidefinite matrix for every positive definite Y and hence its trace equals its trace-norm. On the other hand, one can compute

$$\operatorname{tr}\left(\log\frac{Y+I}{2} - \frac{1}{2}\log Y\right)$$
$$= \log\det\frac{A^{-1/2}(B+A)A^{-1/2}}{2} - \frac{1}{2}\log\det B + \frac{1}{2}\log\det A$$
$$= \log\det\frac{A+B}{2} - \log\det A - \frac{1}{2}\log\det B + \frac{1}{2}\log\det A = S_{JS}(A,B).$$

We now present our main result which is a far reaching generalization of the mentioned structural theorems obtained in [**62**].

THEOREM 2.1. (Molnár, Szokol [**63**])
Let N be a unitarily invariant norm on $\mathbb{M}_n$. Assume $f\colon]0,\infty[\to \mathbb{R}$ is a continuous function such that

(a1) $f(y) = 0$ *holds if and only if* $y = 1$*;*
(a2) *there exists a number* $K > 1$ *such that*

$$|f(y^2)| \geq K|f(y)|, \quad y \in]0,\infty[.$$

Define, as above, $d_{N,f}\colon \mathbb{P}_n \times \mathbb{P}_n \to [0,\infty[$ *by*

$$d_{N,f}(A,B) = N(f(A^{-1/2}BA^{-1/2})), \quad A, B \in \mathbb{P}_n.$$

Assume that $n \geq 3$*. If* $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ *is a surjective map which leaves* $d_{N,f}(.,.)$ *invariant, i.e., which satisfies*

$$d_{N,f}(\phi(A), \phi(B)) = d_{N,f}(A, B), \quad A, B \in \mathbb{P}_n, \tag{2.4}$$

then there exist an invertible matrix $T \in \mathbb{M}_n$ *and a real number* c *such that* ϕ *is of one of the following forms*

(f1) $\phi(A) = (\det A)^c TAT^*, \quad A \in \mathbb{P}_n$;
(f2) $\phi(A) = (\det A)^c TA^{-1}T^*, \quad A \in \mathbb{P}_n$;
(f3) $\phi(A) = (\det A)^c TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n$;
(f4) $\phi(A) = (\det A)^c T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n$.

Apparently, the function $d_{N,f}(.,.)$ appearing in the theorem is a generalized distance measure in the sense we introduced above.

As we have already observed, the function f in (2.4) which corresponds to the metric d_N in (2.2) is the logarithmic function while the function f corresponding to S_{JS} is the one defined by $f(y) = \log((y+1)/(2\sqrt{y}))$, $y > 0$. It is easy to check that both functions have the properties (a1), (a2) listed in the theorem (the constant K being 2 in both cases).

In what follows we point out that Theorem 2.1 applies for many other generalized distance measures. First of all, we mention the Stein's loss. One can easily see that for any $A, B \in \mathbb{P}_n$ we have

$$l(A, B) = \mathrm{tr}(Y^{-1} - \log Y^{-1} - I) = \|Y^{-1} - \log Y^{-1} - I\|_1,$$

where $Y = A^{-1/2}BA^{-1/2}$. The latter equality follows from the fact that the matrix $Y^{-1} - \log Y^{-1} - I$ is positive semidefinite for every positive definite Y which is the consequence of the inequality $y^{-1} - \log y^{-1} - 1 \geq 0$, $y > 0$. Therefore, we can write $l(A, B) = d_{N,f}(A, B)$, where N is the trace-norm and $f(y) = y^{-1} - \log y^{-1} - 1$, $y > 0$. One can check that this function satisfies the conditions (a1), (a2) (with constant $K = 2$) in Theorem 2.1.

Beside the Jensen-Shannon symmetrization S_{JS} of the Stein's loss l appearing above, in the literature they have investigated in details the so-called Jeffrey's Kullback-Leibler divergence defined by

$$S_{JKL}(A, B) = \frac{l(A, B) + l(B, A)}{2}, \quad A, B \in \mathbb{P}_n$$

which represents the most natural symmetrization of the function l. The advantages offered by this generalized distance measure (which is not a true metric) are similar to those by S_{JS} (more precisely, by δ_S): it has many of the properties of the geodesic distance δ_R but its calculation does not require matrix eigenvalue computations, or logarithms, see [**20**]. It can be easily seen that for any $A, B \in \mathbb{P}_n$ we have

$$S_{JKL}(A,B) = \operatorname{tr}\left(\frac{Y+Y^{-1}-2I}{2}\right) = \left\|\frac{Y+Y^{-1}-2I}{2}\right\|_1,$$

where $Y = A^{-1/2}BA^{-1/2}$. Again, to see the last equality we note that the matrix $(Y+Y^{-1}-2I)/2$ is positive semidefinite for every positive definite Y. Therefore, we can write $S_{JKL}(A,B) = d_{N,f}(A,B)$, where N is the trace-norm and $f(y) = (y+y^{-1}-2)/2$, $y > 0$. Easy computations show that f satisfies the conditions (a1), (a2) (with constant $K = 2$) in Theorem 2.1.

To present further examples, we recall that in the paper [**18**] Chebbi and Moakher introduced and studied a one-parameter family of divergences which is related to the Stein's loss. For any parameter $-1 < \alpha < 1$ they defined the so-called log-determinant α-divergence D^α_{LD} by

$$D^\alpha_{LD}(A,B) = \frac{4}{1-\alpha^2}\log\frac{\det\left(\frac{1-\alpha}{2}A + \frac{1+\alpha}{2}B\right)}{(\det A)^{(1-\alpha)/2}(\det B)^{(1+\alpha)/2}}, \quad A, B \in \mathbb{P}_n.$$

For $\alpha = \pm 1$ they defined

$$D^{-1}_{LD}(A,B) = \operatorname{tr}(A^{-1}B - I) - \log\det(A^{-1}B), \quad A, B \in \mathbb{P}_n;$$
$$D^{1}_{LD}(A,B) = \operatorname{tr}(B^{-1}A - I) - \log\det(B^{-1}A), \quad A, B \in \mathbb{P}_n.$$

We clearly have

$$D^{-1}_{LD}(A,B) = l(B,A) = \|Y - \log Y - I\|_1$$

and

$$D^{1}_{LD}(A,B) = l(A,B) = \|Y^{-1} - \log Y^{-1} - I\|_1,$$

where $Y = A^{-1/2}BA^{-1/2}$. Furthermore, for $-1 < \alpha < 1$, one can easily check that

$$D^\alpha_{LD}(A,B) = \frac{4}{1-\alpha^2}\operatorname{tr}\left(\log\frac{(1-\alpha)I + (1+\alpha)Y}{2} - \frac{1+\alpha}{2}\log Y\right)$$

holds with $Y = A^{-1/2}BA^{-1/2}$. It can be shown by elementary calculus that

$$\log\left(\frac{(1-\alpha)+(1+\alpha)y}{2y^{(1+\alpha)/2}}\right) \geq 0 \tag{2.5}$$

for all $y > 0$. Therefore, the matrix

$$\log\frac{(1-\alpha)I + (1+\alpha)Y}{2} - \frac{1+\alpha}{2}\log Y$$

is positive semidefinite for any positive definite Y and we obtain that D_{LD}^{α} can be written as $D_{LD}^{\alpha} = d_{N,f}$, where N is the trace-norm and f is the function of the real variable y that appears in (2.5). It is not difficult to check that this f also satisfies the conditions (a1), (a2) (again, with constant $K = 2$). To sum up, above we have shown that the field of possible applications of Theorem 2.1 is really large, a number of generalized distance measures fulfill its assumptions.

Also relating to the applications of our main theorem, we must point out that in the particular choices of the unitarily invariant norm N and real function f, after the use of Theorem 2.1 one may need to make further steps in order to determine the precise structure of particular distance measure preservers. In accordance with this we present the complete structural result for the measures we have discussed above.

THEOREM 2.2. (Molnár, Szokol [**63**])
Let $div(.,.)$ denote any of the functions $l(.,.)$, $D_{LD}^{\alpha}(.,.)$, $-1 < \alpha < 1$. A surjective map $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ preserves $div(.,.)$, i.e., satisfies

$$div(\phi(A), \phi(B)) = div(A, B), \quad A, B \in \mathbb{P}_n,$$

if and only if there exists an invertible matrix $T \in \mathbb{M}_n$ such that ϕ is of one of the forms

$$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n.$$

A surjective map $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ preserves $S_{JKL}(.,.)$, if and only if there exists an invertible matrix $T \in \mathbb{M}_n$ such that ϕ is of one of the forms

$$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = TA^{-1}T^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n.$$

The proof follows from Theorem 2.1 and from rather simple calculations hence we shall not present it.

In connection with the problem of defining the geometric mean of a finite collection of positive definite matrices, in [**53**] Moakher studied the submanifold $\mathbb{P}_n^1$ of $\mathbb{P}_n$ which consists of all $n \times n$ positive definite matrices with determinant 1. Moreover, in the paper [**32**] the authors examined the same structure for its interesting connections to the space of so-called diffusion tensors. In fact, they also studied the set $\mathbb{P}_n^c$ of all positive definite matrices with constant determinant c which, for any positive c, is a so-called totally geodesic submanifold of $\mathbb{P}_n$. These facts motivate us to complete our main result by describing the corresponding generalized distance measure preservers also on $\mathbb{P}_n^c$. In fact, following the approach given in [**62**] we first determine the structure of all continuous Jordan triple endomorphisms of $\mathbb{P}_n^1$ (i.e., continuous maps respecting the Jordan triple product ABA). Finally, in our last result we shall describe the structure of all surjective transformations on $\mathbb{P}_n^1$ which leave a given generalized distance measure $d_{N,f}$ invariant.

THEOREM 2.3. (Molnár, Szokol [**63**])
Assume $n \geq 3$. Let $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ be a continuous map which is a Jordan triple endomorphism, i.e., ϕ is a continuous map which satisfies

$$\phi(ABA) = \phi(A)\phi(B)\phi(A), \quad A, B \in \mathbb{P}_n^1.$$

Then there is a unitary matrix $U \in \mathbb{M}_n$ such that ϕ is of one of the following forms

(g1) $\phi(A) = UAU^*, \quad A \in \mathbb{P}_n^1$;
(g2) $\phi(A) = UA^{-1}U^*, \quad A \in \mathbb{P}_n^1$;
(g3) $\phi(A) = UA^{\mathrm{tr}}U^*, \quad A \in \mathbb{P}_n^1$;
(g4) $\phi(A) = U(A^{\mathrm{tr}})^{-1}U^*, \quad A \in \mathbb{P}_n^1$;
(g5) $\phi(A) = I, \quad A \in \mathbb{P}_n^1$.

The theorem immediately gives us the following structural result on the continuous Jordan triple automorphisms of $\mathbb{P}_n^1$.

COROLLARY 2.4. *Assume $n \geq 3$. Let $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ be a continuous Jordan triple automorphism, i.e., a continuous bijective map which satisfies*

$$\phi(ABA) = \phi(A)\phi(B)\phi(A), \quad A, B \in \mathbb{P}_n^1.$$

Then ϕ is of one of the forms (g1)-(g4).

Our result on the form of surjective transformations of $\mathbb{P}_n^1$ leaving a generalized distance measure $d_{N,f}$ invariant reads as follows.

THEOREM 2.5. (Molnár, Szokol [**63**])
Let N be a unitarily invariant norm on $\mathbb{M}_n$ and $f\colon]0,\infty[\to \mathbb{R}$ be a continuous function which satisfies the conditions (a1), (a2). Assume that $n \geq 3$. Let $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ be a surjective map which preserves $d_{N,f}(.,.)$, i.e., which satisfies

$$d_{N,f}(\phi(A), \phi(B)) = d_{N,f}(A, B), \quad A, B \in \mathbb{P}_n^1.$$

Then there exists an invertible matrix T with $|\det T| = 1$ such that ϕ is of one of the following forms

(h1) $\phi(A) = TAT^*, \quad A \in \mathbb{P}_n^1$;

(h2) $\phi(A) = TA^{-1}T^*, \quad A \in \mathbb{P}_n^1$;
(h3) $\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n^1$;
(h4) $\phi(A) = T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n^1$.

Using this theorem one can easily obtain the structure of $d_{N,f}$-preserving surjective maps of the spaces $\mathbb{P}_n^c$ as follows. Observe that for any $d_{N,f}$-preserving surjective map ϕ of $\mathbb{P}_n^c$ and for the number $\lambda = \sqrt[n]{c}$, the transformation ψ defined by $\psi(A) = (1/\lambda)\phi(\lambda A)$, $A \in \mathbb{P}_n^1$ is a $d_{N,f}$-preserving surjective map of $\mathbb{P}_n^1$. Hence, Theorem 2.5 applies and we have the following corollary.

COROLLARY 2.6. *Let N, f be as in the previous theorem and assume $n \geq 3$ and c is a positive real number. If $\phi\colon \mathbb{P}_n^c \to \mathbb{P}_n^c$ is a surjective map which preserves $d_{N,f}(.,.)$, i.e., which satisfies*

$$d_{N,f}(\phi(A), \phi(B)) = d_{N,f}(A, B), \quad A, B \in \mathbb{P}_n^c,$$

then there exists an invertible matrix T with $|\det T| = 1$ such that ϕ is of one of the following forms

$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n^c$;
$\phi(A) = \lambda^2 TA^{-1}T^*, \quad A \in \mathbb{P}_n^c$;
$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n^c$;
$\phi(A) = \lambda^2 T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n^c$,

where $\lambda = \sqrt[n]{c}$.

2.2. Proofs

In this section we present the proofs of our results. We begin with some auxiliary statements. The most important one among them, Proposition 2.11, shows that on certain substructures of groups surjective transformations that preserve a given generalized distance measure d which is compatible with the group operation, necessarily preserve locally the so-called inverted Jordan triple product (i.e., they respect the operation $xy^{-1}x$). We emphasize that results of this kind, which can be considered as noncommutative versions of the famous Mazur-Ulam theorem (see Theorem 0.3), first appeared in the paper [**38**]. In fact, below we closely follow the approach presented in Sections 2 and 3 of that paper but here we have to make several small modifications according to our present need.

In what follows, after a simple definition we shall exhibit statements that are similar to Lemma 2.3 and Theorem 2.4 in [**38**] and then we shall introduce conditions similar to the ones $B(.,.)$ and $C(.,.)$ in Definitions 3.2 and 3.4 in that paper. Finally, we shall obtain Proposition 2.11, a statement similar to Corollary 3.10 in [**38**] which is the basic tool in the proof of our main result.

DEFINITION 2.7. *Let X be a set and $d : X \times X \to [0,\infty[$ be any function. We say that a map $\varphi \colon X \to X$ is d-preserving if*

$$d(\varphi(x), \varphi(y)) = d(x, y)$$

holds for any $x, y \in X$. We say that φ is d-reversing if

$$d(\varphi(x), \varphi(y)) = d(y, x)$$

holds for any $x, y \in X$.

LEMMA 2.8. *Let X be a set and $d : X \times X \to [0,\infty[$ be an arbitrary function. Assume $\varphi : X \to X$ is a bijective d-reversing map, $b \in X$, and $K > 1$ is a constant such that*

$$d(x, \varphi(x)) \geq K d(x, b), \quad x \in X.$$

If $\sup\{d(x, b) | x \in X\} < \infty$, then for every bijective d-reversing map $f \colon X \to X$ we have $d(f(b), b) = 0$.

PROOF. Let

$$\lambda = \sup\{d(f(b), b) | f \colon X \to X \text{ is a bijective } d\text{-reversing map}\}.$$

Then $0 \leq \lambda < \infty$. For an arbitrary bijective d-reversing map $f \colon X \to X$, consider $\tilde{f} = f^{-1} \circ \varphi \circ f$. Then $\tilde{f}$ is also a bijective d-reversing transformation and

$$\lambda \geq d(\tilde{f}(b), b) = d(f(b), \varphi(f(b)) \geq K d(f(b), b).$$

By the definition of λ we get $\lambda \geq K\lambda$ which implies that $\lambda = 0$ and this completes the proof. □

PROPOSITION 2.9. *Let X be a set, $d : X \times X \to [0,\infty[$ be any function. Let $a, b \in X$ and assume that $\varphi \colon X \to X$ is a bijective d-reversing map such that $\varphi(b) = b$ and $\varphi \circ \varphi$ is the identity on X. We set*

$$L = \{x \in X | d(a, x) = d(x, \varphi(a)) = d(a, b)\}.$$

Suppose that $\sup\{d(x, b) | x \in L\} < \infty$ and there exists a constant $K > 1$ such that

$$d(x, \varphi(x)) \geq K d(x, b), \quad x \in L.$$

If $T \colon X \to X$ is a bijective d-preserving map, $\psi \colon X \to X$ is a bijective d-reversing map, moreover $\psi(T(a)) = T(\varphi(a))$ and $\psi(T(\varphi(a))) = T(a)$ hold, then we have

$$d(\psi(T(b)), T(b)) = 0.$$

PROOF. Since $\varphi(b) = b$ and φ is a d-reversing map, we have

$$d(a, b) = d(\varphi(b), \varphi(a)) = d(b, \varphi(a)),$$

which implies that $b \in L$. Let

$$L' = \{y \in X | d(T(a), y) = d(y, T(\varphi(a))) = d(a, b)\}.$$

Using the bijectivity and the d-preserving property of T one can easily check that $T(L) = L'$. Furthermore, in a similar way, by the bijectivity and the d-reversing property of the maps φ, ψ we obtain that $\varphi(L) = L$ and $\psi(L) = L$. Consider now the transformation $\tilde{T} = T^{-1} \circ \psi \circ T$. Plainly, the restrictions of the maps $\tilde{T}$ and φ to L are bijective d-reversing maps of L. Since $\sup\{d(x, b) | x \in L\} < \infty$, applying the previous lemma we deduce that

$$0 = d(\tilde{T}(b), b) = d(\psi(T(b)), T(b)).$$

□

In the following we need some notions. Let G be a group. The operation $(x, y) \mapsto xy^{-1}x$ is called inverted Jordan triple product. A non-empty subset X of G is called a twisted subgroup if it is closed under that operation, i.e. $xy^{-1}x \in X$ holds for every pair $x, y \in X$. We say that X is 2-divisible if for each $a \in X$ the equation $x^2 = a$ has a solution $x \in X$. We say that X is 2-torsion free if the unit element e of G belongs to X and the equality $x^2 = e$ implies $x = e$.

We shall need the following technical lemma. We remark that its proof has appeared as a part of the proof of Corollary 3.10 in [**38**].

LEMMA 2.10. *Let X be a twisted subgroup of a group which is 2-divisible and 2-torsion free and let $c \in X$. The only solution $x \in X$ of the equation $cx^{-1}c = x$ is $x = c$.*

PROOF. Since X is 2-divisible there exists an element $g \in X$ such that $g^2 = c$. From $g^2x^{-1}g^2 = x$ it follows that $g^2x^{-1}g^2x^{-1} = e$ and then multiplying by g^{-1} from the left and by g from the right, we have

$$e = gx^{-1}g^2x^{-1}g = (gx^{-1}g)^2.$$

By the 2-torsion free property of X we deduce that $gx^{-1}g = e$. This implies $x^{-1} = g^{-2}$ and hence $x = g^2 = c$. □

We next introduce some conditions for a pair a, b of elements that belong to a twisted subgroup of a group. We shall use them in the next proposition.

Let X be a twisted subgroup of a group G, let $d : X \times X \to [0, \infty[$ be any function and pick $a, b \in X$. We say that the pair a, b satisfies the condition

(b1) if the equality

$$d(bx^{-1}b, by^{-1}b) = d(y, x)$$

holds for any $x, y \in X$;

(b2) if $\sup\{d(x, b)|x \in L_{a,b}\} < \infty$, where

$$L_{a,b} = \{x \in X|d(a, x) = d(x, ba^{-1}b) = d(a, b)\};$$

(b3) if there exists a constant $K > 1$ such that

$$d(x, bx^{-1}b) \geq Kd(x, b), \quad x \in L_{a,b};$$

(b4) if there exists an element $c \in X$ with $ca^{-1}c = b$ such that

$$d(cx^{-1}c, cy^{-1}c) = d(y, x)$$

holds for any $x, y \in X$.

Now we present our basic tool in the proof of Theorem 2.1.

PROPOSITION 2.11. *Let G be a group and $X \subset G$ a twisted subgroup which is 2-divisible and 2-torsion free. Assume that the function $d : X \times X \to [0, \infty[$ is a generalized distance measure, i.e., it has the property that for any $x, y \in X$ we have $d(x, y) = 0$ if and only if $x = y$. Let $T \colon X \to X$ be a surjective d-preserving map. Pick $a, b \in X$ such that the pair a, b satisfies the conditions (b1)-(b3) and the pair $T(a), T(ba^{-1}b)$ satisfies the condition (b4). Then we have*

$$T(ba^{-1}b) = T(b)T(a)^{-1}T(b).$$

PROOF. First observe that any d-preserving function is automatically injective. Let $\varphi(x) = bx^{-1}b$ for every $x \in X$. Then φ is a bijective d-reversing map on X and it satisfies the conditions appearing in Proposition 2.9, i.e., it fixes b and $\varphi \circ \varphi$ is the identity. Since (b4) holds for the pair $T(a), T(ba^{-1}b)$, there exists an element $c \in X$ such that

$$cT(a)^{-1}c = T(ba^{-1}b) \tag{2.6}$$

and $d(cx^{-1}c, cy^{-1}c) = d(y, x)$ holds for all $x, y \in X$. Let the map $\psi \colon X \to X$ be defined by $\psi(x) = cx^{-1}c$ for every $x \in X$. Clearly, ψ is a bijective d-reversing map on X and by (2.6) we have that $\psi(T(a)) = T(\varphi(a))$ and also that $\psi(T(\varphi(a))) = T(a)$ holds. Now we are in a position to apply Proposition 2.9 and we get that $d(\psi(T(b)), T(b)) = 0$ which implies $T(b) = cT(b)^{-1}c$. Using Lemma 2.10 we infer that $c = T(b)$. Finally, by (2.6) we obtain

$$T(ba^{-1}b) = T(b)T(a)^{-1}T(b).$$

□

After these preliminaries we can present the proof of Theorem 2.1.

PROOF OF THEOREM 2.1. Let N, f be as in the formulation of theorem and let $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ be a surjective map which preserves the generalized distance measure $d_{N,f}(.,.)$, i.e., assume

$$d_{N,f}(\phi(A), \phi(B)) = d_{N,f}(A, B), \quad A, B \in \mathbb{P}_n.$$

We are going to apply Proposition 2.11. To do this, we show that all conditions appearing there are satisfied for $\mathbb{P}_n$ and for any pair A, B of its elements.

First, $X = \mathbb{P}_n$ is a twisted subgroup of the group of all invertible matrices which is clearly 2-divisible and 2-torsion free. Next, we assert that the equalities

$$\begin{aligned} d_{N,f}(A^{-1}, B^{-1}) &= d_{N,f}(B, A); \\ d_{N,f}(TAT^*, TBT^*) &= d_{N,f}(A, B) \end{aligned} \tag{2.7}$$

hold for all $A, B \in \mathbb{P}_n$ and invertible matrix $T \in \mathbb{M}_n$. Indeed, let $A, B \in \mathbb{P}_n$ and consider the polar decomposition $B^{-1/2}A^{1/2} = U|B^{-1/2}A^{1/2}|$. We see that $|A^{1/2}B^{-1/2}|^2 = U|B^{-1/2}A^{1/2}|^2U^*$ and then compute

$$\begin{aligned} d_{N,f}(A^{-1}, B^{-1}) &= N(f(A^{1/2}B^{-1}A^{1/2})) = N(f(|B^{-1/2}A^{1/2}|^2)) \\ &= N(f(U^*|A^{1/2}B^{-1/2}|^2U)) = N(U^*f(|A^{1/2}B^{-1/2}|^2)U) \\ &= N(f(B^{-1/2}AB^{-1/2})) = d_{N,f}(B, A). \end{aligned}$$

Now, for an arbitrary invertible matrix $T \in \mathbb{M}_n$ we deduce

$$\begin{aligned} &((TAT^*)^{-1/2}TBT^*(TAT^*)^{-1/2})^2 \\ &= (TAT^*)^{-1/2}TBT^*(TAT^*)^{-1}TBT^*(TAT^*)^{-1/2}. \end{aligned}$$

For $X = A^{-1/2}BT^*(TAT^*)^{-1/2}$ we have

$$XX^* = A^{-1/2}BA^{-1}BA^{-1/2} = (A^{-1/2}BA^{-1/2})^2.$$

Hence, using the polar decomposition $X = V|X|$, we compute

$$\begin{aligned}&(TAT^*)^{-1/2}TBT^*(TAT^*)^{-1/2}\\ &= ((TAT^*)^{-1/2}TBT^*(TAT^*)^{-1}TBT^*(TAT^*)^{-1/2})^{1/2}\\ &= ((TAT^*)^{-1/2}TBA^{-1}BT^*(TAT^*)^{-1/2})^{1/2}\\ &= (X^*X)^{1/2} = |X| = V^*|X^*|V = V^*(A^{-1/2}BA^{-1/2})V.\end{aligned}$$

It readily follows that $d_{N,f}(TAT^*, TBT^*) = d_{N,f}(A, B)$ holds for any $A, B \in \mathbb{P}_n$ completing the proof of (2.7).

Let us now select two arbitrary elements A, B of $\mathbb{P}_n$. By (2.7), the condition (b1) is satisfied for the pair A, B. As for condition (b2), let us consider the set $\mathcal{H}$ of those elements $X \in \mathbb{P}_n$ for which we have

$$\begin{aligned}d_{N,f}(A, X) &= N(f(A^{-1/2}XA^{-1/2}))\\ &= N(f(A^{-1/2}BA^{-1/2})) = d_{N,f}(A, B).\end{aligned}$$

(Clearly, $L_{A,B} \subset \mathcal{H}$.) We show that the corresponding set of numbers

$$\begin{aligned}N(f(X^{-1/2}BX^{-1/2})) &= d_{N,f}(X, B)\\ = d_{N,f}(B^{-1}, X^{-1}) &= N(f(B^{1/2}X^{-1}B^{1/2}))\end{aligned}$$

is bounded. Indeed, since $N(f(A^{-1/2}XA^{-1/2}))$ is constant on $\mathcal{H}$ and N is equivalent to the operator norm $||.||$, the set

$$\{\|f(A^{-1/2}XA^{-1/2})\| \colon X \in \mathcal{H}\}$$

is bounded. It is easy to see that (a1), (a2) imply

$$\lim_{y\to 0} f(y), \lim_{y\to\infty} f(y) \in \{-\infty, \infty\}.$$

Then it follows easily that there are positive numbers m, M such that $mI \leq A^{-1/2}XA^{-1/2} \leq MI$ holds for all $X \in \mathcal{H}$. Clearly, we then have another pair m', M' of positive numbers such that $m'I \leq X \leq M'I$ and finally another one m'', M'' such that $m''I \leq B^{1/2}X^{-1}B^{1/2} \leq M''I$ holds for all $X \in \mathcal{H}$. By continuity, f is bounded on the interval $[m'', M'']$ and this implies that the set

$$\{N(f(B^{1/2}X^{-1}B^{1/2})) \colon X \in \mathcal{H}\}$$

is bounded. We conclude that the condition (b2) is fulfilled.

Relating to condition (b3) we assert that $N(f(C^2)) \geq KN(f(C))$ holds for every $C \in \mathbb{P}_n$. To see this, we recall the famous fact that any unitarily invariant norm on $\mathbb{M}_n$ is induced by some symmetric gauge function on $\mathbb{R}^n$. By a well-known result of Ky Fan [**27**], for given finite sequences $0 \le a_n \le \ldots \le a_1$ and $0 \le b_n \le \ldots \le b_1$ of numbers we have $\Phi(a_1,\ldots,a_n) \le \Phi(b_1,\ldots,b_n)$ for all symmetric gauge functions Φ on $\mathbb{R}^n$ if and only if the inequality

$$\sum_{i=1}^{k} a_k \le \sum_{i=1}^{k} b_k$$

holds for every $1 \le k \le n$. By (a2) it then follows that

$$\Psi\left(|f(\lambda_1^2)|,\ldots,|f(\lambda_n^2)|\right) \ge K\Psi\left(|f(\lambda_1)|,\ldots,|f(\lambda_n)|\right), \tag{2.8}$$

where Ψ is the symmetric gauge function corresponding to N and $\lambda_1,\ldots,\lambda_n$ are the eigenvalues of an arbitrary positive definite matrix $C \in \mathbb{P}_n$. Consequently, we obtain the desired inequality $N(f(C^2)) \ge KN(f(C))$.

Next, selecting any $X \in \mathbb{P}_n$ and setting $Y = X^{-1/2}BX^{-1/2}$, we easily deduce that

$$d_{N,f}(X, BX^{-1}B) = N(f(X^{-1/2}BX^{-1}BX^{-1/2}))$$

$$= N(f(Y^2)) \ge KN(f(Y)) = KN(f(X^{-1/2}BX^{-1/2})) = Kd_{N,f}(X,B).$$

Therefore, the condition (b3) is also satisfied. Consequently, all assumptions (b1)-(b3) are fulfilled for any pair $A, B \in \mathbb{P}_n$.

We assert that the same holds in relation with condition (b4), too. To see this, observe that for any pair $A, B \in \mathbb{P}_n$ we can find $C \in \mathbb{P}_n$ such that $CA^{-1}C = B$. Indeed, the geometric mean of A and B, that is, the positive definite matrix $C = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$ is a solution of that equation. The remaining invariance property of $d_{N,f}$ in (b4) has already been verified in (2.7).

Taking all the information what we have into account, we can now apply Proposition 2.11 and obtain that $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ is a bijective map which satisfies

$$\phi(BA^{-1}B) = \phi(B)\phi(A)^{-1}\phi(B)$$

for all $A, B \in \mathbb{P}_n$. We prefer to write

$$\phi(AB^{-1}A) = \phi(A)\phi(B)^{-1}\phi(A), \quad A, B \in \mathbb{P}_n.$$

Consider the transformation $\psi\colon \mathbb{P}_n \to \mathbb{P}_n$ defined by

$$\psi(A) = \phi(I)^{-1/2}\phi(A)\phi(I)^{-1/2}, \quad A \in \mathbb{P}_n.$$

It is easy to see that ψ is a bijective map on $\mathbb{P}_n$ which satisfies

$$\psi(AB^{-1}A) = \psi(A)\psi(B)^{-1}\psi(A), \quad A, B \in \mathbb{P}_n$$

and has the additional property that $\psi(I) = I$. Substituting $A = I$ in the above displayed equation we obtain that $\psi(B^{-1}) = \psi(B)^{-1}$ which implies that ψ is a Jordan triple automorphism of $\mathbb{P}_n$, i.e., a bijective map satisfying

$$\psi(ABA) = \psi(A)\psi(B)\psi(A), \quad A, B \in \mathbb{P}_n.$$

We next prove that ψ is continuous in the operator norm. Clearly, ψ preserves $d_{N,f}(.,.)$ which is a consequence of the second invariance property in (2.7). Let (X_n) be a sequence in $\mathbb{P}_n$ which tends to $X \in \mathbb{P}_n$ with respect to the operator norm topology. Then $X^{-1/2}X_nX^{-1/2} \to I$, and hence

$$d_{N,f}(X, X_n) = N(f(X^{-1/2}X_nX^{-1/2})) \to N(f(I)) = 0.$$

Since ψ preserves the generalized distance measure $d_{N,f}(.,.)$, we infer that

$$N(f(\psi(X)^{-1/2}\psi(X_n)\psi(X)^{-1/2})) = d_{N,f}(\psi(X), \psi(X_n)) \to 0.$$

It follows that $f(\psi(X)^{-1/2}\psi(X_n)\psi(X)^{-1/2}) \to 0$ in the operator norm. By the continuity of f and the property (a1), it is easy to verify that we necessarily have

$$\psi(X)^{-1/2}\psi(X_n)\psi(X)^{-1/2} \to I,$$

i.e., $\psi(X_n) \to \psi(X)$ in the operator norm and we obtain the continuity of ψ.

The structure of continuous Jordan triple automorphisms of $\mathbb{P}_n$ has been determined in [**62**]. Applying Corollary 2 in that paper we have a unitary matrix U and a scalar $c \neq -1/n$ such that ψ is of one of the forms

(i) $\psi(A) = (\det A)^c UAU^*, \quad A \in \mathbb{P}_n$;
(ii) $\psi(A) = (\det A)^c UA^{-1}U^*, \quad A \in \mathbb{P}_n$;
(iii) $\psi(A) = (\det A)^c UA^{\mathrm{tr}}U^*, \quad A \in \mathbb{P}_n$;
(iv) $\psi(A) = (\det A)^c U(A^{\mathrm{tr}})^{-1}U^*, \quad A \in \mathbb{P}_n$.

By the definition of the transformation ψ we get that ϕ is necessarily of one of the forms (f1)-(f4) and the proof of the theorem is complete. □

As mentioned before, when trying to determine the precise structure of bijective maps of $\mathbb{P}_n$ preserving a generalized distance measure with particular N and f, one should not stop at applying Theorem 2.1 but proceed further and check which ones of the possibilities (f1)-(f4) and for which parameters c and T give transformations that really have the desired preserver property (2.4). In fact, as for T, we can tell that for any invertible matrix $T \in \mathbb{M}_n$ the map $A \mapsto TAT^*$ satisfies (2.4). This follows from the second equality in (2.7). Concerning the inverse operation $A \mapsto A^{-1}$, there are cases where it does not show up. In fact, by the first equality in (2.7) that map is $d_{N,f}$-reversing, hence when $d_{N,f}$ is not symmetric, the inverse is surely not $d_{N,f}$-preserving. For example, this is the case with the Stein's loss $l(.,.)$. However, the transpose is always $d_{N,f}$-preserving. Indeed, it follows from the facts that the transpose operation commutes with the inverse operation, with the square root, with the map $A \mapsto f(A)$, and furthermore $N(C^{\mathrm{tr}}) = N(C)$ holds for every self-adjoint matrix C. For the above reasons, the map $A \mapsto (A^{\mathrm{tr}})^{-1}$ sometimes shows up, sometimes does not. This is the case with the determinant function too as can be seen, for example, in Theorem 3 in [**62**].

We now turn to the proof of Theorem 2.3. The structure of continuous Jordan triple endomorphisms of $\mathbb{P}_n$ has been described in Theorem 1 in [**62**]. We are going to apply that result in the proof below.

PROOF OF THE THEOREM 2.3. Let $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ be a continuous Jordan triple endomorphism of $\mathbb{P}_n^1$, i.e., a continuous map which satisfies

$$\phi(ABA) = \phi(A)\phi(B)\phi(A), \quad A, B \in \mathbb{P}_n^1.$$

Consider the transformation $\psi\colon \mathbb{P}_n \to \mathbb{P}_n$ defined by

$$\psi(A) = \sqrt[n]{\det(A)}\phi\left(\frac{A}{\sqrt[n]{\det(A)}}\right), \quad A \in \mathbb{P}_n.$$

One can check trivially that ψ is a Jordan triple endomorphism of $\mathbb{P}_n$ which extends ϕ. Applying Theorem 1 in [**62**], it follows that there exist a unitary matrix $U \in \mathbb{M}_n$ and a real number c such that ψ is of one of the forms (i)-(iv) appearing at the end of the previous proof, or there exist a set $\{P_1, \dots, P_n\}$ of mutually orthogonal rank-one projections in $\mathbb{M}_n$ and a set $\{c_1, \dots, c_n\}$ of real numbers such that ψ is of the form

(v) $\psi(A) = \sum_{i=1}^n (\det A)^{c_i} P_i, \quad A \in \mathbb{P}_n.$

Clearly, in this latter case ψ sends matrices with unit determinant to the identity. This implies that ϕ is really of one of the forms (g1)-(g5). The proof of the theorem is complete. □

In what remains we present the key steps of the proof of Theorem 2.5. In fact, we use an approach very similar to the one we followed in the proof of Theorem 2.1 above hence the details are omitted.

SKETCH OF THE PROOF OF THEOREM 2.5. Let $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ be a surjective map which preserves the generalized distance measure $d_{N,f}(.,.)$, i.e., which satisfies

$$d_{N,f}(\phi(A), \phi(B)) = d_{N,f}(A, B), \quad A, B \in \mathbb{P}_n^1.$$

We claim that all conditions appearing in Proposition 2.11 are satisfied.

Clearly, the set $\mathbb{P}_n^1$ is a 2-divisible and 2-torsion free twisted subgroup of the group of all invertible matrices. Referring back to the proof of Theorem 2.1, the invariance properties (2.7) of $d_{N,f}$ hold true on the set $\mathbb{P}_n^1$, too. Similarly, the conditions (b1)-(b4) are satisfied for every pair A, B of elements of the subset $\mathbb{P}_n^1$ of $\mathbb{P}_n$. This means that we can apply Proposition 2.11 and we then obtain that ϕ is an inverted Jordan triple automorphism of $\mathbb{P}_n^1$, i.e.,

$$\phi(AB^{-1}A) = \phi(A)\phi(B)^{-1}\phi(A), \quad A, B \in \mathbb{P}_n^1.$$

Next, we consider the transformation $\psi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ defined by

$$\psi(A) = \phi(I)^{-1/2}\phi(A)\phi(I)^{-1/2}, \quad A \in \mathbb{P}_n^1.$$

It turns to be a Jordan triple automorphism of $\mathbb{P}_n^1$ which also preserves the generalized distance measure $d_{N,f}$. Following the argument presented in the proof of Theorem 2.1 gives that ψ is continuous. Therefore, by Corollary 2.4, we get that ψ is of one of the forms (g1)-(g4). Finally we conclude that ϕ is of one of the forms (h1)-(h4) and this completes the proof. $\square$

Finally, we note that in several applications the set of symmetric positive definite real matrices plays more important role than that of the positive definite complex matrices. See, e.g., [**53**] and [**32**]. In accordance with this, we remark that the main results of this chapter, Theorem 2.1 and Theorem 2.5, along with Theorem 2.2 and Corollary 2.4, Corollary 2.6 remain valid also in the real case. Indeed, a careful examination of our arguments above shows that all steps in the proofs can be unaltered, the only thing we really need to deal with is the structure of all continuous Jordan triple automorphisms of the set of all $n \times n$ symmetric positive definite real matrices ($n \geq 3$). In the complex case, those transformations have been described by Molnár in Corollary 2 in [**62**]. In the real case, we can follow steps similar to the ones given in the proofs of Lemmas 5-7 and Theorem 1 in that paper. In fact, the mentioned lemmas can be shown in the same way as in [**62**] (the proof of Lemma 7 is given in [**54**]), but as for Theorem 1 and its Corollary 2 we need to use the result of Chan and Lim which describes the structure of all bijective commutativity preserving linear maps on the space of $n \times n$ symmetric real matrices [**17**]. Apparently, this means that in the real case we have a structural result only for Jordan triple automorphisms and not for all continuous Jordan triple endomorphisms.

3

Surjective isometries of the space of all generalized distribution functions

3.1. Introduction and statement of the results

In this chapter we determine the structure of surjective isometries of the space of all (continuous) generalized probability distribution functions with respect to the Kolmogorov-Smirnov metric. The corresponding preserver results can be found in [**61**].

Motivated by the famous Banach-Stone theorem, in the paper [**26**] G. Dolinar and L. Molnár have described the general forms of surjective isometries of the space of all probability distribution functions with respect to the Kolmogorov-Smirnov metric. By a distribution function here we mean a mapping $d : \mathbb{R} \to \mathbb{R}$ which is monotone increasing, continuous from the right, and has limit 0 at $-\infty$ and 1 at ∞. We note that the set of all such functions plays an important and fundamental role in probability theory and statistics. Throughout the present chapter this set will be denoted by $D(\mathbb{R})$. For any two elements f and g of $D(\mathbb{R})$ the Kolmogorov-Smirnov distance between them is defined by the formula

$$\rho(f, g) = \sup_{t \in \mathbb{R}} |f(t) - g(t)|,$$

which shows that it originates from the uniform norm on the Banach space of all real-valued bounded functions of a real variable. We note that the significance of this metric lies in its applications in the Kolmogorov-Smirnov statistics and test.

As mentioned before in [**26**] the authors have determined all surjective isometries of $D(\mathbb{R})$ with respect to Kolmogorov-Smirnov metric. It turned out that similarly to the conclusion in Banach-Stone theorem the

Kolmogorov-Smirnov isometries are closely related to composition operators. This result reads as follows.

THEOREM 3.1. (Dolinar, Molnár [**26**])
Let $\phi : D(\mathbb{R}) \to D(\mathbb{R})$ be a surjective isometry with respect to the Kolmogorov-Smirnov metric, i.e. assume that ϕ is a bijective map which satisfies that

$$\rho(\phi(f), \phi(g)) = \rho(f, g), \quad f, g \in D(\mathbb{R}).$$

Then either there exists a strictly increasing bijection $\varphi : \mathbb{R} \to \mathbb{R}$ such that ϕ is of the form

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in D(\mathbb{R})$$

or there exists a strictly decreasing bijection $\psi : \mathbb{R} \to \mathbb{R}$ such that ϕ is of the form

$$\phi(f)(t) = 1 - f(\psi(t)-), \quad t \in \mathbb{R}, f \in D(\mathbb{R}).$$

Here $f(t-)$ denotes the left limit of the distribution function f at the point $t \in \mathbb{R}$. As $D(\mathbb{R})$ is equipped with the metric coming from the uniform norm the result can also be viewed as a Banach-Stone type theorem for the function space $D(\mathbb{R})$. However, we emphasize that unlike in the Banach-Stone theorem, in Theorem 3.1 the linearity of the isometries in question is not assumed as the underlying space is not a linear space. The key step of the proof of Theorem 3.1 has been a metric characterization of Dirac distribution functions (i.e. distribution functions corresponding to one point mass measures).

In an other recent paper [**58**] Molnár has studied the surjective Kolmogorov-Smirnov isometries of important subspaces of $D(\mathbb{R})$, too. The main result has been the description of the structure of all surjective isometries of the space $D_c(\mathbb{R})$ of all continuous distribution functions. It reads as follows.

THEOREM 3.2. (Molnár [**58**])
Let $\phi : D_c(\mathbb{R}) \to D_c(\mathbb{R})$ be a surjective isometry with respect to the Kolmogorov-Smirnov metric. Then either there exists a strictly increasing bijection $\varphi : \mathbb{R} \to \mathbb{R}$ such that ϕ is of the form

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in D_c(\mathbb{R})$$

or there exists a strictly decreasing bijection $\psi : \mathbb{R} \to \mathbb{R}$ such that ϕ is of the form

$$\phi(f)(t) = 1 - f(\psi(t)), \quad t \in \mathbb{R}, f \in D_c(\mathbb{R}).$$

Since the Dirac distribution functions are not present in the space $D_c(\mathbb{R})$, the author has used an approach which is completely different from the argument applied in the proof of Theorem 3.1.

The aim of the present chapter is to extend Theorem 3.1 to the larger space $\Delta(\mathbb{R})$ of all generalized probability distribution functions and Theorem 3.2 to the set $\Delta_c(\mathbb{R})$ of all continuous elements of $\Delta(\mathbb{R})$. This problem has been raised by M. Barczy. By a generalized distribution function we mean a function from $\mathbb{R}$ to $[0, 1]$ which is monotone increasing and continuous from the right without restrictions on its limits at $\pm\infty$. Generalized distribution functions appear naturally when one considers random variables taking values not only in $\mathbb{R}$ but in the extended real line which setting proves useful in several investigations. In fact, some practical reasons are coming from measure theory (e.g., if one deals with extended real-valued functions, then in Beppo Levi theorem there is no need to assume the convergence of the series of non-negative measurable functions). Beside them we recall serious applications in the theory of renewal processes (see [**29**] where the word "defective" is used for random variables taking the value ∞ with positive probability and also for the relating distribution functions) or Helly's fundamental theorem on the weak sequential compactness of the space of generalized distribution functions (see, e.g., [**80**]) that plays an important role in probability theory. The Kolmogorov-Smirnov metric on $\Delta(\mathbb{R})$ are defined by the same formula as on $D(\mathbb{R})$.

Our result that follows describes the structure of surjective Kolmogorov-Smirnov isometries of the space of generalized distribution functions and shows that this structure is formally the same as that of the surjective isometries of $D(\mathbb{R})$. Namely, any surjective isometry of $\Delta(\mathbb{R})$ is induced either by a strictly increasing bijection or by a strictly decreasing bijection of $\mathbb{R}$.

THEOREM 3.3. (Molnár, Szokol [**61**])
Let $\phi\colon \Delta(\mathbb{R}) \to \Delta(\mathbb{R})$ be a surjective isometry with respect to the Kolmogorov-Smirnov metric. Then either there exists a strictly increasing bijection $\varphi\colon \mathbb{R} \to \mathbb{R}$ such that ϕ is of the form

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in \Delta(\mathbb{R})$$

or there exists a strictly decreasing bijection $\psi\colon \mathbb{R} \to \mathbb{R}$ such that ϕ is of the form

$$\phi(f)(t) = 1 - f(\psi(t)-), \quad t \in \mathbb{R}, f \in \Delta(\mathbb{R}).$$

Now we present the extension of Theorem 3.2 to the case of the surjective isometries of $\Delta_c(\mathbb{R})$ equipped with the Kolmogorov-Smirnov metric.

In the continuous case we also have that the result is formally the same as the one concerning the surjective isometries of $D_c(\mathbb{R})$.

THEOREM 3.4. (Molnár, Szokol [**61**])
Let $\phi\colon \Delta_c(\mathbb{R}) \to \Delta_c(\mathbb{R})$ *be a surjective isometry with respect to the Kolmogorov-Smirnov metric. Then either there exists a strictly increasing bijection* $\varphi\colon \mathbb{R} \to \mathbb{R}$ *such that* ϕ *is of the form*

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in \Delta_c(\mathbb{R})$$

or there exists a strictly decreasing bijection $\psi\colon \mathbb{R} \to \mathbb{R}$ *such that* ϕ *is of the form*

$$\phi(f)(t) = 1 - f(\psi(t)), \quad t \in \mathbb{R}, f \in \Delta_c(\mathbb{R}).$$

As mentioned before in the paper [**58**] results on the Kolmogorov-Smirnov isometries of the spaces of all absolute continuous or singular distribution functions have also been presented. However, since those concepts (absolute continuity and singularity) are not unambiguously defined for functions in $\Delta(\mathbb{R})$, we do not discuss the corresponding extensions of the results in [**58**] for the setting of generalized distribution functions.

3.2. Proofs

In this section we are going to present the proof of Theorem 3.3 and the sketch of the proof of Theorem 3.4.

Before the proof of Theorem 3.3, we emphasize that, as explained in [**26**], every transformation which appears on the right-hand side of the two displayed formulas appearing in Theorem 3.3 is a surjective isometry of $\Delta(\mathbb{R})$. The content of our result is, in fact, that the reverse statement is also true: every surjective isometry of $\Delta(\mathbb{R})$ is necessarily of one of those two forms.

PROOF OF THEOREM 3.3. The basic idea of our proof is to find a way to deduce the statement from Theorem 3.1. All that follow are done for that purpose.

We start with presenting a metric characterization of the constant 0 and constant 1 functions. For any real number $c \in [0,1]$ let $\underline{c}$ denote the constant c function. The closed ball with center $f \in \Delta(\mathbb{R})$ and radius $r \geq 0$ is denoted by $B_r(f)$. We now assert that for any $f, g \in \Delta(\mathbb{R})$ we have that $B_{1/2}(f) \cap B_{1/2}(g)$ is a singleton if and only if $\{f, g\} = \{\underline{0}, \underline{1}\}$. The sufficiency is obvious. To see the necessity, assume that $f, g \in \Delta(\mathbb{R})$ are such

that $B_{1/2}(f) \cap B_{1/2}(g)$ is a singleton. We distinguish some cases.
(I) If

$$|f(x) - g(x)| = 1, \quad x \in \mathbb{R},$$

then we easily get that $f = \underline{0}$ and $g = \underline{1}$, or that $f = \underline{1}$ and $g = \underline{0}$.
(II) Suppose that there exists an $x \in \mathbb{R}$ such that $|f(x) - g(x)| < 1$ but for all such x we have $f(x) = g(x)$. This means that for every real number x either we have $|f(x)-g(x)| = 1$ or we have $f(x) = g(x)$ and there do exist a real number x_0 such that the second condition holds. In what follows, by an identity point of f and g we mean a real number x such that $f(x) = g(x)$ and by an extreme point of the function $h \in \Delta(\mathbb{R})$ we mean a point $x \in \mathbb{R}$ such that $h(x) = 0$ or $h(x) = 1$ holds. Now, if every identity point of f and g is an extreme point of f and g, then in the present case we obtain that every point is an extreme point of both f and g. This implies that f (and also g) must be a Dirac distribution function, or $\underline{0}$, or $\underline{1}$. On the other hand, if there exists an identity point of f and g where the corresponding common value of the functions is in $]0, 1[$, then one can easily verify that $f = g$ which obviously contradicts the fact that $B_{1/2}(f) \cap B_{1/2}(g)$ is a singleton. We arrive at the same contradiction when both f and g are Dirac distribution functions, and also when one of them is a Dirac distribution function and the other one is either $\underline{0}$ or $\underline{1}$. It follows that $\{f, g\} \subset \{\underline{0}, \underline{1}\}$ and since $f \neq g$, the desired conclusion follows in the case (II).
(III) It remains to consider the case where there exists a real number x_0 such that

$$|f(x_0) - g(x_0)| < 1 \text{ and } f(x_0) \neq g(x_0).$$

Without loss of generality we may and do suppose that $0 < f(x_0)-g(x_0) < 1$. Obviously,

$$h(x) \doteq \frac{f(x) + g(x)}{2}, \quad x \in \mathbb{R}$$

is an element of $B_{1/2}(f) \cap B_{1/2}(g)$. We are going to show that this function can be modified on a short interval such that the so obtained function also belongs to $B_{1/2}(f) \cap B_{1/2}(g)$. Since f, g and h are continuous from the right, there exists $\epsilon > 0$ such that

$$h(x_0 + \epsilon) < f(x_0), \quad g(x_0 + \epsilon) < h(x_0),$$

$$f(x_0 + \epsilon) - \frac{1}{2} < h(x_0), \quad h(x_0 + \epsilon) < g(x_0) + \frac{1}{2}.$$

(a) Suppose that $h(x_0) < h(x_0 + \epsilon)$. Let h_1 be any generalized distribution function with $h_1(x_0) = h(x_0)$ which may differ from the function h only on the interval $[x_0, x_0+\epsilon[$. From the inequalities above it follows that for all $t \in [x_0, x_0+\epsilon[$ we have

$$0 < h(x_0) - g(x_0+\epsilon) \leq h_1(t) - g(t) \leq h(x_0+\epsilon) - g(x_0) < \frac{1}{2},$$

and

$$0 < f(x_0) - h(x_0+\epsilon) \leq f(t) - h_1(t) \leq f(x_0+\epsilon) - h(x_0) < \frac{1}{2},$$

which means that $h_1 \in B_{1/2}(f) \cap B_{1/2}(g)$. Observe that there are infinitely many such functions h_1.
(b) In what follows we analyze the case where

$$h(x_0) = h(x_0+\epsilon).$$

Consider the set of all real numbers x such that $x_0 \leq x$ and $h(x_0) = h(x)$ (we also have $f(x_0) = f(x)$ and $g(x_0) = g(x)$). These points form an interval. In the case where this interval is not bounded, we have that f and g are constant functions on $[x_0, \infty[$. It follows easily that on this interval one can change the constant value of h to other constants such that the so obtained functions are elements of $B_{1/2}(f) \cap B_{1/2}(g)$, a contradiction.

If the interval above is bounded, then denote by x_1 its supremum. If $h(x_1) > h(x_0)$, then the constant value of the function h on $[x_0, x_1[$ can be changed to obtain different elements of $B_{1/2}(f) \cap B_{1/2}(g)$, a contradiction. Assuming $h(x_1) = h(x_0)$, for all $\delta > 0$ we have that $h(x_1) < h(x_1+\delta)$. Then considering x_1 in the place of x_0 above, the argument in (a) applies and it follows that there are more than one elements in the intersection of the balls under consideration, a contradiction again. This completes the proof of the assertion that if $B_{1/2}(f) \cap B_{1/2}(g)$ is a singleton, then we necessarily have either

$$f = \underline{0} \text{ and } g = \underline{1}, \text{ or } f = \underline{1} \text{ and } g = \underline{0}.$$

We proceed with the proof as follows. Since ϕ is an isometry we have that

$$B_{1/2}(\phi(\underline{0})) \cap B_{1/2}(\phi(\underline{1})) = \phi(B_{1/2}(\underline{0}) \cap B_{1/2}(\underline{1}))$$

which is a singleton. This implies that either $\phi(\underline{0}) = \underline{0}$ and $\phi(\underline{1}) = \underline{1}$, or $\phi(\underline{0}) = \underline{1}$ and $\phi(\underline{1}) = \underline{0}$. Let us assume the former case, i.e. that ϕ fixes $\underline{0}$ and $\underline{1}$.

For any $0 \leq a \leq b \leq 1$ denote by $\Delta_{a,b}$ the set

$$\Delta_{a,b} = \{g \in \Delta(\mathbb{R}) : \lim_{t\to-\infty} g(t) = a \text{ and } \lim_{t\to\infty} g(t) = b\}.$$

We assert that ϕ maps $\Delta_{a,b}$ onto itself. Indeed, it follows from the easy fact that $g \in \Delta(\mathbb{R})$ is an element of $\Delta_{a,b}$ if and only if $\rho(g, \underline{1}) = 1 - a$ and $\rho(g, \underline{0}) = b$. Observe that we obtain in particular that ϕ leaves all constant functions invariant.

Now, let $a < b$ be two fixed but arbitrary elements of $[0, 1]$ and denote by $\phi_{a,b}$ the restriction of the transformation ϕ onto the subspace $\Delta_{a,b}$. It is apparent that the transformation

$$f \mapsto \frac{\phi_{a,b}((b-a)f + a) - a}{b - a}, \quad f \in D(\mathbb{R})$$

is a surjective isometry of $D(\mathbb{R})$. It follows from Theorem 3.1 that either there exists a strictly increasing bijection $\varphi_{a,b} : \mathbb{R} \to \mathbb{R}$ such that we have

$$\frac{\phi_{a,b}((b-a)f(t) + a) - a}{b - a} = f(\varphi_{a,b}(t)), \quad t \in \mathbb{R}, f \in D(\mathbb{R}) \tag{3.1}$$

or there exists a strictly decreasing bijection $\psi_{a,b} : \mathbb{R} \to \mathbb{R}$ such that we have

$$\frac{\phi_{a,b}((b-a)f(t) + a) - a}{b - a} = 1 - f(\psi_{a,b}(t)-) \tag{3.2}$$

for all $t \in \mathbb{R}$ and $f \in D(\mathbb{R})$. For an arbitrary element F of $\Delta_{a,b}$ we see that $(F - a)/(b - a)$ is an element of $D(\mathbb{R})$. Insert this function into the corresponding equation (3.1) or (3.2). We get

$$\phi_{a,b}(F)(t) = F(\varphi_{a,b}(t)), \quad t \in \mathbb{R} \tag{3.3}$$

or

$$\phi_{a,b}(F)(t) = a + b - F(\psi_{a,b}(t)-), \quad t \in \mathbb{R}. \tag{3.4}$$

Recall that the numbers a and b above are fixed. Let us say that $\phi_{a,b}$ is of type I if it is of the form (3.3) with a strictly increasing bijection $\varphi_{a,b} : \mathbb{R} \to \mathbb{R}$ and let us say that $\phi_{a,b}$ is of type II if it is of the form (3.4) with a strictly decreasing bijection $\psi_{a,b} : \mathbb{R} \to \mathbb{R}$.

Letting now a and b vary, we show that either all $\phi_{a,b}$'s are of type I, or all of them are of type II. To verify this, first suppose that there exists $0 < b < b' \leq 1$ such that, for example, $\phi_{0,b}$ is of the form (3.3), i.e.

$$\phi_{0,b}(f)(t) = f(\varphi_{0,b}(t)), \quad t \in \mathbb{R}, f \in \Delta_{0,b}$$

and $\phi_{0,b'}$ is of the form (3.4), i.e.

$$\phi_{0,b'}(f)(t) = b' - f(\psi_{0,b'}(t)-), \quad t \in \mathbb{R}, f \in \Delta_{0,b'},$$

where $\varphi_{0,b} : \mathbb{R} \to \mathbb{R}$ is a strictly increasing bijection and $\psi_{0,b'} : \mathbb{R} \to \mathbb{R}$ is a strictly decreasing bijection. Let p and p' be arbitrary real numbers.

We consider Dirac-type generalized distribution functions $d_p^b \in \Delta_{0,b}$ and $d_{p'}^{b'} \in \Delta_{0,b'}$ defined by

$$d_p^b(t) = \begin{cases} 0 & \text{if } t < p \\ b & \text{if } t \geq p \end{cases} \quad \text{and} \quad d_{p'}^{b'}(t) = \begin{cases} 0 & \text{if } t < p' \\ b' & \text{if } t \geq p'. \end{cases}$$

Suppose that $b' - b < b$. We see that $p = p'$ if and only if $\rho(d_p^b, d_{p'}^{b'}) = b' - b$. Easy computation shows that

$$\phi(d_p^b) = \phi_{0,b}(d_p^b) = d^b_{\varphi_{0,b}^{-1}(p)} \quad \text{and} \quad \phi(d_{p'}^{b'}) = \phi_{0,b'}(d_{p'}^{b'}) = d^{b'}_{\psi_{0,b'}^{-1}(p')}.$$

Since the distance between those functions must also be $b' - b$, it follows that $\varphi_{0,b}^{-1}(p') = \psi_{0,b'}^{-1}(p')$ for all $p' \in \mathbb{R}$ which is an obvious contradiction. It is easy to see that we would arrive at a similar contradiction if we assumed that $\phi_{0,b}$ is of type II and $\phi_{0,b'}$ is of type I. It follows that $\phi_{0,b}$ and $\phi_{0,b'}$ are of the same type whenever $b' - b < b$. Assume now that $b' - b \geq b$. Then one can pick a finite sequence $b = b_1 < b_2 < \ldots < b_n = b'$ such that $b_{i+1} - b_i < b_i$ holds for all $i = 1, \ldots, n - 1$. From what we have proved above, it follows step by step that the maps $\phi_{0,b_1}, \ldots, \phi_{0,b_n}$ are all of the same type. Therefore, we obtain that all maps $\phi_{0,b}$, $b \in [0, 1]$ are of the same type. (Observe that since ϕ leaves the constant functions invariant, hence all $\phi_{a,a}$, $0 \leq a \leq 1$ are both of type I and II.) In a similar fashion one can show that the maps $\phi_{a,1}$, $0 \leq a \leq 1$ are all of the same type, too. Still using the same idea, for any $0 \leq a \leq b < b' \leq 1$ one can verify that $\phi_{a,b}$ and $\phi_{a,b'}$ are of the same type and next that for any $0 \leq a < a' \leq b$ the maps $\phi_{a,b}$ and $\phi_{a',b}$ are of the same type. From these one can easily conclude that all maps $\phi_{a,b}$ are of the same type, namely the type of $\phi_{0,1}$.

However, the above argument shows not only that the types of the maps $\phi_{a,b}$ are all the same but also that the inducing functions $\varphi_{a,b}$ or $\psi_{a,b}$ appearing in (3.3) or (3.4) are identical for all $a, b \in [0, 1], a \leq b$. Indeed, first suppose that all $\phi_{a,b}$ are of type I. Let $0 < b < b' \leq 1$. Then $\phi_{0,b}$ is of the form

$$\phi_{0,b}(f)(t) = f(\varphi_{0,b}(t)), \quad t \in \mathbb{R}, f \in \Delta_{0,b}$$

and $\phi_{0,b'}$ is of the form

$$\phi_{0,b'}(f)(t) = f(\varphi_{0,b'}(t)), \quad t \in \mathbb{R}, f \in \Delta_{0,b'},$$

where the inducing functions $\varphi_{0,b}, \varphi_{0,b'} : \mathbb{R} \to \mathbb{R}$ are strictly monotone increasing bijections. Assume that $b' - b < b$. For the Dirac-type generalized distribution functions d_u^b and $d_u^{b'}$ we have

$$\rho(d_u^b, d_u^{b'}) = b' - b, \quad u \in \mathbb{R}$$

which implies that the distance between their images

$$\phi(d^b_u) = d^b_{\varphi^{-1}_{0,b}(u)}, \quad \phi(d^{b'}_u) = d^{b'}_{\varphi^{-1}_{0,b'}(u)}$$

is also $b' - b$ meaning that

$$\rho(d^b_{\varphi^{-1}_{0,b}(u)}, d^{b'}_{\varphi^{-1}_{0,b'}(u)}) = b' - b, \quad u \in \mathbb{R}.$$

This gives us that $\varphi^{-1}_{0,b}(u) = \varphi^{-1}_{0,b'}(u)$ holds for all $u \in \mathbb{R}$. This yields that the inducing functions $\varphi_{0,b}$ and $\varphi_{0,b'}$ are identical. We can proceed exactly as with the types. Namely, if $b' - b \geq b$, one can pick a finite sequence $b = b_1 < b_2 < \ldots < b_n = b'$ such that $b_{i+1} - b_i < b_i$ holds for all $i = 1, \ldots, n-1$ and infer that the inducing functions of $\phi_{0,b_1}, \ldots, \phi_{0,b_n}$ are identical. As a consequence, we obtain that the inducing functions $\varphi_{0,b}$ are the same for all $b \in [0,1]$. In a similar fashion one can show that for every $a \in [0,1]$ the strictly monotone increasing bijections $\varphi_{a,1}$ appearing in the form of $\phi_{a,1}$ are identical. Still using the same idea, for any $0 \leq a \leq b < b' \leq 1$ one can verify that $\varphi_{a,b}$ and $\varphi_{a,b'}$ are the same and next that for any $0 \leq a < a' \leq b$ the functions $\varphi_{a,b}$ and $\varphi_{a',b}$ are identical. From these one can easily conclude that the inducing functions of $\phi_{a,b}$ are all the same, namely, they equal the inducing function of $\phi_{0,1}$. Obviously, one can follow a similar argument in the case where all restricted maps $\phi_{a,b}$ are of type II.

Therefore, we have proved that either there exists a strictly increasing bijection $\varphi : \mathbb{R} \to \mathbb{R}$ such that

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in \Delta(\mathbb{R}) \tag{3.5}$$

or there exists a strictly decreasing bijection $\psi : \mathbb{R} \to \mathbb{R}$ such that

$$\phi(f)(t) = f_{-\infty} + f_{\infty} - f(\psi(t)-), \quad t \in \mathbb{R}, f \in \Delta(\mathbb{R}),$$

where $f_{-\infty} = \lim_{t\to-\infty} f(t)$ and $f_{\infty} = \lim_{t\to\infty} f(t)$. However, this latter possibility can be ruled out easily. Indeed, consider for example the following generalized distribution functions:

$$f(t) = \begin{cases} 1/4 & \text{if } t < 0 \\ 1/2 & \text{if } t \geq 0 \end{cases} \quad \text{and} \quad g(t) = \begin{cases} 1/4 & \text{if } t < 1 \\ 3/4 & \text{if } t \geq 1. \end{cases}$$

Then we have $\rho(f,g) = 1/4$ while the distance between $\phi(f)$ and $\phi(g)$ would be 1/2, a contradiction. Therefore, we have verified that in the case where ϕ leaves the constant functions $\underline{0}, \underline{1}$ invariant, the transformation ϕ is necessarily of the form (3.5).

In the remains to treat the case where ϕ interchanges the constant functions $\underline{0}$ and $\underline{1}$. Then consider the transformation $\Psi : \Delta(\mathbb{R}) \to \Delta(\mathbb{R})$ defined by $\Psi(f)(t) = 1 - f((-t)-)$, $t \in \mathbb{R}, f \in \Delta(\mathbb{R})$. It is easily seen to be a surjective isometry of $\Delta(\mathbb{R})$ which interchanges $\underline{0}$ and $\underline{1}$. Therefore, the transformation $\Psi \circ \phi$ is a surjective isometry which leaves $\underline{0}$ and $\underline{1}$ invariant. Applying the first part of the proof for this transformation, it follows that it is of the form (3.5). Then composing by $\Psi^{-1} = \Psi$ from the left, we easily get that ϕ is of the second form that appears in the statement of the theorem. The proof is complete. □

Concerning the Kolmogorov-Smirnov isometries of $\Delta_c(\mathbb{R})$ our argument is rather similar to what we have applied in the proof of Theorem 3.3. Namely, we try to deduce the result from the corresponding statement Theorem 3.2 concerning the space $D_c(\mathbb{R})$ of all continuous distribution functions. We only sketch the argument.

SKETCH OF THE PROOF THEOREM 3.4. Similarly to the proof of Theorem 3.3, one can verify that $B_{1/2}(f) \cap B_{1/2}(g)$ is a singleton if and only if $\{f, g\} = \{\underline{0}, \underline{1}\}$. In fact, the proof of this characterization can go along the similar lines as in the corresponding part of the proof of Theorem 3.3 with the difference that in the present case the situation is easier, several subcases do not appear due to the continuity of functions under consideration. Hence we obtain that ϕ either leaves the functions $\underline{0}$ and $\underline{1}$ invariant or interchanges them. Assume that we have the former case, i.e. ϕ fixes $\underline{0}$ and $\underline{1}$. We can show that for arbitrary $a, b \in [0, 1], a \leq b$ the restriction $\phi_{a,b}$ of ϕ onto $\Delta_c(\mathbb{R}) \cap \Delta_{a,b}$ is a surjective isometry and then apply Theorem 3.2 to deduce that either there exists a strictly increasing bijection $\varphi_{a,b} : \mathbb{R} \to \mathbb{R}$ such that $\phi_{a,b}$ is of the form

$$\phi_{a,b}(F)(t) = F(\varphi_{a,b}(t)), \quad t \in \mathbb{R}, F \in \Delta_c(\mathbb{R}) \cap \Delta_{a,b} \tag{3.6}$$

or there exists a strictly decreasing bijection $\psi_{a,b} : \mathbb{R} \to \mathbb{R}$ such that $\phi_{a,b}$ is of the form

$$\phi_{a,b}(F)(t) = b + a - F(\psi_{a,b}(t)), \quad t \in \mathbb{R}, F \in \Delta_c(\mathbb{R}) \cap \Delta_{a,b}. \tag{3.7}$$

In the case (3.6) we say that $\phi_{a,b}$ is of type I while in the case (3.7) we say that it is of type II. In the proof of Theorem 3.3 we have employed Dirac-type generalized distribution functions to see that either all $\phi_{a,b}$ are of type I or all of them are of type II, and moreover, that the inducing function appearing in the forms of $\phi_{a,b}$'s are all the same. At this point the present proof must be a bit changed due to the fact that the Dirac-type functions are not continuous. We argue as follows. Let $0 < b < b' \leq 1$. Suppose, for example, that $\phi_{0,b}$ is of type II and $\phi_{0,b'}$ is of type I with inducing function $\psi_{0,b}$ and $\varphi_{0,b'}$, respectively. Clearly, the former function is a strictly decreasing while the latter one is a strictly increasing bijection of $\mathbb{R}$. Pick two arbitrary real numbers $x < y$. Let $x < z < y$ be such that $b/(z-x) = b'/(y-x)$. Consider the following continuous generalized distribution functions.

(3.8)

$$f(t) = \begin{cases} 0 & \text{if } t < x \\ \frac{b(t-x)}{z-x} & \text{if } x \le t < z \\ b & \text{if } z \le t \end{cases} \quad \text{and} \quad g(t) = \begin{cases} 0 & \text{if } t < x \\ \frac{b'(t-x)}{y-x} & \text{if } x \le t < y \\ b' & \text{if } y \le t, \end{cases}$$

Obviously, the distance between f and g is $b' - b$. We assert that

$$\psi_{0,b}^{-1}(y) \le \varphi_{0,b'}^{-1}(y).$$

Indeed, in the opposite case there were a real number t such that

$$\psi_{0,b}^{-1}(y) > t > \varphi_{0,b'}^{-1}(y)$$

which would imply that $y < \psi_{0,b}(t)$ and $y < \varphi_{0,b'}(t)$. By (3.6), (3.7) we would deduce that the distance between $\phi(f)$ and $\phi(g)$ is b', a clear contradiction. Therefore, we have $\psi_{0,b}^{-1}(y) \le \varphi_{0,b'}^{-1}(y)$ for all real numbers y which is obviously untenable by the different monotonicity properties of $\psi_{0,b}$ and $\varphi_{0,b'}$. One can apply a similar argument in the case where $\phi_{0,b}$ is of type I and $\phi_{0,b'}$ is of type II. Then one can continue showing that the types of $\phi_{a,b}$ are all the same and next that the inducing functions are also identical. We obtain that either there exists a strictly increasing bijection $\varphi : \mathbb{R} \to \mathbb{R}$ such that

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in \Delta_c(\mathbb{R}) \tag{3.9}$$

or there exists a strictly decreasing bijection $\psi : \mathbb{R} \to \mathbb{R}$ such that

$$\phi(f)(t) = f_{-\infty} + f_{\infty} - f(\psi(t)), \quad t \in \mathbb{R}, f \in \Delta_c(\mathbb{R}).$$

The second form for ϕ can be ruled out by choosing a pair of appropriate simple functions as in the proof of Theorem 3.3 modified to functions like f and g in (3.8) above. This completes the proof when ϕ fixes $\underline{0}$ and $\underline{1}$.

In the remaining case where ϕ interchanges the functions $\underline{0}$ and $\underline{1}$, we consider the transformation Ψ on $\Delta_c(\mathbb{R})$ defined by $\Psi(f)(t) = 1 - f(-t)$, $t \in \mathbb{R}, f \in \Delta_c(\mathbb{R})$ which is a surjective isometry interchanging $\underline{0}$ and $\underline{1}$. Composing ϕ by Ψ from the left we obtain a surjective isometry on $\Delta_c(\mathbb{R})$ which leaves $\underline{0}$ and $\underline{1}$ invariant. Applying the first part of the proof for $\Psi \circ \phi$ we get that it is of the form (3.9). Composing by $\Psi^{-1} = \Psi$ from the left again, we obtain the second possible form for the original map ϕ that appears in the formulation of the theorem. □

4

Separation by convex interpolation families

4.1. Introduction and statement of the results

In the last two chapters we are going to present some theorems concerning different separation problems. That kinds of results play a crucial role especially in the field of convex analysis [**50**], [**78**]. It is a well-known result that if a convex and a concave function are given such that the convex function is "above" the concave one, then there exists an affine function between them. Of course, the assumptions on convexity/concavity are *sufficient* but not *necessary* for the existence of an affine separator. However, in [**67**] Nikodem and Wąsowicz proved a nice result which gives a characterization of those pairs of real functions that can be separated by an affine function. A set of continuous functions defined on an interval I is called an n-parameter Beckenbach family, if each n points of $I \times \mathbb{R}$ (with pairwise distinct first coordinates) can be interpolated by a unique element of the set. The aim of the present chapter is to generalize the result of Nikodem and Wąsowicz to this setting, i.e. to characterize such pairs of real valued functions that can be separated by a member of a given convex Beckenbach family of order n. The present chapter is based on the paper [**14**].

Assume that $f, g\colon I \to \mathbb{R}$ are such that can be separated by an affine function $h\colon I \to \mathbb{R}$. Then, as direct calculations show, for all elements x, y of I and $\lambda \in [0, 1]$, the following inequalities hold:

$$\begin{aligned} f\big(\lambda x + (1-\lambda)y\big) &\le \lambda g(x) + (1-\lambda)g(y), \\ g\big(\lambda x + (1-\lambda)y\big) &\ge \lambda f(x) + (1-\lambda)f(y). \end{aligned} \tag{4.1}$$

Due to a result of Nikodem and Wąsowicz [**67**], these inequalities are not merely a *consequence* of the existence of an affine separator, but *characterize* the existence of such separator.

Separation problems of this spirit was studied intensively by several authors and in several contexts. The polynomial case is due to Wąsowicz [**87**] and by Balaj and Wąsowicz [**3**]. Their approach is based on some selection principles of Behrend and Nikodem [**8**] and by Balaj and Nikodem [**2**]. The separation problem was solved by Nikodem and Páles [**66**] for so-called two parameter interpolation families. When the parameter of the family is *arbitrary* but the structure is *linear*, the characterization was presented by Bessenyei and Páles [**12**].

Now we introduce the necessary notion which will be applied in the present chapter.

DEFINITION 4.1. *Let H be a real subset of at least n elements. A set $\mathcal{F}_n(H)$ of real functions is called an n-parameter* interpolation family *over H, if its members are defined on H and, for all points $(x_1, y_1), \dots, (x_n, y_n)$ of $H \times \mathbb{R}$ (with pairwise distinct first coordinates) there exists exactly one element φ of $\mathcal{F}_n(H)$ such that*

$$\varphi(x_1) = y_1, \dots, \varphi(x_n) = y_n. \tag{4.2}$$

An n-parameter interpolation family is said to be a Beckenbach family, *if its members are continuous. A Beckenbach family of order n is denoted by $\mathcal{B}_n(H)$.*

Throughout this chapter, the members of an interpolation family $\mathcal{F}_n(H)$ are termed briefly *generalized lines*. The most important subclass of interpolation families can be obtained via Haar systems. In fact, we shall prove that a linear interpolation family coincides the linear hull of a suitable Haar system.

DEFINITION 4.2. *Let H be a real subset of at least n elements and let $\omega_1, \dots, \omega_n \colon H \to \mathbb{R}$ be given functions. We say that $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)$ is a (positive)* Haar system *on H if, for all elements $x_1 < \dots < x_n$ of H,*

$$\left| \boldsymbol{\omega}(x_1) \quad \dots \quad \boldsymbol{\omega}(x_n) \right| := \begin{vmatrix} \omega_1(x_1) & \dots & \omega_1(x_n) \\ \vdots & \ddots & \vdots \\ \omega_n(x_1) & \dots & \omega_n(x_n) \end{vmatrix} \overset{(>)}{\neq} 0.$$

Under a Chebyshev system *we mean a Haar system of continuous functions.*

The most important example for a Haar system of parameter n is the polynomial system, that is, monomials up to degree $(n-1)$. Indeed, the corresponding determinant in this case is a Vandermonde determinant and hence nonvanishing. In fact, the polynomial system is also a Chebyshev system.

Obviously, the set of all linear combinations of an n-parameter Haar system over H is an interpolation family and it will be denoted by $\Omega_n(H)$. Hence, for our convenience, the terminology "Haar system" is also used for this linear span. In fact, identifying the vectors of functions with their linear hull is not misleading and is widely accepted in the technical literature. Let us mention that Haar and Chebyshev systems play an important role, sometimes indirectly, in numerous fields of mathematics; the book of Karlin and Studden [**44**] contains a rich material and bibliography of the topics for the interested reader.

Now, we are in a position to present the main result of this chapter. It gives a characterization of those pairs of real functions that can be separated by a generalized line belonging to a Beckenbach family which is supposed to be closed under convex combinations. It turns out that the proper separator needs to satisfy a system of inequalities.

THEOREM 4.3. (Bessenyei, Szokol [**14**])
Let $\mathcal{B}_n(I)$ be a Beckenbach family over the real interval I which is closed under convex combinations and $f, g\colon I \to \mathbb{R}$ be given functions. Then the following statements are equivalent:

(i) there exists $h \in \mathcal{B}_n(I)$ such that $f \le h \le g$;
(ii) for all $u \le x_1 < \cdots < x_n \le v$ of I, we have the inequalities

$$\varphi_1(v) \ge f(v), \qquad \psi_1(v) \le g(v); \tag{4.3}$$

and

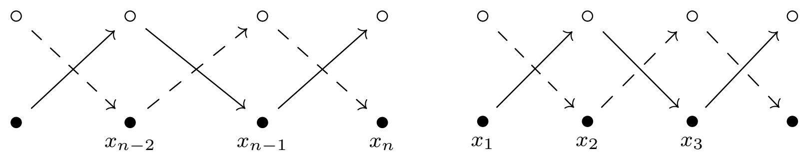
$$\varphi_2(u) \ge f(u), \qquad \psi_2(u) \le g(u), \tag{4.4}$$

where $\varphi_1, \varphi_2, \psi_1, \psi_2 \in \mathcal{B}_n(I)$ are determined by the interpolation properties

$$\begin{aligned}
&\varphi_1(x_k) = g(x_k),\ \psi_1(x_k) = f(x_k), && n-k \in \{0,\dots,n-1\} \cap 2\mathbb{Z};\\
&\varphi_1(x_k) = f(x_k),\ \psi_1(x_k) = g(x_k), && n-k \in \{0,\dots,n-1\} \cap \big(2\mathbb{Z}+1\big);\\
&\varphi_2(x_k) = g(x_k),\ \psi_2(x_k) = f(x_k), && k \in \{1,\dots,n\} \cap \big(2\mathbb{Z}+1\big);\\
&\varphi_2(x_k) = f(x_k),\ \psi_2(x_k) = g(x_k), && k \in \{1,\dots,n\} \cap 2\mathbb{Z}.
\end{aligned}$$

The meaning of the interpolation properties appearing in the theorem, can be illustrated in the following expressive way. For simplicity, the symbols "$\bullet$" and "$\circ$" stand for the values of f and g, respectively. Then, φ_1 and

φ_2 are obtained by "$\to$", while ψ_1 and ψ_2 are obtained by "$\dashrightarrow$".



If the underlying Beckenbach family is a Chebyshev system, then the main result reduces to the next corollary. Its statement remains true also for Haar systems [**12**].

COROLLARY 4.4. *Let I be a real interval, $f, g\colon I \to \mathbb{R}$ be two given functions and assume that $\boldsymbol{\omega} := (\omega_1,\dots,\omega_n)$ is a positive Chebyshev system over I. Then, there exists $h \in \Omega(I)$ such that $f \le h \le g$ if and only if, for all elements $x_1 \le \dots \le x_{n+1}$ of I the next inequalities are satisfied:*

$$\begin{vmatrix} \dots & \boldsymbol{\omega}(x_{n-2}) & \boldsymbol{\omega}(x_{n-1}) & \boldsymbol{\omega}(x_n) & \boldsymbol{\omega}(x_{n+1}) \\ \dots & g(x_{n-2}) & f(x_{n-1}) & g(x_n) & f(x_{n+1}) \end{vmatrix} \le 0;$$

$$\begin{vmatrix} \dots & \boldsymbol{\omega}(x_{n-2}) & \boldsymbol{\omega}(x_{n-1}) & \boldsymbol{\omega}(x_n) & \boldsymbol{\omega}(x_{n+1}) \\ \dots & f(x_{n-2}) & g(x_{n-1}) & f(x_n) & g(x_{n+1}) \end{vmatrix} \ge 0.$$

Two direct consequences of Corollary 4.4 are as follow. The first one is the polynomial setting, the main result of [**3**] and [**87**]. For technical convenience, we use the next concepts: if points $x_0 \le \dots \le x_n$ are fixed elements of an interval I, then denote the Vandermonde determinants built on the system $\{x_0,\dots,x_n\} \setminus \{x_k\}$ by $V_k(x_0,\dots,x_n)$. Furthermore, denote the sets $(2\mathbb{Z}) \cap [0,n]$ and $(2\mathbb{Z}+1) \cap [0,n]$ by $N_0(n)$ and $N_1(n)$, respectively.

COROLLARY 4.5. *Let I be an interval and $f, g\colon I \to \mathbb{R}$ be given functions. Then, there exists a polynomial h of degree at most $(n-1)$ satisfying $f \le h \le g$ if and only if, for all elements $x_0 \le \dots \le x_n$ of I, the following inequalities hold:*

$$\sum_{k \in N_0(n)} f(x_{n-k}) V_{n-k}(x_0,\dots,x_n) \le \sum_{k \in N_1(n)} g(x_{n-k}) V_{n-k}(x_0,\dots,x_n),$$
$$\sum_{k \in N_0(n)} g(x_{n-k}) V_{n-k}(x_0,\dots,x_n) \le \sum_{k \in N_1(n)} f(x_{n-k}) V_{n-k}(x_0,\dots,x_n).$$

The other direct consequence of Corollary 4.4 (and of course, also of Corollary 4.5) is the main result of [**67**], which is one of the motivations of Theorem 4.3.

COROLLARY 4.6. *Let I be an interval and $f, g\colon I \to \mathbb{R}$ be given functions. Then, there exists an affine function $h\colon I \to \mathbb{R}$ if and only if, for all elements x, y of I and $\lambda \in [0,1]$, the inequalities* (4.1) *hold.*

4.2. Proofs

In this section we present the proof of Theorem 4.3 and the sketch of the proof of Corollary 4.4. We begin with a short summary presenting the basic steps of the proof of Theorem 4.3. Unlike the preliminary approaches, our proof does not require selection methods, but it is based on the geometric feature of interpolation families. The key tool is the classical Helly theorem, one of the most important results in convex and combinatorial geometry [**85**]. However, in the lack of linear structure, it cannot be applied *directly*. Therefore, we shall make a "detour" according to the figure below:

$$\begin{array}{ccc} \big(\mathcal{B}_n(I), d\big) & \xrightarrow{\;\Phi\;} & \big(\mathbb{R}^n, \|\cdot\|\big) \\ \big\downarrow & & \big\downarrow \\ \bigcap_{x\in I} \mathcal{K}(x) \neq \emptyset & \xleftarrow{\;\Phi^{-1}\;} & \bigcap_{x\in I} K(x) \neq \emptyset \end{array}$$

It turns out that a Beckenbach family can be considered as a metric space $\big(\mathcal{B}_n(I), d\big)$. Moreover, this metric space is Φ-homeomorphic to the Euclidean space $\big(\mathbb{R}^n, \|\cdot\|\big)$. Denoting the set of generalized lines that separate the given functions at point x by $\mathcal{K}(x)$, one should check that $\bigcap_{x\in I} \mathcal{K}(x)$ is nonempty. Instead of this, we prove that the Φ-image of the intersection, denoted by $\bigcap_{x\in I} K(x)$, is nonempty. To do this, we conclude the sets $K(x)$ are such subsets of $\mathbb{R}^n$ that fulfill the conditions of Helly's theorem. Checking most of the conditions are quite simple; the only difficulty is to prove that each $(n+1)$ member of the collection $\{K(x) \mid x \in I\}$ has a nonempty intersection. This property can be verified via constructing a homotopy based on the systems of those inequalities that are involved in our main (characterization) theorem.

To follow the above steps we need the following three lemmas. The first one states that an interpolation family can be metrized in a quite natural way. Let us emphasize that, in the rest of this chapter, every topological notion about interpolation families is interpreted within this framework.

LEMMA 4.7. *Let $\mathcal{F}_n(H)$ be an interpolation family over the set H of at least n elements and let $x_1, \dots, x_n$ be pairwise distinct elements of H. Then,*

$\big(\mathcal{F}_n(H), d\big)$ *is a complete metric space where*

$$d(\varphi,\psi) := \max\{|\varphi(x_1)-\psi(x_1)|,\dots,|\varphi(x_n)-\psi(x_n)|\}, \quad \varphi,\psi\in\mathcal{F}_n(H).$$

Moreover, $\varphi_m \to \varphi$ *with respect to* d *if and only if* $\varphi_m \to \varphi$ *pointwise on* H.

PROOF. The properties $d(\varphi,\psi)\ge 0$, $d(\varphi,\psi)=d(\psi,\varphi)$ and $d(\varphi,\varphi)=0$ are trivial. Assume $d(\varphi,\psi)=0$. Then, φ and ψ coincide at n pairwise distinct points of H and hence, due to the unique interpolation property, $\varphi=\psi$. The triangle inequality is a direct consequence of the properties of absolute value and of the maximum functions. The completeness of the metric space follows from the completeness of $\mathbb{R}^n$. Indeed, take a Cauchy-sequence (φ_m) in $\mathcal{F}_n(H)$. Then, $\varphi_m(x_k)$ is a Cauchy-sequence in $\mathbb{R}$ for all $k=1,\dots,n$. The completeness of the reals implies that there exists values y_k such that $\varphi_m(x_k)\to y_k$ as $m\to\infty$. Let φ be the unique element of $\mathcal{F}_n(H)$ which is determined by the interpolation properties

$$\varphi(x_k)=y_k, \quad (k=1,\dots,n).$$

According to the construction, $\varphi_m\to\varphi$ with respect to d, which implies the completeness of $\big(\mathcal{F}_n(H), d\big)$. Finally, the unique interpolation property implies, that $\varphi_m\to\varphi$ pointwise on H if $\varphi_m\to\varphi$. □

LEMMA 4.8. *If* $\mathcal{F}_n(H)$ *is an interpolation family over the set* H *of at least* n *elements and* $x_1,\dots,x_n$ *are pairwise distinct elements of* H*, then the mapping* $\Phi\colon \mathcal{F}_n(H)\to\mathbb{R}^n$ *given by* $\Phi(\varphi)=\big(\varphi(x_1),\dots,\varphi(x_n)\big)$ *is a homeomorphism.*

PROOF. The bijectivity of Φ is a straightforward consequence of the unique interpolation property, while the continuity of Φ follows immediately from the definition of the metrics. For the continuity of Φ^{-1}, take a sequence $y_m=(y_{m1},\dots,y_{mn})$ in $\mathbb{R}^n$ that tends to $y=(y_1,\dots,y_n)$ as $m\to\infty$. If $\varphi_m:=\Phi^{-1}(y_m)$ and $\varphi:=\Phi^{-1}(y)$, then, by definition, for all $k=1,\dots,n$,

$$\varphi_m(x_k)=y_{mk}, \quad \text{and} \quad \varphi(x_k)=y_k.$$

Therefore $\varphi_m\to\varphi$. This means that $\Phi^{-1}(y_m)\to\Phi^{-1}(y)$ as $m\to\infty$, showing the desired continuity. □

LEMMA 4.9. *Let* H *be a real subset of at least* n *elements. An interpolation family* $\mathcal{F}_n(H)$ *is linear if and only if there exists a Haar system* $\boldsymbol{\omega}=(\omega_1,\dots,\omega_n)$ *such that* $\mathcal{F}_n(H)=\mathrm{Lin}(\boldsymbol{\omega})$*. Moreover, there exist non-linear interpolation families that are convex.*

PROOF. We shall concentrate only for sufficiency since the necessity is trivial. Fix pairwise distinct elements $x_1, \dots, x_n$ of H and define $\omega_k \colon H \to \mathbb{R}$ of $\mathcal{F}_n(H)$ via the interpolation properties

$$\omega_k(x_l) = \delta_{kl},$$

where δ_{kl} stands for the Kronecker delta symbol. Our aim is to prove that $\boldsymbol{\omega} := (\omega_1, \dots, \omega_n)$ is a Haar system fulfilling $\mathcal{F}_n(H) = \text{Lin}(\boldsymbol{\omega})$. Let φ be an arbitrary element of $\mathcal{F}_n(H)$ and assume that $\varphi(x_l) = \alpha_l$. Since $\mathcal{F}_n(H)$ is closed under linear combinations, the function $\omega := \alpha_1\omega_1 + \cdots + \alpha_n\omega_n$ belongs to $\mathcal{F}_n(H)$. On the other hand,

$$\omega(x_l) = \sum_{j=1}^{n} \alpha_j \omega_j(x_l) = \sum_{j=1}^{n} \alpha_j \delta_{jl} = \alpha_l = \varphi(x_l).$$

Consequently, due to the unique interpolation property, $\omega = \varphi$. That is, $\mathcal{F}_n(H) = \text{Lin}(\boldsymbol{\omega})$. Finally, we have to check, that $\boldsymbol{\omega}$ is a Haar system, indeed. Suppose indirectly, that there exist pairwise distinct elements $t_1,\dots,t_n$ of H such that

$$0 = \begin{vmatrix} \omega_1(t_1) & \dots & \omega_n(t_1) \\ \vdots & \ddots & \vdots \\ \omega_1(t_n) & \dots & \omega_n(t_n) \end{vmatrix}.$$

In this case, there exist nontrivial coefficients $\beta_1, \dots, \beta_n$ such that $\psi = \beta_1\omega_1 + \cdots + \beta_n\omega_n$ vanishes at $t_1, \dots, t_n$. Note that $\psi \not\equiv 0$ since $\psi(x_l) = \beta_l$ and there exists an index l such that $\beta_l \neq 0$. Then, the points $(t_1, 0)$; $\dots$; $(t_n, 0)$ can be interpolated by ψ and by the zero function, which is a contradiction.

For the second statement, take an arbitrary set H^* of at least $(n + 1)$ elements and a Haar system $\mathcal{F}_{n+1}(H^*)$ over this set. Fix $x^* \in H^*$, define $H := H^* \setminus \{x^*\}$, and denote those elements of $\mathcal{F}_{n+1}(H^*)$ that take value 1 at x^* by $\mathcal{F}_n(H)$. Then, $\mathcal{F}_n(H)$ is a convex interpolation family of parameter n over H. On the other hand, $\mathcal{F}_n(H)$ has no linear structure since it is not closed under multiplication by -1. (In more simple: polynomials p of degree at most n on $[0, 1]$ fulfilling $p(-1) = 1$ form a convex but nonlinear Beckenbach family of parameter n.) □

In fact, the second statement of Lemma 4.7 can considerably generalized: Due to a result of Tornheim [**82**], uniform convergence on compact subintervals of the domain is equivalent to the convergence of n pairwise distinct points among generalized lines belonging to a n parameter Beckenbach family.

According to Lemma 4.9, linear interpolation families are linear hulls. However, the counterpart of the lemma shows, that this is not the case if we assume only convex-closedness. This observation guarantees that our main result is a meaningful generalization of the earlier (linear) ones.

PROOF OF THEOREM 4.3. For simplicity, denote φ_1 by φ. To verify $(i) \Rightarrow (ii)$, assume indirectly that $\varphi(v) < f(v)$. Let $\varepsilon > 0$ be arbitrary, and consider the generalized line φ_ε determined by the interpolation properties

$$\varphi_\varepsilon(x_k) = \varphi(x_k) + (-1)^{n-k}\varepsilon \qquad (k = 1, \dots, n.)$$

By Bolzano's theorem, and by the inequalities $f \le h \le g$, there exists $\xi_k \in]x_k, x_{k+1}[$ for all $k = 1, \dots, n-1$ such that $h(\xi_k) = \varphi_\varepsilon(\xi_k)$. Then, $h(t) \ne \varphi_\varepsilon(t)$ if $t \in I \setminus \{\xi_1, \dots, \xi_{n-1}\}$. Moreover, using the conventions $\xi_0 := \inf I$ and $\xi_n := \sup I$, for all $x \in]\xi_{k-1}, \xi_k[$ and $k = 1, \dots, n$ we have the inequalities

$$(-1)^{n-k}\big(\varphi_\varepsilon(x) - h(x)\big) > 0.$$

In particular, $h(v) < \varphi_\varepsilon(v)$. Since $\varphi_\varepsilon \to \varphi$ pointwise as $\varepsilon \to 0$, there exists some positive ε_0 such that $\varphi(v) < \varphi_{\varepsilon_0}(v) < f(v)$. Consequently $h(v) < f(v)$ follows, which contradicts to $f \le h$. The further inequalities can be proved via similar arguments.

For the converse implication first we check that $f \le g$ holds. Let v be an element of I differing from $\inf(I)$, and fix the elements $x_1 < \dots < x_n := v$. Then, according to the definition of φ and applying the corresponding inequality of the second assertion,

$$f(v) \le \varphi(v) = g(v)$$

follows. The case when $u = \inf(I)$ belongs to the interval, both the inequalities of (4.4) lead to $f(u) \le g(u)$, which means that f is majorized by g on the whole interval I. In particular, the subset $\mathcal{K}(x)$ of $\mathcal{B}_n(I)$, defined by

$$\mathcal{K}(x) := \{\omega \in \mathcal{B}_n(I) \mid f(x) \le \omega(x) \le g(x)\}$$

is non-empty for all elements $x \in I$. Let $t_1, \dots, t_n$ be fixed and pairwise distinct elements of I, and consider the homeomorphism $\Phi\colon \mathcal{B}_n(I) \to \mathbb{R}^n$ given by

$$\Phi(\omega) := \big(\omega(t_1), \dots, \omega(t_n)\big).$$

Then, $K(x) := \Phi\big(\mathcal{K}(x)\big) \ne \emptyset$ for all $x \in I$. Let $y_1 = (y_{11}, \dots, y_{1n})$ and $y_2 = (y_{21}, \dots, y_{2n})$ be elements of $K(x)$ and $\lambda \in [0, 1]$. Then, there exist generalized lines ω_1 and ω_2 that belong to $\mathcal{K}(x)$ and fulfill the properties

$$\omega_1(t_k) = y_{1k}, \quad \omega_2(t_k) = y_{2k}, \quad k = 1, \cdots, n.$$

Define $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$. Since $\mathcal{B}_n(I)$ is convex, ω is a generalized line that belongs to $\mathcal{K}(x)$. On the other hand, $\omega(t_k) = \lambda y_{1k} + (1-\lambda)y_{2k}$ by construction, showing that $\Phi(\omega) = \lambda y_1 + (1-\lambda)y_2$. In other words, $K(x)$ is convex.

Let $y_m = (y_{m1}, \dots, y_{mn})$ be a sequence in $K(x)$ that tends to $y = (y_1, \dots, y_n)$. Then, there exist a sequence (ω_m) in $\mathcal{K}(x)$ and an element ω of $\mathcal{B}_n(I)$ such that

$$\omega_m(t_k) = y_{mk}, \qquad \omega(t_k) = y_k.$$

Since $\omega_m \to \omega$ at the points of the set $\{t_1, \dots, t_n\}$, $\omega_m \to \omega$ on the whole interval I as $m \to \infty$. In particular, $\lim_{m\to\infty} \omega_m(x) = \omega(x)$. On the other hand, the property $\omega_m \in \mathcal{K}(x)$ implies

$$f(x) \le \omega_m(x) \le g(x);$$

passing to infinity, $\omega \in \mathcal{K}(x)$ follows. That is, the set $K(x)$ is closed.

Let $x_1 < \cdots < x_n$ be elements of I and take the homeomorphism $\Psi\colon \mathcal{B}_n(I) \to \mathbb{R}^n$ defined by the usual way

$$\Psi(\omega) := \big(\omega(x_1), \dots, \omega(x_n)\big).$$

The closed set $\Psi\big(\mathcal{K}(x_1) \cap \ldots \cap \mathcal{K}(x_n)\big)$ is bounded evidently in the maximum norm of $\mathbb{R}^n$ and hence, by the Heine–Borel Theorem, is compact. Since Φ is a homeomorphism,

$$\begin{aligned} K(x_1) \cap \ldots \cap K(x_n) &= \Phi\big(\mathcal{K}(x_1)\big) \cap \ldots \cap \Phi\big(\mathcal{K}(x_n)\big) \\ &= \Phi\big(\mathcal{K}(x_1) \cap \ldots \cap \mathcal{K}(x_n)\big) \\ &= \big(\Phi \circ \Psi^{-1}\big)\Big(\Psi\big(\mathcal{K}(x_1) \cap \ldots \cap \mathcal{K}(x_n)\big)\Big). \end{aligned}$$

Applying the continuity of $\Phi \circ \Psi^{-1}$ and the compactness of the argument in the latter term, we get that each n member subcollection of the family $\{K(x) \mid x \in I\}$ has a compact intersection.

Finally we show, that each $(n+1)$ member subcollection of the family $\{K(x) \mid x \in I\}$ has a nonempty intersection. Clearly, it is enough to check the analogous property for the members of the family $\{\mathcal{K}(x) \mid x \in I\}$. Take elements $x_1 < \cdots < x_{n+1} =: v$ of I and define the generalized lines $\varphi := \varphi_1$ and $\psi := \psi_1$ as in assertion (ii). If $f(v) \le \varphi(v) \le g(v)$ or $f(v) \le \psi(v) \le g(v)$ holds, then $h = \varphi$ or $h = \psi$ belongs to the intersection $\mathcal{K}(x_1) \cap \ldots \cap \mathcal{K}(x_{n+1})$. In the opposite case, due to the inequalities (4.3), $\varphi(v) > g(v)$ and $\psi(v) < f(v)$ follow. Define the homotopy $H\colon [0,1] \times I \to$

$\mathbb{R}$ as the convex combination of φ and ψ:

$$H(t,x) := (1-t)\varphi(x) + t\psi(x).$$

The convexity of $\mathcal{B}_n(I)$ guarantees that $H(t,\cdot)$ is a generalized line for all $t \in [0,1]$. By Bolzano's theorem, there exists $t_0 \in]0,1[$ such that $f(v) \leq H(t_0,v) \leq g(v)$. In this case, $h = H(t_0,\cdot)$ is a generalized line belonging to the intersection in question.

To sum up the aboves, $\{K(x) \mid x \in I\}$ is such a collection of nonempty, convex, closed subsets of $\mathbb{R}^n$, in which there exist finite many sets of compact intersection and each $(n+1)$ member subcollection has a nonempty intersection. Helly's Theorem guarantees, that the entire intersection

$$\bigcap \{K(x) \mid x \in I\}$$

is also nonempty. Therefore, there exists an element h of $\bigcap\{\mathcal{K}(x) \mid x \in I\}$, which, by definition of the sets $\mathcal{K}(x)$, is a proper separator for f and g. □

As it turns out from the proof, inequalities (4.4) guarantee the order between f and g at the left endpoint of the domain if the endpoint belongs to I. Similarly, the order at the right endpoint (if it makes sense) is determined by (4.3). Therefore, in case of open intervals, only one of the pairs of inequalities is needed.

Now we turn to the sketch of the proof of Corollary 4.4.

HINT OF THE PROOF COROLLARY 4.4. Fix points $x_1 < \ldots < x_n$ of I and let $x_{n+1} := v \geq x_n$ be arbitrary. For our convenience, denote the generalized line φ_1 of the main result by φ. Using the representation $\varphi = \alpha_1\omega_1 + \cdots + \alpha_n\omega_n$ and the interpolation properties $\varphi(x_n) = g(x_n), \varphi(x_{n-1}) = f(x_{n-1}), \ldots$ reduce to a system of inhomogeneous linear equations with respect to $\alpha_1, \ldots, \alpha_n$. Applying Cramer's Rule,

$$\alpha_k = (-1)^{n-k}\frac{D_{f,g}(x_1,\ldots,x_n)}{D(x_1,\ldots,x_n)}$$

where $D(x_1,\ldots,x_n)$ is the determinant built from the columns $\boldsymbol{\omega}(x_1), \ldots, \boldsymbol{\omega}(x_n)$, and the other determinant $D_{f,g}(x_1,\ldots,x_n)$ is obtained for this, replacing its last row by $\ldots, f(x_{n-1}), g(x_n)$. Then, rearranging the inequality $f(x_{n+1}) \leq \varphi(x_{n+1})$ and applying the expansion theorem of determinants, we get the first inequality of the corollary. If the base points are not pairwise distinct, the inequality holds obviously. The proof of the other case is similar. □

5

Convex separation by regular pairs

5.1. Introduction and statement of the results

As a continuation of the previous chapter we recall that similarly to the characterization of those pairs of functions that can be separated by an affine function, there exists a characterization of the existence of a convex separator between two real-valued functions via single inequality. This result is proved by Baron, Matkowski and Nikodem. On the other hand, the notion of convexity can be generalized applying regular pairs (in other words, two dimensional Chebyshev systems). The main goal of the present chapter is to extend the above mentioned result to this setting. We note that results of the present chapter are appearing in [**13**].

Let I be a real interval and assume that functions $f, g\colon I \to \mathbb{R}$ are given such that there exists a convex function $h\colon I \to \mathbb{R}$ fulfilling $f \le h \le g$. Simple calculations show, that in this case, for all elements x, y of I and $\lambda \in [0,1]$, we have the inequality

$$f\big(\lambda x + (1-\lambda)y\big) \le \lambda g(x) + (1-\lambda)g(y). \tag{5.1}$$

Surprisingly, in [**6**] the Authors proved that this inequality is not merely a *necessary* but also a *sufficient* condition for the existence of a convex separator between two given functions.

It is well-known that the notion of standard convexity can be extended using Chebyshev systems (see, for example, [**44**]). The geometrical meaning of the convexity notion induced by a Chebyshev system, roughly speaking, is that each member of the system interpolating the function's graph, intersects the graph alternately. More precisely, we have the following definition:

DEFINITION 5.1. *Let I be a real interval and $\boldsymbol{\omega} := (\omega_1, \dots, \omega_n)$ be a positive Chebyshev system on I. A function $f\colon I \to \mathbb{R}$ is said to be* $\boldsymbol{\omega}$-convex *if,* *for all elements $x_0 \leq \cdots \leq x_n$ of I, we have the inequality*

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \dots & \omega_1(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \omega_n(x_0) & \omega_n(x_1) & \dots & \omega_n(x_n) \\ f(x_0) & f(x_1) & \dots & f(x_n) \end{vmatrix} \geq 0.$$

The notion of $\boldsymbol{\omega}$*-concavity* can be defined via the reversed inequality above. Clearly, a function f is $\boldsymbol{\omega}$-concave if and only if $(-f)$ is $\boldsymbol{\omega}$-convex. A function ω is termed $\boldsymbol{\omega}$*-affine*, if the determinant above (replacing f by ω) vanishes. This property is equivalent to the simultaneous $\boldsymbol{\omega}$-convexity and $\boldsymbol{\omega}$-concavity of ω; or, to the property that ω belongs to the linear hull of the components of $\boldsymbol{\omega}$.

Beside the polynomial system, appearing in the previous chapter, we note that the solution set of an nth order linear homogeneous differential equations with constant coefficients is also a typical example for Chebyshev systems. In this case, a function (having enough regularities) is convex with respect to the system if and only if it satisfies the corresponding differential inequality. For further details, consult [**75**] and [**15**].

We note that not only Chebyshev systems, but every Beckenbach family induces a (generalized) convexity notion. Not claiming completeness, we quote here the works of Beckenbach [**7**], Hopf [**42**], Popoviciu [**77**] and Tornheim [**82**]. For further details, consult the introduction of [**10**].

Throughout this chapter by a *regular pair* we mean a two dimensional positive Chebyshev system. These kinds of systems play a distinguished role among Chebyshev systems. For instance, standard convexity is induced by the regular pair $\boldsymbol{\omega} = (1, \mathrm{id})$; another important example is the pair $\boldsymbol{\omega} = (1, \exp)$ which generates the so-called *log-convexity*. The notion of *relative convexity* can also be interpreted in this framework (see [**64**]).

Now we present the main result of this chapter, which gives a characterization of the existence of an $\boldsymbol{\omega}$-convex separator between two given real functions for every regular pair $\boldsymbol{\omega}$. It states that the characterization can be given via a determinant inequality which is analogous to that of (5.1).

THEOREM 5.2. (Bessenyei, Szokol [**13**])
If $\boldsymbol{\omega} = (\omega_1, \omega_2)$ is a regular pair on a real interval I and $f, g\colon I \to \mathbb{R}$, then there exists an $\boldsymbol{\omega}$-convex function $h\colon I \to \mathbb{R}$ with $f \leq h \leq g$ if and only if,

for all elements $x_0 \le x_1 \le x_2$ *of* I,

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ g(x_0) & f(x_1) & g(x_2) \end{vmatrix} \ge 0. \tag{5.2}$$

We have to note that this result appears in [**12**], but as a *consequence* of the Baron–Matkowski–Nikodem theorem. Hence, in that approach, it cannot be considered as a *real* generalization. However, some ideas presented in [**6**] can be adapted into this setting with some suitable modifications.

It turns out that, just as with the standard convexity, the lower $\boldsymbol{\omega}$-convex envelope of the epigraph of the "upper" function results an $\boldsymbol{\omega}$-convex function. To prove this property, a characterization of $\boldsymbol{\omega}$-convex functions is applied. The inequality of the main result that corresponds to (5.1) guarantees that the $\boldsymbol{\omega}$-convex envelope possesses the required separating property. To use this inequality, the representation of $\boldsymbol{\omega}$-convex hulls is needed; this representation can be given with the help of a version of the well-known Carathéodory theorem.

Clearly, the existence of a concave separator between two given functions is equivalent to the existence of a convex separator between the negative of the functions to be separated. This observation, combined with Theorem 5.2 immediately reduces to the following result.

THEOREM 5.3. (Bessenyei, Szokol [**13**])
If $\boldsymbol{\omega} = (\omega_1, \omega_2)$ *is a regular pair on a real interval* I *and* $f, g \colon I \to \mathbb{R}$, *then there exists a* $\boldsymbol{\omega}$*-concave function* $h \colon I \to \mathbb{R}$ *with* $f \le h \le g$ *if and only if, for all elements* $x_0 \le x_1 \le x_2$ *of* I,

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ f(x_0) & g(x_1) & f(x_2) \end{vmatrix} \le 0. \tag{5.3}$$

Note also, that similar characterizations remain true when the convexity notion is induced by two-parameter Beckenbach families [**66**]. The approach followed therein is analogous to that of [**6**] (or that of presented here) with the difference that, instead of Carathéodory's theorem, a result of Kakutani is applied.

The notion of standard approximate convexity can be extended to the $\boldsymbol{\omega}$-convexity setting via the next definition. It is easy to see that the definition leads to the standard one in the particular case $\boldsymbol{\omega} = (1, \mathrm{id})$ and $\omega \equiv \varepsilon/2$.

Note that positive $\boldsymbol{\omega}$-affine functions always exist [**9**], even in the higher-dimensional setting [**44**]. In the rest of this chapter, I° stands for the interior of I.

DEFINITION 5.4. *Let $\boldsymbol{\omega} = (\omega_1, \omega_2)$ be a regular pair on a real interval I and let ω be an $\boldsymbol{\omega}$-affine function which is positive on I°. We say that $\varphi\colon I \to \mathbb{R}$ is* approximately $\boldsymbol{\omega}$-convex *with error term ω, if, for all elements $x_0 \le x_1 \le x_2$ of I,*

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ (\varphi+\omega)(x_0) & (\varphi-\omega)(x_1) & (\varphi+\omega)(x_2) \end{vmatrix} \ge 0. \tag{5.4}$$

By well-known results [**43**], standard convexity is stable. That is, approximate convex functions can be decomposed to the sum of a convex and a "small" part. Similar statement remains true for approximately $\boldsymbol{\omega}$-convex functions. In fact, this result is presented in [**9**]. However, that proof is based on the classical theorem; now, using Theorem 5.2, an independent approach can be elaborated.

THEOREM 5.5. (Bessenyei, Szokol [**13**])
If $\boldsymbol{\omega} = (\omega_1, \omega_2)$ is a regular pair on a real interval I and ω is an $\boldsymbol{\omega}$-affine function which is positive on I°, then $\varphi\colon I \to \mathbb{R}$ is approximately $\boldsymbol{\omega}$-convex with error term ω if and only if $\varphi = h + \psi$, where h is $\boldsymbol{\omega}$-convex and $|\psi(t)| \le \omega(t)$ for all $t \in I$.

Those regular pairs play a specific role that contain constant functions. The $\boldsymbol{\omega}$-convexity induced by this kind of pairs is called *relative convexity* in the technical literature (consult, for example, [**64**, pp 91–96]). In particular, multiplicative convexity can also be formulated in terms of relative convexity. Let us present here first the corresponding separation and the stability results in this context.

COROLLARY 5.6. *Let I be a real interval, $\alpha\colon I \to \mathbb{R}$ be a continuous, strictly monotone increasing function and $f, g\colon I \to \mathbb{R}$. Then, there exists a function $h\colon I \to \mathbb{R}$ satisfying $f \le h \le g$ and, for all elements $x_0 \le x_1 \le x_2$ of I the inequality*

$$\begin{aligned}\big(\alpha(x_2)-\alpha(x_0)\big)h(x_1) \le \big(\alpha(x_2)-\alpha(x_1)\big)&h(x_0)\\ + \big(\alpha(x_1)-\alpha(x_0)\big)&h(x_2)\end{aligned} \tag{5.5}$$

holds, if and only if, for all $x_0 \le x_1 \le x_2$ of I, we have

$$\big(\alpha(x_2)-\alpha(x_0)\big)f(x_1) \le \big(\alpha(x_2)-\alpha(x_1)\big)g(x_0) + \big(\alpha(x_1)-\alpha(x_0)\big)g(x_2).$$

COROLLARY 5.7. *Let I be a real interval, $\alpha\colon I \to \mathbb{R}$ be a continuous, strictly monotone increasing function and $\varepsilon > 0$. Then, a function $\varphi\colon I \to \mathbb{R}$ satisfies the inequality*

$$\begin{aligned}\big(\alpha(x_2)-\alpha(x_0)\big)\varphi(x_1) \;&\le\; \big(\alpha(x_2)-\alpha(x_1)\big)\varphi(x_0)\\ &+\; \big(\alpha(x_1)-\alpha(x_0)\big)\varphi(x_2)+2\varepsilon\big(\alpha(x_2)-\alpha(x_0)\big)\end{aligned}$$

for all elements $x_0 \le x_1 \le x_2$ of I if and only if there exist functions $h,\psi\colon I \to \mathbb{R}$ such that $\varphi = h+\psi$, where h fulfills (5.5) *and $\|\psi\| \le \varepsilon$. (Here $\|\cdot\|$ stands for the supremum norm.)*

Observe, that in the standard setting, when $\boldsymbol{\omega} = (1,\mathrm{id})$ Corollary 5.6 reduces to the main result of [**6**]; similarly, Corollary 5.7 gives the one dimensional particular case of the stability of classical convexity [**43**]. Moreover, these corollaries involve the case of log-convexity (that is, when $\boldsymbol{\omega} = (1,\exp)$). For some other aspects of stability, we refer to the papers Páles [**70**] and Házy–Páles [**39**].

It can easily be checked that the pair $\boldsymbol{\omega} = (\cos,\sin)$ is a regular one on the interval $]-\pi/2,\pi/2[$. This pair and the induced convexity notion is important in the study of convex curves and the behavior of analytic functions in certain domains. For details, see [**76**]. As further applications, we give examples for Theorem 5.2 and Theorem 5.5 in this setting.

COROLLARY 5.8. *If $I \subset]-\frac{\pi}{2},\frac{\pi}{2}[$ is a real interval, $f,g\colon I \to \mathbb{R}$, then there exists a $(\cos,\sin)$-convex function $h\colon I \to \mathbb{R}$ with $f \le h \le g$ if and only if, for all $\lambda \in [0,1]$ and $x \le y$ of I,*

$$\sin(y-x)f\big(\lambda x+(1-\lambda)y\big) \le \sin\big(\lambda(y-x)\big)g(x)+\sin\big((1-\lambda)(y-x)\big)g(y).$$

COROLLARY 5.9. *If $I \subset]-\frac{\pi}{2},\frac{\pi}{2}[$ is a real interval, then $\varphi\colon I \to \mathbb{R}$ is approximately $(\cos,\sin)$-convex with error term $\varepsilon\cdot\cos$ if and only if $\varphi = h+\psi$, where $h\colon I \to \mathbb{R}$ is $(\cos,\sin)$-convex, and $\psi\colon I \to \mathbb{R}$ satisfies $|\psi(t)| \le \varepsilon\cdot\cos(t)$ for all $t \in I$.*

The pair $(\cosh,\sinh)$ form a regular one over arbitrary real intervals. In this case, the analogue statements of Corollary 5.8 and Corollary 5.9 can also be formulated, replacing the involved functions $\cos$ and $\sin$ by $\cosh$ and $\sinh$, respectively.

5.2. Proofs

In this section our first aim is to prove Theorem 5.2. To do this we need three auxiliary lemmas that are applied in the proof of the main result. These lemmas are based on suitable generalizations of some notions of convex geometry. In the particular setting $\boldsymbol{\omega} = (1, \mathrm{id})$, the definition below leads to usual convex combination, convex hull and convex sets of the plane.

DEFINITION 5.10. *Let I be a real interval and $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)$ be a Chebyshev system on I. A point (x, y) of $I \times \mathbb{R}$ is said to be the $\boldsymbol{\omega}$-*convex combination *of some given points $(x_0, y_0), \dots, (x_m, y_m)$ of $I \times \mathbb{R}$ with coefficients $\lambda_0, \dots, \lambda_m$, if the coefficients are nonnegative, and*

$$\sum_{k=0}^{m} \lambda_k \boldsymbol{\omega}(x_k) = \boldsymbol{\omega}(x), \qquad \sum_{k=0}^{m} \lambda_k y_k = y.$$

*If $H \subset I \times \mathbb{R}$, then the $\boldsymbol{\omega}$-*convex hull $\mathrm{conv}_{\boldsymbol{\omega}}(H)$ *is the set of all $\boldsymbol{\omega}$-convex combinations of the elements of H. We say that H is $\boldsymbol{\omega}$-convex, if $H =$* $\mathrm{conv}_{\boldsymbol{\omega}}(H)$.

The motivation of the definition is the following. Replace the affine segments of $I \times \mathbb{R}$ by the $\boldsymbol{\omega}$-affine segments, that is, the linear hull of a regular pair $\boldsymbol{\omega}$. Fix points $p_0 = (x_0, y_0)$ and $p_1 = (x_1, y_1)$. Simple calculations show that the $\boldsymbol{\omega}$-affine segment joining p_0 and p_1 and the $\boldsymbol{\omega}$-convex combinations of p_0 and p_1 *coincide* since their points $p = (x, y)$ simultaneously fulfill the same identity

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x) & \omega_1(x_1) \\ \omega_2(x_0) & \omega_2(x) & \omega_2(x_1) \\ y_0 & y & y_1 \end{vmatrix} = 0.$$

The first technical lemma is an analogue of the classical Carathéodory theorem [**47**]. Although they are not a direct consequence of each other, their proofs are quite similar.

LEMMA 5.11. *If $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)$ is a Chebyshev system over a real interval I, $H \subset I \times \mathbb{R}$, and $(x, y) \in \mathrm{conv}_{\boldsymbol{\omega}}(H)$, then there exist points $(x_0, y_0) \dots, (x_n, y_n)$ of H such that*

$$(x, y) \in \mathrm{conv}_{\boldsymbol{\omega}}\{(x_k, y_k) \mid k = 0, \dots, n\}.$$

PROOF. Let (x, y) be an element of $\mathrm{conv}_{\boldsymbol{\omega}}(H)$. Then, there exist nonnegative coefficients $\lambda_0, \dots, \lambda_m$ and elements $(x_0, y_0), \dots, (x_m, y_m) \in H$ such that

$$\sum_{k=0}^{m} \lambda_k \boldsymbol{\omega}(x_k) = \boldsymbol{\omega}(x), \qquad \sum_{k=0}^{m} \lambda_k y_k = y.$$

Without loss of generality (and for our convenience) we may assume that $m = n + 1$. Our aim is to show that the parameter $(n + 1)$ can be reduced to at most n. Denote the vectors $(\boldsymbol{\omega}(x_k), y_k)$ and $(\boldsymbol{\omega}(x), y)$ by a_k and b, respectively. Since $a_0, \dots, a_{n+1}$ belong to the euclidean $(n+1)$-space, there exist nontrivial scalars $\mu_0, \dots, \mu_{n+1}$ such that $\mu_0 a_0 + \cdots + \mu_{n+1} a_{n+1} = 0$. Multiplying both sides by an arbitrary parameter α and taking into consideration the fact that b can be expressed as a linear combination of the vectors $a_0, \dots, a_{n+1}$, we arrive at

$$\sum_{k=0}^{n+1} (\lambda_k - \alpha\mu_k) a_k = b.$$

We may assume that there exists at least one coefficient μ_k which is positive. Therefore, in particular, we can specify α as

$$\alpha := \min\left\{\frac{\lambda_k}{\mu_k} \mid \mu_k > 0\right\}.$$

Then, $\lambda_k - \alpha\mu_k \geq 0$, and there exists an index j such that $\lambda_j - \alpha\mu_j = 0$. Interchanging the indices (if necessary) suppose $j = n+1$. Then, $b = \nu_0 a_0 + \cdots + \nu_n a_n$; that is,

$$\sum_{k=0}^{n} \nu_k \boldsymbol{\omega}(x_k) = \boldsymbol{\omega}(x), \qquad \sum_{k=0}^{n} \nu_k y_k = y,$$

yielding the desired reduction. □

The second technical lemma gives a characterization of functions that are $\boldsymbol{\omega}$-convex with respect to a regular pair. In fact, many properties are known that characterize these kind of functions. For details, just as for the proof of the lemma, consult [**11**] or [**9**].

LEMMA 5.12. *Let $\boldsymbol{\omega} = (\omega_1, \omega_2)$ be a regular pair on the nonempty interval I such that ω_1 is positive on I°. Then, $h\colon I \to \mathbb{R}$ is $\boldsymbol{\omega}$-convex if and only if, for all $x, x_1, x_2 \in I$ and $\lambda_1, \lambda_2 \geq 0$ satisfying the conditions $\lambda_1\boldsymbol{\omega}(x_1) + \lambda_2\boldsymbol{\omega}(x_2) = \boldsymbol{\omega}(x)$ we have that*

$$h(x) \leq \lambda_1 h(x_1) + \lambda_2 h(x_2).$$

Note also, that the assumption of the positivity on the first component does not yield serious restriction [**11**]. Namely, every regular pair $\boldsymbol{\omega}$ given on an interval I can be replaced by another one denoted by $\boldsymbol{\omega}^*$ such that the first member of $\boldsymbol{\omega}^*$ is positive on the interior of I; moreover, $\boldsymbol{\omega}$ and $\boldsymbol{\omega}^*$ generate the same convexity notion.

The third technical lemma is a quite simple observation. However, its $m = 1$ case has some role in the proof of the main result. This particular case says that if (x, y) is an $\boldsymbol{\omega}$-convex combination of points (x_0, y_0) and (x_1, y_1), then x is between x_0 and x_1.

LEMMA 5.13. *If* $\boldsymbol{\omega} = (\omega_1, \omega_2)$ *is a regular pair on a real interval* I *and*

$$(x, y) \in \mathrm{conv}_{\boldsymbol{\omega}}\{(x_0, y_0), \ldots, (x_m, y_m)\},$$

then

$$\min\{x_0, \ldots, x_m\} \le x \le \max\{x_0, \ldots, x_m\}.$$

PROOF. Assume indirectly that $x < \min\{x_0, \ldots, x_n\}$ where (x, y) is the $\boldsymbol{\omega}$-convex combination of the points (x_k, y_k) with nonnegative coefficients λ_k. Without loss of generality, we may assume that at least one of the coefficients is positive. Since $\boldsymbol{\omega}$ is a special Chebyshev system, we have the inequality $\omega_1(x)\omega_2(x_k) > \omega_1(x_k)\omega_2(x)$ for all $k = 0, \ldots, m$. Hence

$$\omega_1(x) \sum_{k=0}^{m} \lambda_k \omega_2(x_k) > \omega_2(x) \sum_{k=0}^{m} \lambda_k \omega_1(x_k)$$

follows. On the other hand, by the definition of $\boldsymbol{\omega}$-convex combinations, both sides have the common value $\omega_1(x)\omega_2(x)$, which is a contradiction. The indirect assumption to the other case can be treated similarly. □

Now, we are in a position to prove Theorem 5.2. For our technical convenience, we shall use in the proof the next abbreviation:

$$D(u, v) := \begin{vmatrix} \omega_1(u) & \omega_2(u) \\ \omega_1(v) & \omega_2(v) \end{vmatrix}.$$

PROOF OF THEOREM 5.2. Assume h is a generalized convex separator between f and g. Then replacing f and g in the last row of the determinant appearing in the left-hand side of the inequality (5.2) by h, we get a lower estimation. Moreover, the modified determinant, according to the $\boldsymbol{\omega}$-convex property of h, is nonnegative.

Conversely, assume that f and g satisfy the condition of the theorem. Substitute $x = x_0 = x_1$ and $y = x_2$ into (5.2), and then expand the determinant with respect to the third row. Then, the coefficient of $g(y)$ vanishes; rearranging and then simplifying the obtained inequality with the common positive coefficient, $f(x) \le g(x)$ follows. Denote the $\boldsymbol{\omega}$-convex hull of the epigraph of g by A; that is,

$$A := \mathrm{conv}_{\boldsymbol{\omega}}\{(x, y) \in I \times \mathbb{R} \mid g(x) \le y\}.$$

Let $(x, y) \in A$ be fixed. Applying the special case $n = 2$ of Lemma 5.11, (x, y) belongs to a generalized $\boldsymbol{\omega}$-simplex which vertices (x_0, y_0), (x_1, y_1),

(x_2, y_2) are elements of the epigraph of g:

$$(x, y) \in \operatorname{conv}_{\boldsymbol{\omega}}\{(x_0, y_0), (x_1, y_1), (x_2, y_2)\} =: S_{\boldsymbol{\omega}}.$$

Define

$$z_0 := \inf\{z \in \mathbb{R} \mid (x, z) \in S_{\boldsymbol{\omega}}\}.$$

Then, $z_0 \leq y$ and (x, z_0) belongs to the boundary of $S_{\boldsymbol{\omega}}$. For simplicity we may assume that $x_0 < x_1 < x_2$. Then, (x, z_0) is the $\boldsymbol{\omega}$-convex combination of (x_0, y_0) and (x_2, y_2). Let $\omega\colon I \to \mathbb{R}$ be the generalized line interpolating these points. Then, $\omega = \alpha_1\omega_1 + \alpha_2\omega_2$, where

$$\alpha_1 = \frac{1}{D(x_0, x_2)} \begin{vmatrix} y_0 & \omega_2(x_0) \\ y_2 & \omega_2(x_2) \end{vmatrix}; \qquad \alpha_2 = \frac{1}{D(x_0, x_2)} \begin{vmatrix} \omega_1(x_0) & y_0 \\ \omega_1(x_2) & y_2 \end{vmatrix}.$$

Suppose indirectly that $z_0 < f(x)$. Since $z_0 = \omega(x)$ and the determinant $D(x_0, x_2)$ is positive, $D(x_0, x_2)f(x) > D(x_0, x_2)\omega(x)$ follows. By Lemma 5.13, we have $x_0 < x < x_2$ follows, which guarantees the positivity of $D(x_0, x)$ and $D(x, x_2)$. Hence, applying the representation of ω, we arrive at

$$\begin{aligned} D(x_0, x_2)f(x) &> \begin{vmatrix} y_0 & \omega_2(x_0) \\ y_2 & \omega_2(x_2) \end{vmatrix} \omega_1(x) + \begin{vmatrix} \omega_1(x_0) & y_0 \\ \omega_1(x_2) & y_2 \end{vmatrix} \omega_2(x) \\ &= (\omega_1(x)\omega_2(x_2) - \omega_2(x)\omega_1(x_2))y_0 \\ &\quad +(\omega_1(x_0)\omega_2(x) - \omega_2(x_0)\omega_1(x))y_2 \\ &= D(x, x_2)y_0 + D(x_0, x)y_2 \\ &\geq D(x, x_2)g(x_0) + D(x_0, x)g(x_2). \end{aligned}$$

Thus

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x) & \omega_2(x_2) \\ g(x_0) & f(x) & g(x_2) \end{vmatrix} < 0,$$

which contradicts to (5.2). Consider now the function $h\colon I \to \mathbb{R}$ given by the formula

$$h(x) := \inf\{z \in \mathbb{R} \mid (x, z) \in A\}.$$

The previous arguing shows, that h is a function, indeed. Moreover, according to the aboves again, $f \leq h$ also holds, while the other inequality $h \leq g$ follows immediately by the construction.

In the rest of the proof, we shall verify the $\boldsymbol{\omega}$-convexity of h. Let (x_1, y_1) and (x_2, y_2) be arbitrary elements of A, and consider a point (x, y) fulfilling the equations

$$\begin{aligned} \lambda_1\omega_1(x_1) + \lambda_2\omega_1(x_2) &= \omega_1(x), \\ \lambda_1\omega_2(x_1) + \lambda_2\omega_2(x_2) &= \omega_2(x), \\ \lambda_1 y_1 + \lambda_2 y_2 &= y \end{aligned}$$

with nonnegative coefficients λ_1, λ_2. The $\boldsymbol{\omega}$-convexity of A implies $(x, y) \in A$ and hence $h(x) \le y$. Taking infimum in y_1 and y_2, we get $h(x) \le \lambda_1 h(x_1) + \lambda_2 h(x_2)$. In view of Lemma 5.12, this implies the $\boldsymbol{\omega}$-convexity of h, and the proof is completed. □

As we mentioned in the introduction the existence of a concave separator between two given functions is equivalent to the existence of a convex separator between the negative of the functions to be separated. Therefore we skip the proof of Theorem 5.3. Moreover, we note that Theorem 5.3 can be proved directly, applying analogous arguments appearing in the proof of Theorem 5.2.

Using Theorem 5.2 we are in a position to prove Theorem 5.5.

PROOF OF THEOREM 5.5. For necessity, assume that $\varphi = h + \psi$, where h is $\boldsymbol{\omega}$-convex and $|\psi(t)| \le \omega(t)$ for all $t \in I$. Then, $\psi + \omega \ge 0$ and $\psi - \omega \le 0$. Hence, using the representation $\varphi = h + \psi$ and the $\boldsymbol{\omega}$-convexity of h, we arrive at

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ (\varphi+\omega)(x_0) & (\varphi-\omega)(x_1) & (\varphi+\omega)(x_2) \end{vmatrix} \ge$$

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ h(x_0) & h(x_1) & h(x_2) \end{vmatrix} \ge 0.$$

For sufficiency, assume that φ is approximately $\boldsymbol{\omega}$-convex with error term ω. Then, the functions $f := \varphi - \omega$ and $g := \varphi + \omega$ satisfy (5.2) and hence there exists an $\boldsymbol{\omega}$-convex separator h fulfilling $f \le h \le g$. In other words, $\varphi - \omega \le h \le \varphi + \omega$. Therefore the function $\psi := \varphi - h$ is a proper choice for the required decomposition. □

Among corollaries we present only the proofs of the ones concerning relative convexity (Corollary 5.6 and Corollary 5.7).

PROOF OF COROLLARY 5.6. The properties of α guarantee that $\boldsymbol{\omega} = (1, \alpha)$ is a regular pair over I. Hence the statement immediately follows from the definition of $\boldsymbol{\omega}$-convexity and Theorem 5.2, expanding the determinants involved. □

PROOF OF COROLLARY 5.7. The properties of α guarantee again that $\boldsymbol{\omega} = (1, \alpha)$ is a regular pair over I whose linear hull contains the positive (constant) element ε. Simple calculations show that the inequality of the Corollary can be written into the form (5.4) with $\omega \equiv \varepsilon$. Hence Theorem 5.5 can be applied using the fact that $|\psi(t)| \leq \varepsilon$ is equivalent to $\|\psi\| \leq \varepsilon$. □

5.3. Concluding remarks

Finally, we pose some open problems concerning theorems that appear in the last two chapters. Firstly, we mention a problem which gives a connection between the convex and affine separation problems. As we have learnt at the beginning of Chapter 4, there exists a characterization theorem for the existence of an affine separator between two given functions [**67**]. This characterization is given via a double inequality (4.1). This method works in more general settings and reduces to analogous results; for example, as we have seen in Corollary 4.5 and Corollary 4.6, when the convexity notion is induced by the polynomial system [**87**] or by an arbitrary Chebyshev system [**3**]. The paper [**12**] is also devoted to investigate affine separation problems for Chebyshev systems. Denoting the column $\big(\omega_1(x), \dots, \omega_n(x)\big)$ by $\boldsymbol{\omega}(x)$, its main result reads as follows.

THEOREM 5.14. (Bessenyei, Páles)
Let H be a real subset of at least n elements and $f, g\colon H \to \mathbb{R}$. Then, the following statements are equivalent:

(i) There exists an $\boldsymbol{\omega}$-affine function $\omega\colon H \to \mathbb{R}$ such that $f \leq \omega \leq g$;
(ii) there exists an $\boldsymbol{\omega}$-concave function $\varphi\colon H \to \mathbb{R}$ and an $\boldsymbol{\omega}$-convex function $\psi\colon H \to \mathbb{R}$ satisfying the inequalities $f \leq \varphi \leq g$ and $f \leq \psi \leq g$;
(iii) for all elements $x_0 \leq \dots \leq x_n$ of H,

$$\begin{vmatrix} \dots & \boldsymbol{\omega}(x_{n-3}) & \boldsymbol{\omega}(x_{n-2}) & \boldsymbol{\omega}(x_{n-1}) & \boldsymbol{\omega}(x_n) \\ \dots & g(x_{n-3}) & f(x_{n-2}) & g(x_{n-1}) & f(x_n) \end{vmatrix} \leq 0;$$

$$\begin{vmatrix} \dots & \boldsymbol{\omega}(x_{n-3}) & \boldsymbol{\omega}(x_{n-2}) & \boldsymbol{\omega}(x_{n-1}) & \boldsymbol{\omega}(x_n) \\ \dots & f(x_{n-3}) & g(x_{n-2}) & f(x_{n-1}) & g(x_n) \end{vmatrix} \geq 0.$$

These kinds of separation theorems have a particular importance in convex analysis [**50**]. If $H = I$ is a real interval and the Chebyshev system is a regular pair, the inequalities of the third item are exactly (5.2) and (5.3). Hence the theorem above states that the simultaneous inequalities characterize the existence of an ω-affine separator, while, as a counterpart, Theorem 5.2 and Theorem 5.3 show that the separated inequalities (in the two dimensional case) are responsible for the existence of convex/concave separations. Therefore the question arises, quite evidently, whether this phenomenon remains true in the general case. Till now, the answer is not known and could be the topic of further research.

Now we present some further open problems concerning Theorem 4.3. By a result of Nikodem and Páles it is known that if we have two parameters, then the separation can be characterized under no further assumption on the underlying Beckenbach family. Hence the natural question arises, if the convex-closedness in Theorem 4.3 is redundant or not.

In fact, the main result of [**12**] considering affine separation problems for Chebyshev systems remains valid also for Haar systems. This suggest that Theorem 4.3 could be studied within the framework of interpolation families instead of Beckenbach families.

Let us point out here to the work of Krzyszkowski. In his paper [**45**], he introduced generalized convex sets of $I \times \mathbb{R}$ using two parameter Beckenbach families. By Lemma 4.9 regular pairs can be considered as Beckenbach families that are closed under linear combinations, and the generalized convexity notion, appearing in the present chapter, is a special case to that of due to Krzyszkowski.

The above mentioned Author also defined some stability notion and obtained results [**46**]. However, the connection between his and our stability notions is not direct; the clarification of this connection is over the framework of the present note. The results of Krzyszkowski deeply influenced also the paper of Nikodem and Páles [**66**].

Summary

The present dissertation contains results about preserver problems on different mathematical structures and separation problems. It consists of an introduction, five chapters, a summary (both in English and in Hungarian) and a bibliography. In the introduction, we present some important and well-known preserver results which are closely related to the theorems appearing in the thesis. Moreover, we introduce the so-called separation problems and we collect the fundamental theorems which give the motivation for our results on separation problems. Chapters 1-3 are dealing with certain preserver results on different kinds of mathematical structures. In what follows we summarize them.

In Chapter 1 we study preserver transformations on density operators (i.e. positive operators with unit trace) that play an important role in quantum information theory. To present our results we introduce some notation. Let H be a finite dimensional Hilbert space and $B(H)$ denote the set all bounded linear operators acting on H. We denote by $B(H)^+$ the cone of all positive semi-definite operators on H. Finally, $S(H)$ stands for the set of all density operators. Relative entropy is a fundamental notion in quantum information theory which has several versions. The most common one is the Umegaki relative entropy which is defined by

$$S(A||B) = \begin{cases} \operatorname{tr} A(\log A - \log B), & \operatorname{supp} A \subset \operatorname{supp} B \\ \infty, & \text{otherwise,} \end{cases}$$

for all $A, B \in S(H)$. In [**56**] L. Molnár described the structure of all surjective transformations that leave the Umegaki relative entropy invariant. He proved that the corresponding transformations are induced by a unitary or antiunitary operator. In Chapter 1 we show that his result remains true without assuming that the transformation is surjective.

THEOREM. (Molnár, Szokol)
Let $\phi\colon S(H) \to S(H)$ be a transformation which preserves the Umegaki

relative entropy, i.e. which satisfies

$$S(\phi(A)||\phi(B)) = S(A||B), \quad A, B \in S(H).$$

Then there exists either a unitary or an antiunitary operator U on H such that ϕ is of the form

$$\phi(A) = UAU^*, \quad A \in S(H).$$

Motivated by the previous theorems in the main theorem of Chapter 1 we describe all transformations on $S(H)$ that preserve the quantum f-divergence with respect to an arbitrary strictly convex function f defined on the non-negative real line. It is well-known that with a particular choice of the function f the definition of quantum f-divergence leads to the notion of Umegaki relative entropy. Hence, the main result gives a far-reaching generalization of the theorem concerning the structure of all Umegaki relative entropy preserving maps. Let $f\colon [0, \infty[\to \mathbb{R}$ be a function which is continuous on $]0, \infty[$ and the limit

$$\alpha := \lim_{x\to\infty} \frac{f(x)}{x}$$

exists in $[-\infty, \infty]$. Let $A, B \in B(H)^+$ and for any $\lambda \in \mathbb{R}$ denote by P_λ, respectively by Q_λ the projection on H projecting onto the kernel of $A-\lambda I$, respectively onto the kernel of $B - \lambda I$. The quantum f-divergence between A and B can be given by the formula

$$S_f(A||B) = \sum_{a\in\sigma(A)} \left(\sum_{b\in\sigma(B)\setminus\{0\}} bf\left(\frac{a}{b}\right) \operatorname{tr} P_a Q_b + \alpha a \operatorname{tr} P_a Q_0 \right),$$

where $\sigma(.)$ stands for the spectrum of elements in $B(H)$ and the convention $0\cdot(-\infty) = 0\cdot\infty = 0$ is used. Now, we are in a position to present the main result of Chapter 1.

THEOREM. (Molnár, Nagy, Szokol)
Assume that $f\colon [0, \infty[\to \mathbb{R}$ is a strictly convex function and $\phi\colon S(H) \to S(H)$ is a transformation satisfying

$$S_f(\phi(A)||\phi(B)) = S_f(A||B), \quad A, B \in S(H).$$

Then there is either a unitary or an antiunitary operator U on H such that ϕ is of the form

$$\phi(A) = UAU^*, \quad A \in S(H).$$

In Chapter 2 we substantially extend and unify former results on the structure of surjective isometries of spaces of positive definite matrices obtained in the paper [**62**]. The novelty in our result is that we consider not only true metrics but so-called generalized distance measures which are parameterized by unitarily invariant norms and continuous real functions satisfying certain conditions. By a generalized distance measure we mean a function $d : X \times X \to [0, \infty[$ (X is any set) which has the definiteness property (for arbitrary $x, y \in X$ we have $d(x, y) = 0$ if and only if $x = y$), but neither the symmetry of d nor the triangle inequality for d is assumed. In the following $\mathbb{M}_n$ stands for the set of all $n \times n$ complex matrices and we denote by $\mathbb{P}_n$ the set of all $n \times n$ positive definite matrices. Moreover, let $\mathbb{P}_n^1$, respectively $\mathbb{P}_n^c$ ($c > 0$) denote the set of all elements in $\mathbb{P}_n$ with unit determinant, respectively with determinant equal to c. Our main result can be formulated in the following way.

THEOREM. (Molnár, Szokol)
Let N be a unitarily invariant norm on $\mathbb{M}_n$. Assume $f\colon]0, \infty[\to \mathbb{R}$ is a continuous function such that

(a1) *$f(y) = 0$ holds if and only if $y = 1$;*
(a2) *there exists a number $K > 1$ such that*

$$|f(y^2)| \geq K|f(y)|, \quad y \in]0, \infty[.$$

Define $d_{N,f}\colon \mathbb{P}_n \times \mathbb{P}_n \to [0, \infty[$ by

$$d_{N,f}(A, B) = N(f(A^{-1/2}BA^{-1/2})), \quad A, B \in \mathbb{P}_n. \tag{6.1}$$

Assume that $n \geq 3$. If $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ is a surjective map which leaves $d_{N,f}(.,.)$ invariant, i.e., which satisfies

$$d_{N,f}(\phi(A), \phi(B)) = d_{N,f}(A, B), \quad A, B \in \mathbb{P}_n,$$

then there exist an invertible matrix $T \in \mathbb{M}_n$ and a real number c such that ϕ is of one of the following forms

$$\phi(A) = (\det A)^c TAT^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = (\det A)^c TA^{-1}T^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = (\det A)^c TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = (\det A)^c T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n.$$

Apparently, the function $d_{N,f}(.,.)$ appearing in the theorem is a generalized distance measure in the sense we introduced above. Moreover, we note that the metrics appearing in [**62**] can be obtained as particular cases of generalized distance measures and the functions f in (6.1) which correspond to those metrics have the properties (a1), (a2) listed in the theorem.

We emphasize that our main result applies for many other generalized distance measures. For any $A, B \in \mathbb{P}_n$ let $Y_{A,B}$ denote the positive definite matrix $A^{-1/2}BA^{-1/2}$. Then these quantities can be obtained as follows.

(i) Stein's loss: $l(A, B) = \|Y_{A,B}^{-1} - \log Y_{A,B}^{-1} - 1\|_1$;

(ii) Jeffrey's Kullback-Leibler divergence:

$$S_{JKL}(A, B) = \left\| \frac{Y_{A,B} + Y_{A,B}^{-1} - 2I}{2} \right\|_1 ;$$

(iii) log-determinant α-divergence (for any parameter $-1 < \alpha < 1$):

$$D_{LD}^{\alpha}(A, B) =$$

$$\frac{4}{1-\alpha^2} \left\| \log \frac{(1-\alpha)I + (1+\alpha)Y_{A,B}}{2} - \frac{1+\alpha}{2} \log Y_{A,B} \right\|_1 ,$$

where $\|.\|_1$ stands for the trace-norm, which is unitarily invariant. One can check easily that for the previous examples our theorem applies. We must point out that in the particular choices of the unitarily invariant norm N and real function f, after the use of our theorem one may need to make further steps in order to determine the precise structure of particular distance measure preservers. In accordance with this we present the complete structural result for the measures we have discussed above.

THEOREM. (Molnár, Szokol)
Let $div(.,.)$ denote any of the functions $l(.,.)$, $D_{LD}^{\alpha}(.,.)$, $-1 \leq \alpha \leq 1$. A surjective map $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ preserves $div(.,.)$, i.e., satisfies

$$div(\phi(A), \phi(B)) = div(A, B), \qquad A, B \in \mathbb{P}_n,$$

if and only if there exists an invertible matrix $T \in \mathbb{M}_n$ such that ϕ is of one of the forms

$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n$;
$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n$.

A surjective map $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ preserves $S_{JKL}(.,.)$, if and only if there exists an invertible matrix $T \in \mathbb{M}_n$ such that ϕ is of one of the forms

$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n$;
$\phi(A) = TA^{-1}T^*, \quad A \in \mathbb{P}_n$;
$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n$;
$\phi(A) = T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n$.

In Chapter 3 we also present results concerning similar preserver transformations defined on the subset $\mathbb{P}_n^1$ or $\mathbb{P}_n^c$, respectively. In fact, to prove those results we need to determine the structure of all continuous Jordan triple automorphisms of $\mathbb{P}_n^1$ (i.e., continuous bijections respecting the Jordan triple product ABA).

THEOREM. (Molnár, Szokol)
Assume $n \geq 3$. Let $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ be a continuous map which is a Jordan triple automorphism, i.e., ϕ is a continuous bijective map which satisfies

$$\phi(ABA) = \phi(A)\phi(B)\phi(A), \qquad A, B \in \mathbb{P}_n^1.$$

Then there is a unitary matrix $U \in \mathbb{M}_n$ such that ϕ is of one of the following forms

$$\begin{aligned}
&\phi(A) = UAU^*, \quad A \in \mathbb{P}_n^1;\\
&\phi(A) = UA^{-1}U^*, \quad A \in \mathbb{P}_n^1;\\
&\phi(A) = UA^{\mathrm{tr}}U^*, \quad A \in \mathbb{P}_n^1;\\
&\phi(A) = U(A^{\mathrm{tr}})^{-1}U^*, \quad A \in \mathbb{P}_n^1.
\end{aligned}$$

Our result on the form of surjective transformations of $\mathbb{P}_n^1$ leaving a generalized distance measure $d_{N,f}$ invariant reads as follows.

THEOREM. (Molnár, Szokol)
Let N be a unitarily invariant norm on $\mathbb{M}_n$ and $f\colon]0,\infty[\to \mathbb{R}$ be a continuous function which satisfies the conditions (a1), (a2). Assume that $n \geq 3$. Let $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ be a surjective map which preserves $d_{N,f}(.,.)$. Then there exists an invertible matrix T with $|\det T| = 1$ such that ϕ is of one of the following forms

$$\begin{aligned}
&\phi(A) = TAT^*, \quad A \in \mathbb{P}_n^1;\\
&\phi(A) = TA^{-1}T^*, \quad A \in \mathbb{P}_n^1;\\
&\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n^1;\\
&\phi(A) = T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n^1.
\end{aligned}$$

From this theorem we easily deduce the following corollary.

COROLLARY. *Let N, f be as in the previous theorem and assume $n \geq 3$ and c is a positive real number. If $\phi\colon \mathbb{P}_n^c \to \mathbb{P}_n^c$ is a surjective map which preserves $d_{N,f}(.,.)$, then there exists an invertible matrix T with $|\det T| = 1$ such that ϕ is of one of the following forms*

$$\begin{aligned}
&\phi(A) = TAT^*, \quad A \in \mathbb{P}_n^c;\\
&\phi(A) = \lambda^2 TA^{-1}T^*, \quad A \in \mathbb{P}_n^c;\\
&\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n^c;\\
&\phi(A) = \lambda^2 T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n^c,
\end{aligned}$$

where $\lambda = \sqrt[n]{c}$.

The study of linear isometries of linear function spaces has also been an extensive research area in functional analysis. However, there are some important metric spaces of functions which are not linear spaces. For example, the set of all probability distribution functions on $\mathbb{R}$ that plays so fundamental role in probability theory and statistics is not a linear space. In [**26**] the general forms of surjective isometries of the space $D(\mathbb{R})$ of all probability distribution function on $\mathbb{R}$ are determined with respect to Kolmogorov-Smirnov metric. In Chapter 3 we extend the mentioned result to a larger space $\Delta(\mathbb{R})$ of all generalized probability distribution functions. By a generalized distribution function we mean a function from $\mathbb{R}$ to $[0,1]$ which is monotone increasing and continuous from the right without restrictions on its limits at $\pm\infty$. For any pair f, g of $D(\mathbb{R})$ the Kolmogorov-Smirnov distance between them is defined by

$$\rho(f,g) = \sup_{t\in\mathbb{R}} |f(t) - g(t)|.$$

The first result of Chapter 3 shows that the structure of surjective Kolmogorov-Smirnov isometries of $\Delta(\mathbb{R})$ is formally the same as that of the surjective isometries of $(D(\mathbb{R}), \rho)$.

THEOREM. (Molnár, Szokol)
Let $\phi\colon \Delta(\mathbb{R}) \to \Delta(\mathbb{R})$ be a surjective isometry with respect to the Kolmogorov-Smirnov metric, i.e. assume that ϕ is a bijective map with the property that

$$\rho(\phi(f), \phi(g)) = \rho(f,g)$$

holds for all $f, g \in \Delta(\mathbb{R})$. Then either there exists a strictly increasing bijection $\varphi\colon \mathbb{R} \to \mathbb{R}$ such that ϕ is of the form

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in \Delta(\mathbb{R})$$

or there exists a strictly decreasing bijection $\psi\colon \mathbb{R} \to \mathbb{R}$ such that ϕ is of the form

$$\phi(f)(t) = 1 - f(\psi(t)-), \quad t \in \mathbb{R}, f \in \Delta(\mathbb{R}).$$

Moreover, in [**58**] Molnár studied the surjective Kolmogorov-Smirnov isometries of the space of all continuous elements of $D(\mathbb{R})$. Motivated by this result we also described the structure of all surjective isometries of the set $\Delta_c(\mathbb{R})$ of all continuous generalized distribution functions. The corresponding result reads as follows.

THEOREM. (Molnár, Szokol)
Let $\phi\colon \Delta_c(\mathbb{R}) \to \Delta_c(\mathbb{R})$ be a surjective isometry with respect to the Kolmogorov-Smirnov metric. Then either there exists a strictly increasing bijection

$\varphi\colon \mathbb{R}\to\mathbb{R}$ *such that* ϕ *is of the form*

$$\phi(f)(t) = f(\varphi(t)), \quad t\in\mathbb{R}, f\in\Delta_c(\mathbb{R})$$

or there exists a strictly decreasing bijection $\psi\colon \mathbb{R}\to\mathbb{R}$ *such that* ϕ *is of the form*

$$\phi(f)(t) = 1 - f(\psi(t)), \quad t\in\mathbb{R}, f\in\Delta_c(\mathbb{R}).$$

Chapter 4 and Chapter 5 are devoted to the investigation of so-called separation problems. It is a well-known separation theorem that if a convex and a concave function are given such that the convex function is "above" the concave one, then there exists an affine function between them. Moreover, those pairs of real functions that can be separated by an affine function was characterized via double inequalities. This result is due to Nikodem and Wąsowicz and it appeared in [**67**]. In Chapter 4 we generalize this result. More precisely, we characterize such pairs of real valued functions that can be separated by a member of a given convex Beckenbach family of order n.

Let I be a real interval. A set $\mathcal{B}_n(I)$ of continuous real functions is called an n-parameter Beckenbach family over I, if its members are defined on I and, for all points $(x_1,y_1),\dots,(x_n,y_n)$ of $I\times\mathbb{R}$ (with pairwise distinct first coordinates) there exists exactly one element φ of $\mathcal{B}_n(I)$ such that

$$\varphi(x_1)=y_1,\dots,\varphi(x_n)=y_n.$$

The main result of this chapter gives a characterization of those pairs of real functions that can be separated by an element of a given Beckenbach family which is supposed to be closed under convex combinations.

THEOREM. (Bessenyei, Szokol)
Let $\mathcal{B}_n(I)$ *be a Beckenbach family over the real interval* I *which is closed under convex combinations and* $f,g\colon I\to\mathbb{R}$ *be given functions. Then the following statements are equivalent:*

(i) *there exists* $h\in\mathcal{B}_n(I)$ *such that* $f\le h\le g$*;*
(ii) *for all* $u\le x_1<\cdots<x_n\le v$ *of* I*, we have the inequalities*

$$\varphi_1(v)\ge f(v),\quad \psi_1(v)\le g(v);\quad \textit{and}\quad \varphi_2(u)\ge f(u),\quad \psi_2(u)\le g(u),$$

where $\varphi_1,\varphi_2,\psi_1,\psi_2\in\mathcal{B}_n(I)$ *are determined by the interpolation properties*

$$\varphi_1(x_k)=g(x_k),\ \psi_1(x_k)=f(x_k),\quad n-k\in\{0,\dots,n-1\}\cap 2\mathbb{Z};$$
$$\varphi_1(x_k)=f(x_k),\ \psi_1(x_k)=g(x_k),\quad n-k\in\{0,\dots,n-1\}\cap \big(2\mathbb{Z}+1\big);$$
$$\varphi_2(x_k)=g(x_k),\ \psi_2(x_k)=f(x_k),\quad k\in\{1,\dots,n\}\cap \big(2\mathbb{Z}+1\big);$$
$$\varphi_2(x_k)=f(x_k),\ \psi_2(x_k)=g(x_k),\quad k\in\{1,\dots,n\}\cap 2\mathbb{Z}.$$

In fact, the characterization of those pairs of real functions that can be separated by an affine function was preceded by the characterization of the existence of a convex separator between two real valued functions. The corresponding result is due to Baron, Matkowski and Nikodem [**6**]. In Chapter 5 we give an analogous result in the case when the convexity notion induced by so-called regular pairs.

To present the mentioned result we need to define the notion of Chebyshev systems that can be regarded as particular cases of Beckenbach families. Let I be a real interval and $\omega_1,\dots,\omega_n\colon I\to\mathbb{R}$ be given continuous functions. We say that $\boldsymbol{\omega} := (\omega_1,\dots,\omega_n)$ is a positive Chebyshev system if, for all elements $x_1 < \cdots < x_n$ of I, the determinant of the matrix $\big[\omega_i(x_j)\big]_{i,j=1,\dots,n}$ is positive. A function $f\colon I\to\mathbb{R}$ is said to be $\boldsymbol{\omega}$-convex if, for all elements $x_0\le\cdots\le x_n$ of I, we have the inequality

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \dots & \omega_1(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \omega_n(x_0) & \omega_n(x_1) & \dots & \omega_n(x_n) \\ f(x_0) & f(x_1) & \dots & f(x_n) \end{vmatrix} \ge 0.$$

Moreover, by a regular pair we mean a two-parameter positive Chebyshev system. In our main result it turns out that the existence of an $\boldsymbol{\omega}$-convex separator between two given real functions can be characterized via a determinant inequality.

THEOREM. (Bessenyei, Szokol)
If $\boldsymbol{\omega} = (\omega_1,\omega_2)$ is a regular pair on a real interval I and $f,g\colon I\to\mathbb{R}$, then there exists an $\boldsymbol{\omega}$-convex function $h\colon I\to\mathbb{R}$ with $f\le h\le g$ if and only if, for all elements $x_0\le x_1\le x_2$ of I, we have

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ g(x_0) & f(x_1) & g(x_2) \end{vmatrix} \ge 0.$$

We note, that this result appears in [**12**], but as a consequence of the Baron–Matkowski–Nikodem theorem. Hence, in that approach, it cannot be considered as a real generalization.

In Chapter 5 we also consider the so-called approximate $\boldsymbol{\omega}$-convex functions. The notion of approximate $\boldsymbol{\omega}$-convexity, which originates from that of standard approximate convexity, can be defined as follows.

Let $\boldsymbol{\omega} = (\omega_1,\omega_2)$ be a regular pair on a real interval I and let ω be an $\boldsymbol{\omega}$-affine function which is positive on I°. We say that $\varphi\colon I\to\mathbb{R}$ is approximately $\boldsymbol{\omega}$-convex with error term ω, if, for all elements $x_0 \le x_1 \le x_2$ of I,

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ (\varphi+\omega)(x_0) & (\varphi-\omega)(x_1) & (\varphi+\omega)(x_2) \end{vmatrix} \ge 0.$$

Just as with the standard case it can be shown that every approximate $\boldsymbol{\omega}$-convex function can be decomposed to the sum of a $\boldsymbol{\omega}$-convex function and a "small" part.

THEOREM. (Bessenyei, Szokol)
If $\boldsymbol{\omega} = (\omega_1, \omega_2)$ is a regular pair on a real interval I and ω is an $\boldsymbol{\omega}$-affine function which is positive on I°, then $\varphi\colon I \to \mathbb{R}$ is approximately $\boldsymbol{\omega}$-convex with error term ω if and only if $\varphi = h + \psi$, where h is $\boldsymbol{\omega}$-convex and $|\psi(t)| \le \omega(t)$ for all $t \in I$.

Finally, we consider some particular cases of regular pairs and we present the separation and stability results in those setting.

COROLLARY. *Let I be a real interval, $\alpha\colon I \to \mathbb{R}$ be a continuous, strictly monotone increasing function and $f, g\colon I \to \mathbb{R}$. Then, there exists a function $h\colon I \to \mathbb{R}$ satisfying $f \le h \le g$ and, for all elements $x_0 \le x_1 \le x_2$ of I the inequality*

$$\begin{aligned} \big(\alpha(x_2) - \alpha(x_0)\big)h(x_1) \le \big(\alpha(x_2) - \alpha(x_1)\big)h(x_0) \\ + \big(\alpha(x_1) - \alpha(x_0)\big)h(x_2) \end{aligned} \tag{6.2}$$

holds, if and only if, for all $x_0 \le x_1 \le x_2$ of I, we have

$$\big(\alpha(x_2)-\alpha(x_0)\big)f(x_1) \le \big(\alpha(x_2)-\alpha(x_1)\big)g(x_0) + \big(\alpha(x_1)-\alpha(x_0)\big)g(x_2).$$

COROLLARY. *Let I be a real interval, $\alpha\colon I \to \mathbb{R}$ be a continuous, strictly monotone increasing function and $\varepsilon > 0$. Then, a function $\varphi\colon I \to \mathbb{R}$ satisfies the inequality*

$$\begin{aligned} \big(\alpha(x_2) - \alpha(x_0)\big)\varphi(x_1) \;&\le\; \big(\alpha(x_2) - \alpha(x_1)\big)\varphi(x_0) \\ &+\; \big(\alpha(x_1) - \alpha(x_0)\big)\varphi(x_2) + 2\varepsilon\big(\alpha(x_2) - \alpha(x_0)\big) \end{aligned}$$

for all elements $x_0 \le x_1 \le x_2$ of I if and only if there exist functions $h, \psi\colon I \to \mathbb{R}$ such that $\varphi = h + \psi$, where h fulfills (6.2) *and $\|\psi\| \le \varepsilon$. (Here $\|\cdot\|$ stands for the supremum norm.)*

COROLLARY. *If* $I \subset]-\frac{\pi}{2}, \frac{\pi}{2}[$ *is a real interval,* $f, g\colon I \to \mathbb{R}$*, then there exists a* $(\cos, \sin)$*-convex function* $h\colon I \to \mathbb{R}$ *with* $f \le h \le g$ *if and only if, for all* $\lambda \in [0,1]$ *and* $x \le y$ *of* I*,*

$$\sin(y-x) f\big(\lambda x+(1-\lambda)y\big) \le \sin\big(\lambda(y-x)\big)g(x)+\sin\big((1-\lambda)(y-x)\big)g(y).$$

COROLLARY. *If* $I \subset]-\frac{\pi}{2}, \frac{\pi}{2}[$ *is a real interval, then* $\varphi\colon I \to \mathbb{R}$ *is approximately* $(\cos, \sin)$*-convex with error term* $\varepsilon \cdot \cos$ *if and only if* $\varphi = h + \psi$*, where* $h\colon I \to \mathbb{R}$ *is* $(\cos, \sin)$*-convex, and* $\psi\colon I \to \mathbb{R}$ *satisfies* $|\psi(t)| \le \varepsilon \cdot \cos(t)$ *for all* $t \in I$*.*

Összefoglalás

Ez a disszertáció különböző matematikai struktúrákon értelmezett megőrzési problémákat, valamint szeparációs tételeket tartalmaz. Egy bevezetésből, öt fejezetből, egy magyar és egy angol nyelvű összefoglalóból és irodalomjegyzékből áll. A bevezetésben szerepel néhány fontos és jól ismert megőrzési probléma, melyek szoros kapcsolatban vannak a disszertációban szereplő tételekkel. Továbbá a szeparációs tételeinket motiváló alapvető tételeket is felsorakoztatjuk a bevezetésben. Az 1-3 fejezetekben különböző matematikai struktúrák bizonyos megőrzési problémáit vizsgáljuk.

Az első fejezetben azon transzformációk szerkezete kerül leírásra, melyek invariánsan hagynak egy, a kvantum-információelméletben fontos szerepet játszó mennyiséget. Eredményeink megfogalmazásához bevezetünk néhány jelölést. Legyen H egy véges dimenziós Hilbert tér és jelölje $B(H)$ a H-n értelmezett korlátos lineáris operátorok algebráját. Továbbá jelölje $B(H)^+$ a H Hilbert téren értelmezett, pozitív szemidefinit operátorok kúpját, illetve $S(H)$ a H sűrűségoperátorainak (azaz $B(H)^+$ 1-trace-ű elemeinek) halmazát. A relatív entrópia egy alapvető fogalom a kvantum-információelméletben, melynek több változata is létezik. Ezek közül a legismertebb az Umegaki relatív entrópia, mely tetszőleges $A, B \in S(H)$ pár esetén a következőképpen van definiálva:

$$S(A||B) = \begin{cases} \operatorname{tr} A(\log A - \log B), & \operatorname{supp} A \subset \operatorname{supp} B \\ \infty, & \text{egyébként.} \end{cases}$$

A [**56**] cikkben Molnár leírta azon szürjektív transzformációk szerkezetét, melyek invariánsan hagyják az Umegaki relatív entrópiát. Kiderült, hogy ezen transzformációk alakja igen egyszerű. Az első fejezetben megmutatjuk, hogy az állítás abban az esetben is igaz marad, ha elhagyjuk a szürjektivitás feltételét.

TÉTEL. (Molnár, Szokol)
Legyen $\phi\colon S(H) \to S(H)$ *egy olyan transzformáció, amely megőrzi az*

Umegaki relatív entrópiát. Ekkor létezik olyan unitér vagy antiunitér U operátor H-n, amellyel ϕ az

$$\phi(A) = UAU^*, \quad A \in S(H)$$

alakba írható.

Ezen tétel motiválja az első fejezet fő eredményét, melyben a sűrűség-operátorokon értelmezett azon transzformációk szerkezetét írjuk le, melyek egy adott, szigorúan konvex f függvény esetén invariánsan hagyják az ún. kvantum f-divergenciát. Jól ismert tény, hogy speciális f függvény választásával a kvantum f-divergencia definíciója az Umegaki relatív entrópia fogalmához vezet. Így a következőkben megfogalmazásra kerülő eredmény jelentős általánosítása az előző, Umegaki relatív entrópiát megőrző leképezések szerkezetére vonatkozó tételnek.

Legyen $f\colon [0,\infty[\to \mathbb{R}$ egy olyan, a $]0,\infty[$ intervallumon folytonos függvény, amely esetén létezik a

$$\alpha := \lim_{x\to\infty} \frac{f(x)}{x}$$

határérték és $\alpha \in [-\infty,\infty]$. Ha $A, B \in B(H)^+$, akkor tetszőleges $\lambda \in \mathbb{R}$ esetén jelölje P_λ, illetve Q_λ H azon projekcióit, melyek az $A - \lambda I$, illetve $B - \lambda I$ magjára vetítenek. Ekkor a kvantum f-divergencia A és B között a következő formulával adható meg:

$$S_f(A||B) = \sum_{a\in\sigma(A)} \left(\sum_{b\in\sigma(B)\setminus\{0\}} bf\left(\frac{a}{b}\right) \operatorname{tr} P_a Q_b + \alpha a \operatorname{tr} P_a Q_0 \right),$$

ahol $\sigma(.)$ jelöli $B(H)$ elemeinek spektrumát és megállapodunk abban, hogy $0\cdot(-\infty) = 0\cdot\infty = 0$. Az első fejezet fő eredménye ezek után a következőképpen fogalmazható meg.

TÉTEL. (Molnár, Nagy, Szokol)
Legyen $f\colon [0,\infty[\to \mathbb{R}$ egy szigorúan konvex függvény és tegyük fel, hogy $\phi\colon S(H) \to S(H)$ olyan transzformáció, amely kielégíti a

$$S_f(\phi(A)||\phi(B)) = S_f(A||B), \quad A, B \in S(H).$$

egyenletet minden $A, B \in S(H)$ pár esetén. Ekkor létezik olyan U unitér vagy antiunitér operátor H-n, mellyel ϕ a

$$\phi(A) = UAU^*, \quad A \in S(H).$$

alakba írható.

A második fejezetben sikerült közös keretbe foglalni és lényegesen általánosítani a [**62**] cikkben szereplő eredményeket. Nevezetesen, a korábban tárgyalt, valódi metrikákra vonatkozó szürjektív izometriák leírása után sikerült meghatározni azon szürjektív transzformációk szerkezetét, melyek megőriznek egy adott, unitér-invariáns normával illetve bizonyos feltételeknek eleget tevő folytonos valós függvénnyel paraméterezett, úgynevezett általánosított távolság mértéket. Általánosított távolság mérték alatt olyan $d : X \times X \to [0, \infty[$ függvényt értünk (X egy adott halmaz), amelyre teljesül, hogy minden $x, y \in X$ pár esetén, a $d(x, y)$ távolság pontosan akkor 0, ha $x = y$. Azonban d-ről nem tesszük fel sem a szimmetrikusságot, sem azt, hogy eleget tesz a háromszög-egyenlőtlenségnek. A következőkben jelölje $\mathbb{M}_n$ az $n \times n$-es komplex mátrixok algebráját és $\mathbb{P}_n$ az $n \times n$-es pozitív definit mátrixok halmazát. Továbbá jelölje $\mathbb{P}_n^1$ és $\mathbb{P}_n^c$ $(c > 0)$ a $\mathbb{P}_n$ azon elemeinek halmazát, melyek determinánsa rendre 1-gyel, illetve c-vel egyenlő. A szükséges fogalmak és jelölések bevezetése után megfogalmazható a második fejezet fő eredménye.

TÉTEL. (Molnár, Szokol)
Legyen N egy unitér-invariáns norma $\mathbb{M}_n$-en és $f\colon]0, \infty[\to \mathbb{R}$ egy olyan folytonos függvény, amelyre teljesülnek az alábbi feltételek:

(a1) $f(y) = 0$ *akkor és csak akkor, ha* $y = 1$*;*
(a2) *létezik egy* $K > 1$ *konstans úgy, hogy*

$$|f(y^2)| \geq K|f(y)|, \quad y \in]0, \infty[.$$

Definiáljuk a $d_{N,f}\colon \mathbb{P}_n \times \mathbb{P}_n \to [0, \infty[$ leképezést a következőképpen

$$d_{N,f}(A, B) = N(f(A^{-1/2}BA^{-1/2})), \quad A, B \in \mathbb{P}_n. \tag{7.1}$$

Tegyük fel, hogy $n \geq 3$. Ha $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ egy olyan szürjektív leképezés, mely megőrzi a $d_{N,f}(.,.)$ mennyiséget, azaz minden $A, B \in \mathbb{P}_n$ esetén eleget tesz az

$$d_{N,f}(\phi(A), \phi(B)) = d_{N,f}(A, B)$$

egyenletnek, akkor létezik olyan invertálható $T \in \mathbb{M}_n$ mátrix és c valós szám, hogy ϕ az alábbi alakok valamelyikébe írható:

$$\phi(A) = (\det A)^c TAT^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = (\det A)^c TA^{-1}T^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = (\det A)^c TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = (\det A)^c T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n.$$

Megjegyezzük, hogy a tételben szereplő $d_{N,f}(.,.)$ leképezés valóban egy általánosított távolság mérték a korábban bevezetett értelemben. Könnyen ellenőrizhető továbbá, hogy a [**62**] cikkben vizsgált metrikák olyan speciális általánosított távolság mértékek, melyeknek (a (7.1) alapján) megfelelő f függvények teljesítik az (a1) és (a2) feltételeket. A fenti tételünk azonban további általánosított távolság mértékek esetén is alkalmazható. Tetszőleges $A, B \in \mathbb{P}_n$ esetén jelölje $Y_{A,B}$ az $A^{-1/2}BA^{-1/2}$ pozitív definit mátrixot. Ekkor az alábbi formulákkal újabb általánosított távolság mértékek definiálhatóak:

(i) Stein's loss: $l(A, B) = \|Y_{A,B}^{-1} - \log Y_{A,B}^{-1} - 1\|_1$;

(ii) Jeffrey's Kullback-Leibler eltérés:

$$S_{JKL}(A, B) = \left\| \frac{Y_{A,B} + Y_{A,B}^{-1} - 2I}{2} \right\|_1 ;$$

(iii) log-determináns α-eltérés (tetszőleges $-1 < \alpha < 1$ paraméterre):

$$D_{LD}^{\alpha}(A, B) =$$

$$\frac{4}{1-\alpha^2} \left\| \log \frac{(1-\alpha)I + (1+\alpha)Y_{A,B}}{2} - \frac{1+\alpha}{2} \log Y_{A,B} \right\|_1 ,$$

ahol $\|.\|_1$ az unitér-invariáns trace-normát jelöli. Könnyen ellenőrizhető, hogy tételünk alkalmazható a fenti általánosított távolság mértékekre. Továbbá megjegyezzük, hogy az N unitér-invariáns norma és az f valós függvény speciális választása esetén ahhoz, hogy meghatározzuk azon transzformációk szerkezetét, melyek valóban megőrzik a kérdéses általánosított távolság mértéket, a tétel alkalmazása után további vizsgálatok szükségesek. Ezzel kapcsolatos a következő tételünk.

TÉTEL. (Molnár, Szokol)
Jelölje $div(.,.)$ az $l(.,.)$, illetve a $D_{LD}^{\alpha}(.,.)$, $-1 \leq \alpha \leq 1$ függvények valamelyikét. Ekkor egy $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ transzformáció pontosan akkor őrzi meg a $div(.,.)$ leképezést, azaz tesz eleget az

$$div(\phi(A), \phi(B)) = div(A, B), \qquad A, B \in \mathbb{P}_n,$$

egyenletnek, ha létezik olyan invertálható $T \in \mathbb{M}_n$ mátrix, mellyel ϕ az alábbi alakok valamelyikébe írható:

$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n$;
$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n$.

Egy $\phi\colon \mathbb{P}_n \to \mathbb{P}_n$ szürjektív leképezés pontosan akkor őrzi meg a $S_{JKL}(.,.)$ eltérést, ha létezik olyan invertálható $T \in \mathbb{M}_n$ mátrix, mellyel ϕ az alábbi alakok valamelyikébe írható:

$$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = TA^{-1}T^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n;$$
$$\phi(A) = T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n.$$

A fő eredményhez hasonló eredményt sikerült igazolni abban az esetben is, amikor a ϕ transzformáció az egy determinánsú, pozitív definit mátrixok $\mathbb{P}_n^1$ halmazán van definiálva. Ehhez szükségünk volt a $\mathbb{P}_n^1$ összes folytonos automorfizmusának alakjára a Jordan hármas szorzatra (ABA) vonatkozóan, (azaz az összes folytonos bijekció szerkezetére, amely megőrzi a Jordan hármas szorzatot).

TÉTEL. (Molnár, Szokol)
Legyen $n \geq 3$ és tegyük fel, hogy $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ egy folytonos automorfizmus a Jordan hármas szorzatra vonatkozóan, azaz egy olyan folytonos bijekció, amely teljesíti az

$$\phi(ABA) = \phi(A)\phi(B)\phi(A)$$

egyenlőséget minden $A, B \in \mathbb{P}_n^1$ esetén. Ekkor létezik olyan $U \in \mathbb{M}_n$ unitér mátrix, mellyel ϕ az alábbi alakok valamelyikébe írható:

$$\phi(A) = UAU^*, \quad A \in \mathbb{P}_n^1;$$
$$\phi(A) = UA^{-1}U^*, \quad A \in \mathbb{P}_n^1;$$
$$\phi(A) = UA^{\mathrm{tr}}U^*, \quad A \in \mathbb{P}_n^1;$$
$$\phi(A) = U(A^{\mathrm{tr}})^{-1}U^*, \quad A \in \mathbb{P}_n^1.$$

Ezek után már leírható a $\mathbb{P}_n^1$ összes szürjektív, $d_{N,f}$ általánosított távolság mértéket megőrző transzformációinak szerkezete.

TÉTEL. (Molnár, Szokol)
Legyen N az $\mathbb{M}_n$ egy unitér-invariáns normája és $f :]0, \infty[\to \mathbb{R}$ egy olyan folytonos függvény, mely eleget tesz az (a1), (a2) feltételeknek. Tegyük fel, hogy $n \geq 3$. Legyen $\phi\colon \mathbb{P}_n^1 \to \mathbb{P}_n^1$ egy olyan szürjektív leképezés, amely megőrzi a $d_{N,f}(.,.)$ mértéket. Ekkor létezik olyan T invertálható mátrix, melyre $|\det T| = 1$ és amellyel ϕ az alábbi alakok valamelyikébe írható:

$$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n^1;$$
$$\phi(A) = TA^{-1}T^*, \quad A \in \mathbb{P}_n^1;$$
$$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n^1;$$
$$\phi(A) = T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n^1.$$

Az előző tételből könnyen adódik az alábbi következmény.

KÖVETKEZMÉNY. *Legyen N és f olyan, mint az előző tételben és tegyük fel, hogy $n \geq 3$ és c egy pozitív valós szám. Ha $\phi\colon \mathbb{P}_n^c \to \mathbb{P}_n^c$ egy olyan szürjektív leképezés, mely megőrzi a $d_{N,f}(.,.)$ mértéket, akkor létezik olyan T invertálható mátrix, melyre $|\det T| = 1$ úgy, hogy ϕ az alábbi alakok valamelyikébe írható:*

$$\phi(A) = TAT^*, \quad A \in \mathbb{P}_n^c;$$
$$\phi(A) = \lambda^2 TA^{-1}T^*, \quad A \in \mathbb{P}_n^c;$$
$$\phi(A) = TA^{\mathrm{tr}}T^*, \quad A \in \mathbb{P}_n^c;$$
$$\phi(A) = \lambda^2 T(A^{\mathrm{tr}})^{-1}T^*, \quad A \in \mathbb{P}_n^c,$$

ahol $\lambda = \sqrt[n]{c}$.

Lineáris függvényterek lineáris izometriáinak vizsgálata szintén egy jelentős kutatási terület a funkcionálanalízisben. Azonban van néhány olyan fontos függvénytér, amely nem lineáris. A valószínűségszámításban alapvető szerepet játszó eloszlásfüggvények tere szintén nem lineáris. A [**26**] cikkben a szerzők leírták az $\mathbb{R}$-en értelmezett eloszlásfüggvények $D(\mathbb{R})$ tere szürjektív izometriáinak szerkezetét a Kolmogorov-Smirnov metrikára vonatkozóan. A harmadik fejezetben az előbb említett eredményt terjesztettük ki az úgynevezett általánosított eloszlásfüggvények $\Delta(\mathbb{R})$ terére. Általánosított eloszlásfüggvényen olyan $f\colon \mathbb{R} \to [0,1]$ függvényt értünk, amely monoton növekvő és jobbról folytonos (a $\pm\infty$-beli határérték feltételek nincsenek megkövetelve). Ha f és g két tetszőleges eloszlásfüggvény, akkor Kolmogorov-Smirnov távolságuk:

$$\rho(f,g) = \sup_{t\in\mathbb{R}} |f(t) - g(t)|.$$

A harmadik fejezet első eredménye azt állítja, hogy a $\Delta(\mathbb{R})$ tér szürjektív izometriáinak alakja megegyezik a $D(\mathbb{R})$ tér szürjektív izometriáinak alakjával.

TÉTEL. (Molnár, Szokol)
Legyen $\phi\colon \Delta(\mathbb{R}) \to \Delta(\mathbb{R})$ egy szürjektív izometria a Kolmogorov-Smirnov metrikára vonatkozóan. Ekkor vagy létezik olyan $\varphi\colon \mathbb{R} \to \mathbb{R}$ szigorúan monoton növekvő bijekció, mellyel ϕ az

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in \Delta(\mathbb{R})$$

alakba írható, vagy létezik olyan $\psi\colon \mathbb{R} \to \mathbb{R}$ szigorúan monoton csökkenő bijekció, mellyel ϕ az

$$\phi(f)(t) = 1 - f(\psi(t)-), \quad t \in \mathbb{R}, f \in \Delta(\mathbb{R})$$

alakba írható.

Később a [**58**] cikkben Molnár a folytonos eloszlásfüggvények $D_c(\mathbb{R})$ terének szürjektív Kolmogorov-Smirnov izometriáit határozta meg. Ez az eredmény motiválta a harmadik fejezet második eredményét, melyben leírtuk a folytonos általánosított eloszlásfüggvények $\Delta_c(\mathbb{R})$ halmaza összes szürjektív izometriájának szerkezetét a Kolmogorov-Smirnov metrikára vonatkozóan.

TÉTEL. (Molnár, Szokol)
Legyen $\phi\colon \Delta_c(\mathbb{R}) \to \Delta_c(\mathbb{R})$ *egy szürjektív izometria a Kolmogorov-Smirnov metrikára vonatkozóan. Ekkor vagy létezik olyan* $\varphi\colon \mathbb{R} \to \mathbb{R}$ *szigorúan monoton növekvő bijekció, mellyel* ϕ *az*

$$\phi(f)(t) = f(\varphi(t)), \quad t \in \mathbb{R}, f \in \Delta_c(\mathbb{R})$$

alakba írható, vagy létezik olyan $\psi\colon \mathbb{R} \to \mathbb{R}$ *szigorúan monoton csökkenő bijekció, mellyel* ϕ *az*

$$\phi(f)(t) = 1 - f(\psi(t)), \quad t \in \mathbb{R}, f \in \Delta_c(\mathbb{R})$$

alakba írható.

A negyedik és ötödik fejezetekben szeparációs problémákat vizsgálunk. Egy jól ismert szeparációs tétel szerint, ha egy konvex függvény egy konkáv "felett" helyezkedik el, akkor létezik a kettő között egy affin függvény. Sőt, két tetszőleges függvény esetén adható az affin szeparáció jellemzésére egy szükséges és elégséges feltétel. Ezen eredményt Nikodem és Wąsowicz bizonyították a [**67**] cikkben. A negyedik fejezetben a fenti tételt általánosítottuk oly módon, hogy jellemeztük azon valós függvény párokat, melyek egy n-ed rendű konvex Beckenbach család tagjával szeparálhatóak.

Legyen I egy valós intervallum. Egy folytonos, valós függvényekből álló $\mathcal{B}_n(I)$ halmazt egy I feletti n-paraméteres Beckenbach családnak nevezünk, ha a függvények értelmezési tartománya I és minden $I \times \mathbb{R}$-beli páronként különböző első koordinátájú (x_1, y_1), …, (x_n, y_n) pont esetén létezik pontosan egy olyan $\varphi \in \mathcal{B}_n(I)$, melyre

$$\varphi(x_1) = y_1, \dots, \varphi(x_n) = y_n.$$

A negyedik fejezet fő eredménye azon valós függvényeket karakterizálja, melyek szeparálhatóak egy adott, a konvexitásra zárt Beckenbach család valamely elemével.

TÉTEL. (Bessenyei, Szokol)
Legyen $\mathcal{B}_n(I)$ *egy valós* I *intervallum felett értelmezett, konvexitásra zárt*

Beckenbach család és $f, g\colon I \to \mathbb{R}$ adott függvények. Ekkor a következő állítások ekvivalensek:

(i) *létezik olyan $h \in \mathcal{B}_n(I)$, melyre $f \le h \le g$;*
(ii) *minden I-beli $u \le x_1 < \cdots < x_n \le v$ elem esetén a következő egyenlőtlenségek teljesülnek*

$$\varphi_1(v) \ge f(v), \quad \psi_1(v) \le g(v); \quad \textit{és} \quad \varphi_2(u) \ge f(u), \quad \psi_2(u) \le g(u),$$

ahol $\varphi_1, \varphi_2, \psi_1, \psi_2 \in \mathcal{B}_n(I)$ az alábbi interpolációs tulajdonságok által meghatározottak:

$$\varphi_1(x_k) = g(x_k),\ \psi_1(x_k) = f(x_k), \quad n-k \in \{0, \dots, n-1\} \cap 2\mathbb{Z};$$
$$\varphi_1(x_k) = f(x_k),\ \psi_1(x_k) = g(x_k), \quad n-k \in \{0, \dots, n-1\} \cap \big(2\mathbb{Z}+1\big);$$
$$\varphi_2(x_k) = g(x_k),\ \psi_2(x_k) = f(x_k), \quad k \in \{1, \dots, n\} \cap \big(2\mathbb{Z}+1\big);$$
$$\varphi_2(x_k) = f(x_k),\ \psi_2(x_k) = g(x_k), \quad k \in \{1, \dots, n\} \cap 2\mathbb{Z}.$$

Nikodem és Wąsowicz eredményét valójában Baron, Matkowski és Nikodem azon tétele motiválta [**6**], melyben jellemezve vannak az olyan valós függvénypárok, melyek egy konvex függvénnyel szeparálhatóak. Az ötödik fejezetben az úgynevezett reguláris párok által indukált konvexitási fogalommal kapcsolatban igazoltunk egy, az említett tétellel analóg állítást.

A tétel megfogalmazásához bevezetjük a Csebisev-rendszer fogalmát, mely a Beckenbach családok egy speciális osztályát adja. Legyen $I \subset \mathbb{R}$ egy intervallum és $\omega_1, \dots, \omega_n\colon I \to \mathbb{R}$ folytonos függvények. Azt mondjuk, hogy az $\boldsymbol{\omega} := (\omega_1, \dots, \omega_n)$ egy pozitív Csebisev-rendszer, ha minden I-beli $x_1 < \cdots < x_n$ elem esetén, a $\big[\omega_i(x_j)\big]_{i,j=1,\dots,n}$ mátrix determinánsa pozitív. Egy $f\colon I \to \mathbb{R}$ függvényt $\boldsymbol{\omega}$-konvexnek nevezünk, ha minden I-beli $x_0 \le \cdots \le x_n$ elem esetén teljesül az alábbi egyenlőtlenség:

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \dots & \omega_1(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \omega_n(x_0) & \omega_n(x_1) & \dots & \omega_n(x_n) \\ f(x_0) & f(x_1) & \dots & f(x_n) \end{vmatrix} \ge 0.$$

Reguláris pár alatt két-paraméteres pozitív Csebisev-rendszert értünk. A fő eredményünkben kiderül, hogy egy determinánst tartalmazó egyenlőtlenség segítségével szükséges és elégséges feltétel adható meg két tetszőleges valós függvény közötti $\boldsymbol{\omega}$-konvex szeparátor létezésére.

TÉTEL. (Bessenyei, Szokol)
Legyen $\boldsymbol{\omega} = (\omega_1, \omega_2)$ egy reguláris pár az I valós intervallum felett és

$f, g\colon I \to \mathbb{R}$ adott függvények. Pontosan akkor létezik olyan $h\colon I \to \mathbb{R}$ $\boldsymbol{\omega}$-konvex függvény, melyre $f \le h \le g$, ha minden I-beli $x_0 \le x_1 \le x_2$ elem esetén

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ g(x_0) & f(x_1) & g(x_2) \end{vmatrix} \ge 0.$$

Megjegyezzük, hogy ezen eredményt már a [**12**] cikkben bebizonyították a szerzők a Baron–Matkowski–Nikodem tétel segítségével. Azonban nekünk sikerült egy független bizonyítást adni erre az eredményre.

A standard esethez analóg módon definiálhatunk közelítőleg $\boldsymbol{\omega}$-konvex függvényeket a következőképpen.

Legyen $\boldsymbol{\omega} = (\omega_1, \omega_2)$ egy reguláris pár az I valós intervallum felett és legyen ω egy $\boldsymbol{\omega}$-affin függvény, amely pozitív I°-n. A $\varphi\colon I \to \mathbb{R}$ függvényt közelítőleg $\boldsymbol{\omega}$-konvexnek nevezzük ω hibával, ha minden I-beli $x_0 \le x_1 \le x_2$ elem esetén

$$\begin{vmatrix} \omega_1(x_0) & \omega_1(x_1) & \omega_1(x_2) \\ \omega_2(x_0) & \omega_2(x_1) & \omega_2(x_2) \\ (\varphi+\omega)(x_0) & (\varphi-\omega)(x_1) & (\varphi+\omega)(x_2) \end{vmatrix} \ge 0.$$

A standard esethez hasonlóan, minden közelítőleg $\boldsymbol{\omega}$-konvex függvény felbontható egy $\boldsymbol{\omega}$-konvex függvény és egy "kicsi" rész összegére.

TÉTEL. (Bessenyei, Szokol)
Legyen $\boldsymbol{\omega} = (\omega_1, \omega_2)$ egy reguláris pár az I valós intervallum felett és legyen ω egy $\boldsymbol{\omega}$-affin függvény, amely pozitív I°-n. Ekkor $\varphi\colon I \to \mathbb{R}$ pontosan akkor közelítőleg $\boldsymbol{\omega}$-konvex ω hibával, ha felírható $\varphi = h + \psi$ alakban, ahol h egy $\boldsymbol{\omega}$-konvex függvény és $|\psi(t)| \le \omega(t)$ minden $t \in I$ esetén.

Az ötödik fejezet végén a fő tételek következményeként, speciális reguláris párok esetén fogalmaztuk meg a nekik megfelelő szeparációs, illetve stabilitási tételeket.

KÖVETKEZMÉNY. *Legyen I egy valós intervallum, $\alpha\colon I \to \mathbb{R}$ egy folytonos, szigorúan monoton növekvő függvény. Legyenek továbbá $f, g\colon I \to \mathbb{R}$ tetszőlegesen adott függvények. Pontosan akkor létezik olyan $h\colon I \to \mathbb{R}$ függvény, amelyre $f \le h \le g$ és amely teljesíti az alábbi egyenlőtlenséget*

$$\big(\alpha(x_2) - \alpha(x_0)\big)h(x_1) \le \big(\alpha(x_2) - \alpha(x_1)\big)h(x_0) + \big(\alpha(x_1) - \alpha(x_0)\big)h(x_2) \tag{7.2}$$

minden $x_0 \le x_1 \le x_2 \in I$ elemre, ha tetszőleges I-beli $x_0 \le x_1 \le x_2$ esetén az alábbi egyenlőtlenség fennáll:

$$\big(\alpha(x_2)-\alpha(x_0)\big)f(x_1) \le \big(\alpha(x_2)-\alpha(x_1)\big)g(x_0)+\big(\alpha(x_1)-\alpha(x_0)\big)g(x_2).$$

KÖVETKEZMÉNY. *Legyen I egy valós intervallum, $\alpha\colon I \to \mathbb{R}$ egy folytonos, szigorúan monoton növekvő függvény és $\varepsilon > 0$. Ekkor egy $\varphi\colon I \to \mathbb{R}$ függvény pontosan akkor teljesíti a*

$$\begin{aligned}\big(\alpha(x_2)-\alpha(x_0)\big)\varphi(x_1) \;&\le\; \big(\alpha(x_2)-\alpha(x_1)\big)\varphi(x_0)\\ &+\; \big(\alpha(x_1)-\alpha(x_0)\big)\varphi(x_2)+2\varepsilon\big(\alpha(x_2)-\alpha(x_0)\big)\end{aligned}$$

egyenlőtlenséget minden I-beli $x_0 \le x_1 \le x_2$ esetén, ha léteznek olyan $h, \psi\colon I \to \mathbb{R}$ függvények, melyekkel $\varphi = h + \psi$, ahol h eleget tesz a (7.2) *egyenlőtlenségnek és $\|\psi\| \le \varepsilon$. (Itt $\|\cdot\|$ a szuprémum normát jelöli.)*

KÖVETKEZMÉNY. *Legyen $I \subset]-\frac{\pi}{2}, \frac{\pi}{2}[$ egy valós intervallum, $f, g\colon I \to \mathbb{R}$. Pontosan akkor létezik olyan $h\colon I \to \mathbb{R}$ $(\cos, \sin)$-konvex függvény, amelyre $f \le h \le g$, ha minden $\lambda \in [0,1]$ és I-beli $x \le y$ esetén*

$$\sin(y-x)f\big(\lambda x+(1-\lambda)y\big) \le \sin\big(\lambda(y-x)\big)g(x)+\sin\big((1-\lambda)(y-x)\big)g(y).$$

KÖVETKEZMÉNY. *Legyen $I \subset]-\frac{\pi}{2}, \frac{\pi}{2}[$ egy valós intervallum. Ekkor egy $\varphi\colon I \to \mathbb{R}$ függvény pontosan akkor közelítőleg $(\cos, \sin)$-konvex $\varepsilon \cdot \cos$ hibával, ha φ előáll $\varphi = h + \psi$ alakban, ahol $h\colon I \to \mathbb{R}$ egy $(\cos, \sin)$-konvex függvény és $\psi\colon I \to \mathbb{R}$ minden $t \in I$ esetén teljesíti az $|\psi(t)| \le \varepsilon \cdot \cos(t)$ egyenlőtlenséget.*

Talks held by the author

(1) *Pozitív szemidefinit operátorokon értelmezett, relatív entrópiát megőrző leképezések*, Debreceni Egyetem TEK Természettudományi és Technológiai Kar 2009. Őszi Tudományos Diákköri Konferencia.

(2) *Maps on positive semidefinite operators preserving relative entropy*, The 6th International Students' Conference on Analysis, Síkfőkút, Hungary, 2010 January 28 – February 1.

(3) *Kolmogorov–Smirnov isometries of the space of generalized distribution functions*, The 11th Katowice–Debrecen Winter Seminar, Wisła-Malinka, Poland, 2011 February 2 – 5.

(4) *Properties of quantum relative entropy*, The 7th International Students' Conference on Analysis, Wisła, Poland, 2011 February 5 – 8.

(5) *Maps on positive semidefinite operators preserving relative entropy*, XXX. Jubileumi Országos Tudományos Diákköri Konferencia, Fizika, Földtudományok és Matematika Szekció, Nyíregyházi Főiskola, Hungary, 2011 April 27 – 29.

(6) *Az általánosított eloszlásfüggvények terén értelmezett Kolmogorov-Szmirnov izometriák*, Analízis Tanszék Szeminárium, Síkfőkút, Hungary, 2011 May 28.

(7) *Convex and affine separation problems I.*, The 12th Debrecen-Katowice Winter Seminar on Functional Equations and Inequalities, Hajdúszoboszló, Hungary, 2012 January 25 – 28.

(8) *Jordan triple automorphisms and inverted Jordan triple automorphisms*, The 8th International Students' Conference on Analysis, Síkfőkút, Hungary, 2012 January 28 – February 1.

(9) *Konvex szeparáció reguláris párokkal*, Analízis Tanszék Szeminárium, Síkfőkút, Hungary, 2012 May 25 – 28.

(10) *Convex separation by regular pairs*, The 50th International Conference on Functional Equations, Hajdúszoboszló, Hungary, 2012 June 17 – 24.

(11) *Maps on density operators preserving f-divergences*, The International Workshop on Functional Analysis, Temesvár, Romania, 2012 October 12 – 14.

(12) *Maps on density operators preserving quantum f-divergences*, The 13th Katowice-Debrecen Winter Seminar on Functional Equations and Inequalities, Zakopane, Poland, 2013 January 30 – February 2.

(13) *Separation by convex interpolation families,* The 9th International Students' Conference on Analysis, Ustron, Poland, 2013 February 2 – 5.

(14) *Kvantum f-divergenciát megőrző leképezések sűrűségoperátorokon*, "Lendület" FIFA'13 Mini-Konferencia, Debrecen, Hungary, 2013 April 26 – 27.

(15) *Szeparáció konvex interpolációs családokkal*, Analízis Tanszék Szeminárium, Síkfőkút, Hungary, 2012 May 25 – 28.

(16) *Separation by convex interpolation families*, The 51th International Symposium on Functional Equations, Rzeszów, Poland, 2013 June 17 – 23.

(17) *Transformations on density operators leaving f-divergences invariant*, Workshop on Functional Analysis and its Applications in Mathematical Physics and Optimal Control, Nemecká, Slovak Republic, 2013 September 9 – 14.

(18) *Transformations on density operators preserving quantum f - divergences*, Kaohsiung, Taiwan, 2013 October 30.

(19) *Transformations on positive definite matrices preserving generalized distance measures*, The 14th Debrecen-Katowice Winter Seminar on Functional Equations and Inequalities, Hajdúszoboszló, Hungary, 2014 January 29 – February 1.

(20) *Transformations on positive definite matrices preserving generalized distance measures*, Young Functional Analysts' Meeting 2014, Debrecen, Hungary, 2014 April 11 – 13.

(21) *Maps preserving geodesics and their connection with relative entropies*, The 52nd International Symposium on Functional Equations, Innsbruck, Austria, 2014 June 22 – 29.

(22) *Maps preserving geodesics and their connection with relative entropy and geometric mean*, CSM - The Third Conference of PhD Students in Mathematics, Szeged, Hungary, 2014 June 30 – July 2.

(23) *Maps preserving numerical quantities of geodesics in space of positive definite matrices*, Conference on Inequalities and Applications '14, Hajdúszoboszló, Hungary, 2014 September 7– 13.

(24) *Maps on positive definite matrices preserving generalized distance measures*, Szemináriumi előadás, National Sun Yat-sen University, Kaohsiung, Taiwan, 2014 November 6.

(25) *Transformations preserving norms of means of positive operators and nonnegative functions*, The 15th Katowice-Debrecen Winter Seminar on Functional Equations and Inequalities, Będlewo, Poland, 2015 January 28 – 31.

(26) *Surjective isometries of the space of generalized distibution functions,* The 11th International Students' Conference on Analysis, Ustron, Poland, 2015 January 31 – February 3.

(27) *Transformations preserving norms of means of positive operators and nonnegative functions*, Szemináriumi előadás, National Sun Yat-sen University, Kaohsiung, Taiwan, 2015 October 22.

Publications of the author

1. L. Molnár and P. Szokol, *Maps on states preserving the relative entropy II*, Linear Algebra Appl., **432** (2010), 3343–3350.
2. L. Molnár, G. Nagy and P. Szokol, *Maps on density operators preserving quantum f-divergences*, Quantum Inf. Process. **12** (2013), 2309–2323.
3. M. Bessenyei, P. Szokol, *Convex separation by regular pairs*, J. Geom., **104** (2013), 45–56.
4. M. Bessenyei, P. Szokol, *Separation by convex interpolation families*, J. Convex Anal., **20** (2013), 937–946.
5. L. Molnár, P. Szokol, *Kolmogorov-Smirnov isometries of the space of generalized distribution functions*, Math. Slovaca, **64** (2014), 433–444.
6. L. Molnár, P. Szokol, *Transformations on positive definite matrices preserving generalized distance measures,* Linear Algebra Appl., **466** (2015), 141–159.
7. P. Szokol, M.-C. Tsai, J. Zhang, *Preserving problems of geodesic affine maps and related topics on positive definite matrices,* Linear Algebra Appl., **483** (2015), 293–308.
8. L. Molnár, P. Szokol, *Transformations preserving norms of means of positive operators and nonnegative functions,* Integr. Equ. Oper. Theory., to appear.
9. G. Dolinar, B. Kuzma, G. Nagy and P. Szokol, *Restricted skew-morphisms on matrix algebras,* Linear Algebra Appl., to appear.

Bibliography

[1] S.M. Ali and S.D. Silvey, *A general class of coefficients of divergence of one distribution from another,* J. Roy. Stat. Soc. Ser. B, **28** (1966), 131–142.

[2] M. Balaj and K. Nikodem, *Remarks on Bárány's theorem and affine selections*, Discrete Math., **224** (2000), 259–263.

[3] M. Balaj and Sz. Wąsowicz, *Haar spaces and polynomial selections*, Math. Pannon., **14** (2003), 63–70.

[4] S. Banach, *Théorie des opérations linéaires,* Warsaw, 1932.

[5] V. Bargmann, *Note on Wigner's theorem on symmetry operations,* J. Math. Phys., **5** (1964), 862–868.

[6] K. Baron, J. Matkowski, K. Nikodem, *A sandwich with convexity,* Math. Pannon., **5** (1994), 139–144.

[7] E. F. Beckenbach, *Generalized convex functions*, Bull. Amer. Math. Soc., **43** (1937), 363–371.

[8] E. Behrends and K. Nikodem, *A selection theorem of Helly type and its applications*, Studia Math., **116** (1995), 43–48.

[9] M. Bessenyei, *Hermite-Hadamard-type inequalities for generalized convex functions,* JIPAM. J. Inequal. Pure Appl. Math., **9**, Article 63, (2008) p. 51.

[10] M. Bessenyei, *The Hermite–Hadamard inequality in Beckenbach's setting*, J. Math. Anal. Appl., **364** (2010), 366–383.

[11] M. Bessenyei and Zs. Páles, *Hadamard-type inequalities for generalized convex functions,* Math. Inequal. Appl., **6**, (2003) 379–392.

[12] M. Bessenyei and Zs. Páles, *Separation by linear interpolation families*, J. Nonlinear Conv. Anal., **13** (2012), 49–56.

[13] M. Bessenyei, P. Szokol, *Convex separation by regular pairs*, J. Geom., **104** (2013), 45–56.

[14] M. Bessenyei, P. Szokol, *Separation by convex interpolation families*, J. Convex Anal., **20** (2013), 937–946.

[15] F.F. Bonsall, *The characterization of generalized convex functions* Quart. J. Math. Oxford Ser., **2** (1950), 100–111.

[16] P. Busch, *Stochastic Isometries in Quantum Mechanics,* Math. Phys., Analysis and Geometry, **2** (1999), 83–106.

[17] G.H. Chan, M.H. Lim, *Linear transformations on symmetric matrices that preserve commutativity,* Linear Algebra Appl., **47** (1982), 11–22.

[18] Z. Chebbi and M. Moakher, *Means of hermitian positive-definite matrices based on the log-determinant α-divergence function,* Linear Algebra Appl., **436** (2012), 1872–1889.

[19] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos, *Efficient similarity search for covariance matrices via the Jensen-Bregman LogDet divergence,* IEEE International Conference on Computer Vision (ICCV), (2011), 2399–2406.

[20] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos, *Jensen-Bregman LogDet divergence with application to efficient similarity search for covariance matrices,* IEEE Trans. Pattern Anal. Mach. Intell., **35** (2012), 2161–2174.

[21] J. B. Conway, *A Course in Functional Analysis,* Springer-Verlag, New York, 1985.

[22] G. Corach, H. Porta and L. Recht, *Geodesics and operator means in the space of positive operators,* Internat. J. Math., **4** (1993), 193–202.

[23] G. Corach, H. Porta and L. Recht, *Convexity of the geodesic distance on spaces of positive operators,* Illinois J. Math., **38** (1994), 87–94.

[24] G. Corach and A.L. Maestripieri, *Differential and metrical structure of positive operators,* Positivity, **3** (1999), 297–315.

[25] I. Csiszár, *Information type measure of difference of probability distributions and indirect observations,* Studia Sci. Math. Hungar., **2** (1967), 299–318.

[26] G. Dolinar and L. Molnár, *Isometries of the space of distribution functions with respect to the Kolmogorov-Smirnov metric,* J. Math. Anal. Appl., **348** (2008), 494–498.

[27] K. Fan, *Maximum properties and inequalities for the eigenvalues of completely continuous operator,* Proc. Nat. Acad. Sci. USA, **37** (1951), 760–766.

[28] C.A. Faure, *An elementary proof of the fundamental theorem of projective geometry,* Geom. Dedicata., **90** (2002), 145–151.

[29] Feller, W., *An Introduction to Probability Theory and its Applications, Vol. II*, John Wiley and Sons, New York, 1971.

[30] Fleming, R. J., Jamison, J. E., *Isometries on Banach spaces: function spaces. Vol. 1.*, Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics 129, Boca Raton, 2002.

[31] Fleming, R. J., Jamison, J. E., *Isometries on Banach spaces: Vector-valued function spaces. Vol. 2.*, Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics 138, Boca Raton, 2007.

[32] P.T. Fletcher and J. Sarang, *Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors,* Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis. Springer Berlin Heidelberg, **3117** (2004), 87–98.

[33] G. Frobenius, *Über die Darstellung der endlichen Gruppen durch lineare Substitutionen,* Sitzungsber. Königl. Preuss. Akad. Wiss. Berlin (1897), 994–1015.

[34] J.I. Fujii, *Differential geometries and their metrics on the positive operators,* RIMS Kôkyuroku No., **1632** (2009), 28–37.
http://www.kurims.kyoto-u.ac.jp/ kyodo/kokyuroku/contents/pdf/1632-04.pdf

[35] S. Furuichi, K. Yanagi and K. Kuriyama, *Fundamental properties of Tsallis relative entropy,* J. Math. Phys., **45** (2004), 4868–4877.

[36] Gy. P. Gehér, *An elementary proof for the non-bijective version of Wigner's theorem,* Phys. Lett. A, **378** (2014), 2054–2057.

[37] M. Győry, *A new proof of Wigner's theorem,* Rep. Math. Phys., **54** (2004), 159–167.

[38] O. Hatori, G. Hirasawa, T. Miura and L. Molnár, *Isometries and maps compatible with inverted Jordan triple products on groups,* Tokyo J. Math., **35** (2012), 385–410.

[39] A. Házy, Zs. Páles, *On approximately midconvex functions,* Bull. London Math. Soc., **36** (2004), 339–350.

[40] F. Hiai, M. Mosonyi, D. Petz and C. Bény, *Quantum f-divergences and error correction,* Rev. Math. Phys., **23** (2011), 691–747.

[41] F. Hiai and D. Petz, *The proper formula for relative entropy and its asymptotics in quantum probability,* Comm. Math. Phys., **143** (1991), 99–114.

[42] E. Hopf, *Über die Zusammenhänge zwischen gewissen höheren Differenzenquotienten reeller Funktionen einer reellen Variablen und deren Differenzierbarkeitseigenschaften*, Ph.D. thesis, Friedrich–Wilhelms–Universität Berlin, 1926.

[43] D.H. Hyers, S.M. Ulam, *Approximately convex functions,* Proc. Am. Math. Soc., **3** (1952), 821–828.

[44] S. Karlin and W. J. Studden, *Tchebycheff systems: With applications in analysis and statistics*, Pure and Applied Mathematics, Vol. XV, Interscience Publishers John Wiley & Sons, New York-London-Sydney, 1966.

[45] Krzyszkowski, J.: *Generalized convex sets*, Rocznyk Nauk.-Dydaktik Prace Mat., **14** (1997), 59–68.

[46] Krzyszkowski, J.: *Approximately generalized convex functions*, Math. Pannon., **12** (2001), 93–104.

[47] Lay, S.R.: *Convex Sets and Their Applications*, Wiley-Interscience, New York, (1982).

[48] F. Liese and I. Vajda, *Convex Statistical Distances,* B.G. Teubner Verlagsgesellschaft, Leipzig, 1987.

[49] J. S. Lomont and P. Mendelson, *The Wigner unitary-antiunitary theorem,* Ann. Math., **78** (1963), 548–559.

[50] G. G. Magaril-Il'yaev and V. M. Tikhomirov, *Convex Analysis: Theory and Applications*, Amer. Math. Soc., Providence, Rhode Island, 2003.

[51] S. Mazur and S. Ulam, *Sur les transformation d'espaces vectoriels normé,* C.R. Acad.Sci. Paris, **194** (1932), 946–948.

[52] L. Molnár, *An algebraic approach to Wigner's unitary-antiunitary theorem,* J. Austral. Math. Soc., **65** (1998), 354–369.

[53] M. Moakher, *A differential geometric approach to the geometric mean of symmetric positive-definite matrices,* SIAM. J. Matrix Anal. Appl., **26** (2005), 735–747.

[54] L. Molnár, *A remark on the Kochen-Specker theorem and some characterizations of the determinant on sets of hermitian matrices,* Proc. Amer. Math. Soc., **134** (2006), 2839–2848.

[55] L. Molnár, *Selected Preserver Problems on Algebraic Structures of Linear Operators and on Function Spaces,* Springer, Berlin Heidelberg, 2007.

[56] L. Molnár, *Maps on states preserving the relative entropy,* J. Math. Phys. **49** (2008), 032114.

[57] L. Molnár and P. Szokol, *Maps on states preserving the relative entropy II*, Linear Algebra Appl., **432** (2010), 3343–3350.

[58] L. Molnár, *Kolmogorov-Smirnov isometries and affine automorphisms of spaces of distribution functions*, Cent. Eur. J. Math., **9** (2011), 789–796.

[59] L. Molnár and G. Nagy, *Isometries and relative entropy preserving maps on density operators*, Linear Multilinear Algebra, **60** (2012), 93–108.

[60] L. Molnár, G. Nagy and P. Szokol, *Maps on density operators preserving quantum f-divergences*, Quantum Inf. Process, **12** (2013), 2309–2323.

[61] L. Molnár and P. Szokol, *Kolmogorov-Smirnov isometries of the space of generalized distribution functions*, Math. Slovaca, **64** (2014), 433–444.

[62] L. Molnár, *Jordan triple endomorpisms and isometries of spaces of positive definite matrices*, Linear and Multilinear Algebra, **63** (2015), 12–33.

[63] L. Molnár and P. Szokol, *Transformations on positive definite matrices preserving generalized distance measures,* Linear Algebra Appl., **466** (2015), 141–159.

[64] Niculescu, C.P. and Persson L.-E.:, *Convex Functions and Their Applications*, CMS Books in Mathematics, 23, Springer-Verlag, New York, 2006.

[65] M. A. Nielsen and I. L. Chuang *Quantum Computation and Quantum Information,* Cambridge University Press, Cambridge, 2000.

[66] K. Nikodem and Zs. Páles, *Generalized convexity and separation theorems*, J. Convex Anal., **14** (2007), no. 2, 239–248.

[67] K. Nikodem and Sz. Wąsowicz, *A sandwich theorem and Hyers-Ulam stability of affine functions*, Aequationes Math., **49** (1995), no. 1-2, 160–164.

[68] C. K. Li and S. Pierce, *Linear preserver problems,* Amer. Math. Monthly, **108** (2001), 591–605.

[69] C. K. Li and N. K. Tsing, *Linear preserver problems: A brief introduction and some special techniques,* Linear Algebra Appl., **162-164** (1992), 217–235.

[70] Zs. Páles, *On approximately convex functions*, Proc. Amer. Math. Soc., **131** (2003), 243–252.

[71] D. Petz, *Quasi-entropies for states of a von Neumann algebra,* Publ. RIMS. Kyoto Univ., **21** (1985), 781–800.

[72] D. Petz, *Quasi-entropies for finite quantum systems,* Rep. Math. Phys., **23** (1986), 57–65.

[73] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer, Berlin-Heidelberg, 2008.

[74] D. Petz, *From f-divergence to quantum quasi-entropies and their use,* Entropy, **12** (2010), 304–325.

[75] Pólya, Gy.: *On the mean-value theorem corresponding to a given linear homogeneous differential equation*, Trans. Amer. Math. Soc., **24** (1922), 312–324.

[76] Pólya, Gy.: *Untersuchungen über Lücken und Singularitäten von Potenzreihen*, Math. Z., **29**, 54.

[77] T. Popoviciu, *Les fonctions convexes*, Hermann et Cie, Paris, 1944.

[78] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N. J., 1970.

[79] S. Sra, *Positive definite matrices and the symmetric Stein divergence,* arXiv: [math.FA-1110.1773v2]

[80] Shiryaev, A. N.: *Probability*, Graduate Texts in Mathematics, No. 95, Springer-Verlag, 1996.

[81] J.C. Taylor, *An Introduction to Measure and Probability,* Springer, 1997.

[82] L. Tornheim, *On n-parameter families of functions and associated convex functions*, Trans. Amer. Math. Soc., **69** (1950), 457–467.

[83] U. Uhlhorn, *Representation of symmetry transformations in quantum mechanics*, Ark. Fysik, **23** (1963), 307-340.

[84] J. Väisälä, *A proof of the Mazur-Ulam theorem,* Amer. Math. Monthly, **110** (2003), 633–635 .

[85] M. L. J. van de Vel, *Theory of convex structures*, North-Holland Mathematical Library, vol. 50, North-Holland Publishing Co., Amsterdam, 1993.

[86] A. Vogt, *Maps which preserve equality of distance,* Studia Math., **45** (1973) 43–48.

[87] Sz. Wąsowicz, *Polynomial selections and separation by polynomials*, Studia Math., **120** (1996), no. 1, 75–82.

[88] E. P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektrum,* Fredrik Vieweg und Sohn, 1931.