QUANTUM ENTROPIES, RELATIVE ENTROPIES, AND RELATED PRESERVER PROBLEMS

DÁNIEL VIROSZTEK

Department of Analysis,
Budapest University of Technology and Economics, Hungary

Supervisor: Prof. Dénes Petz

Thesis booklet

2016

1. PRELIMINARIES

1.1. **Introduction.** The classical work [24] of *Andrey Nikolaevich Kolmogorov* laid the foundations of probability theory in 1933. In Kolmogorov's approach, the basic concept of probability theory is the *probability space.* A probability space is a triplet $(X, \mathscr{A}, \mathbf{P})$, where X is an arbitrary set, $\mathscr{A} \subseteq P(X)$ is a σ-algebra — $P(X)$ denotes the power set of X — and $\mathbf{P}$ is a finite measure on $\mathscr{A}$ which is normalized, that is, $\mathbf{P}(X) = 1$. This means that a probability space is nothing else but a measure space with total measure one, so one may consider probability theory as a branch of measure theory. On the other hand, probability theory is a richer structure than measure theory in the sense that several measure theoretical notions gain intuitive meanings from the viewpoint of a probability theorist. Without the requirement of generality, let us mention some of the intuitions which are associated with the notions of measure theory. The most basic concept is that the *measurable sets* — that is, the elements of the σ-algebra $\mathscr{A}$ — are considered to be *events.* A *measurable function* $f : (X, \mathscr{A}) \to (\mathbb{K}, \mathscr{B})$ is called a real/complex *random variable* if $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, respectively. Therefore, the *Lebesgue integral* $\int_X f \mathrm{d}\mathbf{P}$ of the measurable function f is called the *expected value* — if it exists. As $\mathbf{P}$ is a finite measure, it is quite easy to guarantee the existence of the integral of a measurable function. If f is essentially bounded, that is, $\mathbf{P}\left(\{x \in X : \left|f(x)\right| > K\}\right) = 0$ for some $K > 0$, then f is integrable, moreover, any power of f is integrable. This latter fact is remarkable as the integral $\int_X f^k \mathrm{d}\mathbf{P}$ is called the kth *moment* of the random variable f and plays an important role in probability theory. Let us denote by $L^\infty(X, \mathscr{A}, \mathbf{P})$ the set of essentially bounded measurable complex valued functions on the probability space $(X, \mathscr{A}, \mathbf{P})$. Let us introduce the notation

$$L^2(X, \mathscr{A}, \mathbf{P}) = \left\{ f : X \to \mathbb{C} \,\middle|\, f \text{ is measurable and } \int_X |f|^2 \, \mathrm{d}\mathbf{P} < \infty \right\},$$

as well. Clearly, $L^2(X, \mathscr{A}, \mathbf{P})$ is a Hilbert space with the inner product $\langle f, g\rangle = \int_X \overline{f} g \mathrm{d}\mathbf{P}$. Every bounded measurable function $f : X \mapsto \mathbb{C}$ determines a bounded linear operator on the Hilbert space $L^2(X, \mathscr{A}, \mathbf{P})$ in the following way. Set $f \in L^\infty(X, \mathscr{A}, \mathbf{P})$. Let us define the *multiplication operator* M_f by

$$M_f : L^2(X, \mathscr{A}, \mathbf{P}) \to L^2(X, \mathscr{A}, \mathbf{P}), \, g \mapsto M_f(g) := fg.$$

Straightforward computations show that M_f is linear, and the proof of the boundedness of M_f is quite easy, as well. So, $M_f \in \mathscr{B}\left(L^2(X, \mathscr{A}, \mathbf{P})\right)$ for any $f \in L^\infty(X, \mathscr{A}, \mathbf{P})$. Moreover, the operator norm of M_f coincides with the supremum norm of f, that is, $\left\|M_f\right\| = \|f\|_\infty$. This latter fact is also rather easy to prove. The map

$$M : L^{\infty}(X, \mathscr{A}, \mathbf{P}) \to \mathscr{B}\left(L^2(X, \mathscr{A}, \mathbf{P})\right),\ f \mapsto M_f \tag{1}$$

is a canonical isometric embedding of the commutative normed algebra $L^{\infty}(X, \mathscr{A}, \mathbf{P})$ into the normed algebra $\mathscr{B}\left(L^2(X, \mathscr{A}, \mathbf{P})\right)$, which is far from being commutative in general. This embedding is the starting point of the noncommutative generalization of probability theory.

1.2. The basics of noncommutative probability theory.

Definition 1 (Normed algebra)**.** *A unital complex algebra $\mathscr{A}$ endowed with the norm $\|.\|$ is said to be a* normed algebra, *if the norm is submultiplicative, i. e., $\|ab\| \leq \|a\| \|b\|$ for any $a, b \in \mathscr{A}$ and the identity element is of norm one, that is, $\|1_{\mathscr{A}}\| = 1$.*

Definition 2 (Banach algebra)**.** *A normed algebra which is a Banach space — that is, a complete normed space — is called a* Banach algebra.

Definition 3 (Involution)**.** *Let $\mathscr{A}$ be a complex algebra. A map $* : \mathscr{A} \to \mathscr{A}$, $a \mapsto a^*$ is called an* involution *if it satisfies the following properties.*

- *$*$ is antilinear: $(\lambda a + b)^* = \overline{\lambda} a^* + b^*$ for any $a, b \in \mathscr{A}$ and $\lambda \in \mathbb{C}$.*
- *$*^2 = \mathrm{id}$, that is, $(a^*)^* = a$ for any $a \in \mathscr{A}$.*
- *$*$ is an antihomomorphism with recpect to the product: $(ab)^* = b^* a^*$ for any $a, b \in \mathscr{A}$.*

Definition 4 (C^*-algebra)**.** *A Banach algebra endowed with an involution $* : \mathscr{A} \to \mathscr{A}$ which satisfies $\|a^* a\| = \|a\|^2$ for any $a \in \mathscr{A}$ is called a C^*-*algebra.

The above definition of C^*-algebras is rather abstract. However, we do not loose any generality if we consider the elements of a C^*-algebra as bounded operators on an appropriate Hilbert space. Indeed, any C^*-algebra is isomorphic to a closed (in the operator norm topology) unital *-subalgebra (that is, it is closed under the involution) of the operator algebra $\mathscr{B}(\mathscr{H})$ for a suitable Hilbert space $\mathscr{H}$. Furthermore, any commutative C^*-algebra is isomorphic to $C(X)$ for some compact Hausdorff space X. (The symbol $C(X)$ denotes the algebra of all continuous complex-valued functions defined on X endowed with the supremum norm.)

Despite the above remarkable facts, the C^*-algebra is still a bit too general notion to formalize the concepts of noncommutative probability theory. With an extra topological assumption we achieve the desired level of generality.

Definition 5 (von Neumann algebra)**.** *A C^*-algebra which is closed not just in the operator norm but also in the weak operator topology is called a* von Neumann algebra.

Note that the above definition is correct as any C^*-algebra is isomorphic to an algebra of bounded linear operators on a Hilbert space $\mathscr{H}$, hence the condition about the closedness in the weak operator topology makes sense. The weak operator topology on $\mathscr{B}(\mathscr{H})$ is defined by the family of seminorms $\{p_{x,y} : x, y \in \mathscr{H}\}$ where $p_{x,y}(A) = |\langle Ax, y\rangle|$ ($A \in \mathscr{B}(\mathscr{H})$).

Now, we are in the position to answer the question *why we call von Neumann algebra theory sometimes noncommutative probability theory?*

It is clear by Definition 2 that the function space $L^\infty(X, \mathscr{A}, \mathbf{P})$ is a commutative Banach algebra for any probability space $(X, \mathscr{A}, \mathbf{P})$. Furthermore, it is easy to see that these Banach algebras are also C^*-algebras with the complex conjugation as involution. It is folklore that $L^\infty(X, \mathscr{A}, \mathbf{P})$ is the Banach dual of the Banach space $L^1(X, \mathscr{A}, \mathbf{P})$, which is defined as follows:

$$L^1(X, \mathscr{A}, \mathbf{P}) = \left\{ f : X \to \mathbb{C} \,\middle|\, f \text{ is measurable and } \int_X |f| \, \mathrm{d}\mathbf{P} < \infty \right\}.$$

So, $L^\infty(X, \mathscr{A}, \mathbf{P})$ is a commutative C^*-algebra which is the dual of the Banach space $L^1(X, \mathscr{A}, \mathbf{P})$. It follows that $L^\infty(X, \mathscr{A}, \mathbf{P})$ is a commutative von Neumann algebra. The interesting fact is that the converse statement is also true. That is, every abelian von Neumann algebra is isomorphic to $L^\infty(X, \mathscr{S}, \mu)$ for some localizable measure space $(X, \mathscr{S}, \mu)$, see [36]. (A localizable measure space is the direct sum of finite measure spaces.)

We can deduce that every probability space determines a commutative von Neumann algebra — the algebra of the bounded random variables — and every commutative von Neumann algebra determines a probability space, up to harmless normalization. That is the reason why the theory of von Neumann algbras may be considered as noncommutative probability theory.

2. The main results

We finished the previous section with the description of the correspondence between probability spaces and abelian von Neumann algebras. Fortunately, several interesting and useful notions of probability theory can be extended to the general von Neumann algebra setting. We focus on two distinguished concepts of probability theory, namely the *(co)variance* and the *entropy*.

2.1. Decomposition of quantum covariances ([4]).

First, we investigate the following problem. Can we characterize those sets of *observables* for which the induced covariance mapping is a *roof*? (See Def. 7 for the definition of roof.) Note that this question does not make sense in the case of abelian von Neumann algebra for the following reason. It is known that every *pure state* is multiplicative on a commutative von Neumann algebra, see, e.g., [22, 4.4.1. Prop.]. Therefore, the covariance of any two observables is zero in any pure state. So, the covariance mapping is a roof if and only if it is identically zero which is clearly not the case.

Let $\mathscr{A}$ be a von Neumann algebra of type I_n and let ϕ be a — necessarily normal — state on $\mathscr{A}$. The *covariance* of the self-adjoint elements $A, B \in \mathscr{A}$ is defined by

$$\mathrm{Cov}_\phi(A,B) = \phi(AB) - \phi(A)\phi(B).$$

In particular, the variance of the observable (self-adjoint elements are often called observables) A in the state ϕ is given by

$$\mathrm{Var}_\phi(A) = \mathrm{Cov}_\phi(A,A) = \phi\left(A^2\right) - \left(\phi(A)\right)^2.$$

It is rather easy to check that

$$\mathrm{Var}_\phi(A + \lambda I_{\mathscr{A}}) = \mathrm{Var}_\phi(A) \qquad (A \in \mathscr{A}, \lambda \in \mathbb{R})$$

holds for any state ϕ.

It is useful to introduce the *covariance matrix* of several observables. If $A_1, \dots A_r$ are self-adjoint elements of $\mathscr{A}$, then their covariance matrix is defined as

$$\left[\mathbf{Cov}_\phi(A_1,\dots,A_r)\right]_{i,j} := \mathrm{Cov}_\phi\left(A_i, A_j\right) \qquad (1 \le i, j \le r).$$

Observe that the above defined covariance matrix is necessarily self-adjoint as $\phi\left(A_i A_j\right) = \overline{\phi\left(A_j A_i\right)}$.

One of the most important properties of the covariance is that it is a concave map on the set of states, that is, the mapping

$$\mathbf{Cov}_{(.)}(A_1,\dots,A_r) : \mathscr{S}_{\mathscr{A}} \to \mathbf{M}_r^{sa}; \phi \mapsto \mathbf{Cov}_\phi(A_1,\dots,A_r) \tag{2}$$

is concave with respect to the *Loewner ordering* on the final space $\mathbf{M}_r$. (For any $A, B \in \mathbf{M}_r^{sa}$ we say that $A \le B$ if $B - A$ is a positive semidefinite matrix.)

As the von Neumann algebra $\mathscr{A}$ is of type I_n — that is, it is isomorphic to the operator algebra $\mathscr{B}(\mathscr{H})$ for an n-dimensional complex Hilbert space $\mathscr{H}$, — every state is represented by a unique density operator. For the sake of simplicity, we will use the following notation. If the state ϕ is represented by the density operator D, then we define $\mathrm{Cov}_D(.,.) := \mathrm{Cov}_\phi(.,.)$, and so on, $\mathrm{Var}_D(.) := \mathrm{Var}_\phi(.)$ and $\mathbf{Cov}_D(.,.,\dots,.) := \mathbf{Cov}_\phi(.,.,\dots,.)$.

Using this notation, the above declared concavity of the covariance matrix map (2) can be written as

$$\mathbf{Cov}_D(A_1,\dots,A_r) \ge \sum_{k=1}^m \lambda_k \mathbf{Cov}_{D_k}(A_1,\dots,A_r) \quad \text{if} \quad D = \sum_{k=1}^m \lambda_k D_k,$$

where $\lambda_k \geq 0$ and $\sum_{k=1}^m \lambda_k = 1$.

For any inequality, it is an interesting task to investigate the case of equality. For such an investigation, a useful tool is the recently introduced notion of *roof* which is defined as follows.

Definition 6 (Roof point). *Let Ω be a compact convex set contained in a finite dimensional real linear space. Let G be a mapping from Ω into a partially ordered set. A point $\omega \in \Omega$ is called* roof point, *if there are some extremal points $\pi_1, \dots, \pi_m$ of Ω and nonnegative numbers $p_1, \dots, p_m$ with $\sum_{k=1}^m p_k = 1$ such that*

$$\sum_{k=1}^m p_k \pi_k = \omega$$

and

$$\sum_{k=1}^m p_k G(\pi_k) = G(\omega).$$

Definition 7 (Roof). *A mapping G defined on Ω is called* roof *if every $\omega \in \Omega$ is a roof point.*

As $\mathscr{A}$ is finite dimensional, the set of the density operators is a compact convex subset of the real vector space of the self-adjoint elements of $\mathscr{A}$. We are interested in the following question. *Is the concave mapping* (2) *a roof on $\mathscr{S}_{\mathscr{A}}$?* It is well-known that the extremal points of the set of densities are exactly the rank-one projections. So we can reformulate our question. Given an arbitrary density D, can we find rank one projections $P_1, \dots, P_m$ and nonnegative weights $p_1, \dots, p_m$ (with $\sum_{k=1}^m p_k = 1$) such that

$$D = \sum_{k=1}^m p_k P_k \tag{3}$$

and

$$\mathbf{Cov}_D(A_1, \dots, A_r) = \sum_{k=1}^m p_k \mathbf{Cov}_{P_k}(A_1, \dots, A_r)?$$

We say that (3) is an *extremal convex decomposition* of D.

For $r = 1$ the answer is positive, and this is the first result in this topic, made by *Petz* and *Tóth* [35]. An extension of the former result was given by *Petz* and *Léka* in [26]. They proved that the answer is positive even in the case $r = 2$. We give a necessary and sufficient condition for the covariance mapping (2) being a roof in terms of the corresponding observables. Our result applies for any finite collection of observables, and it recovers all the aforementioned results easily.

Recall that our von Neumann algebra $\mathscr{A}$ is (isomorphic to) the operator algebra $\mathscr{B}(\mathscr{H})$, where $\mathscr{H}$ is a Hilbert space of dimension n. For an arbitrary subspace $\mathscr{K} \subset \mathscr{H}$, we denote by $Q^{\mathscr{K}}$ the orthogonal projection onto $\mathscr{K}$. We define

$$A^{\mathscr{K}} := Q^{\mathscr{K}} A Q^{\mathscr{K}}$$

for every element $A \in \mathscr{A}$ and

$$\mathscr{B}(\mathscr{K}) := Q^{\mathscr{K}} \mathscr{B}(\mathscr{H}) Q^{\mathscr{K}}, \qquad \mathscr{B}^{sa}(\mathscr{K}) := Q^{\mathscr{K}} \mathscr{B}^{sa}(\mathscr{H}) Q^{\mathscr{K}},$$
$$\mathscr{B}^{+}(\mathscr{K}) := Q^{\mathscr{K}} \mathscr{B}^{+}(\mathscr{H}) Q^{\mathscr{K}}, \qquad \mathscr{S}(\mathscr{K}) := \{X \in \mathscr{B}^{+}(\mathscr{K}) : \operatorname{Tr} X = 1\}.$$

Definition 8. *Let* $\{A_1, \ldots, A_r\}$ *be a set of self-adjoint elements of* $\mathscr{A} = \mathscr{B}(\mathscr{H})$. *The set* $\{A_1, \ldots, A_r\}$ *is said to be* variance-decomposable *if for every* $D \in \mathscr{S}_{\mathscr{A}}$ *there exists an extremal convex decomposition*

$$D = \sum_{k=1}^{m} \lambda_k P_k$$

of D *such that*

$$\mathbf{Cov}_D(A_1, \ldots, A_r) = \sum_{k=1}^{m} \lambda_k \mathbf{Cov}_{P_k}(A_1, \ldots, A_r)$$

In other words, $\{A_1, \ldots, A_r\}$ is variance-decomposable if and only if the mapping $D \mapsto \mathbf{Cov}_D(A_1, \ldots, A_r)$ is a roof. Our main result reads as follows.

Theorem 9. *The set* $\{A_1, \ldots, A_r\} \subset \mathscr{A}$ *is variance-decomposable if and only if*

$$\dim\left(\operatorname{span}\left\{I^{\mathscr{K}}, A_1^{\mathscr{K}}, \ldots, A_r^{\mathscr{K}}\right\}\right) < (\dim \mathscr{K})^2 \tag{4}$$

for every subspace $\mathscr{K} \subset \mathscr{H}$ *with* $\dim \mathscr{K} > 1$.

2.2. **Inequalities for Tsallis entropy related to the strong subadditivity ([5]).** In this subsection the strong subadditivity inequality of the entropy is investigated. Fairly nontrivial but rather easy computations show that the *Shannon entropy* is strongly subadditive. In my opinion, a much more sophisticated argument shows that its noncommutative counterpart, the *von Neumann entropy* is also strongly subadditive. The latter statement is a celebrated result of Lieb and Ruskai [27]. We consider a one-parameter generalization of the von Neumann entropy which is called *Tsallis entropy*. We show — in particular — that the Tsallis entropy is not strongly subadditive for noncommutative von Neumann algebras in spite of the facts that it is strongly subadditive in the commutative case [18, Thm 3.4] and that it is subadditive in the noncommutative case, as well [9].

Let $\mathscr{A}$ be a von Neumann algebra of type I_n and let us denote by $\mathscr{H}$ the underlying n-dimensional Hilbert space — that is, $\mathscr{A} = \mathscr{B}(\mathscr{H})$. Let ρ be a density operator which represents a state on $\mathscr{A}$. Note that in this case $\rho \in \mathscr{A}$ and the expression $f(\rho)$ makes sense by the continuous functional calculus for any complex function f which is continuous on the spectrum of ρ. The von Neumann entropy of the density operator ρ is defined by

$$S(\rho) = -\mathrm{Tr}\rho \ln \rho \tag{5}$$

see, e.g., [12, 20, 34]. Let the Hilbert space $\mathscr{H}$ be the tensor product of three finite dimensional Hilbert spaces, that is, $\mathscr{H} := \mathscr{H}_1 \otimes \mathscr{H}_2 \otimes \mathscr{H}_3$. Let $\rho_{123} \in \mathscr{B}(\mathscr{H})$ be a density operator. The *reduced densities* are defined by partial traces. Let us use the following notation.

$$\rho_{12} := \mathrm{Tr}_3\, \rho_{123}, \qquad \rho_2 := \mathrm{Tr}_1\, \rho_{12}, \qquad \rho_{23} := \mathrm{Tr}_1\, \rho_{123}. \tag{6}$$

As in our case the states and the density operators are in one-to-one correspondence, densities will be called sometimes states, and we will refer to reduced densities sometimes by the expression *reduced state*.

One of the most important results in quantum information theory is the strong subadditivity of the von Neumann entropy, which is the following inequality.

$$S(\rho_{123}) + S(\rho_2) \le S(\rho_{12}) + S(\rho_{23})\,.$$

This result was made by E. Lieb and M. B. Ruskai in 1973 [27, 34]. Our aim is to generalize this inequality in various ways. The key object of our investigations is a certain generalization of the von Neumann entropy which is called *Tsallis entropy*.

The Tsallis entropy is a one-parameter extension of the von Neumann entropy. For any real q, one can define the deformed logarithm (or q-logarithm) function $\ln_q : (0,\infty) \to \mathbb{R}$ by

$$\ln_q x := \int_1^x t^{q-2}\mathrm{d}t = \begin{cases} \frac{x^{q-1}-1}{q-1} & \text{if } q \neq 1\,, \\ \ln x & \text{if } q = 1\,. \end{cases} \tag{7}$$

The corresponding entropy

$$S_q(\rho) = -\mathrm{Tr}\rho \ln_q \rho$$

is called Tsallis entropy [8, 16]. It is reasonable to restrict ourselves to the $0 < q$ case, because $\lim_{x \to 0^+} -x \ln_q x = 0$ if and only if $0 < q$. If we introduce the notation $f_q(x) = x \ln_q x$ we can write $S_q(\rho) = -\mathrm{Tr} f_q(\rho)$.

2.2.1. *The Tsallis entropy is subadditive, but not strongly subadditive.* Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be finite dimensional Hilbert spaces. If ρ_{12} is a state on a Hilbert space $\mathscr{H}_1 \otimes \mathscr{H}_2$ — that is, $\rho_{12} \in \mathscr{B}(\mathscr{H}_1 \otimes \mathscr{H}_2)$ such that $0 \le \rho_{12}$ and $\mathrm{Tr}\, \rho_{12} = 1$, — then it has reduced states $\rho_1 := \mathrm{Tr}_2\, \rho_{12}$ and $\rho_2 := \mathrm{Tr}_1\, \rho_{12}$ on the spaces $\mathscr{H}_1$ and $\mathscr{H}_2$, respectively. The subadditivity inequality of the Tsallis entropy is

$$S_q(\rho_{12}) \le S_q(\rho_1) + S_q(\rho_2), \tag{8}$$

and it has been proved for $q > 1$ by Audenaert in 2007 [9].

However, the strong subadditivity inequality

$$S_q(\rho_{123}) + S_q(\rho_2) \le S_q(\rho_{12}) + S_q(\rho_{23}) \tag{9}$$

does not hold in general.

Theorem 10. *The only strongly subadditive Tsallis entropy is the von Neumann entropy, that is, the strong subadditivity of the Tsallis entropy holds if and only if* $q = 1$.

Therefore, our goal is to find an inequality

$$S_q(\rho_{123}) + S_q(\rho_2) \le S_q(\rho_{12}) + S_q(\rho_{23}) + g_q(\rho_{123}), \tag{10}$$

where $g_1(\rho_{123}) = 0$. Such a result may be considered as a generalization of the strong subadditivity inequality.

The strong subadditivity of the von Neumnann entropy can be derived from the monotonicity of the Umegaki relative entropy, which is a particular quasi-entropy [15, 33]. Therefore, it seems to be useful to reformulate the strong subadditivity of the Tsallis entropy as an inequality of certain quasi-entropies.

Theorem 11. *Let* ρ_{123} *be an element of* $\mathscr{B}^{++}(\mathscr{H}_1 \otimes \mathscr{H}_2 \otimes \mathscr{H}_3)$. *The strong subadditivity inequality of the Tsallis entropy (9) is equivalent to*

$$S^U_{-\ln_q}\left(\rho_{123} \,||\, \rho_{12} \otimes I_3\right) \ge S^V_{-\ln_q}\left(\rho_{23} \,||\, \rho_2 \otimes I_3\right), \tag{11}$$

where

$$U = \rho_{123}^{\frac{1}{2}(q-1)}, \qquad V = \rho_{23}^{\frac{1}{2}(q-1)}. \tag{12}$$

Using the previous statement, the following theorem provides an inequality which is of the form (10).

Theorem 12. *For any* $0 < q \le 2$ *the inequality*

$$S_q(\rho_{12}) + S_q(\rho_{23}) - S_q(\rho_{123}) - S_q(\rho_2)$$
$$\ge (q-1)\left(S_{\ln_q}^{(-\ln_q \rho_{123})^{\frac{1}{2}}}\left(\rho_{123} || \rho_{12} \otimes I_3\right) - S_{\ln_q}^{(-\ln_q \rho_{23})^{\frac{1}{2}}}\left(\rho_{23} \,||\, \rho_2 \otimes I_3\right)\right)$$

holds.

Moreover, we can find a sufficient condition concerning the structure of the state ρ_{123} which ensures the strong subadditivity.

Theorem 13. *If* ρ_{123} *and* $I_1 \otimes \rho_{23}$ *commute, and (using the notation* $\rho_{123} = \sum_j \lambda_j \left|\varphi_j\right\rangle\left\langle\varphi_j\right|$ *and* $\rho_{12} \otimes I_3 = \sum_k \mu_k \left|\psi_k\right\rangle\left\langle\psi_k\right|$*) we have* $\lambda_j \le \mu_k$ *whenever* $\left\langle\psi_k|\varphi_j\right\rangle \ne 0$, *then for any* $1 \le q \le 2$ *the strong subadditivity inequality*

$$S_q(\rho_{123}) + S_q(\rho_2) \le S_q(\rho_{12}) + S_q(\rho_{23})$$

holds.

Note that if ρ_{123} is a classical probability distribution (that is, $\rho_{123} = \mathrm{Diag}(\{p_{jkl}\})$), then the conditions of Theorem 13 are clearly satisfied.

2.3. **Joint convexity Bregman divergences ([6]).** In this subsection we introduce the *Bregman divergences* which may be considered as certain generalizations of the *Umegaki relative entropy*. We characterize those Bregman divergences which are jointly convex, and we use this result to derive a sharp inequality for Tsallis entropy which can be considered as a generalization of the strong subadditivity inequality of the von Neumann entropy.

In applications that involve measuring the dissimilarity between two objects (numbers, vectors, matrices, functions and so on) the definition of a divergence becomes essential. One such measure is a distance function, but there are many important measures which do not satisfy the properties of distance. For instance, the square loss function has been used widely for regression analysis, Kullback-Leibler divergence [25] has been applied to compare two probability density functions, the Itakura-Saito divergence [21] is used as a measure of the perceptual difference between spectra, or the Mahalonobis distance [28] is to measure the dissimilarity between two random vectors of the same distribution. The Bregman divergence was introduced by Lev Bregman [13] for convex functions $\phi : \mathbb{R}^d \to \mathbb{R}$ with gradient $\nabla\phi$, as the ϕ-dependent nonnegative measure of discrepancy

$$D_\phi(p,q) = \phi(p) - \phi(q) - \langle \nabla\phi(q), p-q \rangle \tag{13}$$

of d-dimensional vectors $p, q \in \mathbb{R}^d$. Originally his motivation was the problem of convex programming, but it became widely researched both from theoretical and practical viewpoints. For example the remarkable fact that all the aforementioned divergences are special cases of the Bregman divergence shows its importance [10]. In some literature it is applied under the name Bregman distance, in spite of that it is not in general a metric. Indeed, D_ϕ is definite, but does not satisfy the triangle inequality nor symmetry.

2.3.1. *Definition and basic properties.* Let the Hilbert space $\mathscr{H}$ be finite dimensional, as usual. Let $f : (0,\infty) \to \mathbb{R}$ be a convex function. Then the induced map

$$\varphi_f : \mathscr{B}^{++}(\mathscr{H}) \to \mathbb{R}, \;\; X \mapsto \varphi_f(X) := \operatorname{Tr} f(X)$$

is convex, as well [15]. A differentiable convex function is bounded from below by its first-order Taylor polynomial, no matter what the base point is. Therefore, the expression

$$\varphi_f(X) - \varphi_f(Y) - \mathbf{D}\varphi_f[Y](X-Y),$$

where $\mathbf{D}\varphi_f[Y]$ denotes the Fréchet derivative of φ_f at the point Y, is non-negative for any $X, Y \in \mathscr{B}^{++}(\mathscr{H})$. By the linearity of the trace, for any $Y \in \mathscr{B}^{++}(\mathscr{H})$ we have $\mathbf{D}\varphi_f[Y] = \operatorname{Tr} \circ \mathbf{D}f[Y]$, where $\mathbf{D}f[Y]$ denotes the Fréchet derivative of the standard operator function $f : \mathscr{B}^{++}(\mathscr{H}) \to \mathscr{B}^{sa}(\mathscr{H})$ at Y. Let us define the central object of this investigation precisely.

Definition. *Let $f \in C^1((0,\infty))$ be a convex function and $X, Y \in \mathscr{B}^{++}(\mathscr{H})$. The Bregman f-divergence of X and Y is defined by*

$$H_f(X,Y) = \operatorname{Tr}\left(f(X) - f(Y) - \mathbf{D}f[Y](X-Y)\right). \tag{14}$$

We investigate the Bregman f-divergence from the viewpoint of joint convexity, which is essential in the further applications. Since f is convex, it is clear that the Bregman divergence is convex in the first variable. For the original Bregman divergence (13) Bauschke and Borwein show [11] that D_ϕ is jointly convex - i. e.

$$D_\phi(tp_1 + (1-t)p_2, tq_1 + (1-t)q_2) \le tD_\phi(p_1,q_1) + (1-t)D_\phi(p_2,q_2),$$

where $p_1, p_2, q_1, q_2 \in \mathbb{R}^d$, $t \in [0,1]$ - if and only if the inverse of the Hessian of ϕ is concave in Loewner sense. Particularly, if ϕ is an $\mathbb{R} \supset I \to \mathbb{R}$ convex function, then D_ϕ is jointly convex if and only if $1/\phi''$ is concave. From this viewpoint the next characterization is rather interesting.

Theorem 14. *Let $f \in C^2((0,\infty))$ be a convex function with $f'' > 0$ on $(0,\infty)$. Then the following conditions are equivalent.*

(1) *The map*

$$\mathscr{B}^{++}(\mathscr{H}) \to \mathscr{B}\left(\mathscr{B}^{sa}(\mathscr{H})\right); \quad X \mapsto \left(\mathbf{D}f'[X]\right)^{-1}$$

is operator concave.

(2) *The Bregman f-divergence*

$$H_f : \mathscr{B}^{++}(\mathscr{H}) \times \mathscr{B}^{++}(\mathscr{H}) \to [0,\infty); \quad (X,Y) \mapsto H_f(X,Y)$$

is jointly convex.

Moreover, we can provide a sufficient condition for the joint convexity of the Bregman f-divergence.

Theorem 15. *Let $f \in C^2((0,\infty))$ be a convex function. If f'' is operator convex and numerically non-increasing, then the Bregman f-divergence*

$$H_f : \mathscr{B}^{++}(\mathscr{H}) \times \mathscr{B}^{++}(\mathscr{H}) \to [0,\infty); \quad (X,Y) \mapsto H_f(X,Y)$$

is jointly convex.

As an application of the previous theorem, we derived a sharp inequality for Tsallis entropies which generalizes the strong subadditivity of the von Neumann entropy.

Theorem 16. *If $\mathscr{H}_i$ is a finite dimensional Hilbert space for any $i \in \{1,2,3\}$, $d_i = \dim \mathscr{H}_i$, $1 \leq q \leq 2$, then for any $\rho_{123} \in \mathscr{B}^+(\mathscr{H}_1 \otimes \mathscr{H}_2 \otimes \mathscr{H}_3)$ the inequality*

$$d_3^{1-q} \operatorname{Tr} \rho_{12}^q + d_1^{1-q} \operatorname{Tr} \rho_{23}^q \leq \operatorname{Tr} \rho_{123}^q + (d_1 d_3)^{1-q} \operatorname{Tr} \rho_2^q. \tag{15}$$

holds, where notations like ρ_{12} denote the appropriate reduced operators.

2.4. **Preservers of Bregman and Jensen divergences ([1]).** It is quite easy to see that the Bregman divergences of positive definite operators are invariant under unitary conjugations. It is also not so hard to show that unitary conjugations are not the only transformations of the positive definite cone which preserve the Bregman divergences. It is a very natural goal to determine all the transformations on the set of positive definite operators which leave the Bregman divergences invariant. This question leads us to the topic of *preserver problems.*

A preserver problem consists of the following ingredients. Let H be a set. Let $\phi : H \to H$ be a mapping. Let m be a positive integer, let K be a set and let $X : H^m \to K$ be a map. We say that the transformation ϕ *preserves* X, if either

$$X\left(\phi(A_1), \ldots, \phi(A_m)\right) = X(A_1, \ldots, A_m) \qquad (A_1, \ldots, A_m \in H), \tag{16}$$

or

$$X\left(\phi(A_1), \ldots, \phi(A_m)\right) = \phi(X(A_1, \ldots, A_m)) \qquad (A_1, \ldots, A_m \in H) \tag{17}$$

holds, depending on the nature of the map X. (The equation (17) may play the role of the preserver equation only if $K = H$.) For any given sets H, K and mapping X, the solution of the preserver problem is the description of the structure of all the transformations ϕ which preserve X.

The following table enumerates some preserver problems.

H	m	K	X	Equation	Name of the problem
$\mathbb{R}^n$	2	$[0,\infty)$	$(a,b)\mapsto \|a-b\|$	(16)	isometries of $\mathbb{R}^n$
$\mathbf{M}_n$	1	$\mathbb{C}$	$A\mapsto \mathrm{Det}A$	(16)	determinant preserving maps
$\mathbf{M}_n^{sa}$	2	$\{0,1\}$	$(A,B)\mapsto \mathbb{1}_{A\le B}$	(16)	order preserving maps
$\mathbf{M}_n^{++}$	2	$\mathbf{M}_n^{++}$	$(A,B)\mapsto ABA$	(17)	triple product preserving maps
$\mathbf{M}_n^{+}$	2	$[-\infty,\infty]$	$(A,B)\mapsto S_f(A,B)$	(16)	preservers of the quantum f-divergence
$\mathbf{M}_n^{++}$	m	$\mathbf{M}_n^{++}$	$(A_1,\dots,A_m)\mapsto M_G(A_1,\dots,A_m)$	(17)	preservers of the multivariable geometric mean

The above table makes it transparent that the topic of preserver problems covers a large area of mathematics. An exhaustive description such problems — including *Frobenius' theorem* on determinant preserving maps, the *Mazur-Ulam theorem* on isometries of real normed spaces and *Wigner's theorem* on the symmetry transformations of pure states with respect to the *transition probability* — can be found of in the monography [31] written by *Lajos Molnár*.

Let $\mathscr{H}$ be a finite dimensional Hilbert space, as usual. For a differentiable convex function f on $(0,\infty)$, the Bregman f-divergence on $\mathscr{B}^{++}(\mathscr{H})$ is defined by

$$H_f(X,Y)=\mathrm{Tr}\left(f(X)-f(Y)-f'(Y)(X-Y)\right),\quad X,Y\in\mathscr{B}^{++}(\mathscr{H})$$

If $\lim_{x\to0+}f(x)$ and $\lim_{x\to0+}f'(x)$ exist, then f,f' have continuous extensions onto $[0,\infty)$ and the Bregman f-divergence is well-defined and finite for any pair of positive semidefinite operators, too. For a convex function f on $(0,\infty)$ and for given $\lambda\in(0,1)$, the Jensen $\lambda-f$-divergence on $\mathscr{B}^{++}(\mathscr{H})$ is defined by

$$J_{f,\lambda}(X,Y)=\mathrm{Tr}\left(\lambda f(X)+(1-\lambda)f(Y)-f(\lambda X+(1-\lambda)Y)\right).$$

If $\lim_{x\to0+}f(x)$ exists, then the Jensen $\lambda-f$-divergence is also well-defined and finite for any pair of positive semidefinite operators.

Our results about the preservers of Bregman and Jensen divergences read as follows.

Theorem 17. *Let f be a differentiable convex function on $(0,\infty)$ such that f' is bounded from below and unbounded from above. Let $\phi : \mathscr{B}^{++}(\mathscr{H}) \to \mathscr{B}^{++}(\mathscr{H})$ be a bijective map which satisfies*

$$H_f\left(\phi(A),\phi(B)\right) = H_f(A,B), \quad A,B \in \mathscr{B}^{++}(\mathscr{H}).$$

Then there exists a unitary or antiunitary operator $U : \mathscr{H} \to \mathscr{H}$ such that ϕ is of the form

$$\phi(A) = UAU^*, \quad A \in \mathscr{B}^{++}(\mathscr{H}).$$

Theorem 18. *Let f be a differentiable strictly convex function on $(0,\infty)$, assume $\lim_{x\to 0+} f(x)$ exists and finite and f' is unbounded from above. Pick $\lambda \in (0,1)$. If $\phi : \mathscr{B}^{++}(\mathscr{H}) \to \mathscr{B}^{++}(\mathscr{H})$ is a surjective map which satisfies*

$$J_{f,\lambda}\left(\phi(A),\phi(B)\right) = J_{f,\lambda}(A,B), \quad A,B \in \mathscr{B}^{++}(\mathscr{H}),$$

then there exists a unitary or antiunitary operator $U : \mathscr{H} \to \mathscr{H}$ such that ϕ is of the form

$$\phi(A) = UAU^*, \quad A \in \mathscr{B}^{++}(\mathscr{H}).$$

2.5. **Jordan triple endomorphisms ([2]).** Now we turn to another preserver problem on the cone of positive operators. Namely, we describe the structure of the Jordan triple endomorphisms on the cone of positive definite operators acting on a two-dimensional Hilbert space. These endomorphisms are maps which are morphisms with respect to the operation of the Jordan triple product $(A,B) \mapsto ABA$ which is a well-known operation in ring theory. Our main reason for investigating these maps comes from the fact that they naturally appear in the study of surjective isometries and surjective maps preserving generalized distance measures between positive definite cones. For details see [30, 29, 32].

The main theorem reads as follows.

Theorem 19. *Let $\mathscr{H}$ be a two-dimensional Hilbert space. Let $\phi : \mathscr{B}^{++}(\mathscr{H}) \to \mathscr{B}^{++}(\mathscr{H})$ be a continuous Jordan-triple endomorphism. Then we have the following possibilities:*

(b1) *there is a unitary operator $U \in \mathscr{B}(\mathscr{H})$ and a real number c such that*

$$\phi(A) = (\mathrm{Det}A)^c UAU^*, \quad A \in \mathscr{B}^{++}(\mathscr{H});$$

(b2) *there is a unitary operator $V \in \mathscr{B}(\mathscr{H})$ and a real number d such that*

$$\phi(A) = (\mathrm{Det}A)^d VA^{-1}V^*, \quad A \in \mathscr{B}^{++}(\mathscr{H});$$

(b3) *there is a unitary operator $W \in \mathscr{B}(\mathscr{H})$ and real numbers c_1, c_2 such that*

$$\phi(A) = W\mathrm{Diag}[(\mathrm{Det}A)^{c_1}, (\mathrm{Det}A)^{c_2}]W^*, \quad A \in \mathscr{B}^{++}(\mathscr{H}).$$

The following structural result concerning the continuous Jordan triple automorphisms of $\mathscr{B}^{++}(\mathscr{H})$ follows from the proof of Theorem 19.

Theorem 20. *Assume that* $\dim(\mathscr{H}) = 2$. *If* $\phi : \mathscr{B}^{++}(\mathscr{H}) \to \mathscr{B}^{++}(\mathscr{H})$ *is a continuous Jordan triple automorphism, then* ϕ *is of one of the following two forms:*

(c1) *there is a real number* $c \neq -1/2$ *and* $U \in \mathbf{SU}(2)$ *such that*

$$\phi(A) = (\mathrm{Det} A)^c U A U^*, \quad A \in \mathscr{B}^{++}(\mathscr{H});$$

(c2) *there is a real number* $d \neq 1/2$ *and* $V \in \mathbf{SU}(2)$ *such that*

$$\phi(A) = (\mathrm{Det} A)^d V A^{-1} V^*, \quad A \in \mathscr{B}^{++}(\mathscr{H}).$$

The result above has the following immediate consequence. In the case where $\dim(\mathscr{H}) \geq 3$, in [32, Theorem 1] a general result was obtained describing the possible structure of surjective maps on $\mathscr{B}^{++}(\mathscr{H})$ which preserve a generalized distance measure of a certain quite general kind. It is easy to see that, following the proof of [32, Theorem 1] and applying Theorem 20, the result in [32] remains valid also in the case where $\dim(\mathscr{H}) = 2$.

Effects play an important role in certain parts of quantum mechanics, for instance, in the quantum theory of measurement [14]. Mathematically, effects are represented by positive semi-definite Hilbert space operators which are bounded (in the natural order $\leq$ among self-adjoint operators) by the identity. The set of all Hilbert space effects are called the Hilbert space effect algebra (although it is clearly not an algebra in the classical algebraic sense). In [19] Gudder and Nagy introduced the operation $\circ$ called sequential product on effects which has an important physical a meaning and which is closely related the Jordan triple product. Namely, they defined

$$A \circ B = A^{1/2} B A^{1/2}$$

for arbitrary Hilbert space effects A, B. The corresponding endomorphisms, i.e., maps ϕ on Hilbert space effects which satisfy

$$\phi(A \circ B) = \phi(A) \circ \phi(B)$$

for all pairs A, B of effects are called sequential endomorphisms.

Now, we present an application of Theorem 19 for the description of so-called sequential endomorphisms of effect algebras.

Theorem 21. *Assume that* $\dim(\mathscr{H}) = 2$ *and* $\phi : \mathbb{E}_2 \to \mathbb{E}_2$ *is a continuous sequential endomorphism. Then we have the following four possibilities:*

(d1) *there exists a unitary* $U \in \mathscr{B}(\mathscr{H})$ *and a non-negative real number* c *such that*

$$\phi(A) = (\mathrm{Det} A)^c U A U^*, \quad A \in \mathbb{E}_2;$$

(d2) *there exists a unitary* $V \in \mathscr{B}(\mathscr{H})$ *such that*

$$\phi(A) = V(\operatorname{adj} A)V^*, \quad A \in \mathbb{E}_2;$$

(d3) *there exists a unitary* $V \in \mathscr{B}(\mathscr{H})$ *and a real number* $d > 1$ *such that*

$$\phi(A) = \begin{cases} (\operatorname{Det} A)^d VA^{-1}V^*, & \textit{if } A \in \mathbb{E}_2 \textit{ is invertible;} \\ 0, & \textit{otherwise;} \end{cases}$$

(d4) *there exists a unitary* $W \in \mathscr{B}(\mathscr{H})$ *and non-negative real numbers* c_1, c_2 *such that*

$$\phi(A) = W\operatorname{Diag}[(\operatorname{Det} A)^{c_1}, (\operatorname{Det} A)^{c_2}]W^*, \quad A \in \mathbb{E}_2.$$

Here, we mean $0^0 = 1$.

2.6. **Endomorphisms of the Einstein gyrogroup (**[3]**).** Velocity addition was defined by Einstein in his famous paper of 1905 which founded the special theory of relativity. In fact, the whole theory is essentially based on Einstein velocity addition law, see [17]. The algebraic structure corresponding to this operation is a particular example of so-called gyrogroups the general theory of which has been developed by Ungar [37].

The Einstein gyrogroup of dimension three is the pair $(\mathbf{B}, \oplus)$, where $\mathbf{B} = \{\mathbf{u} \in \mathbb{R}^3 : \|\mathbf{u}\| < 1\}$ and $\oplus$ is the binary operation on $\mathbf{B}$ given by

$$\oplus : \mathbf{B} \times \mathbf{B} \to \mathbf{B};\ (\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} \oplus \mathbf{v} := \frac{1}{1 + \langle \mathbf{u}, \mathbf{v} \rangle} \left(\mathbf{u} + \frac{1}{\gamma_{\mathbf{u}}}\mathbf{v} + \frac{\gamma_{\mathbf{u}}}{1 + \gamma_{\mathbf{u}}} \langle \mathbf{u}, \mathbf{v} \rangle\, \mathbf{u} \right), \tag{18}$$

where $\gamma_{\mathbf{u}} = \left(1 - \|\mathbf{u}\|^2\right)^{-\frac{1}{2}}$ is the so-called Lorentz factor. The operation $\oplus$ is called Einstein velocity addition or relativistic sum (cf. [7, 23]).

The main result is obtained as an application of the result on the Jordan triple endomorphisms of positive definite operators acting on two-dimensional Hilbert spaces. The other ingredient of our argument is the result [23, Theorem 3.4] of Kim.

The main statement reads as follows.

Theorem 22. *Let* $\beta : \mathbf{B} \to \mathbf{B}$ *be a continuous map. We have* β *is an algebraic endomorphism with respect to the operation* $\oplus$*, i.e.,* β *satisfies*

$$\beta(\mathbf{u} \oplus \mathbf{v}) = \beta(\mathbf{u}) \oplus \beta(\mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in \mathbf{B}$$

if and only if

(i) *either there is an orthogonal matrix* $O \in \mathbf{M}_3(\mathbb{R})$ *such that*

$$\beta(\mathbf{v}) = O\mathbf{v}, \quad \mathbf{v} \in \mathbf{B};$$

(ii) *or we have*

$$\beta(\mathbf{v}) = 0, \quad \mathbf{v} \in \mathbf{B}.$$

REFERENCES

Related own publications

[1] L. Molnár, J. Pitrik and D. Virosztek, *Maps on positive definite matrices preserving Bregman and Jensen divergences*, Linear Algebra Appl. **495** (2016), 174–189.
[2] L. Molnár and D. Virosztek, *Continuous Jordan triple endomorphisms of* $\mathbb{P}_2$, J. Math. Anal. Appl. **438(2)** (2016), 828-839.
[3] L. Molnár and D. Virosztek, *On algebraic endomorphisms of the Einstein gyrogroup,* J. Math. Phys. **56**, 082302 (2015).
[4] D. Petz and D. Virosztek, *A characterization theorem for matrix variances*, Acta Sci. Math. (Szeged) **80** (2014), 681-687.
[5] D. Petz and D. Virosztek, *Some inequalities for quantum Tsallis entropy related to the strong subadditivity*, Math. Inequal. Appl. **18(2)**(2015), 555-568.
[6] J. Pitrik and D. Virosztek, *On the joint convexity of the Bregman divergence of matrices,* Lett. Math. Phys. **105** (2015), 675-692.

Related publications by other authors

[7] T. Abe, *Gyrometric preserving maps on Einstein gyrogroups, Möbius gyrogroups and proper velocity gyrogroups*, Nonlinear Functional Analysis and Applications **19** (2014), 1-17.
[8] J. Aczél and Z. Daróczy, *On Measures of Information and Their Characterizations*, Academic Press, San Diego, 1975.
[9] K. M. R. Audenaert, *Subadditivity of q-entropies for $q>1$*, J. Math. Phys. **48**(2007), 083507.
[10] A. Banerjee et al., *Clustering with Bregman divergences*, J. Mach. Learn. Res. **6**(2005), 1705-1749.
[11] H. Bauschke and J. Borwein, *Joint and separate convexity of the Bregman distance*, Inherently Parallel Algorithms in Feasibility and Optimization and their Applications (Haifa 2000), D. Butnariu, Y. Censor, S. Reich (editors), Elsevier, pp. 23-36, 2001.
[12] R. Bhatia, *Matrix Analysis,* Springer-Verlag, New York, 1996.
[13] L. M. Bregman, *The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming,* USSR Computational Mathematics and Mathematical Physics **7(3)**(1967), 200-217.
[14] P. Busch, P.J. Lahti and P. Mittelstaedt, *The Quantum Theory of Measurement,* Springer-Verlag, 1991.
[15] E. Carlen, *Trace inequalities and quantum entropy: an introductory course,* Contemp. Math. **529** (2010), 73-140.
[16] Z. Daróczi, *General information functions*, Information and Control **16**(1970), 36 - 51.
[17] A. Einstein, *Einstein's Miraculous Years: Five Papers That Changed the Face of Physics,* Princeton University, Princeton, NJ, 1998.
[18] S. Furuichi, *Information theoretical properties of Tsallis entropies*, J. Math. Phys. **47**, 023302 (2006)
[19] S. Gudder and G. Nagy, *Sequential quantum measurements,* J. Math. Phys. **42** (2001), 5212–5222.
[20] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications*, Hindustan Book Agency and Springer Verlag, 2014.

[21] F. Itakura and S. Saito, *Analysis synthesis telephony based on the maximum likelihood method*, in 6th Int. Congr. Acoustics, Tokyo, Japan., pp. C-17-C-20 (1968)

[22] R. V. Kadison and J. R. Ringrose, *Fundamentals of the Theory of Operator Algebras*, Volumes I and II, Academic Press, Orlando, 1983 and 1986.

[23] S. Kim, *Distances of qubit density matrices on Bloch sphere*, J. Math. Phys. **52**, 102303 (2011).

[24] A. N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer Verlag, Berlin, 1933; English translation: *Foundations of the Theory of Probability*, Chelsea Publishing Co., New York, 1956.

[25] S. Kullback and R.A: Leibler, *On information and sufficiency*, Ann. Math. Statist. **22(1)**(1951), 79 - 86.

[26] Z. Léka and D. Petz, *Some decompositions of matrix variances*, Probability and Mathematical Statistics, **33**(2013), 191-199.

[27] E. Lieb and M. B. Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy*, J. Math. Phys. **14**(1973), 1938-1941.

[28] P.C. Mahalonobis, *On the generalized distance in statistics*, Proceedings of National Institute of Science of India, **12**(1936), 49 - 55.

[29] L. Molnár, *General Mazur-Ulam type theorems and some applications,* in Operator Semigroups Meet Complex Analysis, Harmonic Analysis and Mathematical Physics, W. Arendt, R. Chill, Y. Tomilov (Eds.), Operator Theory: Advances and Applications, Vol. 250, pp. 311-342, Birkhäuser, 2015.

[30] L. Molnár, *Jordan triple endomorphisms and isometries of spaces of positive definite matrices,* Linear Multilinear Alg. **63** (2015), 12–33.

[31] L. Molnár, *Selected Preserver Problems on Algebraic Structures of Linear Operators and on Function Spaces,* Lecture Notes in Mathematics, Vol. 1895, p. 236, Springer, 2007.

[32] L. Molnár and P. Szokol, *Transformations on positive definite matrices preserving generalized distance measures,* Linear Algebra Appl. **466** (2015), 141–159.

[33] M. Nielsen and D. Petz, *A simple proof of the strong subadditivity inequality*, Quantum Information & Computation, **6**(2005), 507 - 513.

[34] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer-Verlag, Heidelberg, 1993. Second edition 2004.

[35] D. Petz and G. Tóth, *Matrix variances with projections*, Acta Sci. Math. (Szeged), **78**(2012), 683–688.

[36] M. Rédei and S.J. Summers, *Quantum probability theory,* Studies in the History and Philosophy of Modern Physics **38** (2007), 390-417.

[37] A.A. Ungar, *Analytic Hyperbolic Geometry and Albert Einstein's Special Theory of Relativity,* World Scientific, Singapore, 2008.