JORDAN TRIPLE ENDOMORPHISMS AND ISOMETRIES OF SPACES OF POSITIVE DEFINITE MATRICES

LAJOS MOLNÁR

ABSTRACT. In this paper we determine the structure of certain algebraic morphisms and isometries of the space $\mathbb{P}_n$ of all $n \times n$ complex positive definite matrices. In the case $n \geq 3$ we describe all continuous Jordan triple endomorphisms of $\mathbb{P}_n$ which are continuous maps $\phi : \mathbb{P}_n \to \mathbb{P}_n$ satisfying

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

It has recently been discovered that surjective isometries of certain substructures of groups equipped with metrics which are in a way compatible with the group operations have algebraic properties that relate them rather closely to Jordan triple morphisms. This makes us possible to use our structural results to describe all surjective isometries of $\mathbb{P}_n$ that correspond to any member of a large class of metrics generalizing the geodesic distance in the natural Riemannian structure on $\mathbb{P}_n$. Finally, we determine the isometry group of $\mathbb{P}_n$ relative to a very recently introduced metric that originates from the divergence called Stein's loss.

1. INTRODUCTION AND STATEMENT OF THE MAIN RESULTS

The study of analytical and geometrical properties of spaces of positive definite matrices plays an important role in several areas of pure and applied mathematics due to the wide spreading applications. One can get an adequate picture of investigations in that direction and find a lot of information relating to applications in the monograph [1] by R. Bhatia.

In the present paper we consider the set $\mathbb{P}_n$ of all $n \times n$ complex positive definite matrices from certain algebraic and metrical points of view. As for the former one, $\mathbb{P}_n$ becomes an algebraic structure under the Jordan triple product $(A, B) \mapsto ABA$. We recall that this product is rather extensively investigated in pure ring theory and also in the theory of operator algebras and its applications. In the case $n \geq 3$, the structure of all continuous Jordan triple automorphisms of $\mathbb{P}_n$ has been determined in our paper [20] (see Theorem 1 there). Here we substantially strengthen that former result and describe all continuous Jordan triple endomorphisms of $\mathbb{P}_n$ (which are continuous maps that simply respect the Jordan triple product but not necessarily bijective).

There is another algebraic operation on $\mathbb{P}_n$ which is closely related to the Jordan triple product. This is what we have called in our recent paper [13] inverted Jordan

triple product and defined as $(A, B) \mapsto AB^{-1}A$. The reason to introduce that operation is the following. In [13] we have made attempts to generalize Mazur-Ulam theorem in a non-commutative context. That famous result states that the surjective isometries (i.e., surjective distance preserving maps) between normed real linear spaces are automatically affine. Our aim has been to find extensions of this theorem concerning normed spaces to the setting of fairly general non-commutative metric groups. In the paper [13] we have presented some results that show that under certain conditions surjective isometries between metric groups or between certain substructures of groups equipped with metrics that are in some sense compatible with the algebraic structure necessarily posses an algebraic property: they locally preserve the inverted Jordan triple product. We have then utilized this fact and determined the isometries of different non-linear structures. For example, in [14] we have described the surjective isometries of the unitary group over a complex Hilbert space equipped with the operator norm. This result has been generalized for the context of C^*-algebras in [15]. In [23] we have considered the group of unitary matrices and determined their surjective isometries under general unitarily invariant norms and also under a class of recently defined metrics that originate from certain considerations in quantum information processing.

In the present paper we continue that line of investigations. Namely, making use of our results on Jordan triple endomorphisms of $\mathbb{P}_n$ we determine the isometry groups of $\mathbb{P}_n$ relative to a large class of metrics. A particular member of that class is the geodesic distance corresponding to the most natural Riemannian structure on $\mathbb{P}_n$. That metric is studied and applied extensively, see, for example, Chapter 6 in [1]. Another particular metric that we are considering is a brand new distance measure that comes from the symmetric Stein divergence. In fact, in the recent paper [26] Sra has proved that the square root of that sort of divergence is a true metric and presented a number of its nice and interesting properties. Below we determine all surjective isometries of $\mathbb{P}_n$ relative to that metric.

Before presenting our results, we fix the notation. In what follows we denote by $\mathbb{M}_n$ the space of all $n \times n$ complex matrices. Whenever we mention metrical or topological properties or notions concerning matrices without specifying the metric, we always have in mind the usual operator norm (or, in another word, spectral norm) $\|.\|$ on $\mathbb{M}_n$ ($\|A\|$ equals the largest singular value of A). The real linear space of all self-adjoint (or, in another word, Hermitian) elements of $\mathbb{M}_n$ is denoted by $\mathbb{H}_n$ and $\mathbb{U}_n$ stands for the group of all unitaries in $\mathbb{M}_n$.

In the first main result of the paper that follows we give the complete description of the structure of all continuous Jordan triple endomorphisms of $\mathbb{P}_n$ ($n \geq 3$) and hence we generalize the finite dimensional part of Theorem 1 in [20] significantly.

Theorem 1. *Assume $n \geq 3$. Let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be a continuous map which is a Jordan triple endomorphism, i.e., ϕ is a continuous map which satisfies*

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

Then there exist a unitary matrix $U \in \mathbb{U}_n$, a real number c, a set $\{P_1, \ldots, P_n\}$ of mutually orthogonal rank-one projections in $\mathbb{M}_n$, and a set $\{c_1, \ldots, c_n\}$ of real numbers such that ϕ is of one of the following forms:

  (e1) $\phi(A) = (\det A)^c U A U^*, \quad A \in \mathbb{P}_n;$
  (e2) $\phi(A) = (\det A)^c U A^{-1} U^*, \quad A \in \mathbb{P}_n;$
  (e3) $\phi(A) = (\det A)^c U A^{tr} U^*, \quad A \in \mathbb{P}_n;$

(e4) $\phi(A) = (\det A)^c U A^{tr^{-1}} U^*$, $A \in \mathbb{P}_n$;

(e5) $\phi(A) = \sum_{j=1}^n (\det A)^{c_j} P_j$, $A \in \mathbb{P}_n$.

Above and throughout the paper projection means self-adjoint idempotent. As an immediate corollary of our first theorem we have the following result which was originally obtained in [20, Theorem 1].

Corollary 2. *Assume $n \geq 3$. Let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be a continuous Jordan triple automorphism, i.e., a continuous bijective map which satisfies*

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

Then there exist a unitary $U \in \mathbb{U}_n$ and a number $c \neq -1/n$ such that ϕ is of one of the following forms:

(a1) $\phi(A) = (\det A)^c U A U^*$, $A \in \mathbb{P}_n$;

(a2) $\phi(A) = (\det A)^c U A^{-1} U^*$, $A \in \mathbb{P}_n$;

(a3) $\phi(A) = (\det A)^c U A^{tr} U^*$, $A \in \mathbb{P}_n$;

(a4) $\phi(A) = (\det A)^c U A^{tr^{-1}} U^*$, $A \in \mathbb{P}_n$.

The structure of the morphisms above and results what we are about to obtain on the way leading to that will be utilized to describe the surjective isometries of $\mathbb{P}_n$ with respect to members of a large class of metrics.

Particular elements of that class have strong differential geometrical origins and connections. The set $\mathbb{P}_n$ of positive definite matrices is an open subset of the space $\mathbb{H}_n$, hence it is a differentiable manifold which can naturally be equipped with a Riemannian structure in the following way. For any $A \in \mathbb{P}_n$ the tangent space of $\mathbb{P}_n$ at A can be identified with $\mathbb{H}_n$ on which we define an inner product by

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

The corresponding norm is given by

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

Here $\|.\|_{HS}$ stands for the Hilbert-Schmidt norm (or, in another word, Frobenius norm) which is defined by $\|T\|_{HS}^2 = \mathrm{Tr}(T^*T)$, $T \in \mathbb{M}_n$. In that way we obtain a Riemannian space which has long been studied in the literature for many reasons. For example, it provides probably the most important example of a manifold of non-positive curvature. The geometry of this manifold is intimately connected with some matrix inequalities and hence it has a wide range of applications in matrix analysis. We also point out its connections to problems relating to matrix means, a really vivid topic in recent days. As to our present results, it is important to recall that the geodesic distance $\delta_R(A, B)$ between the points $A, B \in \mathbb{P}_n$ in this Riemannian space is given by

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

For details we refer to, e.g., [2] or Chapter 6 in [1].

Connections between means, geodesics, and inequalities were explored in several interesting papers by G. Corach and his coauthors. In fact, in the 1990's they defined and studied a Finsler structure on the manifold of all invertible positive elements of a general C^*-algebra (see, among others, [6], [7], [8]). In the present setting of matrices this means the following. At any point A in $\mathbb{P}_n$, on the tangent space $\mathbb{H}_n$ they defined the Finsler metric (norm) by

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

where $\|.\|$ is again the usual operator norm (spectral norm). Among revealing many interesting and important properties of the so-obtained Finsler space they obtained that the geodesic distance between A and B is given by $\|\log A^{-1/2}BA^{-1/2}\|$. Let us make a short remark here. It is rather surprising that this last quantity appears in a different context, too. Namely, $\|\log A^{-1/2}BA^{-1/2}\|$ equals the distance between A and B relative to the so-called Thompson metric that was defined as a useful modification of the Hilbert projective metric in a setting much more general than that of C^*-algebras. We refer to our paper [22] where we have determined the surjective Thompson isometries of the space of all invertible positive operators on a complex Hilbert space which result has recently been generalized for the setting of general C^*-algebras in [15].

Above we have mentioned norms on $\mathbb{H}_n$ as a tangent space that correspond either to the Hilbert-Schmidt norm or to the operator norm. In the paper [10] Fujii has presented a common generalization of the above two approaches for the setting of finite dimensional C^*-algebras. In the case of the algebra $\mathbb{M}_n$ this means the following (cf. Section 6.4 in [1]). Consider an arbitrary unitarily invariant norm N on $\mathbb{M}_n$ and define

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

for each point $A \in \mathbb{P}_n$ and every vector X from $\mathbb{H}_n$. By Theorem 1 in [10], this formula determines a Finsler metric on $\mathbb{P}_n$ and Theorem 5 in the same paper tells that the shortest path length $d_N(A, B)$ between $A, B \in \mathbb{P}_n$ is given by

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

In our next theorem we determine the surjective isometries of $\mathbb{P}_n$ relative to any of the metrics d_N.

Theorem 3. *Suppose $n \geq 2$. Let N be a unitarily invariant norm on $\mathbb{M}_n$ and $\phi : \mathbb{P}_n \to \mathbb{P}_n$ a surjective isometry relative to the metric d_N. Assume $n \geq 3$ and N is not a scalar multiple of the Hilbert-Schmidt norm. If $n \neq 4$, then there exists an invertible matrix $T \in \mathbb{M}_n$ such that ϕ is of one of the following forms:*

   (t1) $\phi(A) = TAT^*$, $A \in \mathbb{P}_n$;
   (t2) $\phi(A) = TA^{-1}T^*$, $A \in \mathbb{P}_n$;
   (t3) $\phi(A) = TA^{tr}T^*$, $A \in \mathbb{P}_n$;
   (t4) $\phi(A) = TA^{tr^{-1}}T^*$, $A \in \mathbb{P}_n$.

If $n = 4$, then beside (t1)-(t4) the following additional possibilities can occur:

   (d1) $\phi(A) = (\det A)^{-2/n}TAT^*$, $A \in \mathbb{P}_n$;
   (d2) $\phi(A) = (\det A)^{2/n}TA^{-1}T^*$, $A \in \mathbb{P}_n$;
   (d3) $\phi(A) = (\det A)^{-2/n}TA^{tr}T^*$, $A \in \mathbb{P}_n$;
   (d4) $\phi(A) = (\det A)^{2/n}TA^{tr^{-1}}T^*$, $A \in \mathbb{P}_n$.

In the case where $n \geq 3$ and N is a scalar multiple of the Hilbert-Schmidt norm, ϕ is of one of the forms (t1)-(t4), (d1)-(d4). Finally, if $n = 2$, then ϕ can necessarily be written in one of the forms (t1)-(t4).

We note that the remarkable fact that in the case $n = 4$ we have some special additional possibilities follows from a beautiful general result describing the linear isometries of symmetric gauge functions which was obtained by Đoković, Li and Rodman in [9].

We have already mentioned that positive definite matrices play an important role in several areas of pure and applied mathematics. In many of the applications the metrical structure of $\mathbb{P}_n$ is of particular interest. For example, in optimization problems relating to $\mathbb{P}_n$ measuring distances between elements is a key task and a very nontrivial one when the distance function must respect a non-Euclidean geometry on $\mathbb{P}_n$. The already mentioned distance measure $\delta_R(A, B) = \|\log A^{-1/2} B A^{-1/2}\|_{HS}$ is a particularly important example which is computationally very demanding and also complicated to use. In order to allay those difficulties, in the paper [26] (also see [25]), Sra has introduced a new metric on $\mathbb{P}_n$ which not only respects non-Euclidean geometry but offers faster computation than the previous one and it is also much less complicated to use. In the mentioned papers several results have been presented that shed light on the advantages of the new metric and relate it to δ_R in order to justify it is a good proxy for δ_R. Moreover, some experimental results have also been given to demonstrate the usefulness of the new metric which is defined in the following way.

For any pair $A, B \in \mathbb{P}_n$ of positive definite matrices their symmetric Stein divergence is defined by

$$S(A, B) = \log \det \left(\frac{A + B}{2} \right) - \frac{1}{2} \log \det(AB).$$

Actually, it is just the Jensen-Shannon symmetrization of the divergence called Stein's loss (see the first two sections in [26]). In fact, driven by the computational concerns related to the use of δ_R, the measure S has originally been introduced in [4]. The authors of that work claimed that

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

is not a metric while the authors in [3] conjectured that it is. The problem has got a solution in [26, Theorem 5], where it has been proved that δ_S is a true metric on $\mathbb{P}_n$ (also see [25]). Beside presenting several interesting results concerning the properties of the new metric, in the same paper [26] Sra has initiated the study of the metric space $(\mathbb{P}_n, \delta_S)$ from further aspects. We aim to contribute to his program by the following theorem in which we determine the precise structure of the isometry group of $(\mathbb{P}_n, \delta_S)$. This is the last main result of the present paper.

Theorem 4. *Assume $n \geq 2$. Let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be a surjective isometry relative to the metric δ_S. Then there is an invertible matrix $T \in \mathbb{M}_n$ such that ϕ is of one of the following forms:*
- (s1) $\phi(A) = TAT^*, \quad A \in \mathbb{P}_n$;
- (s2) $\phi(A) = TA^{-1}T^*, \quad A \in \mathbb{P}_n$;
- (s3) $\phi(A) = TA^{tr}T^*, \quad A \in \mathbb{P}_n$;
- (s4) $\phi(A) = TA^{tr^{-1}}T^*, \quad A \in \mathbb{P}_n$.

2. Proofs

In this section we present the proofs of our main results.

We begin with collecting some elementary algebraic properties of Jordan triple endomorphisms of $\mathbb{P}_n$. So let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be such a transformation (a Jordan triple map in short), i.e., assume that

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

We have $\phi(I) = \phi(I)^3$ which implies $\phi(I)^2 = I$ and we then obtain $\phi(I) = I$. It is now easy to see that $\phi(A^k) = \phi(A)^k$ holds for all $A \in \mathbb{P}_n$ and positive integer k. We show that ϕ respects the inverse operation, too. For an $A \in \mathbb{P}_n$, let $B \in \mathbb{P}_n$ be such that $B^2 = A$. We compute

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

which implies that

$$\phi(A^{-1}) = \phi(B)^{-2} = \phi(B^2)^{-1} = \phi(A)^{-1}.$$

Therefore, it follows that $\phi(A^k) = \phi(A)^k$ holds for every integer k and $A \in \mathbb{P}_n$. If k is a positive integer and $A \in \mathbb{P}_n$, then we have $\phi(A^{1/k})^k = \phi(A)$ implying that $\phi(A^{1/k}) = \phi(A)^{1/k}$. Consequently, we obtain that $\phi(A^r) = \phi(A)^r$ holds for all $A \in \mathbb{P}_n$ and rational number r.

In what follows we first prove Theorem 1 concerning the structure of continuous Jordan triple endomorphisms of $\mathbb{P}_n$. Our strategy is similar in spirit to the proof of Theorem 1 in [23] that describes the continuous Jordan triple endomorphisms of the unitary group $\mathbb{U}_n$. Namely, we first prove that the continuity of a Jordan triple map $\phi : \mathbb{P}_n \to \mathbb{P}_n$ implies the (local) Lipschitz property and using that we can show that ϕ gives rise to a linear transformation $f : \mathbb{H}_n \to \mathbb{H}_n$ that preserves commutativity. We can then employ a structural result concerning that kind of transformations and deduce that ϕ is necessarily of one of the forms that appear in Theorem 1. However, although the strategy is similar, the work-out differs at a number of points as one can see below.

We begin with the following lemma.

Lemma 5. *Assume $n \geq 2$. Let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be a continuous Jordan triple endomorphism. Then ϕ is a Lipschitz function in a neighborhood of the identity (e.g., in the closed ball with center I and radius $1/5$).*

Proof. We mention that a result on the Lipschitz property of group endomorphisms of linear groups was proved in [18] (p. 177, Satz 1). That argument could closely be followed in the present situation too but here we present a shorter and hopefully somewhat more transparent argument to verify our lemma.

We begin with the following important observation. For any $A \in \mathbb{P}_n$ which is close enough to I we have

$$(1) \qquad\qquad \frac{1}{2}\|A - I\| \leq \|\log A\| \leq 2\|A - I\|.$$

Indeed, this follows easily from the inequalities

$$\|e^H - I\| \leq e^{\|H\|} - 1, \quad H \in \mathbb{H}_n,$$

and

$$\|\log A\| \leq -\log(1 - \|A - I\|), \quad A \in \mathbb{P}_n, \|A - I\| < 1,$$

and from elementary properties of the exponential and logarithm functions of a real variable.

For temporary use, let $\mathcal{G}_r$ denote the closed ball in $\mathbb{P}_n$ with center I and radius $0 < r < 1$. We assert that there exists a positive real number L such that $\|\phi(A) - I\| \leq L\|A - I\|$ holds for all $A \in \mathcal{G}_{1/2}$. Assume on the contrary that we have a sequence (A_k) of elements of $\mathcal{G}_{1/2}$ and a strictly increasing sequence (c_k) of positive integers such that

$$(2) \qquad\qquad \|\phi(A_k) - I\| > c_k\|A_k - I\|$$

holds for every $k \in \mathbb{N}$. By the compactness of $\mathcal{G}_{1/2}$ it follows that (A_k) has a convergent subsequence. Without serious loss of generality we may and do assume that already the original sequence is convergent. If its limit were different from I, by (2) we would obtain that the numerical sequence $\|\phi(A_k) - I\|$ tends to infinity which contradicts to the boundedness of ϕ on the compact set $\mathcal{G}_{1/2}$. Therefore, $A_k \to I$ and hence $\phi(A_k) \to I$ as $k \to \infty$. Observe that we can also assume that

$$\|\phi(A_k) - I\| = \epsilon_k, \quad \epsilon_k < 1$$

holds for all $k \in \mathbb{N}$. Choose positive integers l_k such that

$$1/(l_k + 1) \le \epsilon_k < 1/l_k.$$

By (1), for large enough k we have $\|\log A_k\| \le 2\|A_k - I\|$ and hence obtain

$$\|\log A_k^{l_k}\| = l_k \|\log A_k\| \le 2l_k \|A_k - I\| < 2l_k(\epsilon_k/c_k) < 2/c_k \to 0$$

as $k \to \infty$. It follows that $A_k^{l_k} \to I$ and hence $A_k^{l_k+1} \to I$. Therefore, we infer $\phi(A_k^{l_k+1}) \to \phi(I) = I$. However, using (1) again, for large enough k we also have

$$\frac{1}{2} \le \frac{\epsilon_k(l_k+1)}{2} = \frac{l_k+1}{2}\|\phi(A_k) - I\| \le$$
$$(l_k+1)\|\log \phi(A_k)\| = \|\log \phi(A_k^{l_k+1})\| \to 0$$

which is a contradiction. Consequently, there does exist a positive real number L such that $\|\phi(A) - I\| \le L\|A - I\|$ holds for all $A \in \mathcal{G}_{1/2}$.

To complete the proof pick arbitrary $C, D \in \mathbb{P}_n$ which belong to the closed ball $\mathcal{G}_{1/5}$. Choose $B \in \mathbb{P}_n$ with $B^2 = C$ and let $A = B^{-1}DB^{-1}$. It is easy to check that $A \in \mathbb{P}_n$ has distance at most $1/2$ from I. Indeed, this follows from the following inequalities

$$\frac{2}{3}I = \frac{4}{5}\frac{5}{6}I \le \frac{4}{5}B^{-2} \le B^{-1}DB^{-1} \le \frac{6}{5}B^{-2} \le \frac{6}{5}\frac{5}{4}I = \frac{3}{2}I.$$

Assuming $C \ne D$ we compute

$$\frac{\|\phi(D) - \phi(C)\|}{\|D - C\|} = \frac{\|\phi(B)\phi(A)\phi(B) - \phi(B)^2\|}{\|BAB - B^2\|} \le$$
$$\frac{\|B^{-1}\|^2\|\phi(B)\|^2\|\phi(A) - I\|}{\|A - I\|} = \|C^{-1}\|\|\phi(C)\|\frac{\|\phi(A) - I\|}{\|A - I\|} \le L\|C^{-1}\|\|\phi(C)\|.$$

Clearly, the function $C \mapsto \|C^{-1}\|\|\phi(C)\|$ is bounded on the closed ball $\mathcal{G}_{1/5}$ and thus we obtain the desired Lipschitz property of ϕ in a neighborhood of I. □

The following lemma which shows that every continuous Jordan triple endomorphism of $\mathbb{P}_n$ is the exponential of a commutativity preserving linear map on $\mathbb{H}_n$ composed by the logarithmic function plays an essential role in the proof of our first main result. Recall that we say a linear transformation f on $\mathbb{H}_n$ preserves commutativity if for any pair $T, S \in \mathbb{H}_n$ of commuting matrices we have that $f(T), f(S)$ commute, too.

Lemma 6. *Assume $n \ge 2$. Let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be a continuous Jordan triple endomorphism. Then there exists a commutativity preserving linear transformation $f : \mathbb{H}_n \to \mathbb{H}_n$ such that*

$$(3) \qquad \phi(A) = e^{f(\log A)}, \quad A \in \mathbb{P}_n.$$

Proof. We define $f : \mathbb{H}_n \to \mathbb{H}_n$ by

$$f(T) = \log \phi(e^T), \quad T \in \mathbb{H}_n.$$

We clearly have (3) and need to show that f is linear and preserves commutativity.

We recall that in the beginning of the section we have shown that $\phi(A^r) = \phi(A)^r$ holds for all $A \in \mathbb{P}_n$ and rational number r. By the continuity of ϕ we obtain that $\phi(A^t) = \phi(A)^t$ is true for every real number t, too. This easily implies that f is homogeneous.

We next prove that $f : \mathbb{H}_n \to \mathbb{H}_n$ is additive. Pick $T, S, H \in \mathbb{H}_n$. We compute

$$\frac{e^{(t/2)T}e^{tS}e^{(t/2)T} - e^{tH}}{t} =$$

$$\frac{(e^{(t/2)T} - I)e^{tS}e^{(t/2)T} + (e^{tS} - I)e^{(t/2)T} + (e^{(t/2)T} - I) - (e^{tH} - I)}{t} \to$$

$$T/2 + S + T/2 - H = T + S - H$$

as $t \to 0$. It follows that

$$\lim_{t \to 0} \frac{e^{(t/2)T}e^{tS}e^{(t/2)T} - e^{tH}}{t} = 0 \Longleftrightarrow H = T + S.$$

If $H = T + S$, then using (3) and the Lipschitz property of ϕ in a neighborhood of I that has been proved in Lemma 5 we have

$$\frac{e^{(t/2)f(T)}e^{tf(S)}e^{(t/2)f(T)} - e^{tf(H)}}{t} =$$

$$\frac{\phi(e^{(t/2)T})\phi(e^{tS})\phi(e^{(t/2)T}) - \phi(e^{tH})}{t} =$$

$$\frac{\phi(e^{(t/2)T}e^{tS}e^{(t/2)T}) - \phi(e^{tH})}{t} \to 0$$

as $t \to 0$. On the other hand, just as above we deduce

$$\frac{e^{(t/2)f(T)}e^{tf(S)}e^{(t/2)f(T)} - e^{tf(H)}}{t} \to f(T) + f(S) - f(H).$$

This gives us that $f(T) + f(S) - f(T + S) = 0$, i.e., f is additive.

To verify the commutativity preserving property of f first observe that we have

$$\phi(\sqrt{A}B\sqrt{A}) = \phi(\sqrt{A})\phi(B)\phi(\sqrt{A}) = \sqrt{\phi(A)}\phi(B)\sqrt{\phi(A)}$$

for every $A, B \in \mathbb{P}_n$. We now recall the following notion and fact. Given positive semidefinite matrices or positive Hilbert space operators D, F, the operation defined by $\sqrt{D}F\sqrt{D}$ is called the sequential product of D and F, see [12]. This operation has the interesting property [12, Corollary 3] that commutativity with respect to the sequential product is equivalent to the commutativity relative to the usual product. If we pick commuting matrices $T, S \in \mathbb{H}_n$, then it follows that

(4)
$$\sqrt{e^{f(T)}}e^{f(S)}\sqrt{e^{f(T)}} = \phi(\sqrt{e^T}e^S\sqrt{e^T}) =$$
$$\phi(\sqrt{e^S}e^T\sqrt{e^S}) = \sqrt{e^{f(S)}}e^{f(T)}\sqrt{e^{f(S)}}.$$

Therefore, we infer from (4) that

$$e^{f(T)}e^{f(S)} = e^{f(S)}e^{f(T)}.$$

Since the logarithms of commuting matrices also commute, we obtain $f(T)f(S) = f(S)f(T)$ verifying that f indeed preserves commutativity. $\qquad\square$

The following lemma describes the structure of scalar valued continuous Jordan triple endomorphisms of $\mathbb{P}_n$ and hence it provides a particular case of Theorem 1. The statement immediately follows from [19, Theorem 2], therefore we omit the proof.

Lemma 7. *Assume $n \geq 3$. Let $\varphi : \mathbb{P}_n \to]0, \infty[$ be a continuous Jordan triple functional, i.e., assume that φ is continuous and satisfies*

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

Then there is a real number c such that

$$\varphi(A) = (\det A)^c, \quad A \in \mathbb{P}_n.$$

After these preparations we are now in a position to prove our first main theorem.

Proof of Theorem 1. By Lemma 6, there exists a commutativity preserving linear transformation $f : \mathbb{H}_n \to \mathbb{H}_n$ such that

$$(5) \qquad \phi(e^T) = e^{f(T)}, \quad T \in \mathbb{H}_n.$$

It is assumed that $n \geq 3$. In that case the structure of commutativity preserving linear maps on $\mathbb{H}_n$ is known and was described in [5]. The result [5, Theorem 2] tells us that we have the following possibilities for f:

(a) either the range of f is commutative;
(b) or there exist a unitary matrix $U \in \mathbb{U}_n$, a linear functional $l : \mathbb{H}_n \to \mathbb{R}$ and a nonzero scalar $d \in \mathbb{R}$ such that f is of the form

$$(6) \qquad f(T) = dUTU^* + l(T)I, \quad T \in \mathbb{H}_n$$

or of the form

$$(7) \qquad f(T) = dUT^{tr}U^* + l(T)I, \quad T \in \mathbb{H}_n.$$

First consider the case (b) where f is of the form (6). We have

$$\phi(e^T) = e^{dUTU^* + l(T)I}, \quad T \in \mathbb{H}_n.$$

Apparently, it follows that

$$U^*\phi(e^T)U = e^{dT + l(T)I}, \quad T \in \mathbb{H}_n.$$

Clearly, $U^*\phi(.)U$ is a continuous Jordan triple endomorphism of $\mathbb{P}_n$. Hence, in the remaining steps of the proof we may and do assume without serious loss of generality that

$$(8) \qquad \phi(e^T) = e^{dT + l(T)I} = e^{l(T)}e^{dT}$$

holds for all $T \in \mathbb{H}_n$. If f is of the form (7), then considering the map $(U^*\phi(.)U)^{tr}$ we can again assume that (8) holds. Moreover, composing the transformation by the inverse operation if necessary, it can further be supposed that the number d is positive.

We observe that $\phi(A)$ is a scalar multiple of A^d for every $A \in \mathbb{P}_n$. By the Jordan triple multiplicativity of ϕ we deduce that $(ABA)^d$ and $A^d B^d A^d$ are linearly dependent for every $A, B \in \mathbb{P}_n$. Letting B converge to any rank-one projection $P \in \mathbb{M}_n$, it follows that $(APA)^d$ and $A^d P A^d$ are linearly dependent. We now consider matrices as linear operators on the n-dimensional complex Hilbert space $\mathcal{H}$. For any pair of vectors $x, y \in \mathcal{H}$ we denote by $x \otimes y$ the operator defined by $(x \otimes y)z = \langle z, y \rangle x$, $z \in \mathcal{H}$. Pick an arbitrary unit vector $x \in \mathcal{H}$. Then $P = x \otimes x$ is a rank-one

projection, $(APA)^d$ is a scalar multiple of $Ax \otimes Ax$ and $A^d P A^d = A^d x \otimes A^d x$. Hence $Ax \otimes Ax$ and $A^d x \otimes A^d x$ are linearly dependent. Consequently, we obtain that for all $A \in \mathbb{P}_n$ and for all $x \in \mathcal{H}$, the vectors Ax and $A^d x$ are linearly dependent. It is easy to verify that for positive d this can happen only if $d = 1$. (Consider, for example, the operator A represented by the $n \times n$ diagonal matrix $\mathrm{diag}(1, \ldots, 1, 2)$ and a vector $x \in \mathcal{H}$ represented by the column matrix $(1, \ldots, 1)^{tr}$.) Next, we can write ϕ as $\phi(A) = \varphi(A)A$, $A \in \mathbb{P}_n$. Clearly, φ is necessarily a continuous Jordan triple functional and hence Lemma 7 applies and implies that φ is a power of the determinant function. Remember now the reductions what we may have applied for ϕ above. We may have composed it with an inner automorphism $U^*(.)U$ and/or with the transpose operation and/or with the inverse operation. Having these in mind we deduce the following. In the case (b) what we are considering, dropping all reductions concerning the original transformation ϕ we obtain that it is of one of the forms (e1)-(e4) that appear in the formulation of Theorem 1.

Let us now examine the case (a), i.e., assume that the range of the linear transformation $f : \mathbb{H}_n \to \mathbb{H}_n$ satisfying

$$\phi(e^T) = e^{f(T)}, \quad T \in \mathbb{H}_n$$

(see (5)) is commutative. It follows that the elements of that range can be diagonalized simultaneously. Therefore, we have a commuting set $\{P_1, \ldots, P_n\}$ of rank-one projections in $\mathbb{M}_n$ and a collection $\varphi_1, \ldots, \varphi_n : \mathbb{P}_n \to]0, \infty[$ of functions such that

$$\phi(A) = \sum_{j=1}^{n} \varphi_j(A) P_j, \quad A \in \mathbb{P}_n.$$

Clearly, every φ_i is a continuous Jordan triple functional, i.e., satisfies

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

Applying Lemma 7 we obtain that there are real numbers $c_1, \ldots, c_n$ such that

$$\phi(A) = \sum_{j=1}^{n} (\det A)^{c_j} P_j, \quad A \in \mathbb{P}_n.$$

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

Proof of Corollary 2. Clearly, from Theorem 1 we have only the first four possibilities (e1)-(e4). One can easily check that the transformations which appear there are bijective precisely when $c \neq -1/n$. $\qquad\square$

We now turn to the proofs of our results on isometries. For the proofs of Theorems 3 and 4 we need to recall the following assertion from the paper [13] which states that under some conditions the surjective isometries between certain substructures of groups equipped with metrics compatible with the group operations locally preserve the inverted Jordan triple product. The proposition below appears as Corollary 3.10. in [13] and is recalled as Proposition 5 in [14].

Proposition 8. *Let G_i be groups, $X_i \subset G_i$ subsets such that $yx^{-1}y \in X_i$ holds for all $x, y \in X_i$, $i = 1, 2$. Suppose that X_2 is 2-torsion free meaning that the unit element e of G_2 belongs to X_2 and for any $x \in X_2$ the equality $x^2 = e$ implies $x = e$. We also assume that X_2 is 2-divisible, i.e., for every $x \in X_2$ there is a $y \in X_2$ such that $y^2 = x$. Let X_i be equipped with metric d_i, $i = 1, 2$.*

Assume $\phi : X_1 \to X_2$ is a surjective isometry. Pick $a, b \in X_1$ with the following properties:

(p1) *for all $x, y \in X_1$ we have*
$$d_1(bx^{-1}b, by^{-1}b) = d_1(x, y);$$

(p2) *there exists a constant $K > 1$ such that we have*
$$d_1(bx^{-1}b, x) \geq K d_1(x, b)$$
for all $x \in L_{a,b} = \{x \in X_1 : d_1(a, x) = d_1(ba^{-1}b, x) = d_1(a, b)\}$;

(p3) *there exists an element $c \in X_2$ with $c\phi(a)^{-1}c = \phi(ba^{-1}b)$ such that the equality $d_2(cx^{-1}c, cy^{-1}c) = d_2(x, y)$ holds for all $x, y \in X_2$.*

Then we have
$$\phi(ba^{-1}b) = \phi(b)\phi(a)^{-1}\phi(b).$$

Beside the above proposition, for the proof of Theorem 3 we need to recall the correspondence between unitarily invariant norms on $\mathbb{M}_n$ and symmetric gauge functions on $\mathbb{R}^n$. A unitarily invariant norm on $\mathbb{M}_n$ is a norm N that satisfies $N(UAV) = N(A)$ for any $A \in \mathbb{M}_n$ and $U, V \in \mathbb{U}_n$. To define the concept of a symmetric gauge function we first need the concept of a generalized permutation matrix. This is a matrix in $\mathbb{M}_n$ with the property that in its each row and each column there appears exactly one nonzero element and it is either 1 or -1. The collection of all generalized permutation matrices in $\mathbb{M}_n$ is denoted by GP_n. A symmetric gauge function on $\mathbb{R}^n$ is a function $\Phi : \mathbb{R}^n \to [0, \infty[$ which is a norm that satisfies $\Phi(Px) = \Phi(x)$ for any vector $x \in \mathbb{R}^n$ and generalized permutation matrix $P \in GP_n$. It is well known that there is a one-to-one correspondence between unitarily invariant norms on $\mathbb{M}_n$ and symmetric gauge functions on $\mathbb{R}^n$. For every unitarily invariant norm N on $\mathbb{M}_n$ there is a symmetric gauge function Φ on $\mathbb{R}^n$ such that the following holds: for any matrix $A \in \mathbb{M}_n$ with singular values $\lambda_1, \ldots, \lambda_n$ we have $N(A) = \Phi(\lambda_1, \ldots, \lambda_n)$. Conversely, every symmetric gauge function gives rise to a unitarily invariant norm in that way.

After these preparations we can now present the proof of Theorem 3.

Proof of Theorem 3. Let N be a unitarily invariant norm on $\mathbb{M}_n$ and let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be a surjective isometry relative to the metric d_N. We are going to apply Proposition 8. In order to do that, we need to check if all conditions listed in there are satisfied.

First, we have $X_1 = X_2 = \mathbb{P}_n$ and hence X_2 is obviously 2-torsion free and 2-divisible.

Next, observe that the equalities

(9) $$d_N(A^{-1}, B^{-1}) = d_N(A, B), \quad d_N(TAT^*, TBT^*) = d_N(A, B)$$

hold for all $A, B \in \mathbb{P}_n$ and invertible matrix $T \in \mathbb{M}_n$. In fact, the first equality is easy to see, while the second one follows from Lemma 3 and Theorem 5 in [10]. Using this second invariance property of the metric d_N we see that considering the transformation $\phi(I)^{-1/2}\phi(.)\phi(I)^{-1/2}$ we get a unital map which is again a surjective isometry of $\mathbb{P}_n$. Therefore, without serious loss of generality we may and do assume that our original isometry ϕ satisfies $\phi(I) = I$.

Pick $A, B \in \mathbb{P}_n$. By the equalities in (9), the condition (p1) is satisfied. As for the property (p3), we find that for any $x, y \in X_2$ there exists $c \in X_2$ such that $cx^{-1}c = y$. In fact, this follows from the 2-divisibility of X_2 (see the discussion after

Definition 3.4 in [13]). For curiosity we admit that in our particular case $X_2 = \mathbb{P}_n$ this c is the geometric mean of x and y, i.e., $c = x^{1/2}(x^{-1/2}yx^{-1/2})^{1/2}x^{1/2}$. The metrical condition in (p3) is a consequence of the already mentioned invariance properties of the metric d_N. It remains to verify (p2). For any $X \in \mathbb{P}_n$ we have

(10)
$$d_N(BX^{-1}B, X) = d_N(X, BX^{-1}B) =$$
$$N(\log((X^{-1/2}BX^{-1/2})^2)) = 2N(\log(X^{-1/2}BX^{-1/2})) = 2d_N(X, B).$$

Therefore, all assumptions in Proposition 8 are fulfilled for any $A, B \in \mathbb{P}_n$. Hence it follows that $\phi : \mathbb{P}_n \to \mathbb{P}_n$ is a bijective unital map which satisfies

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

Plugging $A = I$ we obtain that $\phi(B^{-1}) = \phi(B)^{-1}$, $B \in \mathbb{P}_n$ and then we can immediately conclude that ϕ is a Jordan triple automorphism of $\mathbb{P}_n$.

We assert that ϕ is continuous. It was mentioned in [22] that the Thompson metric (i.e., the metric $d_{\|.\|}$) induces the same topology as the operator norm $\|.\|$ (see [22], p. 3854). Since N is equivalent to $\|.\|$, we deduce that d_N induces the same topology as the Thompson metric. It now follows that the Jordan triple automorphism ϕ of $\mathbb{P}_n$ is continuous relative to the topology of the operator norm.

Assume now that $n \geq 3$ in which case we can apply Corollary 2. We have a unitary matrix $U \in \mathbb{U}_n$ and a scalar $c \neq -1/n$ such that ϕ is of one of the forms (a1)-(a4) that appear in Corollary 2. We need to determine the possible values of the scalar c. Observe that the unitary similarity transformation $U^*(.)U$ and the inverse operation are both isometries of $\mathbb{P}_n$ (these follow again from (9)). It is easy to check that the transpose operation is an isometry, too. Therefore, if ϕ is one of the forms (a1), (a3), then we obtain that $A \mapsto (\det A)^c A$ is an isometry of $\mathbb{P}_n$ while if ϕ is of the form (a2) or (a4), then we deduce that $A \mapsto (\det A)^{-c} A$ is an isometry of $\mathbb{P}_n$. We examine only the former case (the latter one then follows readily). In that case we have

$$N(\log((\det A^{-1/2}BA^{-1/2})^c(A^{-1/2}BA^{-1/2}))) = N(\log(A^{-1/2}BA^{-1/2}))$$

for all $A, B \in \mathbb{P}_n$. Since $A^{-1/2}BA^{-1/2}$ runs through the whole set $\mathbb{P}_n$ and the formula $\log \det C = \operatorname{Tr} \log C$ holds for all $C \in \mathbb{P}_n$, it follows that the above displayed equality is equivalent to

$$N(c\operatorname{Tr}(T)I + T) = N(T), \quad T \in \mathbb{H}_n.$$

Inserting $T = I$ we have $N((cn + 1)I) = N(I)$ and this yields $|cn + 1| = 1$. It follows that either $c = 0$ or $c = -2/n$.

Let us examine if the possibility $c = -2/n$ can really occur, i.e., if it can happen that

$$N(T - (2/n)\operatorname{Tr}(T)I) = N(T), \quad T \in \mathbb{H}_n.$$

Denoting by Φ the symmetric gauge function on $\mathbb{R}^n$ corresponding to N, the above equality can be reformulated by saying that the transformation

(11) $$(\lambda_1, \ldots, \lambda_n)^{tr} \mapsto (\lambda_1 - (2/n)\sum_{i=1}^{n}\lambda_i, \ldots, \lambda_n - (2/n)\sum_{i=1}^{n}\lambda_i)^{tr}$$

on $\mathbb{R}^n$ is an isometry relative to Φ.

The complete description of linear isometries of symmetric gauge functions is given in [9] and surveyed in [16]. Assume Φ is not a scalar multiple of the usual Euclidean norm. It follows from Theorem 2.5 in [16] (also see Theorem 4.2 in [9])

that in the case where $n \neq 4$ the isometry group corresponding to Φ equals the group GP_n of generalized permutation matrices. Therefore, the transformation (11) is not an isometry and, consequently, the possibility $c = -2/n$ is ruled out. However, if $n = 4$, the situation is different, there does exist a symmetric gauge function on $\mathbb{R}^4$ with respect to which the transformation (11) is an isometry. To present an example, we follow Example 4.4 in [9] and the discussion preceding it. Pick any bounded set $S \subset \mathbb{R}^4$. Define

$$\Phi(x) = \max\{|\langle x, Py\rangle| \,:\, y \in S, P \in GP_n\}, \quad x \in \mathbb{R}^4.$$

Then Φ is a symmetric gauge function on $\mathbb{R}^4$. As written in [9], the matrix $M \in M_4(\mathbb{R})$ gives rise to an isometry of $\mathbb{R}^4$ relative to Φ if and only if M^{tr} leaves the set $\mathcal{S} = \{Py \,:\, y \in S, P \in GP_n\}$ invariant. We now define

$$S = \{(1, 0, 0, 0)^{tr}, (1/2)(1, 1, 1, 1)^{tr}\}.$$

One can easily verify that the matrix of the transformation (11) is symmetric and leaves the corresponding set $\mathcal{S}$ invariant. This implies that the transformation (11) is an isometry relative to the above defined particular gauge function. Consequently, the possibility $c = -2/4$ can really appear, the transformation $A \mapsto (\det A)^{-1/2}A$ is an isometry of $\mathbb{P}_4$ under a certain unitarily invariant norm on $\mathbb{M}_4$. As for the case where Φ is a scalar multiple of the Euclidean norm, one can easily check that the transformation in (11) is an isometry. We are done whenever $n \geq 3$.

It remains to treat the case $n = 2$. We can argue as follows. First of all, observe that we have even in that case that ϕ is a continuous Jordan triple automorphism of $\mathbb{P}_n$ (see the first part of the proof before employing the assumption $n \geq 3$). It follows from Lemma 6 that there is a commutativity preserving bijective linear transformation $f : \mathbb{H}_n \to \mathbb{H}_n$ such that $\phi(A) = e^{f(\log A)}$, $A \in \mathbb{P}_n$.

We have

$$N(f(T)) = d_N(I, e^{f(T)}) = d_N(\phi(I), \phi(e^T)) = d_N(I, e^T) = N(T), \quad T \in \mathbb{H}_n$$

and this shows that f is an isometry of $\mathbb{H}_n$ with respect to the norm N. The structure of isometries of $\mathbb{H}_n$ under any unitary similarity invariant norm was determined in [17]. Here we apply the result for the case $n = 2$. Assume N is not induced by any inner product on $\mathbb{H}_2$. Theorem 2 in [17] tells us that we have the following two possibilities:

   (i) There is a unitary matrix $U \in \mathbb{U}_2$ and a scalar $d \in \{-1, 1\}$ such that $f : \mathbb{H}_2 \to \mathbb{H}_2$ is either of the form $f(T) = dUTU^*$, $T \in \mathbb{H}_2$ or of the form $f(T) = dUT^{tr}U^*$, $T \in \mathbb{H}_2$.
   (ii) $N(A)$ is a function of $|\operatorname{Tr} A|$ and $\|A\|_{HS}$ and f is a unitary transformation on $\mathbb{H}_2$ satisfying $f(I) = I$ or $f(I) = -I$.

The case (i) brings nothing new, it implies that we have one of the possibilities (t1)-(t4) for ϕ. In the case (ii) f is unitary on $\mathbb{H}_2$ and we can assume that it sends I to I (otherwise we would consider $-f$, i.e. we compose ϕ by the inverse operation). The fact that f is unitary means that it preserves the standard inner product on $\mathbb{H}_2$ (the one defined via the trace functional). Moreover, f maps I to I and hence we can compute

$$\operatorname{Tr} A = \operatorname{Tr}(A\,I) = \operatorname{Tr}(f(A)f(I)) = \operatorname{Tr} f(A), \quad A \in \mathbb{H}_2.$$

Consequently, f is trace preserving. Pick an arbitrary rank-one projection $P \in \mathbb{M}_2$. Then $f(P)$ has trace 1 and we have

$$0 = \mathrm{Tr}(P(I - P)) = \mathrm{Tr}(f(P)(I - f(P)))$$

implying that the trace of $f(P)^2$ is also 1. It follows easily that the eigenvalues of $f(P)$ are precisely 0 and 1 which means that $f(P)$ is a rank-one projection. Consequently, $f : \mathbb{H}_2 \to \mathbb{H}_2$ is a linear bijection which maps projections to projections. It is well known (cf., [21], Appendix) that any linear transformation of $\mathbb{H}_n$ (for any integer $n \geq 2$) which sends each projection to a projection is necessarily either zero or a so-called Jordan *-automorphism of $\mathbb{H}_n$. The structure of those automorphisms is known, they are unitary similarity transformations or unitary similarity transformations composed by the transposition operation. Thus we have a unitary matrix $U \in \mathbb{U}_2$ such that either

$$f(T) = UTU^*, \quad T \in \mathbb{H}_2$$

or

$$f(T) = UT^{tr}U^*, \quad T \in \mathbb{H}_2$$

holds true. Apparently, this means that for ϕ we have one of the possibilities (t1), (t3). If $f(I) = -I$, we obviously obtain the remaining possibilities (t2), (t4).

Finally, we need to consider the case where the unitarily invariant norm N (originally defined on $\mathbb{M}_2$) when restricted to $\mathbb{H}_2$ is induced by an inner product on $\mathbb{H}_2$. It is mentioned in [17] (see p. 216) that N is necessarily of the form

$$N(T)^2 = \alpha(\mathrm{Tr}\, T)^2 + \beta\, \mathrm{Tr}(nT - (\mathrm{Tr}\, T)I)^2, \quad T \in \mathbb{H}_2.$$

In fact, that sort of representation holds also in the general case $n \geq 2$. In the particular case $n = 2$, it is easy to deduce from the above displayed formula that the gauge function corresponding to N is necessarily a scalar multiple of the Euclidean norm on $\mathbb{R}^2$, i.e., N is a scalar multiple of $\|.\|_{HS}$. Therefore, f is a unitary operator on $\mathbb{H}_2$. By the commutativity preserving property of f it follows that $f(I)$ is a real scalar multiple of the identity which then implies that either $f(I) = I$ or $f(I) = -I$ holds. After this we can follow the reasoning applied above to treat the case (ii) and once again obtain one of the possibilities (t1)-(t4).

The proof of the theorem is complete.   □

In the last part of the paper we present the proof of Theorem 4. In fact, we carry it out via a few lemmas. In what follows we shall use several times the equality $\log \det A = \mathrm{Tr} \log A$ which holds for all $A \in \mathbb{P}_n$.

Lemma 9. *Let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be a surjective isometry relative to the metric δ_S. Then ϕ satisfies*

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

Proof. As in the proof of Theorem 3, we shall apply Proposition 8. First we check if all the conditions listed there are satisfied. As already mentioned in the proof of Theorem 3, since $X_1 = X_2 = \mathbb{P}_n$, hence X_2 is obviously 2-torsion free and 2-divisible.

By Lemma 2 in [26], the symmetric Stein divergence satisfies

(12) $$S(A^{-1}, B^{-1}) = S(A, B), \quad S(TAT^*, TBT^*) = S(A, B)$$

for all $A, B \in \mathbb{P}_n$ and invertible matrix $T \in \mathbb{M}_n$. It follows that the property (p1) in Proposition 8 is fulfilled.

Concerning (p3) recall that, as we have already referred to it in the proof of Theorem 3, (p3) follows from the 2-divisibility of X_2 (see the discussion after Definition 3.4 in [13]) and from the invariance properties (12) of the divergence S implying similar properties for the metric $\delta_S = \sqrt{S}$.

It remains to check (p2). We assert that

$$(13) \qquad\qquad S(BX^{-1}B, X) \geq 2S(X, B)$$

holds for every $B, X \in \mathbb{P}_n$. Indeed, we compute

$$\log \det \left(\frac{(X^{-1/2}BX^{-1/2})^2 + I}{2} \right) + \log \det X - \log \det B =$$

$$(14) \qquad \log \det \left(\frac{X^{-1/2}(BX^{-1}B + X)X^{-1/2}}{2} \right) + \log \det X - \log \det B =$$

$$\log \det \left(\frac{BX^{-1}B + X}{2} \right) - \log \det B =$$

$$S(BX^{-1}B, X).$$

On the other hand, we have

$$\log \det \left(\frac{X^{-1/2}BX^{-1/2} + I}{2} \right) + \frac{1}{2}(\log \det X - \log \det B) =$$

$$(15) \qquad \log \det \left(\frac{X^{-1/2}(B + X)X^{-1/2}}{2} \right) + \frac{1}{2}(\log \det X - \log \det B) =$$

$$\log \det \left(\frac{B + X}{2} \right) - \frac{1}{2}(\log \det X + \log \det B) =$$

$$S(B, X).$$

We assert that for every $Y \in \mathbb{P}_n$ we have

$$(16) \qquad\qquad \log \det \left(\frac{Y^2 + I}{2} \right) \geq 2 \log \det \left(\frac{Y + I}{2} \right).$$

Indeed, the function $\log((y^2 + 1)/2) - 2\log((y + 1)/2)$ is nonnegative on $]0, \infty[$. This implies that for any $Y \in \mathbb{P}_n$ we have

$$\log \frac{Y^2 + I}{2} \geq \log \left(\frac{Y + I}{2} \right)^2.$$

Since Tr is a positive functional, we obtain

$$\mathrm{Tr} \log \left(\frac{Y^2 + I}{2} \right) \geq 2 \, \mathrm{Tr} \log \left(\frac{Y + I}{2} \right)$$

and this is exactly what has been asserted in (16). Using that inequality, (14), and (15) we obtain (13).

Therefore, all assumptions in Proposition 8 hold for any $A, B \in \mathbb{P}_n$ and hence we obtain the statement of the lemma. $\qquad\square$

We proceed with the following assertion.

Lemma 10. *The topologies induced by the metric δ_S and the operator norm $\|.\|$ on $\mathbb{P}_n$ coincide.*

Proof. As in (15) we have

$$S(A, B) = \log \det \left(\frac{A^{-1/2}BA^{-1/2} + I}{2} \right) - \frac{1}{2} \log \det A^{-1/2}BA^{-1/2}.$$

So, if (B_k) is a sequence in $\mathbb{P}_n$ such that $B_k \to A$ in the operator norm topology, then we obtain $A^{-1/2}B_kA^{-1/2} \to I$, and the above formula gives us that $S(A, B_k) \to 0$. Consequently, $B_k \to A$ holds in the metric δ_S.

Conversely, assume that $S(A, B_k) \to 0$. Then the numerical sequence

$$S(A, B_k) = \mathrm{Tr} \left[\log \left(\frac{A^{-1/2}B_kA^{-1/2} + I}{2} \right) - \log(A^{-1/2}B_kA^{-1/2})^{1/2} \right].$$

converges to 0.

Define the scalar function

$$f(y) = \log((y + 1)/2) - \log\sqrt{y}, \quad y > 0.$$

Since f is everywhere nonnegative, it follows that

$$\mathrm{Tr}\, f(Y) = \mathrm{Tr} \left(\log((Y + I)/2) - \log(Y^{1/2}) \right) \geq 0, \quad Y \in \mathbb{P}_n.$$

Now assuming that for a positive definite matrix $Y \in \mathbb{P}_n$ the quantity $\mathrm{Tr}\, f(Y)$ is small, we have that all eigenvalues of $f(Y)$ are small. Having a look at the graph of f and recalling that the set of all eigenvalues of $f(Y)$ equals the image of the set of all eigenvalues of Y under f, we conclude that the eigenvalues of Y are necessarily close enough to 1. It follows that for any given $\delta > 0$ the eigenvalues of $A^{-1/2}B_kA^{-1/2}$ fall between $1 - \delta$ and $1 + \delta$ for large enough k. From the inequalities

$$(1 - \delta)I \leq A^{-1/2}B_kA^{-1/2} \leq (1 + \delta)I$$

we deduce

$$-\delta A \leq B_k - A \leq \delta A$$

which implies that $\|B_k - A\| \leq \delta\|A\|$ holds for large k. It follows that $B_k \to A$ in the operator norm. The proof of the lemma is complete. $\qquad\square$

We now present the proof of Theorem 4. Observe that one could follow the strategy in the proof of Theorem 3, use Theorem 1 in the case $n \geq 3$ and treat the remaining case $n = 2$ separately. However, here we shall apply an argument employing differential calculus (together with the previous two auxiliary results and Lemma 6) that works equally well in all dimensions.

Proof of Theorem 4. Let $\phi : \mathbb{P}_n \to \mathbb{P}_n$ be a surjective isometry relative to the metric δ_S. Considering the transformation $\phi(I)^{-1/2}\phi(.)\phi(I)^{-1/2}$, by (12) we have a unital map which is again a surjective isometry of $\mathbb{P}_n$. Therefore, without serious loss of generality we may and do assume that ϕ is a surjective isometry of $\mathbb{P}_n$ relative to δ_S which maps I to I. Applying Lemma 9, we have that ϕ satisfies

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

Just as in the proof of Theorem 3 we deduce that ϕ is in fact a Jordan triple automorphism of $\mathbb{P}_n$. In Lemma 10 we have learned that δ_S induces the same topology as the operator norm. It follows that ϕ is continuous in this latter topology, too. Consequently, Lemma 6 applies and we have a bijective linear transformation $f : \mathbb{H}_n \to \mathbb{H}_n$ such that

$$(17) \qquad\qquad \phi(A) = e^{f(\log A)}, \quad A \in \mathbb{P}_n.$$

In what follows we show that $\phi(A) = (f^{-1})^*(A)$ holds for every $A \in \mathbb{P}_n$, where $*$ stands for the adjoint of a linear transformation on the real Hilbert space $\mathbb{H}_n$ (equipped with the natural inner product defined by the help of the trace functional).

We have

$$S(A, B) = \operatorname{Tr}\left[\log\left(\frac{A+B}{2}\right) - \frac{1}{2}(\log A + \log B)\right], \quad A, B \in \mathbb{P}_n.$$

Using this formula, the isometric property of ϕ and (17), for the linear transformation f we obtain that

(18)
$$\operatorname{Tr}\left[\log\left(\frac{e^{f(\log A)} + e^{f(\log B)}}{2}\right) - \frac{1}{2}(f(\log A) + f(\log B))\right] =$$
$$\operatorname{Tr}\left[\log\left(\frac{A+B}{2}\right) - \frac{1}{2}(\log A + \log B)\right]$$

holds for all $A, B \in \mathbb{P}_n$. Now fix $A \in \mathbb{P}_n$ and consider the left and right hand sides in (18) as functions of the variable $B \in \mathbb{P}_n$. Clearly, those functions are differentiable. In what follows we compute their derivatives at $B = I$.

In the paper [24] Pedersen studied differentiability properties of operator valued functions of one operator variable that correspond to scalar functions via the continuous functional calculus. We follow the notation given in that paper. For any real function g on an open interval in $\mathbb{R}$ consider the matrix valued function $A \mapsto g(A)$ defined on the open set of all self-adjoint matrices whose spectra is a subset of the domain of g. If that function is Fréchet differentiable at some matrix $A \in \mathbb{H}_n$, then its derivative at A (which is a linear transformation from $\mathbb{H}_n$ into $\mathbb{H}_n$) is denoted by $d\,g_A$. In [24], the derivatives of several such functions have been calculated. By the proof of Theorem 3.2 and Paragraph 4 in [24] we have

$$(d\exp_A)(X) = \int_0^1 e^{sA} X e^{(1-s)A}d\,s,$$
$$(d\log_A)(X) = \int_0^\infty (A+sI)^{-1}X(A+sI)^{-1}d\,s$$

In particular, it follows that

(19)
$$(d\log_I)(X) = X.$$

Using the rules of differentiation we hence obtain the following formulae:

$$(d\,(\operatorname{Tr}\circ\log)_A)(X) = \operatorname{Tr}\int_0^\infty (A+sI)^{-1}X(A+sI)^{-1}d\,s =$$
$$\int_0^\infty \operatorname{Tr}\left((A+sI)^{-1}X(A+sI)^{-1}\right)d\,s =$$

(20)
$$\int_0^\infty \operatorname{Tr}\left((A+sI)^{-2}X\right)d\,s =$$
$$\operatorname{Tr}\int_0^\infty (A+sI)^{-2}Xd\,s =$$
$$\operatorname{Tr}((\int_0^\infty (A+sI)^{-2}d\,s)X) = \operatorname{Tr}((d\log_A)(I)X) = \operatorname{Tr}(A^{-1}X)$$

In the last equality we used Proposition 2.2 in [24] stating that for commuting A and X we have $(d\,g_A)(X) = g'(A)X$. Next, by the chain rule we have

$$(d\,(\exp \circ f \circ \log)_A)(X) =$$

(21)
$$\int_0^1 e^{tf(\log A)} f\left[\int_0^\infty (A + sI)^{-1} X(A + sI)^{-1} d\,s\right] e^{(1-t)f(\log A)} d\,t.$$

Consequently,

$$(d\,(\exp \circ f \circ \log)_I)(X) = f(X)$$

Therefore, using the chain rule and the formulae (19), (20), (21), when we differentiate in (18) at $B = I$ (A is being fixed) we obtain that

(22)
$$\mathrm{Tr}\left((I + e^{f(\log A)})^{-1} f(X)\right) - \frac{1}{2}\mathrm{Tr}\,f(X) =$$
$$\mathrm{Tr}\left((A + I)^{-1} X\right) - \frac{1}{2}\mathrm{Tr}\,X$$

holds for every $A \in \mathbb{P}_n$ and $X \in \mathbb{H}_n$.

By Lemma 6 we know that the bijective linear map f in (17) is commutativity preserving. It follows that $f(I) = cI$ holds for some nonzero real number c. We may assume that c is positive, otherwise we consider the map $A \mapsto \phi(A)^{-1}$ to which the linear transformation $-f$ corresponds. Replacing A by λA for any positive λ, from (22) we obtain that

$$\mathrm{Tr}\left((I + \lambda^c e^{f(\log A)})^{-1} f(X)\right) - \frac{1}{2}\mathrm{Tr}\,f(X) = \mathrm{Tr}\left((\lambda A + I)^{-1} X\right) - \frac{1}{2}\mathrm{Tr}\,X$$

holds for every $A \in \mathbb{P}_n$ and $X \in \mathbb{H}_n$. We claim that we necessarily have $c = 1$. Indeed, for example, differentiation in the above formulae with respect to λ gives us that

$$\mathrm{Tr}\left((I + \lambda^c e^{f(\log A)})^{-2} c\lambda^{c-1} e^{f(\log A)} f(X)\right) = \mathrm{Tr}\left((\lambda A + I)^{-2} AX\right).$$

Assuming $0 < c < 1$ or $c > 1$ and letting $\lambda \to 0$, it would follow that $\mathrm{Tr}(AX) = 0$ holds for all $A \in \mathbb{P}_n$, $X \in \mathbb{H}_n$ which is a contradiction. Therefore, $c = 1$ and we have

$$\mathrm{Tr}\left((I + \lambda e^{f(\log A)})^{-2} e^{f(\log A)} f(X)\right) = \mathrm{Tr}\left((\lambda A + I)^{-2} AX\right).$$

Letting $\lambda \to 0$, it follows that

$$\mathrm{Tr}\left(e^{f(\log A)} f(X)\right) = \mathrm{Tr}\left(AX\right).$$

This can be rewritten as

$$\mathrm{Tr}\left(\phi(A)X\right) = \mathrm{Tr}\left(Af^{-1}(X)\right)$$

which implies that

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

Therefore, the linear bijection $F = (f^{-1})^*$ on $\mathbb{H}_n$ is an extension of ϕ. It follows that F is a Jordan triple automorphism on $\mathbb{P}_n$ and by continuity it is a Jordan triple map also on the set of positive semidefinite matrices meaning that $F(ABA) = F(A)F(B)F(A)$ holds for all positive semidefinite matrices $A, B \in \mathbb{M}_n$. In particular, F maps projections to projections. At the end of the proof of Theorem 3 we have already recalled the fact that this property implies that F is a Jordan *-automorphism of $\mathbb{H}_n$. The structure of those transformations is well known. As

mentioned there, we have a unitary matrix $U \in \mathbb{U}_n$ such that F is either of the form

$$F(A) = UAU^*, \quad A \in \mathbb{H}_n$$

or of the form

$$F(A) = UA^{tr}U^*, \quad A \in \mathbb{H}_n.$$

The proof of the theorem can now be completed trivially. $\qquad\square$

Remark 11. We conclude the paper with two open problems.

Our statements concerning isometries are valid for any $n \geq 2$ while the results on the structure of Jordan triple endomorphisms and automorphisms are formulated only for $n \geq 3$. So the problem arises naturally if the conclusions in Theorem 1 or Corollary 2 hold also when $n = 2$. We believe the question is quite difficult and leave it as an open problem. Nevertheless, there is a hope for solution. Namely, in the recent paper [11], Gselmann has managed to describe all Jordan triple functionals of $\mathbb{P}_n$ in the general case $n \geq 2$ even without assuming continuity. Her result states that any such transformation can be obtained as the composition of the determinant function and a multiplicative function of the positive real line (this latter function can behave very badly, its graph may be everywhere dense in the upper right quadrant of the plane). This result generalizes Lemma 7 substantially and gives some hope that one can obtain the structure of Jordan triple maps in the probably much more complicated matrix valued case, too.

The second problem we raise is to investigate the isometries of $\mathbb{P}_n$ relative to the same metrics as what we have considered above but without assuming their surjectivity.

References

[1] R. Bhatia, *Positive Definite Matrices,* Princeton University Press, Princeton Oxford, 2007.

[2] R. Bhatia and J. Holbrook, *Riemannian geometry and matrix geometric means,* Linear Algebra Appl. **423** (2006), 594–618.

[3] Z. Chebbi and M. Moahker, *Means of hermitian positive-definite matrices based on the log-determinant α-divergence function,* Linear Algebra Appl. **436** (2012), 1872-1889.

[4] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos, *Efficient similarity search for covariance matrices via the Jensen-Bregman LogDet divergence,* In International Conference on Computer Vision (ICCV), Nov. 2011.

[5] M.D. Choi, A.A. Jafarian and H. Radjavi, *Linear maps preserving commutativity,* Linear Algebra Appl. **87** (1987), 227–241.

[6] G. Corach, H. Porta and L. Recht, *Geodesics and operator means in the space of positive operators,* Internat. J. Math. **4** (1993), 193–202.

[7] G. Corach, H. Porta and L. Recht, *Convexity of the geodesic distance on spaces of positive operators,* Illinois J. Math. **38** (1994), 87–94.

[8] G. Corach and A.L. Maestripieri, *Differential and metrical structure of positive operators,* Positivity, **3** (1999), 297–315.

[9] D.Ž. Đoković, C.K. Li and L. Rodman, *Isometries of symmetric gauge functions,* Linear Multilinear Algebra **30** (1991), 81–92.

[10] 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

[11] E. Gselmann, *Jordan triple mappings on positive definite matrices,* preprint.

[12] S. Gudder and G. Nagy, *Sequentially independent effects,* Proc. Amer. Math. Soc. **130** (2002), 1125–1130.

[13] 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.

[14] O. Hatori and L. Molnár, *Isometries of the unitary group,* Proc. Amer. Math. Soc. **140** (2012), 2141–2154.

[15] O. Hatori and L. Molnár, *Isometries of the unitary groups and Thompson isometries of the spaces of invertible positive elements in C*-algebras,* J. Math. Anal. Appl., to appear.

[16] C.K. Li, *Some aspects of the theory of norms,* Linear Algebra Appl. **212/213** (1994), 71-100.

[17] C.K. Li and N.K. Tsing, *Linear operators preserving unitary similarity invariant norms,* Linear Multilinear Algebra **27** (1990), 213-224.

[18] W. Maak, *Fastperiodische Funktionen, Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen,* Berlin, 1950.

[19] L. Molnár, *A remark to the Kochen-Specker theorem and some characterizations of the determinant on sets of Hermitian matrices,* Proc. Amer. Math. Soc. **134** (2006) 2839–2848.

[20] L. Molnár, *Non-linear Jordan triple automorphisms of sets of self-adjoint matrices and operators,* Studia Math. **173** (2006), 39–48.

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

[22] L. Molnár, *Thompson isometries of the space of invertible positive operators,* Proc. Amer. Math. Soc. **137** (2009), 3849–3859.

[23] L. Molnár, *Jordan triple endomorphisms and isometries of unitary groups,* preprint.

[24] G.K. Pedersen, *Operator differentiable functions,* Publ. RIMS Kyoto Univ. **36** (2000), 139–157.

[25] S. Sra, *A new metric on the manifold of kernel matrices with application to matrix geometric means,* Adv. Neural Inf. Process. Syst., to appear.

[26] S. Sra, *Positive definite matrices and the symmetric Stein divergence,* arXiv: [math.FA-1110.1773v2]

MTA-DE "Lendület" Functional Analysis Research Group, Institute of Mathematics, University of Debrecen, H-4010 Debrecen, P.O. Box 12, Hungary

E-mail address: `molnarl@science.unideb.hu`

URL: `http://www.math.unideb.hu/~molnarl/`