The University of Manchester

MANCHESTER 1824

# *On the Definition of Two Natural Classes of Scalar Product*

Mackey, D. Steven and Mackey, Niloufer and Tisseur, Francoise

2007

MIMS EPrint: **2007.64**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: http://eprints.maths.manchester.ac.uk/
And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

# ON THE DEFINITION OF TWO NATURAL CLASSES OF SCALAR PRODUCT*

D. STEVEN MACKEY[†], NILOUFER MACKEY*, AND FRANÇOISE TISSEUR[‡]

**Abstract.** We identify two natural classes of scalar product, termed unitary and orthosymmetric, which serve to unify assumptions for the existence of structured factorizations, iterations and mappings. A variety of different characterizations of these scalar product classes is given.

**Key words.** Lie algebra, Jordan algebra, scalar product, bilinear form, sesquilinear form, orthosymmetric, adjoint, structured matrix, Hamiltonian, skew-Hamiltonian, Hermitian, complex symmetric, skew-symmetric, persymmetric, perskew-symmetric, perplectic, symplectic, pseudo-orthogonal.

**1. Introduction.** Matrices that are structured with respect to a scalar product (see Table 1.1) arise in many important applications. Several useful properties, such as the involutory property of the adjoint and the preservation of norm by adjoint, do not hold in every scalar product space. In this note we consider a number of such properties, and show that they cluster together into two groups of equivalent properties, thereby delineating two natural classes of scalar products. We have found that the identification of these classes, which we term *orthosymmetric* and *unitary*, has simplified the development of structured factorizations [6], iterations [1], [2] and mappings [7], [8]. It also helps to clarify existing results in the literature. Note that all the "classical" examples of scalar products listed in Table 1.1 are both orthosymmetric and unitary, as will easily be seen from Theorems 1.6 and 1.8. This short note is a more complete version of [6, App. A].

**1.1. Preliminaries.** We give a very brief summary of the required definitions and notation. For more details, see Mackey, Mackey, and Tisseur [5].

Consider a scalar product on $\mathbb{K}^n$, that is, a bilinear or sesquilinear form $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ defined by any nonsingular matrix $M$: for $x, y \in \mathbb{K}^n$,

$$\langle x, y \rangle_{\mathrm{M}} = \begin{cases} x^T M y, & \text{for real or complex bilinear forms,} \\ x^* M y, & \text{for sesquilinear forms.} \end{cases}$$

Here $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and the superscript $*$ denotes conjugate transpose.

The *adjoint* of $A$ with respect to the scalar product $\langle \cdot, \cdot \rangle_{\mathrm{M}}$, denoted by $A^\star$, is uniquely defined by the property $\langle Ax, y \rangle_{\mathrm{M}} = \langle x, A^\star y \rangle_{\mathrm{M}}$ for all $x$, $y \in \mathbb{K}^n$. It can be shown that the adjoint is given explicitly by

$$A^\star = \begin{cases} M^{-1} A^T M, & \text{for bilinear forms,} \\ M^{-1} A^* M, & \text{for sesquilinear forms.} \end{cases} \tag{1.1}$$

Associated with $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ is an automorphism group $\mathbb{G}$, a Lie algebra $\mathbb{L}$, and a Jordan algebra $\mathbb{J}$, defined by

$$\begin{aligned} \mathbb{G} &:= \left\{ G \in \mathbb{K}^{n \times n} : \langle Gx, Gy \rangle_{\mathrm{M}} = \langle x, y \rangle_{\mathrm{M}} \ \forall x, y \in \mathbb{K}^n \right\} = \left\{ G \in \mathbb{K}^{n \times n} : G^\star = G^{-1} \right\}, \\ \mathbb{L} &:= \left\{ L \in \mathbb{K}^{n \times n} : \langle Lx, y \rangle_{\mathrm{M}} = -\langle x, Ly \rangle_{\mathrm{M}} \ \forall x, y \in \mathbb{K}^n \right\} = \left\{ L \in \mathbb{K}^{n \times n} : L^\star = -L \right\}, \\ \mathbb{J} &:= \left\{ S \in \mathbb{K}^{n \times n} : \langle Sx, y \rangle_{\mathrm{M}} = \langle x, Sy \rangle_{\mathrm{M}} \ \forall x, y \in \mathbb{K}^n \right\} = \left\{ S \in \mathbb{K}^{n \times n} : S^\star = S \right\}. \end{aligned}$$

*Version of April, 5 2007.

[†]Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA ({smackey,nil.mackey}@wmich.edu, http://homepages.wmich.edu/~mackey/).

[‡]School of Mathematics, The University of Manchester, Sackville Street, Manchester, M60 1QD, UK (ftisseur@ma.man.ac.uk, http://www.ma.man.ac.uk/~ftisseur/).

TABLE 1.1
*Structured matrices associated with some orthosymmetric scalar products.*

$$R = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}, \quad J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}, \quad \Sigma_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} \text{ with } p+q=n.$$

| **Space** | **M** | **Automorphism Group** $\mathbb{G} = \{G : G^\star = G^{-1}\}$ | **Jordan Algebra** $\mathbb{J} = \{S : S^\star = S\}$ | **Lie Algebra** $\mathbb{L} = \{K : K^\star = -K\}$ |
|---|---|---|---|---|
| **Bilinear forms** | | | | |
| $\mathbb{R}^n$ | $I$ | Real orthogonals | Symmetrics | Skew-symmetrics |
| $\mathbb{C}^n$ | $I$ | Complex orthogonals | Complex symmetrics | Cplx skew-symmetrics |
| $\mathbb{R}^n$ | $\Sigma_{p,q}$ | Pseudo-orthogonals | Pseudosymmetrics | Pseudoskew-symmetrics |
| $\mathbb{C}^n$ | $\Sigma_{p,q}$ | Cplx pseudo-orthogonals | Cplx pseudo-symm. | Cplx pseudo-skew-symm. |
| $\mathbb{R}^n$ | $R$ | Real perplectics | Persymmetrics | Perskew-symmetrics |
| $\mathbb{R}^{2n}$ | $J$ | Real symplectics | Skew-Hamiltonians | Hamiltonians |
| $\mathbb{C}^{2n}$ | $J$ | Complex symplectics | Cplx $J$-skew-symm. | Complex $J$-symmetrics |
| **Sesquilinear forms** | | | | |
| $\mathbb{C}^n$ | $I$ | Unitaries | Hermitian | Skew-Hermitian |
| $\mathbb{C}^n$ | $\Sigma_{p,q}$ | Pseudo-unitaries | Pseudo-Hermitian | Pseudoskew-Hermitian |
| $\mathbb{C}^{2n}$ | $J$ | Conjugate symplectics | $J$-skew-Hermitian | $J$-Hermitian |

$\mathbb{G}$ is a multiplicative group, while $\mathbb{L}$ and $\mathbb{J}$ are linear subspaces. Table 1.1 shows a sample of well-known structured matrices in $\mathbb{G}$, $\mathbb{L}$ or $\mathbb{J}$ associated with some scalar products.

To demonstrate the equivalence of various scalar product properties, we need a flexible way to detect when a matrix $A$ is a scalar multiple of the identity. It is well known that when $A$ commutes with all of $\mathbb{K}^{n\times n}$, then $A = \alpha I$ for some $\alpha \in \mathbb{K}$. There are many other sets besides $\mathbb{K}^{n\times n}$, though, that suffice to give the same conclusion.

DEFINITION 1.1. *A set of matrices $\mathcal{S} \subseteq \mathbb{K}^{n\times n}$ will be called a* CS-set *for $\mathbb{K}^{n\times n}$ if the centralizer of $\mathcal{S}$ consists only of the scalar multiples*[1] *of $I$; that is,*

$$BS = SB \text{ for all } S \in \mathcal{S} \implies B = \alpha I \text{ for some } \alpha \in \mathbb{K}.$$

The following lemma describes a number of useful examples[2] of CS-sets for $\mathbb{R}^{n\times n}$ and $\mathbb{C}^{n\times n}$. For this lemma, we use $D$ to denote an arbitrary diagonal matrix in $\mathbb{K}^{n\times n}$ with *distinct* diagonal entries, $D_+$ for a diagonal matrix with distinct positive diagonal entries, $N = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$ is the $n \times n$ nilpotent Jordan block, $C = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{bmatrix}$ is the cyclic permutation matrix, $E = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] \oplus I_{n-2}$ and $F = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right] \oplus I_{n-2}$.

LEMMA 1.2. *Suppose $\mathcal{S} \subseteq \mathbb{K}^{n\times n}$. Then*

[1]One may think of "*CS*" as standing for either "Commuting implies Scalar", or "Centralizer equals the Scalars".

[2]Another important source of CS-sets for $\mathbb{C}^{n\times n}$ is the classical "Schur's Lemma" [4], [9] from representation theory: any $\mathcal{S} \subseteq \mathbb{C}^{n\times n}$ for which there is no nontrivial $\mathcal{S}$-invariant subspace in $\mathbb{C}^n$ is a CS-set for $\mathbb{C}^{n\times n}$. Thus the matrices in any irreducible representation of a finite group will be a CS-set.

(a) *$\mathcal{S}$ contains a CS-set $\Rightarrow$ $\mathcal{S}$ is a CS-set.*
(b) *Let $\overline{\mathcal{S}}$ denote $\{\overline{A} : A \in \mathcal{S}\}$. If $\mathcal{S}$ is a CS-set for $\mathbb{C}^{n\times n}$, then so is $\overline{\mathcal{S}}$.*
(c) *Any vector space basis for $\mathbb{K}^{n\times n}$, or algebra generating set for $\mathbb{K}^{n\times n}$, is a CS-set for $\mathbb{K}^{n\times n}$. More generally, any set whose span (either in the vector space sense or the algebra sense) contains a CS-set is a CS-set.*
(d) *Each of the finite sets $\{D, N\}$, $\{D, N+N^T\}$, $\{D_+, 3I+N+N^T\}$, and $\{C, E, F\}$ is a CS-set for $\mathbb{K}^{n\times n}$.*
(e) *Any open subset $\mathcal{S} \subseteq \mathbb{K}^{n\times n}$ is a CS-set. (Indeed any open subset of $\mathbb{R}^{n\times n}$ is a CS-set for $\mathbb{C}^{n\times n}$.)*
(f) *The sets of all unitary matrices, all Hermitian matrices, all Hermitian positive semidefinite matrices and all Hermitian positive definite matrices are each CS-sets for $\mathbb{C}^{n\times n}$. The sets of all real orthogonal matrices and all real symmetric matrices are CS-sets for $\mathbb{R}^{n\times n}$ and for $\mathbb{C}^{n\times n}$.*

*Proof.*
(a) This is an immediate consequence of Definition 1.1.
(b) $B\overline{S} = \overline{S}B$ for all $\overline{S} \in \overline{\mathcal{S}}$ $\Rightarrow$ $\overline{B}S = S\overline{B}$ for all $S \in \mathcal{S}$. But $\mathcal{S}$ is a CS-set, so $\overline{B} = \alpha I$, or equivalently $B = \overline{\alpha} I$. Thus $\overline{\mathcal{S}}$ is a CS-set.
(c) If $B$ commutes with either a vector space basis or an algebra generating set for $\mathbb{K}^{n\times n}$, then it commutes with all of $\mathbb{K}^{n\times n}$, and hence $B = \alpha I$.
(d) Any matrix $B$ that commutes with $D$ must itself be a diagonal matrix, and any diagonal $B$ that commutes with $N$ must have equal diagonal entries, so that $B = \alpha I$. Thus $\mathcal{S} = \{D, N\}$ is a CS-set. Similar arguments show that $\{D, N+N^T\}$ and $\{D_+, 3I+N+N^T\}$ are also CS-sets. To see that $\{C, E, F\}$ is a CS-set, first observe that a matrix $B$ commutes with $C$ iff it is a polynomial in $C$, i.e. iff $B$ is a circulant matrix. But any circulant $B$ that commutes with $E$ must be of the form $B = \alpha I + \beta K$, where $K$ is defined by $K_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } i \neq j \end{cases}$. Finally, $B = \alpha I + \beta K$ commuting with $F$ forces $\beta = 0$, so $B = \alpha I$, showing that $\{C, E, F\}$ is a CS-set.
(e) This follows from (a) and (c), since any open subset of $\mathbb{K}^{n\times n}$ contains a vector space basis for $\mathbb{K}^{n\times n}$.
(f) This follows from (a) and (d), by observing that $\{D, N+N^T\}$ consists of two real symmetric matrices, $\{D_+, 3I+N+N^T\}$ consists of two real symmetric positive definite matrices, and $\{C, E, F\}$ consists of three real orthogonal matrices. $\square$

A second simple result that is needed to show the equivalence of various scalar product properties is the following lemma.

LEMMA 1.3. *Let $M \in \mathbb{K}^{n\times n}$ be a nonzero matrix. Then*
1. *$M^T = \alpha M$ for some $\alpha \in \mathbb{K}$ $\Leftrightarrow$ $M^T = \pm M$.*
2. *$M^* = \alpha M$ for some $\alpha \in \mathbb{K}$ $\Leftrightarrow$ $M^* = \alpha M$ for some $|\alpha| = 1$*
   *$\Leftrightarrow$ $M = \beta H$ for some Hermitian $H$ and $|\beta| = 1$.*
3. *$MM^* = \alpha I$ for some $\alpha \in \mathbb{K}$ $\Leftrightarrow$ $M = \beta U$ for some unitary $U$ and $\beta > 0$.*

*Proof.* Since the proofs of the reverse implications ($\Leftarrow$) in 1, 2, and 3 are immediate, we only include the proofs of the forward implications ($\Rightarrow$) in each case.
1. $M^T = \alpha M \Rightarrow M = (M^T)^T = (\alpha M)^T = \alpha M^T = \alpha^2 M \Rightarrow \alpha^2 = 1 \Rightarrow \alpha = \pm 1$.
2. $M^* = \alpha M \Rightarrow M = (M^*)^* = (\alpha M)^* = \overline{\alpha} M^* = |\alpha|^2 M \Rightarrow |\alpha|^2 = 1 \Rightarrow |\alpha| = 1$. To see the second implication, let $H = \sqrt{\alpha} M$, where $\sqrt{\alpha}$ is either of the two square roots of $\alpha$ on the unit circle. It is easy to check that $H$ is Hermitian, and $M = \beta H$ with $\beta = (\sqrt{\alpha})^{-1}$ on the unit circle.

3. $MM^*$ is positive semidefinite, so $\alpha \geq 0$; then $M \neq 0$ implies $\alpha > 0$. It follows that $U = \frac{1}{\sqrt{\alpha}} M$ is unitary, so $M = \beta U$ with $\beta = \sqrt{\alpha} > 0$. $\square$

This third result is needed for the proof of equivalence of norm preservation conditions defining the class of unitary scalar products.

LEMMA 1.4. *Suppose* $\|\cdot\|$ *is any unitarily invariant norm on* $\mathbb{K}^{n\times n}$ ($\mathbb{K} = \mathbb{R}$ *or* $\mathbb{C}$), *and* $\|\operatorname{diag}(x, 1/x, 1, 1, \ldots, 1)\| = \|I_n\|$ *for some* $x > 0$. *Then* $x = 1$.

*Proof.* The proof proceeds by showing that along the one-parameter set of diagonal matrices $\{\operatorname{diag}(x, 1/x, 1, \ldots, 1) : x > 0\}$, the identity $I_n$ is the *unique* matrix of minimum norm, from which the lemma follows immediately. To facilitate the proof it will be convenient to consider some other sets of diagonal matrices, in the equivalent form of subsets of $\mathbb{R}^n$.

- Let $\mathcal{N} = \{(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n : a_i \geq 0 \text{ and } \sum_{i=1}^n a_i \geq n\}$; clearly $\mathcal{N}$ is a closed, convex subset of the non-negative orthant in $\mathbb{R}^n$.
- Let $\mathcal{H} = \{(a_1, a_2, \ldots, a_n) \in \mathcal{N} : \sum_{i=1}^n a_i = n\} \subset \partial\mathcal{N}$; the set $\mathcal{H}$ is the boundary face of $\mathcal{N}$ closest to the origin. Note that $\mathcal{H}$ is a compact, convex subset of $\mathcal{N}$.
- The matrices of interest correspond to the curve $\mathcal{C} := \{(x, 1/x, 1, \ldots, 1) : x > 0\}$ inside $\mathcal{N}$. Every point of $\mathcal{C}$ lies in the *interior* of $\mathcal{N}$ except for $(1, \ldots, 1) \in \mathcal{H}$, since $x + (1/x) > 2$ for any $x > 0$ with $x \neq 1$.

For brevity we also introduce the notation $\|\operatorname{diag}(a_1, a_2, \ldots, a_n)\| = f(a_1, a_2, \ldots, a_n)$; note that the unitary invariance of $\|\cdot\|$ implies that $f$ is invariant under all permutations of its arguments.

The multiplicative property of norms, i.e. $f(\mu v) = \mu f(v)$ for $\mu > 0$, implies that for any point in the interior of $\mathcal{N}$, e.g. all points of $\mathcal{C}$ except for $(1, 1, \ldots, 1)$, there is a point in $\mathcal{H}$ with a strictly smaller $f$-value. If we can now show that $(1, 1, \ldots, 1)$ attains the minimum $f$-value on $\mathcal{H}$, then it follows that $(1, \ldots, 1)$ must be the unique minimizer of $f$ on $\mathcal{C}$, and the proof will be complete.

Suppose $w$ is any point in $\mathcal{H}$, and consider the average $z = \frac{1}{n!}\left(\sum_{P \in S_n} Pw\right)$ over all permutations in the symmetric group $S_n$. Each coordinate of $w$ gets permuted into any fixed $i$th position by exactly $(n-1)!$ permutations in $S_n$, so this average $z$ is always

$$z = \frac{1}{n!}\left(\sum_{P \in S_n} Pw\right) = \frac{1}{n!}\left(\sum_{j=1}^n (n-1)!\, w_j, \ldots, \sum_{j=1}^n (n-1)!\, w_j\right) = (1, 1, \ldots, 1),$$

since $\sum_{j=1}^n w_j = n$ for any $w \in \mathcal{H}$.

Now let $w \in \mathcal{H}$ be any one of the minimizers of $f$ on $\mathcal{H}$; so $f(w) = \min_{v \in \mathcal{H}} f(v) =: m$. Then the permutation invariance of $f$ implies that $f(Pw) = m$ for every permutation $P \in S_n$. Thus

$$\begin{aligned} f(1, 1, \ldots, 1) = f\left(\frac{1}{n!}\sum_{P \in S_n} Pw\right) &= \frac{1}{n!}\, f\left(\sum_{P \in S_n} Pw\right) \\ &\leq \frac{1}{n!}\sum_{P \in S_n} f(Pw) = \frac{1}{n!}(n!\, m) = m, \end{aligned}$$

and hence $f(1, 1, \ldots, 1)$ is *equal* to $m$, since $m$ is the minimum value. $\square$

**1.2. Orthosymmetric scalar products.** In Shaw [9], scalar products that enjoy property (b) in Theorem 1.6 are called "orthosymmetric". We adopt this name in the following definition.

DEFINITION 1.5 (Orthosymmetric Scalar Product). *A scalar product is said to be* orthosymmetric *if it satisfies any one (and hence all) of the seven equivalent properties in Theorem* 1.6.

THEOREM 1.6. *For a scalar product $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ on $\mathbb{K}^n$, the following are equivalent:*
(a) *Adjoint with respect to $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ is involutory, that is, $(A^\star)^\star = A$ for all $A \in \mathbb{K}^{n\times n}$.*
(a') *$(A^\star)^\star = A$ for all $A$ in some CS-set for $\mathbb{K}^{n\times n}$.*
(b) *Vector orthogonality is a symmetric relation, that is,*
$$\langle x,y\rangle_{\mathrm{M}} = 0 \iff \langle y,x\rangle_{\mathrm{M}} = 0\,,\ \textit{for all}\ x,y\in\mathbb{K}^n.$$
(c) *$\mathbb{K}^{n\times n} = \mathbb{L}\oplus\mathbb{J}$.*
(d) *For bilinear forms, $M^T = \pm M$. For sesquilinear forms, $M^* = \alpha M$ with $\alpha\in\mathbb{C}$, $|\alpha| = 1$; equivalently, $M = \beta H$ with $\beta\in\mathbb{C}$, $|\beta| = 1$ and Hermitian $H$.*
(e) *There exists some CS-set for $\mathbb{K}^{n\times n}$ with the property that every matrix $A$ in this CS-set can be factored as $A = WS$ with $W\in\mathbb{G}$ and $S\in\mathbb{J}$.*
(f) *$\mathbb{L}$ and $\mathbb{J}$ are preserved by arbitrary $\star$-congruence; that is, for $\mathbb{S} = \mathbb{L}$ or $\mathbb{J}$ and $P\in\mathbb{K}^{n\times n}$, $B\in\mathbb{S}\ \Rightarrow\ PBP^\star\in\mathbb{S}$.*

*Proof.* Using (1.1) we have
$$(A^\star)^\star = \begin{cases} (M^{-1}M^T)A(M^{-1}M^T)^{-1} & \text{for bilinear forms,}\\ (M^{-1}M^*)A(M^{-1}M^*)^{-1} & \text{for sesquilinear forms.}\end{cases} \tag{1.2}$$
Hence
$$(A^\star)^\star = A \iff \begin{cases} (M^{-1}M^T)A = A(M^{-1}M^T) & \text{for bilinear forms,}\\ (M^{-1}M^*)A = A(M^{-1}M^*) & \text{for sesquilinear forms.}\end{cases} \tag{1.3}$$

(a) ⇔ (a') ⇔ (d)

(a) ⇒ (a'): Obvious.

(a') ⇒ (d): Equation (1.3) holding for all $A$ in some CS-set means that $M^{-1}M^T = \alpha I$ (resp., $M^{-1}M^* = \alpha I$). The desired conclusion now follows from Lemma 1.3.

(d) ⇒ (a): This follows from a straightforward substitution into (1.2).

(a) ⇔ (c)

(a) ⇒ (c): For any scalar product, $\mathbb{L}\cap\mathbb{J} = \{0\}$; if $B\in\mathbb{L}\cap\mathbb{J}$, then $-B = B^\star = B$, so $B = 0$. Now suppose that (a) holds and consider an arbitrary $A\in\mathbb{K}^{n\times n}$. Define $L = \frac{1}{2}(A - A^\star)$ and $S = \frac{1}{2}(A + A^\star)$ so that $A = L + S$. From $(A^\star)^\star = A$, we conclude that $L^\star = -L$, so that $L\in\mathbb{L}$. Similarly one sees that $S\in\mathbb{J}$. The decomposition $A = L + S$ shows that $\mathbb{K}^{n\times n} = \mathbb{L}+\mathbb{J}$ and because $\mathbb{L}\cap\mathbb{J} = \{0\}$, the sum is direct.

(c) ⇒ (a): $A = L + S \Rightarrow A^\star = L^\star + S^\star = -L + S \Rightarrow (A^\star)^\star = (-L)^\star + S^\star = L + S = A$.

(b) ⇔ (d)

(b) ⇒ (d): Suppose $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ is a bilinear form. Letting $y = Mw$, we have
$$x^Ty = 0 \Leftrightarrow x^TMw = 0 \overset{(b)}{\iff} w^TMx = 0 \Leftrightarrow x^TM^Tw = 0 \Leftrightarrow x^T(M^TM^{-1})y = 0.$$
A similar argument for sesquilinear forms shows that $x^*y = 0 \Leftrightarrow x^*(M^*M^{-1})y = 0$. Thus, property (b) implies that
$$\langle x,y\rangle_{\mathrm{I}} = 0 \Leftrightarrow \langle x,y\rangle_{\mathrm{B}} = 0\,,\ \text{where}\ B = \begin{cases} M^TM^{-1} & \text{for bilinear forms,}\\ M^*M^{-1} & \text{for sesquilinear forms.}\end{cases}$$

Using this relationship we can now probe the entries of $B$ with various pairs $x, y$ such that $\langle x, y \rangle_{\mathrm{I}} = 0$. Let $x = e_i$ and $y = e_j$ with $i \neq j$. Then $B_{ij} = \langle e_i, e_j \rangle_{\mathrm{B}} = 0$, so $B$ must be a diagonal matrix. Next, let $x = e_i + e_j$ and $y = e_i - e_j$ with $i \neq j$. Then

$$0 = \langle e_i + e_j, e_i - e_j \rangle_{\mathrm{B}} = B_{ii} + B_{ji} - B_{ij} - B_{jj} = B_{ii} - B_{jj} \,,$$

so $B_{ii} = B_{jj}$ for all $i \neq j$. Thus $B = \alpha I$ for some nonzero $\alpha \in \mathbb{K}$, and the desired conclusion follows from Lemma 1.3.

(d) $\Rightarrow$ (b): This direction is a straightforward verification. For bilinear forms,

$$\langle x, y \rangle_{\mathrm{M}} = 0 \Leftrightarrow x^T M y = 0 \Leftrightarrow (x^T M y)^T = 0 \Leftrightarrow \pm(y^T M x) = 0 \Leftrightarrow \langle y, x \rangle_{\mathrm{M}} = 0$$

and for sesquilinear forms,

$$\langle x, y \rangle_{\mathrm{M}} = 0 \Leftrightarrow x^* M y = 0 \Leftrightarrow (x^* M y)^* = 0 \Leftrightarrow \alpha(y^* M x) = 0 \Leftrightarrow \langle y, x \rangle_{\mathrm{M}} = 0.$$

(e) $\Leftrightarrow$ (a)

(e) $\Rightarrow$ (a): For all $A$ in our CS-set we have

$$(A^\star)^\star = (S^\star W^\star)^\star = (SW^{-1})^\star = W^{-\star} S^\star = (W^{-1})^{-1} S = WS = A \,,$$

and so (a') holds. That (a') implies (a) was shown earlier.

(a) $\Rightarrow$ (e): The continuity of the eigenvalues of $A^\star A$ implies that there is an open neighborhood $\mathcal{U}$ of the identity in which $A^\star A$ has no eigenvalues on $\mathbb{R}^-$. Thus by [6, Thm. 6.2] every $A$ in the CS-set $\mathcal{U}$ can be factored as $A = WS$ with $W \in \mathbb{G}$ and $S \in \mathbb{J}$.

(a) $\Leftrightarrow$ (f)

(a) $\Rightarrow$ (f): Let $B \in \mathbb{S}$, so that $B^\star = \pm B$. Then $(PBP^\star)^\star = (P^\star)^\star B^\star P^\star = \pm PBP^\star$, and so $PBP^\star \in \mathbb{S}$.

(f) $\Rightarrow$ (a): Consider $\mathbb{S} = \mathbb{J}$ and $B = I \in \mathbb{J}$. Then (f) implies that $PP^\star \in \mathbb{J}$ for any $P \in \mathbb{K}^{n \times n}$, so $PP^\star = (PP^\star)^\star = (P^\star)^\star P^\star$. Since $P^\star$ is nonsingular for any nonsingular $P$, we have $P = (P^\star)^\star$ for every nonsingular $P$. Thus by Lemma 1.2(e) we have property (a'), which was previously shown to be equivalent to (a). ∎

**1.3. Unitary scalar products.** Finally we prove the equivalence of a second set of scalar product space properties. We adopt the name "unitary" for the scalar products satisfying these properties because of (b) and (e) in Theorem 1.8.

Definition 1.7 (Unitary Scalar Product). *A scalar product is said to be* unitary *if it satisfies any one (and hence all) of the six equivalent properties in Theorem* 1.8.

Theorem 1.8. *For a scalar product $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ on $\mathbb{K}^n$, the following are equivalent:*

(a) $(A^*)^\star = (A^\star)^*$ *for all* $A \in \mathbb{K}^{n \times n}$.

(a') $(A^*)^\star = (A^\star)^*$ *for all $A$ in some CS-set for* $\mathbb{K}^{n \times n}$.

(b) *Adjoint preserves unitarity: $U$ unitary $\Rightarrow U^\star$ is unitary.*

(c) *Adjoint preserves Hermitian structure: $H$ Hermitian $\Rightarrow H^\star$ is Hermitian.*

(d) *Adjoint preserves Hermitian positive (semi)definite structure:*
*$H$ Hermitian positive (semi)definite $\Rightarrow H^\star$ is Hermitian positive (semi)definite.*

(e) *$M = \beta U$ for some unitary $U$ and $\beta > 0$.*

(f) *For* **some** *unitarily invariant norm $\|\cdot\|$, $\|A^\star\| = \|A\|$ for all $A \in \mathbb{K}^{n \times n}$.*

(g) *For* **every** *unitarily invariant norm $\|\cdot\|$, $\|A^\star\| = \|A\|$ for all $A \in \mathbb{K}^{n \times n}$.*

*Proof.* From (1.1) it follows that

$$(A^*)^\star = \begin{cases} M^{-1}\overline{A}M & \text{bilinear forms,} \\ M^{-1}AM, & \text{sesquilin. forms} \end{cases} \text{ and } (A^\star)^* = \begin{cases} M^*\overline{A}M^{-*}, & \text{bilinear forms} \\ M^*AM^{-*}, & \text{sesquilin. forms.} \end{cases}$$

Thus for any individual matrix $A \in \mathbb{K}^{n\times n}$ we have

$$(A^*)^\star = (A^\star)^* \iff \begin{cases} \overline{A}(MM^*) = (MM^*)\overline{A} & \text{for bilinear forms} \\ A(MM^*) = (MM^*)A & \text{for sesquilinear forms.} \end{cases} \tag{1.4}$$

<u>(a) ⇔ (a')</u>

(a) ⇒ (a'): This implication is trivial.

(a') ⇒ (a): Suppose $(A^*)^\star = (A^\star)^*$ holds for all $A$ in some CS-set for $\mathbb{K}^{n\times n}$. Then from (1.4) we conclude that $MM^* = \alpha I$, and hence that the two sides of (1.4) hold for all $A \in \mathbb{K}^{n\times n}$.

<u>(a) ⇔ (b)</u>

(a) ⇒ (b): $U^* = U^{-1} \Rightarrow (U^*)^\star = (U^{-1})^\star \overset{(a)}{\Longrightarrow} (U^\star)^* = (U^\star)^{-1} \Rightarrow U^\star$ is unitary.

(b) ⇒ (a): Suppose $U$, and hence also $U^\star$, is unitary. Then we have $(U^\star)^* = (U^\star)^{-1} = (U^{-1})^\star = (U^*)^\star$, showing that $(A^*)^\star = (A^\star)^*$ for all unitary $A$. But from Lemma 1.2 (f), the set of all unitaries is a CS-set for $\mathbb{K}^{n\times n}$, so (a') holds, and hence also (a).

<u>(a) ⇔ (c)</u>

(a) ⇒ (c): $H^* = H \;\Rightarrow\; (H^*)^\star = H^\star \overset{(a)}{\Longrightarrow} (H^\star)^* = H^\star \;\Rightarrow\; H^\star$ is Hermitian.

(c) ⇒ (a): Suppose $H$, and therefore also $H^\star$, is Hermitian. Then we have $(H^\star)^* = H^\star = (H^*)^\star$, and so $(A^*)^\star = (A^\star)^*$ for all Hermitian $A$. But from Lemma 1.2 (f), the set of all Hermitian matrices is a CS-set for $\mathbb{K}^{n\times n}$, so (a') holds, and hence also (a).

<u>(a) ⇔ (d)</u>

(a) ⇒ (d): Because (a) ⇒ (c), we just need to show that positive (semi)definiteness is preserved by adjoint. But for $H$ Hermitian, $H^\star$ and $H$ are similar by definition of the adjoint so the eigenvalues of $H^\star$ and $H$ are the same.

(d) ⇒ (a): This argument is the same as that for (c) ⇒ (a), using the fact that the set of all Hermitian positive (semi)definite matrices is a CS-set for $\mathbb{K}^{n\times n}$.

<u>(a) ⇔ (e)</u>

(a) ⇒ (e): Suppose $(A^*)^\star = (A^\star)^*$ holds for all $A \in \mathbb{K}^{n\times n}$. Then we can conclude from (1.4) that $MM^* = \alpha I$ for some $\alpha \in \mathbb{K}$, and thus from Lemma 1.3 that $M = \beta U$ for some unitary $U$ and $\beta > 0$.

(e) ⇒ (a): $M = \beta U \;\Rightarrow\; MM^* = (\beta U)(\overline{\beta}U^*) = \beta^2 I$. Then by (1.4) we have $(A^*)^\star = (A^\star)^*$ for all $A$.

<u>(e) ⇒ (g) ⇒ (f) ⇒ (e)</u>

(e) ⇒ (g) : Any unitarily invariant norm $\|\cdot\|$ is a function of the singular values [3, p.209–210], so $\|A^T\| = \|A\| = \|A^*\|$ for all $A$. From the formula for the adjoint in (1.1), it follows that $\|A^\star\| = \|A\|$ for all $A$.

(g) ⇒ (f): This direction holds a fortiori.

(f) ⇒ (e): Suppose $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ is bilinear form; with only minor notational changes the same argument works for sesquilinear forms. Let $M = U\Sigma V^*$ be an SVD for the matrix $M$ defining the scalar product. Then

$$\|A^\star\| = \|M^{-1}A^TM\| = \|V\Sigma^{-1}U^*A^TU\Sigma V^*\| = \|\Sigma^{-1}\underbrace{(U^*A^TU)}_{B}\Sigma\|. \tag{1.5}$$

Since $\|A\| = \|A^T\| = \|U^*A^TU\| = \|B\|$, we see that if (f) holds, i.e. if $\|A^\star\| = \|A\|$ for all $A$, then $\Sigma$ has the property that

$$\|\Sigma^{-1}B\Sigma\| = \|B\| \quad \text{for all} \quad B \in \mathbb{K}^{n\times n}. \tag{1.6}$$

Now we choose $B$ to be various permutations in order to probe condition (1.6) and see what constraints it imposes on $\Sigma$. Let $P_{jk}$ (with $j < k$) denote the transposition permutation that interchanges $j$ and $k$. Then $\Sigma^{-1}P_{jk}\Sigma$ differs from the identity $I_n$ only in the $2\times 2$ principal submatrix in the $j$th and $k$th rows and columns; in this submatrix we have $\left[\begin{smallmatrix} 0 & 1/\mu \\ \mu & 0 \end{smallmatrix}\right]$ with $\mu = \sigma_j/\sigma_k$. The unitary invariance of the norm together with $\|\Sigma^{-1}P_{jk}\Sigma\| = \|P_{jk}\|$ now implies that $\|\operatorname{diag}(\mu, 1/\mu, 1, 1, \ldots, 1)\| = \|I\|$. From Lemma 1.4 we can then conclude that $\mu = 1$, so $\sigma_j = \sigma_k$. Since this holds for all $1 \le j < k \le n$, we see that $\Sigma$ must be $\sigma I$ for some $\sigma > 0$. Thus $M = U\Sigma V^* = \sigma UV^*$ which completes the proof. $\square$

REFERENCES


[1] N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur, *Computing the polar decomposition and the matrix sign decomposition in matrix groups*, SIAM J. Matrix Anal. Appl., 25 (2004), pp. 1178–1192.

[2] ———, *Functions preserving matrix groups and iterations for the matrix square root*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 849–877.

[3] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[4] S. Lang, *Linear Algebra*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, third ed., 1987.

[5] D. S. Mackey, N. Mackey, and F. Tisseur, *Structured tools for structured matrices*, Electron. J. Linear Algebra, 10 (2003), pp. 106–145.

[6] ———, *Structured factorizations in scalar product spaces*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 821–850.

[7] ———, *Structured mapping problems for matrices associated with scalar products part I: Lie and Jordan algebras*, MIMS EPrint 2006.44, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Mar. 2006.

[8] ———, *Structured mapping problems for matrices associated with scalar products part II: Automorphism groups*, MIMS EPrint, Manchester Institute for Mathematical Sciences, The University of Manchester, UK. In preparation.

[9] R. Shaw, *Linear algebra and group representations. Vol. I*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1982.