MANCHESTER
1824

The University
of Manchester

On the Definition of Two Natural Classes of Scalar Product

D. Steven Mackey, Niloufer Mackey and
Francoise Tisseur

April 2007

MIMS EPrint: **2007.64**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

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. All the classical examples of scalar products, each giving rise to important classes of structured matrices, are shown to be both orthosymmetric and unitary.

Mathematics Subject Classification (2000). Primary 47A07, 15A63; Secondary 15A57.

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

1. Introduction

Many problems in differential equations and control systems involve operators that are self adjoint or unitary with respect to a symmetric bilinear or Hermitian sesquilinear form. There is an extensive literature going back to the late 1800's in the work of Frobenius [1], Kronecker [9], and Weierstrass [15], and to the seminal work of Krein [7], [8] in the first half of the 1900s. A comprehensive treatment for matrices was given in the influential monograph by Gohberg, Lancaster and Rodman [2], and its recent successor [3].

In this paper we consider scalar products induced by an arbitrary nonsingular matrix M, and then identify what conditions on M lead to useful properties of the scalar product, such as an involutory adjoint or the preservation of norm by the adjoint. We investigate a number of such properties, and show that

The work of the first two authors was supported in part by National Science Foundation Grant DMS-0713799.

they cluster together into two groups of equivalent properties, thereby delineating two natural classes of scalar products, which we term *orthosymmetric* and *unitary*. The identification of these two classes simplifies the development of structured factorizations [11], structured iterations [4], [5] and structured mappings [12], [13], while also helping to clarify existing results in the literature. Our results hold for all forms — real and complex bilinear, as well as complex sesquilinear. All the "classical" examples of scalar products (see Table 1) are shown to be both orthosymmetric and unitary. This short paper is an extended and more complete version of appendix A in [11].

2. Preliminaries

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

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

Here $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and the superscript $*$ denotes conjugate transpose. No restriction other than non-singularity is placed on the matrix M, so that scalar products do not have to *a priori* be either symmetric, Hermitian, or positive definite.

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, $y \in \mathbb{F}^n$, or explicitly by

$$A^\star = \begin{cases} M^{-1}A^T M, & \text{for bilinear forms,} \\ M^{-1}A^* M, & \text{for sesquilinear forms.} \end{cases} \tag{2.1}$$

Associated with each $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ are three classes of structured matrices, an automorphism group $\mathbb{G}$, a Lie algebra $\mathbb{L}$, and a Jordan algebra $\mathbb{J}$, defined by

$$\mathbb{G} := \left\{G \in \mathbb{F}^{n\times n} : \langle Gx, Gy\rangle_{\mathrm{M}} = \langle x, y\rangle_{\mathrm{M}}\ \forall x, y \in \mathbb{F}^n\right\} = \left\{G \in \mathbb{F}^{n\times n} : G^\star = G^{-1}\right\},$$

$$\mathbb{L} := \left\{L \in \mathbb{F}^{n\times n} : \langle Lx, y\rangle_{\mathrm{M}} = -\langle x, Ly\rangle_{\mathrm{M}}\ \forall x, y \in \mathbb{F}^n\right\} = \left\{L \in \mathbb{F}^{n\times n} : L^\star = -L\right\},$$

$$\mathbb{J} := \left\{S \in \mathbb{F}^{n\times n} : \langle Sx, y\rangle_{\mathrm{M}} = \langle x, Sy\rangle_{\mathrm{M}}\ \forall x, y \in \mathbb{F}^n\right\} = \left\{S \in \mathbb{F}^{n\times n} : S^\star = S\right\}.$$

$\mathbb{G}$ is a multiplicative group, while $\mathbb{L}$ and $\mathbb{J}$ are linear subspaces. Table 1 lists a sample of well-known structured matrices in $\mathbb{G}$, $\mathbb{L}$ or $\mathbb{J}$ associated with some familiar scalar products.

Many familiar properties of adjoint that hold in the Euclidean case where M is just the identity matrix, in fact continue to hold in any scalar product space, when M is an arbitrary non-singular matrix:

1. $(A+B)^\star = A^\star + B^\star$
2. $(AB)^\star = B^\star A^\star$
3. $(A^{-1})^\star = (A^\star)^{-1}$
4. $(\alpha A)^\star = \alpha A^\star$ for bilinear forms, $(\alpha A)^\star = \overline{\alpha} A^\star$ for sesquilinear forms
5. $A^\star \sim A$ for bilinear forms, while $A^\star \sim \overline{A}$ for sesquilinear forms, where $\sim$ denotes similarity.

TABLE 1. Structured matrices associated with some orthosymmetric scalar products.

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

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

By contrast, here is a list of some desiderata that do not hold in an arbitrary scalar product space:

6. Adjoint is involutory: $(A^\star)^\star = A$.
7. Vector orthogonality is a symmetric relation: $\langle x, y\rangle_{\mathrm{M}} = 0 \iff \langle y, x\rangle_{\mathrm{M}} = 0$.
8. The right adjoint is also the left adjoint:
   $\langle Ax, y\rangle_{\mathrm{M}} = \langle x, A^\star y\rangle_{\mathrm{M}} \Longrightarrow \langle x, Ay\rangle_{\mathrm{M}} = \langle A^\star x, y\rangle_{\mathrm{M}}$.
9. $\mathbb{F}^{n\times n} = \mathbb{L} \oplus \mathbb{J}$.
10. Adjoint preserves unitarity: U is unitary $\Rightarrow U^\star$ is unitary.
11. Adjoint preserves Hermitianness: H is Hermitian $\Rightarrow H^\star$ is Hermitian.
12. The stars commute: $(A^*)^\star = (A^\star)^*$

Several questions now naturally arise. For example, which matrices M give us an involutory adjoint? Does an involutory adjoint give us a scalar product space in which the stars commute, or one in which orthogonality is a symmetric relation? How are these properties related, and which, if any, of them are equivalent?

The remainder of this section is devoted to laying the groundwork for answering these questions. First, 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 $\mathbb{F}^{n\times n}$, then $A = \alpha I$ for some $\alpha \in \mathbb{F}$. There are many other sets besides $\mathbb{F}^{n\times n}$, though, that suffice to give the same conclusion.

Definition 2.1. A set of matrices $\mathcal{S} \subseteq \mathbb{F}^{n\times n}$ is called a *CS-set* for $\mathbb{F}^{n\times n}$ if the centralizer of $\mathcal{S}$ consists only of the scalar multiples of I, that is,

$$BS = SB \;\text{ for all }\; S \in \mathcal{S} \quad\Longrightarrow\quad B = \alpha I \;\text{ for some }\; \alpha \in \mathbb{F}.$$

One may think of "CS" as standing for either "$\underline{C}$ommuting implies $\underline{S}$calar", or "$\underline{C}$entralizer equals the $\underline{S}$calars".

An important source of CS-sets for $\mathbb{C}^{n\times n}$ is the classical "Schur's Lemma" [10], [14] 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. The following lemma gives a number of other examples of CS-sets for $\mathbb{R}^{n\times n}$ and $\mathbb{C}^{n\times n}$, which will be useful in the context of our investigation. We use D to denote a diagonal matrix in $\mathbb{F}^{n\times n}$ with *distinct* diagonal entries, and D_+ for a diagonal matrix with distinct positive diagonal entries. The $n \times n$ nilpotent Jordan block is $N = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$, whereas $C = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{bmatrix}$ is the cyclic permutation matrix. Finally, let $E = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] \oplus I_{n-2}$ and $F = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right] \oplus I_{n-2}$.

Lemma 2.2. *Suppose $\mathcal{S} \subseteq \mathbb{F}^{n\times n}$. Then*

(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 $\mathbb{F}^{n\times n}$, or algebra generating set for $\mathbb{F}^{n\times n}$, is a CS-set for $\mathbb{F}^{n\times n}$. 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 $\{D, N\}$, $\{D,\ N + N^T\}$, $\{D_+,\ 3I + N + N^T\}$, $\{C, E, F\}$ is a CS-set for $\mathbb{F}^{n\times n}$.*

(e) *Any open subset $\mathcal{S} \subseteq \mathbb{F}^{n\times n}$ 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 2.1.

(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 $\mathbb{F}^{n\times n}$, then it commutes with all of $\mathbb{F}^{n\times n}$, 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 $\mathbb{F}^{n\times n}$ contains a vector space basis for $\mathbb{F}^{n\times n}$.

(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. ∎

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

Lemma 2.3. *Let $M \in \mathbb{F}^{n\times n}$ be a nonzero matrix. Then*

1. $M^T = \alpha M$ *for some* $\alpha \in \mathbb{F} \;\Leftrightarrow\; M^T = \pm M$.
2. $M^* = \alpha M$ *for some* $\alpha \in \mathbb{F} \;\Leftrightarrow\; M^* = \alpha M$ *for some* $|\alpha| = 1$
   $\Leftrightarrow\; M = \beta H$ *for some Hermitian* H *and* $|\beta| = 1$.
3. $MM^* = \alpha I$ *for some* $\alpha \in \mathbb{F} \;\Leftrightarrow\; M = \beta U$ *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 α 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 $\beta = \sqrt{\alpha} > 0$. ∎

Next we prove a result that will be used in Section 4 to establish when the adjoint with respect to a scalar product is norm preserving.

Lemma 2.4. *Suppose $\|\cdot\|$ is any unitarily invariant norm on $\mathbb{F}^{n\times n}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$), and $\|\operatorname{diag}(x, 1/x, 1, 1, \ldots, 1)\| = \|I_n\|$ for some $x > 0$. Then $x = 1$.*

Proof. If we show that the identity I_n is the *unique* matrix of minimum norm along the one-parameter set of diagonal matrices $\{\operatorname{diag}(x, 1/x, 1, \ldots, 1) \,:\, x > 0\}$, the desired result follows immediately. It will be convenient to consider some other sets of real diagonal matrices, in their equivalent form as 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 use 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 ith 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 $\mathcal{H}$, and let $m := f(w) = \min_{v \in \mathcal{H}} f(v)$. 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. ∎

We close this section with the concept of orthogonal companion subspaces introduced in [3, Section 2.2]. We distinguish between left and right orthogonal companion subspaces, with a view to addressing the question of their equality.

Definition 2.5. The *right orthogonal companion* of a nonempty set $S \subseteq \mathbb{F}^n$ with respect to the scalar product $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ is given by

$$S_{\mathrm{M}}^{\perp} := \left\{x \in \mathbb{F}^n : \langle x, y \rangle_{\mathrm{M}} = 0, \quad \text{for all } y \in S\right\},$$

and the *left orthogonal companion* is given by

$$^{\perp}S_{\mathrm{M}} := \left\{x \in \mathbb{F}^n : \langle y, x \rangle_{\mathrm{M}} = 0, \quad \text{for all } y \in S\right\}$$

It is easy to see that the orthogonal companions of any set S are subspaces.

3. Orthosymmetric scalar products

In Shaw [14], scalar products in which vector orthogonality is a symmetric relation (property (b) in Theorem 3.2), are called "orthosymmetric". We adopt this name in the following definition.

Definition 3.1 (Orthosymmetric Scalar Product). A scalar product is said to be *orthosymmetric* if it satisfies any one (and hence all) of the nine equivalent properties in Theorem 3.2.

Theorem 3.2. *For a scalar product $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ on $\mathbb{F}^n$, the following are equivalent:*

(a) *Adjoint with respect to $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ is involutory: $(A^\star)^\star = A$ for all $A \in \mathbb{F}^{n\times n}$.*

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

(b) *Vector orthogonality is a symmetric relation:*
$$\langle x,y\rangle_{\mathrm{M}} = 0 \iff \langle y,x\rangle_{\mathrm{M}} = 0\,,\ \text{for all } x,y\in\mathbb{F}^n.$$

(c) *$\mathbb{F}^{n\times n} = \mathbb{L}\oplus\mathbb{J}$.*

(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 $\mathbb{F}^{n\times n}$ 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: for $\mathbb{S} = \mathbb{L}$ or $\mathbb{J}$ and $P\in\mathbb{F}^{n\times n}$, $B\in\mathbb{S}\Rightarrow PBP^\star\in\mathbb{S}$.*

(g) *Left and right adjoints are equal: for all $A\in\mathbb{F}^{n\times n}$,*
$$\langle Ax,y\rangle_{\mathrm{M}} = \langle x,A^\star y\rangle_{\mathrm{M}} \Longrightarrow \langle x,Ay\rangle_{\mathrm{M}} = \langle A^\star x,y\rangle_{\mathrm{M}}\,,\ \text{for all } x,y\in\mathbb{F}^n.$$

(h) *The left and right orthogonal companion subspaces of any nonempty subset $S\subseteq\mathbb{F}^n$ are always equal.*

Proof. Using (2.1) we have
$$(A^\star)^\star = \begin{cases} (M^{-1}M^T)A(M^{-1}M^T)^{-1} & \text{for bilinear forms,}\\ (M^{-1}M^*)A(M^{-1}M^*)^{-1} & \text{for sesquilinear forms.}\end{cases} \tag{3.1}$$

Hence
$$(A^\star)^\star = A \iff \begin{cases} (M^{-1}M^T)A = A(M^{-1}M^T) & \text{for bilinear forms,}\\ (M^{-1}M^*)A = A(M^{-1}M^*) & \text{for sesquilinear forms.}\end{cases} \tag{3.2}$$

(a) ⇔ (a') ⇔ (d)

(a) ⇒ (a'): Obvious.

(a') ⇒ (d): Equation (3.2) 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 2.3.

(d) ⇒ (a): This follows from a straightforward substitution into (3.1).

(a) ⇔ (c)

(a) $\Rightarrow$ (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 $A \in \mathbb{F}^{n\times n}$. 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 $\mathbb{F}^{n\times n} = \mathbb{L} + \mathbb{J}$ and because $\mathbb{L} \cap \mathbb{J} = \{0\}$, the sum is direct.

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

<u>(b) $\Leftrightarrow$ (d)</u>

(b) $\Rightarrow$ (d): Suppose $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ is a bilinear form. Letting $y = Mw$, we have

$$x^T y = 0 \Leftrightarrow x^T M w = 0 \overset{(b)}{\Longleftrightarrow} w^T M x = 0 \Leftrightarrow x^T M^T w = 0 \Leftrightarrow x^T (M^T M^{-1}) y = 0.$$

A similar argument for sesquilinear forms shows that $x^* y = 0 \Leftrightarrow x^*(M^* M^{-1}) y = 0$. Thus, property (b) implies that

$$\langle x, y \rangle_{\mathrm{I}} = 0 \Leftrightarrow \langle x, y \rangle_{\mathrm{B}} = 0\,, \text{ where } B = \begin{cases} M^T M^{-1} & \text{for bilinear forms,} \\ M^* M^{-1} & \text{for sesquilinear forms.} \end{cases}$$

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 $\alpha \in \mathbb{F}$, and the desired conclusion follows from Lemma 2.3.

(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,

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

<u>(e) $\Leftrightarrow$ (a)</u>

(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 [11, Thm. 6.2] every A in the CS-set $\mathcal{U}$ can be factored as $A = WS$ with $W \in \mathbb{G}$ and $S \in \mathbb{J}$.

<u>(a) $\Leftrightarrow$ (f)</u>

(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) ⇒ (a): Consider $\mathbb{S} = \mathbb{J}$ and $B = I \in \mathbb{J}$. Then (f) implies that $PP^\star \in \mathbb{J}$ for any $P \in \mathbb{F}^{n\times n}$, 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 2.2(e) we have property (a'), which was previously shown to be equivalent to (a).

(g) ⇔ (a): This is straightforward to check.

(b) ⇒ (h): This is also straightforward.

(h) ⇒ (b): For each $y \in \mathbb{F}^n$, use (h) on the set $S = \{y\}$. Thus $\langle x, y\rangle_{\mathrm{M}} = 0 \Leftrightarrow x \in {}^{\perp}\{y\}_{\mathrm{M}} = \{y\}^{\perp}_{\mathrm{M}} \Leftrightarrow \langle y, x\rangle_{\mathrm{M}} = 0$. □

4. Unitary scalar products

We now turn to a second set of equivalent properties in a scalar product space. We adopt the name "unitary" for scalar products satisfying these properties because the adjoint preserves unitarity, and also because the matrix defining the scalar product is a positive scalar multiple of a unitary matrix. (properties (b) and (e) in Theorem 4.2).

Definition 4.1 (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 4.2.

Theorem 4.2. *For a scalar product $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ on $\mathbb{F}^n$, the following are equivalent:*

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

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

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

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

(d) *Adjoint preserves Hermitian positive (semi)definite structure:*
H Hermitian positive (semi)definite ⇒ $H^\star$ 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* $A \in \mathbb{F}^{n\times n}$.

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

Proof. From (2.1) it follows that

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

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

$$(A^*)^\star = (A^\star)^* \iff \begin{cases} \overline{A}(MM^*) = (MM^*)\overline{A} & \text{for bilinear forms} \\ A(MM^*) = (MM^*)A & \text{for sesquilinear forms.} \end{cases} \tag{4.1}$$

<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 $\mathbb{F}^{n\times n}$. Then from (4.1) we conclude that $MM^* = \alpha I$, and hence that the two sides of (4.1) hold for all $A \in \mathbb{F}^{n\times n}$.

(a) ⇔ (b)

(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 2.2 (f), the set of all unitaries is a CS-set for $\mathbb{F}^{n\times n}$, so (a') holds, and hence also (a).

(a) ⇔ (c)

(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 2.2 (f), the set of all Hermitian matrices is a CS-set for $\mathbb{F}^{n\times n}$, so (a') holds, and hence also (a).

(a) ⇔ (d)

(a) ⇒ (d): Because (a) ⇒ (c), the result follows if we 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 $\mathbb{F}^{n\times n}$.

(a) ⇔ (e)

(a) ⇒ (e): Suppose $(A^*)^\star = (A^\star)^*$ holds for all $A \in \mathbb{F}^{n\times n}$. Then we can conclude from (4.1) that $MM^* = \alpha I$ for some $\alpha \in \mathbb{F}$, and thus from Lemma 2.3 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 (4.1) we have $(A^*)^\star = (A^\star)^*$ for all A.

(e) ⇒ (g) ⇒ (f) ⇒ (e)

(e) ⇒ (g) : Any unitarily invariant norm $\|\cdot\|$ is a function of the singular values [6, p.209–210], so $\|A^T\| = \|A\| = \|A^*\|$ for all A. From the formula for the adjoint in (2.1), it follows that $\|A^\star\| = \|A\|$ for all A.

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

(f) ⇒ (e): Suppose $\langle\cdot,\cdot\rangle_M$ 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

$$\|A^\star\| = \|M^{-1}A^TM\| = \|V\Sigma^{-1}U^*A^TU\Sigma V^*\| = \|\Sigma^{-1}\underbrace{(U^*A^TU)}_{B}\Sigma\|. \qquad (4.2)$$

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

$$\|\Sigma^{-1}B\Sigma\| = \|B\| \quad \text{for all} \;\; B \in \mathbb{F}^{n\times n}. \qquad (4.3)$$

Now we choose B to be various permutations in order to probe condition (4.3) and see what constraints it imposes on Σ. 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 2×2 principal submatrix in the jth and kth 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 2.4 we can then conclude that $\mu = 1$, so $\sigma_j = \sigma_k$. Since this holds for all $1 \leq j < k \leq n$, we see that Σ must be σI for some $\sigma > 0$. Thus $M = U \Sigma V^* = \sigma U V^*$ which completes the proof. □

5. Concluding Remarks

By being prepared to work in a general scalar product space, with no conditions on the nature of its defining form other than non-degeneracy, one can critically examine the properties one would like to have in such a space, and analyze the relationships between these properties. Investigating a large list of these desiderata, we showed that they fall into two categories, each comprising equivalent conditions, thus delineating two classes of scalar products, one of which we termed *orthosymmetric* and the other *unitary*.

The equivalences provide strong reasons why these are the two natural classes of scalar products to study. Those established in Theorem 3.2 show that an orthosymmetric scalar product is essentially the same, in the real bilinear and in the complex sesquilinear case, as an indefinite inner product defined in Gohberg, Lancaster and Rodman [3]. It is worth observing that this theorem applies equally to the complex bilinear case, which is now being seen in a variety of applications. The theorem also serves to highlight the centrality of this class of scalar products, lying as it does at the nexus of so many useful properties, now revealed to be equivalent.

And finally, the fact that all the classical examples of scalar products enjoy *both* sets of properties, so are simultaneously orthosymmetric as well as unitary, helps to clarify existing results in the literature. It has guided our development of unified results on structured factorizations, structured iterations and structured mappings.

References

[1] G. Frobenius. Uber lineare substitutionen und bilineare Formen. *J. Reine Angew. Math.*, 84:1–63, 1877.

[2] I. Gohberg, P. Lancaster, and L. Rodman. *Matrices and Indefinite Scalar Products*, volume 8 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, Switzerland, 1983.

[3] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications*. Birkhäuser, Basel, Switzerland, 2005.

[4] Nicholas J. Higham, D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Computing the polar decomposition and the matrix sign decomposition in matrix groups. *SIAM J. Matrix Anal. Appl.*, 25(4):1178–1192, 2004.

[5] Nicholas J. Higham, D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Functions preserving matrix groups and iterations for the matrix square root. *SIAM J. Matrix Anal. Appl.*, 26(3):849–877, 2005.

[6] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[7] M. G. Krein. On a new class of hermitian forms. *Izv. Akad. Nauk. SSR, Ser. Fiz.-Mat.*, 9:1259–75, 1933.

[8] M. G. Krein. *The basic propositions in the theory of λ-zones of stability of a canonical system of linear differential equations with periodic coefficients (Russian)*. Izdat. Akad. Nauk. SSR, Moscow, 1955. English tranl., M. G. Krein: Topics in differential and integral equations and operator theory. Edited by I. Gohberg. Operator theory: Advances and Applications, Vol. 7, Birkhauser Verlag, Basel, 1983.

[9] L. Kronecker. Algebraishe Reduktion der Scharen bilinearer Formen. *Sitzungber. Akad. Berlin*, pages 763–776, 1890.

[10] Serge Lang. *Linear Algebra*. Springer-Verlag, New York, 3rd edition, 1987.

[11] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured factorizations in scalar product spaces. *SIAM J. Matrix Anal. Appl.*, 27(3):821–850, 2006.

[12] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured mapping problems for matrices associated with scalar products Part I: Lie and Jordan algebras. *SIAM J. Matrix Anal. Appl.*, 29(4):1389–1410, 2008.

[13] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured mapping problems for matrices associated with scalar with scalar products Part II: Automorphism groups. Technical report, MIMS EPrint, Manchester, England, 2009. In preparation.

[14] Ronald Shaw. *Linear Algebra and Group Representations. Vol. I*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1982.

[15] K. Weierstrass. Zur Theorie der quadratischen und bilinearen Formen. *Monatsber. Akad. Wiss., Berlin*, pages 310–338, 1868.

D. Steven Mackey
Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA
e-mail: `steve.mackey@wmich.edu`

Niloufer Mackey
Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA
e-mail: `nil.mackey@wmich.edu`

Françoise Tisseur
School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK
e-mail: `ftisseur@ma.man.ac.uk`