Factorization of matrices of quaternions

Terry A. Loring

University of New Mexico, Department of Mathematics and Statistics, Albuquerque, NM, 87131, USA

Abstract

We review known factorization results for quaternion matrices. Specifically, we derive the Jordan canonical form, polar decomposition, singular value decomposition, and QR factorization. We prove that there is a Schur factorization for commuting matrices, and from this derive the spectral theorem. We do not consider algorithms, but do point to some of the numerical literature.

Rather than work directly with matrices of quaternions, we work with complex matrices with a specific symmetry based on the dual operation. We discuss related results regarding complex matrices that are self-dual or symmetric, but perhaps not Hermitian.
© 2012 Elsevier GmbH. All rights reserved.

1. The quaternionic condition

It is possible to prove many factorization results for matrices of quaternions by deriving them from their familiar complex counterparts. The amount of additional work is surprisingly small.

There are more abstract factorization theorems that apply to real C^*-algebras, along the lines of the delightful paper [10] by Pedersen on factorization in (complex) C^*-algebras. We are dealing here with more basic questions, involving only finite-dimensional linear

---

E-mail address: loring@math.unm.edu.

algebra. These are never (or hardly ever?) addressed in basic linear algebra texts, creating the impression that linear algebra over the quaternions is more difficult than it is.

There are serious hazards within linear algebra over the quaternions. One can be lured to difficult questions of determinants and handedness of the spectrum, leaving perhaps victorious, but with the impression that all of linear algebra over the quaternions is going to be difficult.

We mathematicians are well-advised, when upon such uneven ground, to seek guidance from physicists. Such guidance helped select the topics here, and helped suggest notation.

One topic was included for reasons of elegance: the Jordan canonical form. For practical purposes, the Schur decomposition will generally suffice. We prove that as well.

It is assumed that the reader is familiar with such things as the spectral theorem and the functional calculus for normal matrices. It is not assumed that the reader knows about C^*-algebras or physics, but these topics are mentioned.

Let $\mathbb{H}$ denote the algebra of quaternions

$$\mathbb{H} = \left\{ a + b\hat{i} + c\hat{j} + d\hat{k} | a, b, c, d \in \mathbb{R} \right\}.$$

This is an algebra over $\mathbb{R}$. The canonical embedding $\mathbb{C} \hookrightarrow \mathbb{H}$ sending 1 to 1 and i to $\hat{i}$ does not make $\mathbb{H}$ into an algebra over $\mathbb{C}$. The trouble is that the embedding is not central. An additional algebraic operation is the involution

$$\left(a + b\hat{i} + c\hat{j} + d\hat{k} \right)^* = a - b\hat{i} - c\hat{j} - d\hat{k}$$

which satisfies several axioms, including the following:

$$\alpha \in \mathbb{R}, \quad x \in \mathbb{H} \Longrightarrow (\alpha x)^* = \alpha x^*$$
$$x^* x = 0 \Longrightarrow x = 0$$
$$(xy)^* = y^* x^*.$$

Turning to matrices, the algebra $\mathbf{M}_N(\mathbb{H})$ has the expected structure of a unital $\mathbb{R}$-algebra, plus the involution

$$[a_{ij}]^* = \left[a_{ji}^* \right]$$

of the conjugate-transpose. We would like to know that

$$A^* A = 0 \Longrightarrow A = 0$$

and

$$AB = I \Longrightarrow BA = I.$$

We will quickly prove these after we consider a representation of $\mathbf{M}_N(\mathbb{H})$ on $\mathbb{C}^{2N}$.

We have an obvious representation of $\mathbf{M}_N(\mathbb{H})$ on $\mathbb{H}^N$ and quickly notice that since left and right scalar multiplication of $\mathbb{H}^N$ disagree, we have both left eigenvalues $\lambda \in \mathbb{H}$ solving

$$A\mathbf{v} = \lambda \mathbf{v}$$

and right eigenvalues $\mu \in \mathbb{H}$ solving

$$A\mathbf{v} = \mathbf{v}\mu.$$

A glance at the survey [13] reveals that many difficulties arise.

Ponder the situation for real matrices, in $\mathbf{M}_N(\mathbb{R})$. To the chagrin of undergraduates, it is most natural to consider the representation of $\mathbf{M}_N(\mathbb{R})$ on $\mathbb{C}^N$. Given a real orthogonal matrix O we get the full picture of why it might not diagonalize in $\mathbf{M}_N(\mathbb{R})$ when we look at the complex eigenvalues. It is at this point that we are implicitly letting $\mathbf{M}_N(\mathbb{R})$ act on $\mathbb{C}^N$.

We do well in the case of the quaternions to regard $\mathbf{M}_N(\mathbb{H})$ as represented on $\mathbb{C}^{2N}$, or in more modern terms that de-emphasize the role of vectors, via a certain embedding

$$\chi : \mathbf{M}_N(\mathbb{H}) \to \mathbf{M}_{2N}(\mathbb{C}).$$

This is a very old trick. For $N > 1$ it appears to have been noticed first by Lee [6].

Definition 1.1. Given two *complex N-by-N* matrices A and B we set

$$\chi\left(A + B\hat{j}\right) = \begin{bmatrix} A & B \\ -\overline{B} & \overline{A} \end{bmatrix}.$$

We generally study complex matrices in the image of χ and only at the end of our calculations do we draw conclusions about $\mathbf{M}_N(\mathbb{H})$. This is in keeping with applications in quantum mechanics, where Hilbert space is always complex and time-reversal symmetry will often be incorporated in the conjugate linear operation $\mathcal{T}$. We define $\mathcal{T} : \mathbb{C}^{2N} \to \mathbb{C}^{2N}$ by

$$\mathcal{T}\left(\begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix}\right) = \begin{bmatrix} -\overline{\mathbf{w}} \\ \overline{\mathbf{v}} \end{bmatrix} \tag{1}$$

for $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{C}^N$. See [8, Section 2.2] for details on when $\mathcal{T}$ is relevant to time reversal symmetry.

A less generous description of our approach is that we plan to study complex matrices with a certain symmetry that is useful in physics, and then sell the same results to pure mathematicians by re-branding them as theorems about matrices of quaternions.

The operator $\mathcal{T}$ is relatively well behaved, despite being only conjugate linear. It preserves orthogonality, and indeed

$$\langle \mathcal{T}\xi, \mathcal{T}\eta \rangle = \overline{\langle \xi, \eta \rangle}. \tag{2}$$

Also

$$\xi \perp \mathcal{T}\xi \tag{3}$$

which is another one-line calculation.

Lemma 1.2. *The mapping χ is well-defined, is an $\mathbb{R}$-algebra homomorphism, is one-to-one and satisfies*

$$(\chi(Y))^* = \chi\left(Y^*\right).$$

Proof. Notice that every quaternion q can be written as $\alpha + \beta \hat{j}$ with α and β in $\mathbb{C}$. This makes the map well-defined. It is clearly one-to-one and $\mathbb{R}$-linear. Notice that $\beta \hat{j} = \hat{j} \bar{\beta}$ for complex number β and so $B \hat{j} = \hat{j} \bar{B}$. Therefore

$$
\chi \left(A + B \hat{j} \right) \chi \left(C + D \hat{j} \right) = \begin{bmatrix} A & B \\ -\bar{B} & \bar{A} \end{bmatrix} \begin{bmatrix} C & D \\ -\bar{D} & \bar{C} \end{bmatrix}
$$

$$
= \begin{bmatrix} AC - B\bar{D} & AD + B\bar{C} \\ -\bar{B}C - \bar{A}D & -\bar{B}D + \bar{A}C \end{bmatrix}
$$

$$
= \chi \left((AC - B\bar{D}) + (AD + B\bar{C}) \hat{j} \right)
$$

$$
= \chi \left(\left(A + B \hat{j} \right) \left(C + D \hat{j} \right) \right)
$$

and

$$
\left(\chi \left(A + B \hat{j} \right) \right)^* = \begin{bmatrix} A & B \\ -\bar{B} & \bar{A} \end{bmatrix}^* = \chi \left(A^* - B^{\mathrm{T}} \hat{j} \right) = \chi \left(\left(A + B \hat{j} \right)^* \right). \quad \square
$$

We can describe the image of χ in several useful ways. We will need several unary operations on complex matrices. For starters we need the transpose A^{T} and the pointwise conjugate $\bar{A}$ and the conjugate-transpose, or adjoint, $A^* = \bar{A}^{\mathrm{T}} = \overline{A^{\mathrm{T}}}$. Finally, we need a twisted transpose that is useful in physics. This is a "generalized involution" $X^\sharp$ called the *dual operation* that is defined only for X in $\mathbf{M}_{2N}(\mathbb{C})$.

Definition 1.3. For A, B, C and D all complex N-by-N matrices, define

$$
\begin{bmatrix} A & B \\ C & D \end{bmatrix}^\sharp = \begin{bmatrix} D^{\mathrm{T}} & -B^{\mathrm{T}} \\ -C^{\mathrm{T}} & A^{\mathrm{T}} \end{bmatrix}.
$$

Alternatively, we set $X^\sharp = -Z X^{\mathrm{T}} Z$ where

$$
Z = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{4}
$$

Notice that Z could be replaced by a matrix with similar properties and define a variation on the dual operation. Indeed, the choice of Z is not standardized.

Lemma 1.4. *For X in $\mathbf{M}_{2N}(\mathbb{C})$ the following are equivalent:*

1. *X is in the image of χ;*
2. *$X^* = X^\sharp$;*
3. *$X = -\mathcal{T} \circ X \circ \mathcal{T}$ meaning $X\xi = -\mathcal{T}(X\mathcal{T}(\xi))$ for every vector ξ in $\mathbb{C}^{2N}$.*

Proof. Assume

$$
X = \begin{bmatrix} A & B \\ -\bar{B} & \bar{A} \end{bmatrix}.
$$

Then

$$
X^* = \begin{bmatrix} \bar{A}^{\mathrm{T}} & -B^{\mathrm{T}} \\ \bar{B}^{\mathrm{T}} & A^{\mathrm{T}} \end{bmatrix} = X^\sharp.
$$

Conversely $X^* = X^\sharp$ translates into block form as

$$\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} = \begin{bmatrix} D^{\mathrm{T}} & -B^{\mathrm{T}} \\ -C^{\mathrm{T}} & A^{\mathrm{T}} \end{bmatrix}$$

so we have proven (1) $\Longleftrightarrow$ (2).

We compute $-\mathcal{T} \circ X \circ \mathcal{T}$ for X again in block form, and find

$$\begin{aligned}
-\mathcal{T} \circ X \circ \mathcal{T} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} &= -\mathcal{T} \left(\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} -\overline{\mathbf{w}} \\ \overline{\mathbf{v}} \end{bmatrix} \right) \\
&= -\mathcal{T} \left(\begin{bmatrix} -A\overline{\mathbf{w}} + B\overline{\mathbf{v}} \\ -C\overline{\mathbf{w}} + D\overline{\mathbf{v}} \end{bmatrix} \right) \\
&= \begin{bmatrix} -\overline{C}\mathbf{w} + \overline{D}\mathbf{v} \\ \overline{A}\mathbf{w} - \overline{B}\mathbf{v} \end{bmatrix} \\
&= \begin{bmatrix} \overline{D} & -\overline{C} \\ -\overline{B} & \overline{A} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix}
\end{aligned}$$

so the matrix for the linear operator $-\mathcal{T} \circ X \circ \mathcal{T}$ is $\left(X^\sharp \right)^*$. Therefore (2) $\Longleftrightarrow$ (3). $\quad\square$

Definition 1.5. Let us call

$$X^* = X^\sharp$$

the *quaternionic condition* and the matrix X we will call a *quaternionic matrix*.

We now dispose of the implications

$$A^*A = 0 \Longrightarrow A = 0$$

and

$$AB = I \Longrightarrow BA = I$$

for matrices of quaternions. These are true implications for complex matrices, and in particular for quaternionic matrices, and therefore true for matrices of quaternions.

We pause to note some axioms of the dual operation. It behaves a lot like the transpose. It is linear:

$$(X + \alpha Y)^\sharp = X + \alpha Y^\sharp$$

even for complex α. Here X and Y are any $2N$-by-$2N$ complex matrices. The dual reverses multiplication:

$$(XY)^\sharp = Y^\sharp X^\sharp.$$

It undoes itself:

$$\left(X^\sharp \right)^\sharp = X$$

and commutes with the adjoint:

$$\left(X^\sharp \right)^* = \left(X^* \right)^\sharp.$$

Lemma 1.6. *Every matrix G in* $\mathbf{M}_{2N}(\mathbb{C})$ *can we expressed in a unique way as*

$$G = X + iY$$

with X and Y being quaternionic matrices.

Proof. Given G we set

$$X = \frac{1}{2}G^{\sharp *} + \frac{1}{2}G$$

and

$$Y = \frac{i}{2}G^{\sharp *} - \frac{i}{2}G. \quad \square$$

2. Kramers degeneracy and Schur factorization

Eigenvalue doubling is a key feature of self-dual, self-adjoint matrices. In physics this is called the Kramers degeneracy theorem, or the theory of Kramers pairs [5,8,12]. This generalizes in two ways, to give eigenvalue doubling given the symmetry $X = X^{\sharp}$ and conjugate-pairing of eigenvalues given the symmetry $X^* = X^{\sharp}$.

Such a collection of paired eigenvectors will, in good situations, form a unitary matrix. If U is unitary matrix that satisfies the quaternionic condition then it satisfies another symmetry making it symplectic, specifically $U^{\mathsf{T}}ZU = Z$.

Lemma 2.1. *Suppose U is a unitary* 2N*-by-*2N *matrix. The following are equivalent:*

1. *U is symplectic;*
2. $U^* = U^{\sharp}$;
3. $Z\overline{U} = UZ$;
4. $U \circ \mathcal{T} = \mathcal{T} \circ U$;
5. *if* **v** *is column j of U for* $j \leq N$ *then column* $N + j$ *of U is* $\mathcal{T}(\mathbf{v})$.

Proof. This follows easily from Lemma 1.4. $\square$

We need to know that the group of symplectic unitary matrices acts transitively on $\mathbb{C}^n$.

Lemma 2.2. *If* **v** *is any unit vector in* $\mathbb{C}^{2N}$ *then there is a symplectic unitary with* $U\mathbf{e}_1 = \mathbf{v}$.

Proof. Let $\mathbf{v}_1 = \mathbf{v}$. We use (2) and (3) to select in order vectors that are an orthonormal basis for $\mathbb{C}^{2N}$, but of the form

$$\mathbf{v}_1, \mathcal{T}\mathbf{v}_1, \mathbf{v}_2, \mathcal{T}\mathbf{v}_2, \ldots, \mathbf{v}_N, \mathcal{T}\mathbf{v}_N.$$

If we reorder to

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N, \mathcal{T}\mathbf{v}_1, \mathcal{T}\mathbf{v}_2, \ldots, \mathcal{T}\mathbf{v}_N$$

we have the columns of the desired symplectic unitary. $\square$

The most basic result in this realm is an ugly lemma that says that $\mathcal{T}$ maps left eigenvectors of X to right eigenvectors of $X^{\sharp}$, with the same eigenvalue. The second part

of this lemma is more elegant, if less general. It specifies how every quaternionic matrix has a conjugate symmetry in its spectral decomposition.

Lemma 2.3. *Suppose X is in $\mathbf{M}_{2N}(\mathbb{C})$.*

1. *If $X\xi = \lambda\xi$ then*

$$(\mathcal{T}\xi)^* X^\sharp = \lambda \, (\mathcal{T}\xi)^*.$$

2. *If $X^* = X^\sharp$ and $X\xi = \lambda\xi$ then*

$$X(\mathcal{T}\xi) = \overline{\lambda}(\mathcal{T}\xi).$$

Proof. (1) Starting with

$$X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

and

$$\xi = \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix}$$

we find that $X\xi = \lambda\xi$ translates to

$$A\mathbf{v} + B\mathbf{w} = \lambda\mathbf{v}$$
$$C\mathbf{v} + D\mathbf{w} = \lambda\mathbf{w}$$

and $(\mathcal{T}\mathbf{v})^* X^\sharp = \lambda (\mathcal{T}\mathbf{v})^*$ translates to

$$-\mathbf{w}^T D^T - \mathbf{v}^T C^T = -\lambda\mathbf{w}^T$$
$$\mathbf{w}^T B^T + \mathbf{v}^T A^T = \lambda\mathbf{v}^T$$

so these are equivalent conditions.

(2) follows from (1) by taking adjoints. $\square$

Now we are able to extend Kramers degeneracy to a variety of situations, starting with a block diagonalization for commuting quaternionic matrices. Part (2) of Theorem 2.4 appeared in [11].

Theorem 2.4. *Suppose $X_1, \ldots, X_k$ in $\mathbf{M}_{2N}(\mathbb{C})$ commute pairwise and $X_j^* = X_j^\sharp$ for all j.*

1. *There is a single symplectic unitary U such that, for all j,*

$$U^* X_j U = \begin{bmatrix} T_j & S_j \\ -\overline{S_j} & \overline{T_j} \end{bmatrix}$$

with T_j upper-triangular and S_j strictly upper-triangular.

2. *If, in addition, the X_j are normal then there is a single symplectic unitary U such that, for all j,*

$$U^* X_j U = \begin{bmatrix} D_j & 0 \\ 0 & \overline{D_j} \end{bmatrix}$$

with D_j diagonal.

3. *Every symplectic unitary has determinant 1.*

Proof. (1) A finite set of commuting matrices will have a common eigenvector, so let $\mathbf{v}$ be a unit vector such that $X_j\mathbf{v} = \lambda_j\mathbf{v}$. We know then that $\mathcal{T}\mathbf{v}$ is an eigenvalue for X_j with eigenvector $\overline{\lambda_j}$. There is a symplectic unitary U_1 such that $U_1\mathbf{e}_1 = \mathbf{v}$, and then $U_1\mathbf{e}_{N+1} = \mathcal{T}\mathbf{v}$ by Lemma 2.1. Let $Y_j = U_1^*X_jU_1$. Then

$$Y_j\mathbf{e}_1 = \lambda_j\mathbf{e}_1$$

and

$$Y_j\mathbf{e}_{N+1} = \overline{\lambda_j}\mathbf{e}_{N+1}.$$

This means that column 1 and column $N + 1$ are all but zeroed-out,

$$Y_j = \begin{bmatrix} \lambda_j & * & 0 & * \\ 0 & A_j & 0 & C_j \\ 0 & * & \overline{\lambda_j} & * \\ 0 & B_j & 0 & D_j \end{bmatrix}.$$

Up to a non-symplectic change of basis we are looking at block-upper-triangular matrices that commute, so the lower-right corners in that basis commute. This means that the

$$Z_j = \begin{bmatrix} A_j & C_j \\ B_j & D_j \end{bmatrix} \tag{5}$$

all commute, and each must satisfy $Z_j^* = Z_j^\sharp$. By induction, we have proven the first claim.

For (2) we modify the proof just a little. Starting with the X_j normal, we find the Y_j are also normal and so

$$Y_j = \begin{bmatrix} \lambda_j & 0 & 0 & 0 \\ 0 & A_j & 0 & C_j \\ 0 & 0 & \overline{\lambda_j} & 0 \\ 0 & B_j & 0 & D_j \end{bmatrix}.$$

This, after an appropriate basis change, would be block-diagonal, and from this we can conclude that the matrix in (5) is normal. The induction proceeds as before, with the stronger conclusion that $B_j = C_j = 0$ and $A_j = \overline{D_j}$ is diagonal.

(3) Applying (2) we find

$$W = U \begin{bmatrix} D & 0 \\ 0 & \overline{D} \end{bmatrix} U^*$$

where D is a diagonal unitary. Therefore

$$\det(W) = \det(D)\det(\overline{D}) = \det(D)\overline{\det(D)} \geq 0$$

and since a unitary has determinant on the unit circle, we are done. $\square$

Corollary 2.5. *Every matrix in $\mathbf{M}_N(\mathbb{H})$ is unitarily equivalent to an upper-triangular matrix in $\mathbf{M}_N(\mathbb{H})$ that has complex numbers on the diagonal.*

There is an algorithm [1] for the Schur decomposition of quaternionic matrices.

Corollary 2.6. *Every normal matrix in* $\mathbf{M}_N(\mathbb{H})$ *is unitarily equivalent to a diagonal matrix in* $\mathbf{M}_N(\mathbb{C})$. *Every Hermitian matrix in* $\mathbf{M}_N(\mathbb{H})$ *is unitarily equivalent to a diagonal matrix in* $\mathbf{M}_N(\mathbb{R})$.

Corollary 2.7. *Every Hermitian self-dual matrix X in* $\mathbf{M}_{2N}(\mathbb{C})$ *is of the form*

$$X = U \begin{bmatrix} D_j & 0 \\ 0 & D_j \end{bmatrix} U^*$$

for some symplectic unitary U and a diagonal real matrix D_j.

These corollaries cause us to reconsider the concept of left and right eigenvalues in $\mathbb{H}$ and focus on just those that are in $\mathbb{C}$. For details on left eigenvalues and non-complex right eigenvalues, consult [3]. We put the complex right eigenvalues in a simple context with the following.

Lemma 2.8. *Suppose* $\mathbf{v}, \mathbf{w} \in \mathbb{C}^N$ *and* $\lambda \in \mathbb{C}$ *and* $A, B \in \mathbf{M}_N(\mathbb{C})$. *Then*

$$\begin{bmatrix} A & B \\ -\overline{B} & \overline{A} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix}$$

if and only if

$$\left(A + B\hat{j} \right) \left(\mathbf{v} - \hat{j}\mathbf{w} \right) = \left(\mathbf{v} - \hat{j}\mathbf{w} \right) \lambda.$$

Therefore $\lambda \in \mathbb{C}$ is a right eigenvalue of $X \in \mathbf{M}_N(\mathbb{H})$ if and only if λ is an eigenvalue of $\chi(X)$.

Proof. This a short, direct calculation. $\quad\square$

3. Jordan canonical form

Kramers degeneracy extends to the generalized eigenvectors used to find the Jordan canonical form.

Lemma 3.1. *Suppose* $X^\sharp = X^*$. *If*

$$(X - \lambda I)^r \mathbf{v} = 0 \quad and \quad (X - \lambda I)^{r-1} \mathbf{v} \neq 0$$

then

$$\left(X - \overline{\lambda} I \right)^r \mathcal{T}\mathbf{v} = 0 \quad and \quad \left(X - \overline{\lambda} I \right)^{r-1} \mathcal{T}\mathbf{v} \neq 0.$$

Proof. For any Y we have

$$\|Y\mathbf{v}\| = \|\overline{Y}\overline{\mathbf{v}}\| = \|Z\overline{Y}ZZ\overline{\mathbf{v}}\|$$

proving

$$\|Y\mathbf{v}\| = \|Y^{\sharp*}\mathcal{T}\mathbf{v}\|.$$

Since

$$\left((X - \lambda I)^k \right)^{\sharp*} = \left(X - \overline{\lambda} I \right)^k$$

the result follows. $\quad\square$

We see the general idea of a proof [14] of the Jordan decomposition for a quaternionic matrix. When we build a Jordan basis we need to respect $\mathcal{T}$ in two ways. Whatever basis we pick for the subspace corresponding to λ with positive imaginary part, we apply $\mathcal{T}$ to get the basis for the subspace corresponding to $\bar{\lambda}$. When λ is real we need to pick generalized eigenvectors in pairs.

The Jordan form of a quaternionic matrix is not elegant, as each Jordan block larger than 2-by-2 gets spread into all four quadrants of the matrix. We work directly with a Jordan basis and then give our final conclusion in terms of quaternions.

Theorem 3.2. *Suppose $X^\sharp = X^*$. There is a Jordan basis for X consisting of pairs of the form $\mathbf{v}, \mathcal{T}\mathbf{v}$.*

Proof. Let N_λ denote the subspace of all generalized eigenvectors for λ together with the zero vector. Recall that just treating X as a complex matrix, the procedure for selecting a Jordan basis involves selecting, for each λ, a basis for N_λ with the following property: whenever $\mathbf{b}$ is in this basis, then $(X - \lambda I)\mathbf{b}$ is either zero or back in this basis.

For λ in the spectrum with positive imaginary part we make such a choice and then apply $\mathcal{T}$ to get a set of vectors in $N_{\bar{\lambda}}$ that has the correct number of elements to be a basis of $N_{\bar{\lambda}}$. It will also be a linearly independent set since Z and conjugation both preserve linear independence. Since

$$(X - \bar{\lambda} I)\mathcal{T}\mathbf{b} = \mathcal{T}(X - \lambda I)\mathbf{b}$$

this basis of $N_{\bar{\lambda}}$ has the desired property.

For λ in the spectrum that is real we need to modify the procedure for selecting the basis of N_λ. A common procedure selects a basis $\mathbf{b}_{r,1}, \ldots, \mathbf{b}_{r,m_r}$ for

$$\ker(X - \lambda)^r \cap \left(\ker(X - \lambda)^{r-1} \right)^\perp \cap (\mathrm{im}(X - \lambda))^\perp \tag{6}$$

and constructs for N_λ the Jordan basis

$$\left\{ (X - \lambda)^j \, \mathbf{b}_{r,k} \mid 1 \le r \le r_{\max}, \ 0 \le j \le r - 1, \ k = 1, \ldots, m_r \right\}$$

Since

$$\mathcal{T}(X - \lambda)^j \, \mathbf{b}_{r,k} = (X - \lambda)^j \, \mathcal{T}\mathbf{b}_{r,k}$$

we get the desired structure if the subspaces (6) are $\mathcal{T}$-invariant. This follows from the next two lemmas and the equality

$$(\mathrm{im}(X - \lambda))^\perp = \ker\left(X^* - \bar{\lambda}\right).$$

We can now assemble the bases of the various N_λ to get a Jordan basis built from the Kramers pairs. □

Lemma 3.3. *If $X^\sharp = X^*$ then $\ker(X)$ is $\mathcal{T}$-invariant.*

Proof. This is a restatement of the $\lambda = 0$ case of Lemma 2.3(2). □

Lemma 3.4. *If a subspace $\mathcal{H}_0$ is $\mathcal{T}$-invariant then $(\mathcal{H}_0)^\perp$ is also $\mathcal{T}$-invariant.*

Proof. If $\mathbf{v}$ is in $(\mathcal{H}_0)^{\perp}$ then for every $\mathbf{w} \in \mathcal{H}_0$ we have

$$\langle \mathcal{T}\mathbf{v}, \mathbf{w} \rangle = - \langle \mathcal{T}\mathbf{v}, \mathcal{T}\mathcal{T}\mathbf{w} \rangle = -\overline{\langle \mathbf{v}, \mathcal{T}\mathbf{w} \rangle} = 0. \quad \square$$

Corollary 3.5. *Suppose* $X \in \mathbf{M}_n(\mathbb{H})$. *There is an invertible matrix* $S \in \mathbf{M}_n(\mathbb{H})$ *and a complex matrix* J *in Jordan form such that* $X = S^{-1} J S$.

4. Norms

There are two operator norms to consider on X in $\mathbf{M}_N(\mathbb{H})$: that induced by quaternionic Hilbert space and that induced by complex Hilbert space on $\chi(X)$. They end up identical.

Theorem 4.1. *Suppose* X *is in* $\mathbf{M}_N(\mathbb{H})$. *Then using the norms*

$$\|\mathbf{v}\| = \left(\sum_{j=1}^{N} v_j^* v_j \right)^{\frac{1}{2}}$$

on $\mathbb{H}^N$ *and*

$$\|\mathbf{w}\| = \left(\sum_{j=1}^{2N} \overline{w_j} w_j \right)^{\frac{1}{2}}$$

on $\mathbb{C}^{2N}$ *we have*

$$\sup_{\mathbf{v} \neq 0} \frac{\|X\mathbf{v}\|}{\|\mathbf{v}\|} = \sup_{\mathbf{w} \neq 0} \frac{\|\chi(X)\mathbf{w}\|}{\|\mathbf{w}\|}.$$

Proof. Utilizing also the norm on $\mathbb{C}^{2N}$ we calculate the four relevant norms:

$$\left\| \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} \right\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.$$

$$\left\| \mathbf{v} - \hat{j}\mathbf{w} \right\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.$$

$$\left\| \begin{bmatrix} A & B \\ -\overline{B} & \overline{A} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} \right\|^2 = \|A\mathbf{v} + B\mathbf{w}\|^2 + \left\| \overline{A}\mathbf{w} - \overline{B}\mathbf{v} \right\|^2.$$

$$\left\| \left(A + B\hat{j} \right) \left(\mathbf{v} - \hat{j}\mathbf{w} \right) \right\|^2 = \left\| A\mathbf{v} + B\mathbf{w} + \hat{j} \left(\overline{B}\mathbf{v} - \overline{A}\mathbf{w} \right) \right\|^2$$
$$= \|A\mathbf{v} + B\mathbf{w}\|^2 + \left\| \overline{B}\mathbf{v} - \overline{A}\mathbf{w} \right\|^2.$$

The result follows. $\quad \square$

5. Singular value decomposition

As in the complex case, a slick way to prove that there is a singular value decomposition is to work out the polar decomposition and then use the spectral theorem on the positive part.

In [13] it is stated that "a little more work is needed for the singular case" when discussing the polar decomposition. The extra work involves padding out a quaternionic partial isometry to be a quaternionic unitary (and so a symplectic unitary).

We remind the reader that U is a partial isometry when $(U^*U)^2 = U^*U$, or equivalently $UU^*U = U$ or $U^*UU^* = U^*$ or $(UU^*)^2 = UU^*$. If we restrict the domain and range of U we find it is an isometry from $(\ker(U))^\perp$ to $(\ker(U^*))^\perp$.

Lemma 5.1. *Suppose $U^* = U^\sharp$ and U is a partial isometry in $\mathbf{M}_{2N}(\mathbb{C})$. There is a symplectic unitary W in $\mathbf{M}_N(\mathbb{C})$ such that $W\xi = U\xi$ for all $\xi \perp \ker(U)$.*

Proof. If $\mathbf{v}$ is in $\ker(U)$ then by Lemma 2.3, $\mathcal{T}\mathbf{v}$ is also in $\ker(U)$. Since $\mathbf{v}$ and $\mathcal{T}\mathbf{v}$ are orthogonal, we can show that $\ker(U)$ has even dimension $2m$ and that is has a basis for the form

$$\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathcal{T}\mathbf{v}_1, \ldots, \mathcal{T}\mathbf{v}_m.$$

We are working in finite dimensions so the dimension of $\ker(U^*)$ is also $2m$ and we select for it a basis

$$\mathbf{w}_1, \ldots, \mathbf{w}_m, \mathcal{T}\mathbf{w}_1, \ldots, \mathcal{T}\mathbf{w}_m.$$

We can define W to agree with U on $(\ker(U))^\perp$ and to send $\mathbf{v}_j$ to $\mathbf{w}_j$ and $\mathcal{T}\mathbf{v}_j$ to $\mathcal{T}\mathbf{w}_j$ and so get a unitary that commutes with $\mathcal{T}$, which means that it is symplectic. $\square$

Lemma 5.2. *Suppose $X^* = X^\sharp$ in $\mathbf{M}_{2N}(\mathbb{C})$. Then there is a unitary U and a positive semidefinite P with $U^* = U^\sharp$ and $P^* = P^\sharp$ and $X = UP$.*

Proof. Let f be a continuous function on the positive reals with $f(0) = 0$ and $f(\lambda) = \lambda^{-\frac{1}{2}}$ for every nonzero eigenvalue of X^*X. Then let

$$W = Xf(X^*X).$$

The usual calculations in functional calculus tell us W is a partial isometry and that $X = WP$ for $P = (X^*X)^{\frac{1}{2}}$. Working with monomials and polynomials and then taking limits, we can show that

$$(f(Y))^\sharp = f\left(Y^\sharp\right)$$

for any positive operator Y and so

$$W^\sharp = f\left(X^\sharp X^{*\sharp}\right) X^\sharp = f\left(X^*X\right) X^* = W^*.$$

Also $P^\sharp = P = P^*$.

We just showed that the matrices in the minimal polar decomposition are quaternionic. We use Lemma 5.1 to finish the argument. $\square$

As expected, the polar decomposition leads to a singular value decomposition.

Theorem 5.3. *Suppose $X^* = X^\sharp$ in $\mathbf{M}_{2N}(\mathbb{C})$. There are symplectic unitary matrices U and V and a diagonal matrix D with nonnegative real entries and $D^\sharp = D^*$ such that $X = UDV$.*

Proof. We take a quaternionic polar decomposition $X = WP$. Since P is positive, we apply Theorem 2.4 to get symplectic unitary matrices Q and V and a diagonal matrix D such that $P = QDV$. The eigenvalues of P are nonnegative, so the same is true for the diagonal elements of D and we have the needed factorization $X = (WQ)DV$. $\square$

6. QR factorization

It is easy to use Lemma 2.2 to get over $\mathbb{H}$ a QR factorization theorem. Notice that upper-triangular matrices are sent by χ to matrices that are block-upper-triangular.

Theorem 6.1. (1) If $X^* = X^\sharp$ in $\mathbf{M}_{2N}(\mathbb{C})$ then there are symplectic unitary matrices Q and R of the form

$$R = \begin{bmatrix} A & B \\ -\overline{B} & \overline{A} \end{bmatrix}$$

with A and B upper-triangular. (2) If $X \in \mathbf{M}_N(\mathbb{H})$ then there is a unitary Q and upper-triangular matrix R in $\mathbf{M}_N(\mathbb{H})$ such that $X = QR$.

Proof. (2) follows directly from (1), so we prove (1).

We apply Lemma 2.2 to the first column of X and find $X = Q_1 R_1$ where

$$R_1 = \begin{bmatrix} A_1 & B_1 \\ -\overline{B_1} & \overline{A_1} \end{bmatrix}$$

where A_1 and B_1 have zeros in their first columns, except perhaps in the top position. As we did earlier, we can proceed with a proof by induction. $\square$

7. Self-dual matrices

If we study matrices with $X^\sharp = X$ then we are no longer working directly with quaternionic matrices, but as we discuss below, there is a connection. We discuss a Schur factorization and a structured polar decomposition for self-dual matrices. The latter is a bit tricky, so we warm up with a structured polar decomposition for symmetric complex matrices.

For dealing with a single self-dual matrix, there is the efficient Paige/Van Loan algorithm [4,9] to implement the following theorem.

Theorem 7.1. Given $X_1, \ldots, X_k$ in $\mathbf{M}_{2N}(\mathbb{C})$ that commute pairwise and are self-dual, there is a single symplectic unitary U such that, for all j,

$$U^* X_j U = \begin{bmatrix} T_j & C_j \\ 0 & T_j^{\mathrm{T}} \end{bmatrix}$$

with T_j upper-triangular and the C_j skew-symmetric.

Proof. Let $\mathbf{v}$ be a nonzero unit vector such that $X_j \mathbf{v} = \lambda_j \mathbf{v}$ for all j. By Lemma 2.3,

$$X_j^* (\mathcal{T}\mathbf{v}) = \overline{\lambda_j} (\mathcal{T}\mathbf{v}).$$

There is a symplectic unitary U_1 such that $U_1\mathbf{e}_1 = \mathbf{v}$ and $U_1\mathbf{e}_{N+1} = \mathcal{T}\mathbf{v}$. Let $Y_j = U_1^* X_j U_1$. Then

$$Y_j\mathbf{e}_1 = \lambda_j\mathbf{e}_1$$

and

$$Y_j^*\mathbf{e}_{N+1} = \overline{\lambda_j}\mathbf{e}_{N+1}.$$

Since $\mathbf{e}_{N+1}$ is real, we take the adjoint and discover that

$$\mathbf{e}_{N+1}^{\mathrm{T}} Y_j = \lambda_j\mathbf{e}_{N+1}^{\mathrm{T}}.$$

This means the column 1 and row $N + 1$ are all but zeroed-out,

$$Y_j = \begin{bmatrix} \lambda_j & * & * & * \\ 0 & A_j & * & C_j \\ 0 & 0 & \lambda_j & 0 \\ 0 & B_j & * & D_j \end{bmatrix}.$$

Basic facts about block-triangular matrices show that the

$$Z_j = \begin{bmatrix} A_j & C_j \\ B_j & D_j \end{bmatrix}$$

are a commuting family of matrices, and since U_1 was chosen to be symplectic, the Z_j will be self-dual. A simple induction now finishes the proof. $\square$

A promising numerical technique for the joint diagonalization of two commuting self-dual self-adjoint matrices H and K would be to form the normal self-dual matrix $X = A + iB$ and apply Paige/Van Loan to reduce to block-diagonal form. Then apply ordinary Schur decomposition. This technique was used in [4, Section 9] to diagonalize matrices that were exactly self-dual and approximately unitary. This idea is mentioned in [2], Section 6.5.

It is not hard to show that the minimal polar decomposition, the one that is unique and can involve a partial isometry, preserves in some way just about any symmetry thrown at it. This is because the functional calculus interacts well with the dual operation [7], as well as with the transpose. It is a bit harder to figure out what happens for the maximal polar decomposition. By the minimal polar decomposition is meant the factorization that is unique and can involve a partial isometry. By the maximal polar decomposition is meant the factorization that involves a unitary.

We begin with the easier result about the polar decomposition of complex symmetric matrices.

Theorem 7.2. *If X in $\mathbf{M}_{2n}(\mathbb{C})$ satisfies $X^{\mathrm{T}} = X$ then there is a unitary U such that $U^{\mathrm{T}} = U$ and $X = U\,|X|$.*

Proof. Again choose f with $f(0) = 0$ and $f(\lambda) = \lambda^{-\frac{1}{2}}$ for every nonzero eigenvalue of X^*X and let

$$W = Xf(X^*X).$$

As always, W is a partial isometry and $X = WP$ for $P = (X^*X)^{\frac{1}{2}} = |X|$. Now we discover that

$$W^{\mathrm{T}} = f\left(X^{\mathrm{T}}\overline{X}\right)X^{\mathrm{T}} = f\left(XX^*\right)X = Xf\left(X^*X\right) = W.$$

To create a unitary U with $X = U|X|$ we must extend W to map $\ker(W)$ to $\ker(W^*)$. We can arrange $U^{\mathrm{T}} = U$ as follows. Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be an orthonormal basis of $\ker(W)$. Then $\overline{\mathbf{v}_1}, \ldots, \overline{\mathbf{v}_m}$ will be an orthonormal basis of $\ker(W^*) = \ker(\overline{W})$ and we define V to be zero on $(\ker(W))^{\perp}$ and $V\mathbf{v}_j = \overline{\mathbf{v}_j}$. Thus V^* will be zero on $(\ker(W))^{\perp}$ and $V^*\overline{\mathbf{v}_j} = \mathbf{v}_j$, but the same can be said about $\overline{V}$. That means that $V^{\mathrm{T}} = V$ and so $U = W + V$ will be the required symmetric unitary.

Notice there is no structure on $|X|$, but considering it with $|X^*|$ we get the formula

$$\left|X^*\right| = |X|^{\mathrm{T}}. \quad \square \tag{7}$$

For the self-dual situation, we shall see that a similar construction works as long as we respect Kramers degeneracy.

Proposition 7.3. *If a partial isometry W in $\mathbf{M}_{2N}(\mathbb{C})$ is self-dual, then the initial space of W will have even dimension and $\mathcal{T}$ will map the initial space isometrically onto the final space of W.*

Proof. Applying Lemma 2.3 to the self-adjoint matrix W^*W we find that

$$W^*W\mathbf{v} = \mathbf{v} \implies \left(W^*W\right)^{\sharp}\mathcal{T}\mathbf{v} = \mathcal{T}\mathbf{v}$$
$$\implies \left(WW^*\right)\mathcal{T}\mathbf{v} = \mathcal{T}\mathbf{v}.$$

That is, when $\mathbf{v} \in (\ker W)^{\perp}$ we have $\mathcal{T}\mathbf{v} \in (\ker W^*)^{\perp}$ and $\|W^*\mathcal{T}\mathbf{v}\| = \|\mathbf{v}\|$. This is useful because

$$\begin{aligned}
\langle \mathbf{v}, W^*\mathcal{T}\mathbf{v}\rangle &= \langle \mathbf{v}, -W^*Z\overline{\mathbf{v}}\rangle \\
&= \left\langle -W^{\mathrm{T}}Z\mathbf{v}, \overline{\mathbf{v}}\right\rangle \\
&= \langle -ZW\mathbf{v}, \overline{\mathbf{v}}\rangle \\
&= \langle \mathbf{v}, W^*Z\overline{\mathbf{v}}\rangle \\
&= -\langle \mathbf{v}, W^*\mathcal{T}\mathbf{v}\rangle
\end{aligned}$$

which means that $\mathbf{v}$ and $W^*\mathcal{T}\mathbf{v}$ are orthogonal, and

$$\begin{aligned}
W^*\mathcal{T}W^*\mathcal{T}\mathbf{v} &= W^*Z\overline{W^*Z\overline{\mathbf{v}}} \\
&= W^*ZW^{\mathrm{T}}Z\mathbf{v} \\
&= -W^*W^{\sharp}\mathbf{v} \\
&= -W^*W\mathbf{v} \\
&= -\mathbf{v}.
\end{aligned}$$

If we start with a unit vector $\mathbf{v}$ in $(\ker W)^{\perp}$ then we end up with an orthogonal pair of unit vectors $\mathbf{v}$ and $\mathbf{w}$ with $W^*\mathcal{T}\mathbf{v} = \mathbf{w}$ and $W^*\mathcal{T}\mathbf{w} = -\mathbf{v}$.

If $\mathbf{q}$ is a vector in $(\ker W)^{\perp}$ that is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$ then $W^{*}\mathcal{T}\mathbf{q}$ will be orthogonal to both $W^{*}\mathcal{T}\mathbf{v}$ and $W^{*}\mathcal{T}\mathbf{w}$ since $\mathcal{T}$ preserves orthogonality everywhere and W^{*} preserves the orthogonality of vectors in $(\ker W^{*})^{\perp}$. Thus we can create a basis of $(\ker W)^{\perp}$ out of pairs

$$\mathbf{v}_1, W^{*}\mathcal{T}\mathbf{v}_1, \ldots, \mathbf{v}_m, W^{*}\mathcal{T}\mathbf{v}_m. \quad \square$$

We need some examples of self-dual partial isometries. Treating vectors as $2N$-by-1 matrices, if we set

$$V = (\mathcal{T}\mathbf{v})\,\mathbf{w}^{*} - (\mathcal{T}\mathbf{w})\,\mathbf{v}^{*}$$

then this rank 2 (at most) matrix is self-dual since

$$\begin{aligned}
V^{\sharp} &= -Z \left(-Z\overline{\mathbf{v}}\mathbf{w}^{*} + Z\overline{\mathbf{w}}\mathbf{v}^{*}\right)^{\mathrm{T}} Z \\
&= -Z \left(\overline{\mathbf{w}}\mathbf{v}^{*}Z - \overline{\mathbf{v}}\mathbf{w}^{*}Z\right) Z \\
&= Z\overline{\mathbf{w}}\mathbf{v}^{*} - Z\overline{\mathbf{v}}\mathbf{w}^{*} \\
&= -\mathcal{T}(\mathbf{w})\mathbf{v}^{*} + \mathcal{T}(\mathbf{v})\mathbf{w}^{*} \\
&= V.
\end{aligned}$$

If we start with $\mathbf{v}$ and $\mathbf{w}$ orthogonal, then V will be the rank 2 partial isometry taking $\mathbf{v}$ to $\mathcal{T}\mathbf{w}$ and $\mathbf{w}$ to $\mathcal{T}\mathbf{v}$. We record this as a lemma that avoids the ugly notation.

Lemma 7.4. *Suppose $\mathbf{v}$ and $\mathbf{w}$ are orthogonal unit vectors. Then the partial isometry from $\mathbb{C}\mathbf{v} + \mathbb{C}\mathbf{w}$ to $\mathbb{C}\mathcal{T}\mathbf{v} + \mathbb{C}\mathcal{T}\mathbf{w}$ that sends $\mathbf{v}$ to $\mathcal{T}\mathbf{w}$ and $\mathbf{w}$ to $-\mathcal{T}\mathbf{v}$ will be self-dual.*

Theorem 7.5. *If X in $\mathbf{M}_{2n}(\mathbb{C})$ satisfies $X^{\sharp} = X$ then there is a unitary U such that $U^{\sharp} = U$ and $X = U\,|X|$.*

Proof. Once more $W = Xf(X^{*}X)$ works to create a partial isometry W with $X = WP$ for $P = |X|$ and this time we have $W^{\sharp} = W$. We need a partial isometry from $\ker(W)$ to $\ker(W^{*})$ that is self-dual. The dimension of $\ker(W)$ will be even, by Proposition 7.3. Moreover if

$$\mathbf{v}_1, \mathbf{w}_1, \ldots, \mathbf{v}_m, \mathbf{w}_m$$

is an orthogonal basis for $\ker(W)$ then

$$\mathcal{T}\mathbf{v}_1, \mathcal{T}\mathbf{w}_1, \ldots, \mathcal{T}\mathbf{v}_m, \mathcal{T}\mathbf{w}_m$$

will be an orthogonal basis for $\ker(W^{*})$. The needed self-dual partial isometry V will send $\mathbf{v}_j$ to $\mathcal{T}\mathbf{w}_j$ and $\mathbf{w}_j$ to $\mathcal{T}\mathbf{v}_j$ and the self-dual unitary that we use will be $U = W + V$. $\quad \square$

8. The odd particle causes even degeneracy

We hope not to scare the mathematical reader with more discussion of Kramers degeneracy. What Kramers discovered was that for a certain systems involving an odd number of electrons, the Hamiltonian always had all eigenvalues with even multiplicity.

A mathematical manifestation of this is that the tensor product of two dual operations is the transpose operation in disguise. In contrast to that, the tensor product of three dual operations is a large dual operation in disguise. Specifically if we implement an orthogonal change of basis, the dual operation becomes the transpose.

This is essentially the same as the facts that are more familiar to mathematicians, that $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{H} \cong \mathbf{M}_4(\mathbb{R})$ and $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{H} \otimes_{\mathbb{R}} \mathbb{H} \cong \mathbf{M}_4(\mathbb{H})$.

In the following, we let Z_N be as in (4), the matrix that specified the dual operation.

Lemma 8.1. *Consider*

$$U = \frac{1}{\sqrt{2}} \left(I \otimes I - i Z_N \otimes Z_M \right).$$

For all $X \in \mathbf{M}_{2N}(C)$ and $Y \in \mathbf{M}_{2M}(\mathbb{C})$,

$$U^* \left(X^\sharp \otimes Y^\sharp \right) U = \left(U^* \left(X \otimes Y \right) U \right)^{\mathrm{T}}.$$

Proof. Since $Z_K^{\mathrm{T}} = -Z_K$ we see $U = U^{\mathrm{T}}$. Also

$$U^* U = \frac{1}{2} \left(I \otimes I + i Z_N \otimes Z_M \right) \left(I \otimes I - i Z_N \otimes Z_M \right)$$
$$= \frac{1}{2} \left(I \otimes I + (Z_N \otimes Z_M)^2 \right) = I \otimes I$$

so U is a unitary and $\overline{U} = U^{-1}$.

Since $U^2 = -i Z_N \otimes Z_M$ we find

$$U^* \left(X^\sharp \otimes Y^\sharp \right) U = \overline{U} \left(X^\sharp \otimes Y^\sharp \right) U$$
$$= \overline{U} \left(-i Z_N \otimes Z_M \right) \left(X^{\mathrm{T}} \otimes Y^{\mathrm{T}} \right) \left(i Z_N \otimes Z_M \right) U$$
$$= \overline{U} U^2 \left(X^{\mathrm{T}} \otimes Y^{\mathrm{T}} \right) \overline{U}^2 U$$
$$= \left(U^* \left(X \otimes Y \right) U \right)^{\mathrm{T}}. \quad \square$$

We have refrained from discussing C^*-algebras, but here they add clarity. What the lemma is showing is

$$\left(\mathbf{M}_{2N}(\mathbb{C}) \otimes \mathbf{M}_{2M}(\mathbb{C}), \sharp \otimes \sharp \right) \cong \left(\mathbf{M}_{2(M+N)}(\mathbb{C}), \mathrm{T} \right).$$

Lemma 8.2. *For all $X \in \mathbf{M}_{2N}(C)$ and $Y \in \mathbf{M}_{2M}(\mathbb{C})$,*

$$X^{\mathrm{T}} \otimes Y^\sharp = (X \otimes Y)^\sharp$$

where the $\sharp$ on the right is taken with respect to $I \otimes Z_M$.

Proof. This is much simpler, as

$$(X \otimes Y)^\sharp = - (I \otimes Z_M) (X \otimes Y)^{\mathrm{T}} (I \otimes Z_M)$$
$$= - (I \otimes Z_M) \left(X^{\mathrm{T}} \otimes Y^{\mathrm{T}} \right) (I \otimes Z_M)$$
$$= X^{\mathrm{T}} \otimes \left(-Z_M Y^{\mathrm{T}} Z_M \right). \quad \square$$

Together, these lemmas show us that a tensor product of three $\left(\mathbf{M}_{2N_j}(\mathbb{C}), \sharp\right)$ leads to something isomorphic to

$$\left(\mathbf{M}_{2(N_1+N_2+N_3)}(\mathbb{C}), \sharp\right).$$

Acknowledgments

References

[1] A. Bunse-Gerstner, R. Byers, V. Mehrmann, A quaternion QR algorithm, Numer. Math. 55 (1989) 83–95.

[2] A. Bunse-Gerstner, R. Byers, V. Mehrmann, Numerical methods for simultaneous diagonalization, SIAM J. Matrix Anal. Appl. 14 (1993) 927–949.

[3] D.R. Farenick, B.A.F. Pidkowich, The spectral theorem in quaternions, Linear Algebra Appl. 371 (2003) 75–102.

[4] M.B. Hastings, T.A. Loring, Topological insulators and C^*-algebras: theory and numerical practice, Ann. Physics 326 (2011) 1699–1759.

[5] H. Kramers, Théorie générale de la rotation paramagnétique dans les cristaux, Proc. Amst. Acad. 33 (1930) 959–972.

[6] H.C. Lee, Eigenvalues and canonical forms of matrices with quaternion coefficients, Proc. R. Ir. Acad. Sect. A 52 (1949) 253–260.

[7] T.A. Loring, A. Sørensen, Almost commuting self-adjoint matrices—the real and self-dual cases. arxiv:1012.3494.

[8] M. Mehta, Random Matrices, Academic Press, 2004.

[9] C. Paige, C. Van Loan, A Schur decomposition for Hamiltonian matrices, Linear Algebra Appl. 41 (1981) 11–32.

[10] G.K. Pedersen, Factorization in C^*-algebras, Exposition. Math. 16 (1998) 145–156.

[11] N. Wiegmann, Some theorems on matrices with real quaternion elements, Canad. J. Math. 7 (1955) 191–201.

[12] E. Wigner, Uber die operation der zeitumkehr in der quantenmechanik, Nach. Ges. Wiss. Gdtt 32 (1932) 546–559.

[13] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra Appl. 251 (1997) 21–57.

[14] F. Zhang, Y. Wei, Jordan canonical form of a partitioned complex matrix and its application to real quaternion matrices, Comm. Algebra 29 (2001) 2363–2375.