Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

STRUCTURED TOOLS FOR STRUCTURED MATRICES*

D. STEVEN MACKEY† , NILOUFER MACKEY†, AND FRANÇOISE TISSEUR‡

Abstract. An extensive and unified collection of structure-preserving transformations is presented and organized for easy reference. The structures involved arise in the context of a non-degenerate bilinear or sesquilinear form on $\mathbb{R}^n$ or $\mathbb{C}^n$. A variety of transformations belonging to the automorphism groups of these forms, that imitate the action of Givens rotations, Householder reflectors, and Gauss transformations are constructed. Transformations for performing structured scaling actions are also described. The matrix groups considered in this paper are the complex orthogonal, real, complex and conjugate symplectic, real perplectic, real and complex pseudo-orthogonal, and pseudo-unitary groups. In addition to deriving new transformations, this paper collects and unifies existing structure-preserving tools.

Key words. Structured matrices, Matrix groups, Givens rotations, Householder reflections, Complex orthogonal, Symplectic, Complex symplectic, Conjugate symplectic, Real perplectic, Pseudo-orthogonal, Complex pseudo-orthogonal, Pseudo-unitary, Scalar product, Bilinear form, Sesquilinear form, Automorphism groups, Jordan algebra, Lie algebra.

AMS subject classifications. 15A57, 15A63, 65F15, 65F30, 65F35.

1. Introduction. We consider structured matrices arising in the context of a non-degenerate bilinear or sesquilinear form on $\mathbb{R}^n$ or $\mathbb{C}^n$. Every such form engenders three important classes of matrices: an automorphism group, a Lie algebra and a Jordan algebra. There is a fundamental relationship between these three classes: the Lie and Jordan algebras remain invariant under similarities by matrices in the automorphism group. These groups therefore play a leading role in the study and development of structure-preserving transformations and factorizations. The automorphism groups considered in this paper include complex orthogonals, real, complex and conjugate symplectics, real perplectics, the Lorentz group, the real and complex pseudo-orthogonal groups and the pseudo-unitary group. Among the associated algebras are complex symmetric, Hamiltonian, J-symmetric, persymmetric, and pseudo-symmetric matrices. Such matrices naturally arise in engineering, physics and statistics, from problems with intrinsic symmetries; see for example [16] and the references therein.

Givens rotations and Householder reflectors are well-known elementary orthogonal transformations used typically to map one vector to another or to introduce zeros into a vector. They are used extensively in numerical linear algebra, most notably in decompositions such as QR factorizations, tridiagonalizations and Hessenberg re-

*Received by the editors on 25 February 2003. Accepted for publication on 27 April 2003. Handling Editor: Ludwig Elsner. Numerical Analysis Report No. 419, Manchester Centre for Computational Mathematics, Manchester, England, February 2003. This work was supported by Engineering and Physical Sciences Research Council grant GR/S15563.

†Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA (steve.mackey@wmich.edu, nil.mackey@wmich.edu).

‡Department of Mathematics, University of Manchester, Manchester, M13 9PL, England (ftisseur@ma.man.ac.uk, http://www.ma.man.ac.uk/~ftisseur/). This work was supported by Engineering and Physical Sciences Research Council grant GR/L76532.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

ductions, and eigenvalue and singular value decompositions. We describe a variety of matrix tools belonging to automorphism groups other than the orthogonal and unitary group that imitate the action of Givens rotations, Householder reflectors and also, when possible, Gauss transformations. We also investigate scaling actions within these groups.

In addition to deriving new results, this paper provides an extensive and unified collection of matrix tools for structured matrices, organized for easy reference. The treatment is necessarily very detailed; it brings out the similarities between and the differences among the various automorphism groups. It is expected that this work will aid in the derivation of new structure-preserving factorizations, as well as in the development of new structure-preserving algorithms.

The paper is organized as follows. Section 2 introduces concepts and notation needed for the unified development of structure-preserving tools. The two main types of actions – introducing zeros into a vector, and scaling a vector – are introduced in section 3. The basic forms we use to build tools are given in subsection 3.2. The main results of the paper are in section 4, which gives a detailed presentation of the tools tailored to each matrix group, in a form that makes them readily accessible for further use and study. Finally, since 2×2 matrices are the building blocks of so many matrix constructions – including several developed in this paper – we derive explicit parameterizations of the 2×2 automorphism groups in Appendix A.

2. Preliminaries.

2.1. Scalar products.

For the convenience of the reader, we briefly review some definitions and properties of scalar products. A more detailed discussion can be found, for example, in Jacobson [27], Lang [29], or Shaw [43].

Let $\mathbb{K}$ denote either the field $\mathbb{R}$ or $\mathbb{C}$, and consider a map $(x, y) \mapsto \langle x, y\rangle$ from $\mathbb{K}^n \times \mathbb{K}^n$ to $\mathbb{K}$. If such a map is linear in each argument, that is,

$$\begin{aligned} \langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle &= \alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle, \\ \langle x, \beta_1 y_1 + \beta_2 y_2\rangle &= \beta_1\langle x, y_1\rangle + \beta_2\langle x, y_2\rangle, \end{aligned}$$

then it is called a *bilinear form*. If $\mathbb{K} = \mathbb{C}$, and the map $(x, y) \mapsto \langle x, y\rangle$ is conjugate linear in the first argument and linear in the second,

$$\begin{aligned} \langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle &= \overline{\alpha}_1\langle x_1, y\rangle + \overline{\alpha}_2\langle x_2, y\rangle, \\ \langle x, \beta_1 y_1 + \beta_2 y_2\rangle &= \beta_1\langle x, y_1\rangle + \beta_2\langle x, y_2\rangle, \end{aligned}$$

then it is called a *sesquilinear form.*

Given a bilinear form on $\mathbb{K}^n$ (respectively a sesquilinear form on $\mathbb{C}^n$), there exists a unique $M \in \mathbb{K}^{n\times n}$ (respectively $M \in \mathbb{C}^{n\times n}$) such that $\langle x, y\rangle = x^T My, \forall x, y \in \mathbb{K}^n$ (respectively $\langle x, y\rangle = x^* My, \forall x, y \in \mathbb{C}^n$). Here, the superscript $*$ is used for conjugate transpose. M is called the matrix associated with the form. When we have more than one scalar product under consideration, we will denote $\langle x, y\rangle$ by $\langle x, y\rangle_{\mathrm{M}}$, using the subscript M to distinguish the forms under discussion.

A bilinear form is said to be symmetric if $\langle x, y\rangle = \langle y, x\rangle$, and skew-symmetric if $\langle x, y\rangle = -\langle y, x\rangle$. It is easily shown that the matrix M associated with a symmetric

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

form is symmetric; similarly, the matrix of a skew-symmetric form is skew-symmetric. A sesquilinear form is Hermitian if $\langle x, y\rangle = \overline{\langle y, x\rangle}$ and skew-Hermitian if $\langle x, y\rangle = -\overline{\langle y, x\rangle}$. The matrices associated with such forms are Hermitian and skew-Hermitian, respectively.

From now on the term *scalar product* refers only to a nondegenerate bilinear or sesquilinear form on $\mathbb{K}^n$, that is, a form for which the associated matrix M is nonsingular. No a priori assumption about the positive definiteness or indefiniteness of the scalar product is made. We will frequently use the associated quadratic functional

$$q_{\mathrm{M}}(x) \stackrel{\mathrm{def}}{=} \langle x, x\rangle_{\mathrm{M}}$$

and will drop the subscript M when there is no ambiguity. Observe that $q(x)$ is the natural generalization of the quantity x^*x for vectors in $\mathbb{K}^n$ equipped with the Euclidean inner product.

For any matrix $A \in \mathbb{K}^{n\times n}$ there is a unique matrix $A^\star$, the *adjoint* of A with respect to $\langle\cdot,\cdot\rangle$, defined by

$$\langle Ax, y\rangle = \langle x, A^\star y\rangle, \quad \forall x, y \in \mathbb{K}^n.$$

Note that in general $A^\star \neq A^*$. It is easy to obtain an explicit formula for the adjoint $A^\star$. If $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ is bilinear, then

$$\langle Ax, y\rangle = x^T A^T M y = x^T M M^{-1} A^T M y = \langle x, M^{-1}A^T M y\rangle.$$

Thus $A^\star = M^{-1}A^T M$. One can show similarly that $A^\star = M^{-1}A^* M$ when $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ is sesquilinear.

2.2. Lie and Jordan algebras, and automorphism groups. There are three important classes of structured matrices associated with each scalar product:

1. The matrices G which preserve the value of the scalar product

$$\langle Gx, Gy\rangle = \langle x, y\rangle, \quad \forall x, y \in \mathbb{K}^n.$$

2. The matrices S that are *self-adjoint* with respect to the scalar product

$$\langle Sx, y\rangle = \langle x, Sy\rangle, \quad \forall x, y \in \mathbb{K}^n.$$

3. The matrices K that are *skew-adjoint* with respect to the scalar product

$$\langle Kx, y\rangle = -\langle x, Ky\rangle, \quad \forall x, y \in \mathbb{K}^n.$$

These classes can be succinctly described using the adjoint operator:

$$\begin{aligned}
\mathbb{G} &\stackrel{\mathrm{def}}{=} \left\{G \in \mathbb{K}^{n\times n} : \ G^\star = G^{-1}\right\},\\
\mathbb{J} &\stackrel{\mathrm{def}}{=} \left\{S \in \mathbb{K}^{n\times n} : \ S^\star = S\right\},\\
\mathbb{L} &\stackrel{\mathrm{def}}{=} \left\{K \in \mathbb{K}^{n\times n} : \ K^\star = -K\right\}.
\end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

TABLE 2.1
Structured matrices associated with some scalar products.

Space	**Bilinear Form** $\langle x, y\rangle$	**Adjoint** $A \in M_n(\mathbb{K})$	**Automorphism Group** $\mathbb{G} = \{G : G^\star = G^{-1}\}$	**Jordan Algebra** $\mathbb{J} = \{S : S^\star = S\}$	**Lie Algebra** $\mathbb{L} = \{K : K^\star = -K\}$
$\mathbb{R}^n$	$x^T y$ symmetric form	$A^\star = A^T$	Real orthogonals $O(n, \mathbb{R})$	Symmetrics	Skew-symmetrics
$\mathbb{C}^n$	$x^T y$ symmetric form	$A^\star = A^T$	Complex orthogonals $O(n, \mathbb{C})$	Complex symmetrics	Complex skew-symmetrics
$\mathbb{R}^n$	$x^T \Sigma_{p,q} y$ symmetric form	$A^\star = \Sigma_{p,q} A^T \Sigma_{p,q}$	Pseudo-orthogonals $O(p, q, \mathbb{R})$	Pseudo symmetrics	Pseudo skew-symmetrics
$\mathbb{C}^n$	$x^T \Sigma_{p,q} y$ symmetric form	$A^\star = \Sigma_{p,q} A^T \Sigma_{p,q}$	Complex pseudo-orthogonals $O(p, q, \mathbb{C})$	Complex pseudo-symmetrics	Complex pseudo-skew-symmetrics
$\mathbb{R}^n$	$x^T R y$ symmetric form	$A^\star = R A^T R$	Real perplectics $\mathcal{P}(n)$	Persymmetrics	Perskew-symmetrics
$\mathbb{R}^{2n}$	$x^T J y$ skew-symm. form	$A^\star = -J A^T J$	Real symplectics $Sp(2n, \mathbb{R})$	Skew-Hamiltonians	Hamiltonians
$\mathbb{C}^{2n}$	$x^T J y$ skew-symm. form	$A^\star = -J A^T J$	Complex symplectics $Sp(2n, \mathbb{C})$	J-skew-symmetric	J-symmetric

Space	**Sesquilinear Form** $\langle x, y\rangle$	**Adjoint** $A \in M_n(\mathbb{C})$	**Automorphism Group** $\mathbb{G} = \{G : G^\star = G^{-1}\}$	**Jordan Algebra** $\mathbb{J} = \{S : S^\star = S\}$	**Lie Algebra** $\mathbb{L} = \{K : K^\star = -K\}$
$\mathbb{C}^n$	$x^* y$ Hermitian form	$A^\star = A^*$	Unitaries $U(n)$	Hermitian	Skew-Hermitian
$\mathbb{C}^n$	$x^* \Sigma_{p,q} y$ Hermitian form	$A^\star = \Sigma_{p,q} A^* \Sigma_{p,q}$	Pseudo-unitaries $U(p, q)$	Pseudo Hermitian	Pseudo skew-Hermitian
$\mathbb{C}^{2n}$	$x^* J y$ skew-Herm. form	$A^\star = -J A^* J$	Conjugate symplectics $Sp^*(2n, \mathbb{C})$	J-skew-Hermitian	J-Hermitian

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

Although it is not a linear subspace, the set $\mathbb{G}$ always forms a multiplicative group (indeed a Lie group), and will be referred to as the *automorphism group* of the scalar product. By contrast, the sets $\mathbb{J}$ and $\mathbb{L}$ are linear subspaces, but they are not closed under multiplication. Instead $\mathbb{L}$ is closed with respect to the Lie bracket $[K_1, K_2] = K_1K_2 - K_2K_1$, while $\mathbb{J}$ is closed with respect to the Jordan product $\{S_1, S_2\} = S_1S_2 + S_2S_1$. Hence we refer to $\mathbb{L}$ and $\mathbb{J}$ as the *Lie and Jordan algebras*, respectively, of the scalar product. For more on these classes of structured matrices, see [1], [22], [28].

The importance of the automorphism groups is underscored by the following result, establishing a fundamental relationship between matrices in the three classes $\mathbb{G}, \mathbb{L}$, and $\mathbb{J}$.

PROPOSITION 2.1. *Let $\langle\cdot,\cdot\rangle$ be any scalar product on $\mathbb{K}^n$, and $\mathbb{G}$ the corresponding automorphism group. For any $G \in \mathbb{G}$, we have*

$$A \in \mathbb{G} \Rightarrow G^{-1}AG \in \mathbb{G}, \quad S \in \mathbb{J} \Rightarrow G^{-1}SG \in \mathbb{J}, \quad K \in \mathbb{L} \Rightarrow G^{-1}KG \in \mathbb{L}.$$

Proof. The first implication is immediate since $\mathbb{G}$ is a multiplicative group. Now suppose $S \in \mathbb{J}$ and $G \in \mathbb{G}$. Then for all $x, y \in \mathbb{K}^n$, we have

$$\langle G^{-1}SGx, y\rangle = \langle SGx, G^{-\star}y\rangle = \langle Gx, S^\star G^{-\star}y\rangle = \langle x, G^\star S^\star G^{-\star}y\rangle = \langle x, G^{-1}SGy\rangle.$$

Thus $G^{-1}SG \in \mathbb{J}$. The third implication is proved in a similar manner. $\square$

This proposition shows that the automorphism groups form the natural classes of structure-preserving similarities for $\mathbb{G}$, $\mathbb{L}$, and $\mathbb{J}$. Thus they will be central to the development of structure-preserving algorithms involving any of these structured classes of matrices.

2.3. Structured matrices. The automorphism groups $\mathbb{G}$ discussed in this paper are listed in Table 2.1, along with their associated Lie and Jordan algebras $\mathbb{L}$ and $\mathbb{J}$. The underlying scalar products $\langle\cdot,\cdot\rangle_M$ use one of the following matrices for M: the $n \times n$ identity matrix I_n,

$$R \stackrel{\text{def}}{=} \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}, \quad J \stackrel{\text{def}}{=} \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix},$$

$$\Sigma_{p,q} \stackrel{\text{def}}{=} \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} \quad \text{with} \quad p + q = n.$$

Table 2.1 and the definition below introduce notation and terminology for the matrix groups that are the focus of this paper. These groups are all examples of “classical groups”, a term originally coined by Weyl [24], [27], [29].

DEFINITION 2.2.

1. *$A \in \mathbb{R}^{n\times n}$ is orthogonal if $A^{-1} = A^T$.*
2. *$A \in \mathbb{C}^{n\times n}$ is complex orthogonal if $A^{-1} = A^T$.*
3. *$A \in \mathbb{R}^{n\times n}$ is $\Sigma_{p,q}$-orthogonal if $A^{-1} = \Sigma_{p,q}A^T\Sigma_{p,q}$.*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

4. $A \in \mathbb{C}^{n\times n}$ *is complex* $\Sigma_{p,q}$*-orthogonal if* $A^{-1} = \Sigma_{p,q}A^T\Sigma_{p,q}$.
5. $A \in \mathbb{R}^{n\times n}$ *is real perplectic if* $A^{-1} = RA^TR$.
6. $A \in \mathbb{R}^{2n\times 2n}$ *is real symplectic if* $A^{-1} = -JA^TJ$.
7. $A \in \mathbb{C}^{2n\times 2n}$ *is complex symplectic if* $A^{-1} = -JA^TJ$.
8. $A \in \mathbb{C}^{n\times n}$*, is unitary if* $A^{-1} = A^*$.
9. $A \in \mathbb{C}^{n\times n}$ *is* $\Sigma_{p,q}$*-unitary if* $A^{-1} = \Sigma_{p,q}A^*\Sigma_{p,q}$,
10. $A \in \mathbb{C}^{2n\times 2n}$ *is conjugate symplectic if* $A^{-1} = -JA^*J$.

$\Sigma_{p,q}$*-orthogonal and* $\Sigma_{p,q}$*-unitary matrices will also be referred to as pseudo-orthogonal and pseudo-unitary matrices, respectively.*
Note that each condition in Definition 2.2 is just a special case of the common defining property $A \in \mathbb{G} \Leftrightarrow A^{-1} = A^\star$. This relation restricts the values of the determinant for matrices in automorphism groups.

PROPOSITION 2.3. *Suppose* $A \in \mathbb{G}$*, where* $\mathbb{G}$ *is the automorphism group of a bilinear form. Then* $\det A = \pm 1$*. In the case of a sesquilinear form,* $|\det A| = 1$.

Proof. $A \in \mathbb{G} \Rightarrow A^\star A = I \Rightarrow M^{-1}A^TMA = I \Rightarrow (\det A)^2 = 1 \Rightarrow \det A = \pm 1$. For a sesquilinear form, $A^\star A = I \Rightarrow M^{-1}A^*MA = I \Rightarrow \overline{\det A}\det A = 1 \Rightarrow |\det A| = 1$. □

The determinant can sometimes be even more restricted. For example, real and complex symplectic matrices have only $+1$ determinant, and -1 is never realized; for several different proofs of this non-obvious fact, see [32].

3. Actions and basic forms for tools.

3.1. Actions.

The algorithms of numerical linear algebra are mainly built upon one technique used over and over again: putting zeros into matrices.
L. N. TREFETHEN and D. BAU, *Numerical Linear Algebra* (1997), [48, p.191]

The reduction of a structured matrix to a structured condensed form, or its factorization into structured factors, is often achieved by making a sequence of elementary structured matrices act on the original one, either by pre- or by post-multiplication.

In other situations (e.g., reduction to Hessenberg or tridiagonal form) the desired reduction may instead need to be realized by similarity or congruence transformations. In either case, the essential effect of the individual elementary transformations is often based on the action of a matrix on a vector.

We restrict our attention to the action on vectors by structured matrices that come from an automorphism group $\mathbb{G}$ associated with a scalar product $\langle\cdot,\cdot\rangle_{\mathrm{M}}$. We are mainly interested in two types of actions:

1. Introducing zeros into a vector.
2. Scaling a vector, or scaling selected entries of a vector.

We have not included an analysis of the numerical behavior, in floating point arithmetic, of the tools developed in this paper; this will be the subject of future work.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

3.1.1. Making zeros. Recall some well-known and commonly used tools for making zeros, from the groups $O(n,\mathbb{R})$ and $U(n)$.

Orthogonal or unitary *Givens rotations* or *plane rotations* given by

$$G \stackrel{\text{def}}{=} \begin{bmatrix} c & s \\ -\bar{s} & \bar{c} \end{bmatrix} \in \mathbb{K}^{2\times 2}, \qquad |c|^2 + |s|^2 = 1, \quad \mathbb{K} = \mathbb{R}, \mathbb{C}, \tag{3.1}$$

are useful tools to selectively zero out individual entries of a vector $x = [x_1, x_2]^T \in \mathbb{K}^2$. If $y = [y_1, y_2]^T = Gx$, then $y_2 = 0$ whenever

$$c = \frac{\omega \bar{x}_1}{\sqrt{x^* x}}, \quad s = \frac{\omega \bar{x}_2}{\sqrt{x^* x}}, \quad |\omega| = 1, \quad \omega \in \mathbb{K}. \tag{3.2}$$

This yields $y_1 = \omega\sqrt{x^*x}$. The unit modulus ω is arbitrary and can be used to control freely the angular position of c, s, or y_1. For example, $\omega = 1$ is commonly used to make $y_1 \in \mathbb{R}^+$. The choice of ω is discussed by Bindel et al. [7]. Their criterion is based on compatibility with existing implementations of Givens rotations, consistency between definitions for orthogonal and unitary Givens rotations (they should agree on real data), continuity of c, s and y_1 as functions of x_1 and x_2 and, finally, amenability to a fast implementation. These criteria cannot all be satisfied simultaneously, but a good compromise is achieved when taking

$$\omega = \operatorname{sign}(x_1) \stackrel{\text{def}}{=} \begin{cases} x_1/|x_1| & \text{if } x_1 \neq 0, \\ 1 & \text{if } x_1 = 0, \end{cases} \tag{3.3}$$

which, if x_1 is real, simplifies to $\omega = -1$ if $x_1 < 0$ and $\omega = 1$ if $x_1 \geq 0$.

Embedding a 2×2 Givens as a principal submatrix of I_n yields plane rotations in the orthogonal group $O(n,\mathbb{R})$ or the unitary group $U(n)$. For 4×4 and 8×8 analogues of Givens rotations when $\mathbb{K} = \mathbb{R}$, see [20], [31].

Householder reflectors are elementary matrices of the form

$$H(u) \stackrel{\text{def}}{=} I + \beta u u^*, \quad 0 \neq u \in \mathbb{K}^n, \quad 0 \neq \beta \in \mathbb{K}, \tag{3.4}$$

which are symmetric orthogonal if $\mathbb{K} = \mathbb{R}$ and $\beta = -2/(u^T u)$, and unitary if $\mathbb{K} = \mathbb{C}$ and β is on the circle $|\beta - r| = |r|$, where $r = -1/(u^*u)$ (see Theorem 3.3; a complete discussion of all the unitary reflectors can be found in [34]). For any distinct x and y such that $x^*x = y^*y$, $H(u)x = y$ whenever $u = y - x$ and $\beta = 1/(u^*x)$. It follows that Householder reflectors can be used to simultaneously introduce up to $n-1$ zeros into an n-vector. The usual choice is $y = -\operatorname{sign}(x_1)\sqrt{x^*x}\, e_1$, where e_1 is the first column of the identity matrix. This yields a Hermitian $H(u)$, since in this case β is always real. Another choice used by LAPACK [2] is $y = \pm\sqrt{x^*x}\, e_1$, which sends x to a real multiple of e_1. This yields a β that may be complex and therefore $H(u)$ may not be Hermitian. This choice may be advantageous for some tasks, such as the reduction of a Hermitian matrix to tridiagonal form, since the resulting tridiagonal matrix is real symmetric and the real QR algorithm can be employed to compute its eigenvalues. For more details, see Lehoucq [30].

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

Gauss transformations are non-orthogonal unit lower triangular matrices of the form $I - ve_k^T$, where the first k components of the vector v are zero. Such matrices are particularly useful for introducing zeros in components $k+1, \ldots, n$ of a vector [23, p. 95].

In this paper, a *Givens-like action* on a vector x consists of setting one, and in some cases more than one, selected component of x to zero. A *Householder-like action* is to send a vector x (or part of it) to a multiple of e_j. A *Gauss-like action* on a vector is carried out by a triangular matrix and consists of introducing k zeros in the top or bottom part of x. Our main aim is to describe tools in various automorphism groups that perform these three types of zeroing action, whenever these actions are possible.

3.1.2. Scaling. Scaling is often used in numerical linear algebra to improve the stability of algorithms. The usual meaning of the term "scaling" is multiplication by a diagonal matrix. For each automorphism group $\mathbb{G}$, we describe all the scaling actions that can be realized by diagonal matrices in $\mathbb{G}$.

There are, however, some automorphism groups in which the set of diagonal matrices is restricted to diag(± 1), so that the corresponding scaling actions are narrowly circumscribed. In these groups, one can often realize a scaling action that acts uniformly on all coordinates of a given vector by an *arbitrarily* chosen scaling factor. However, this can only be achieved on isotropic vectors, using non-diagonal matrices. Recall that a nonzero vector $v \in \mathbb{K}^n$ is said to be *isotropic* with respect to $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ if $q_{\mathrm{M}}(v) = \langle v, v \rangle_{\mathrm{M}} = 0$. In this paper we show constructively how any given isotropic vector $v \in \mathbb{K}^2$ can be scaled by any desired nonzero factor when $\mathbb{G}$ is $O(2,\mathbb{C})$, $O(1,1,\mathbb{R})$, $O(1,1,\mathbb{C})$ or $U(1,1)$ (see Table 2.1 for the definition of these groups). These tools are used in [35] to derive vector-canonical forms and to give a constructive proof of the structured mapping theorem for the groups $O(n,\mathbb{C})$, $O(p,q,\mathbb{R})$, $O(p,q,\mathbb{C})$ and $U(p,q)$.

More generally, Proposition 3.1 shows that isotropic vectors in $\mathbb{K}^n$ can be arbitrarily scaled by matrices in the automorphism group, while non-isotropic vectors may be scaled only in very restricted ways. This is closely connected with the question of which eigenvalue-eigenvector pairs (λ, v) can occur for matrices in $\mathbb{G}$.

Proposition 3.1. *Suppose $\mathbb{G}$ is the automorphism group of a scalar product $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ on $\mathbb{K}^n$, and (λ, v) is an eigenpair for some $G \in \mathbb{G}$.*

(i) *Suppose v is non-isotropic. If $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ is bilinear, then $\lambda = \pm 1$. If $\langle \cdot, \cdot \rangle_{\mathrm{M}}$ is sesquilinear, then $|\lambda| = 1$.*

(ii) *If v is isotropic, then there is no restriction on the eigenvalue λ, other than $\lambda \neq 0$. That is, for any given isotropic vector $v \in \mathbb{K}^n$, and nonzero $\lambda \in \mathbb{K}$, there is some $G \in \mathbb{G}$ such that $Gv = \lambda v$.*

Proof. First observe that every matrix in $\mathbb{G}$ preserves the value of q_{M}, since $q_{\mathrm{M}}(Gv) = \langle Gv, Gv \rangle_{\mathrm{M}} = \langle v, v \rangle_{\mathrm{M}} = q_{\mathrm{M}}(v) \quad \forall v \in \mathbb{K}^n$. Then if $Gv = \lambda v$, we have

$$q_{\mathrm{M}}(v) = \begin{cases} \lambda^2 q_{\mathrm{M}}(v) & \text{if } \langle \cdot, \cdot \rangle_{\mathrm{M}} \text{ is bilinear,} \\ |\lambda|^2 q_{\mathrm{M}}(v) & \text{if } \langle \cdot, \cdot \rangle_{\mathrm{M}} \text{ is sesquilinear.} \end{cases} \tag{3.5}$$

Part (i) now follows immediately from (3.5), while part (ii) is an immediate consequence of the structured mapping theorem proved in [35]. ∎

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

TABLE 3.1
Some basic forms associated with classical groups. A is partitioned as $\left[\begin{smallmatrix} E & G \\ F & H \end{smallmatrix}\right]$.

Automorphism group $\mathbb{G}$	**Inverse $\mathbf{A^{-1} = A^\star}$**	**(Block) upper triangular**	**2×2's**
Complex orthogonal $O(n, \mathbb{C})$	A^T	$\mathrm{diag}\{\pm 1\}$	$\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ or $\begin{bmatrix} \alpha & \beta \\ \beta & -\alpha \end{bmatrix}$ $\alpha, \beta \in \mathbb{C},\ \alpha^2 + \beta^2 = 1$
Pseudo-orthogonal $O(p, q, \mathbb{R})$	$\begin{bmatrix} E^T & -F^T \\ -G^T & H^T \end{bmatrix}$, $E \in \mathbb{R}^{p \times p}$ $H \in \mathbb{R}^{q \times q}$	$\begin{bmatrix} E & 0 \\ 0 & H \end{bmatrix}$, $E \in O(p, \mathbb{R})$ $H \in O(q, \mathbb{R})$	$\begin{bmatrix} c & s \\ s & c \end{bmatrix} \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}, \theta \in \mathbb{R}$ $c = \cosh\theta, s = \sinh\theta$
Complex pseudo-orthogonal $O(p, q, \mathbb{C})$	$\begin{bmatrix} E^T & -F^T \\ -G^T & H^T \end{bmatrix}$, $E \in \mathbb{C}^{p \times p}$ $H \in \mathbb{C}^{q \times q}$	$\begin{bmatrix} E & 0 \\ 0 & H \end{bmatrix}$, $E \in O(p, \mathbb{C})$ $H \in O(q, \mathbb{C})$	$\begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$ $\alpha^2 - \beta^2 = 1,\ \alpha, \beta \in \mathbb{C}$
Perplectic $\mathcal{P}(n)$	$A^F \stackrel{\text{def}}{=} RA^T R$, $R = \begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix}$	$\begin{bmatrix} E & G \\ 0 & E^{-F} \end{bmatrix}$, $E \in \mathbb{R}^{n \times n}$ nonsingular, $G = ES$ with S perskew-symmetric. $\begin{bmatrix} E & -\alpha E y^F & G \\ 0 & \alpha & y \\ 0 & 0 & E^{-F} \end{bmatrix}$, $E \in \mathbb{R}^{n \times n}$ nonsingular, $\alpha = \pm 1,\ y^T \in \mathbb{R}^n$ and $G = ES - \frac{1}{2} E y^F y$ with S perskew-symmetric.	$\begin{bmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{bmatrix}$ or $\begin{bmatrix} 0 & \beta \\ 1/\beta & 0 \end{bmatrix}$ $\alpha, \beta \in \mathbb{R}$
Real symplectic $Sp(2n, \mathbb{R})$	$\begin{bmatrix} H^T & -G^T \\ -F^T & E^T \end{bmatrix}$, $n \times n$ real blocks	$\begin{bmatrix} E & G \\ 0 & E^{-T} \end{bmatrix}$, E nonsingular, $G = ES$ with S symmetric.	$SL(2, \mathbb{R}) \stackrel{\text{def}}{=} \{A \in \mathbb{R}^{2 \times 2} :$ $\det(A) = +1\}$
Complex symplectic $Sp(2n, \mathbb{C})$	$\begin{bmatrix} H^T & -G^T \\ -F^T & E^T \end{bmatrix}$, $n \times n$ complex blocks	$\begin{bmatrix} E & G \\ 0 & E^{-T} \end{bmatrix}$, E nonsingular, $G = ES$ with S complex symmetric.	$SL(2, \mathbb{C}) \stackrel{\text{def}}{=} \{A \in \mathbb{C}^{2 \times 2} :$ $\det(A) = +1\}$
Pseudo-unitary $U(p, q)$	$\begin{bmatrix} E^* & -F^* \\ -G^* & H^* \end{bmatrix}$, $E \in \mathbb{C}^{p \times p}$ $H \in \mathbb{C}^{q \times q}$	$\begin{bmatrix} E & 0 \\ 0 & H \end{bmatrix}$, $E \in U(p)$ $H \in U(q)$	$\begin{bmatrix} \alpha & \beta \\ \bar\beta & \bar\alpha \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}$ $\theta \in \mathbb{R}$, $\alpha, \beta \in \mathbb{C},\ \lvert\alpha\rvert^2 - \lvert\beta\rvert^2 = 1$
Conjugate sym– plectic $Sp^*(2n, \mathbb{C})$	$\begin{bmatrix} H^* & -G^* \\ -F^* & E^* \end{bmatrix}$, $n \times n$ complex blocks	$\begin{bmatrix} E & G \\ 0 & E^{-*} \end{bmatrix}$, E nonsingular, $G = ES$ with S Hermitian.	$\{e^{i\theta} B :\ \theta \in \mathbb{R},$ $B \in SL(2, \mathbb{R})\}$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

3.2. Basic forms for tools. In this section we present a collection of basic and useful forms for matrices in the classical groups under consideration, omitting the orthogonal and unitary groups.

3.2.1. Inverse, triangular and 2×2 forms. Structured tools for each automorphism group are constructed from the basic forms listed in Table 3.1. We note that the scope and flexibility of the tools depend on the extent to which these forms exist in the group. Since $A^{-1} = A^\star$ for any $A \in \mathbb{G}$, the second column in Table 3.1 is found by calculating $A^\star = M^{-1}A^TM$ for bilinear forms and $A^\star = M^{-1}A^*M$ for sesquilinear forms, where M is the matrix of the form. Block triangular forms are needed when constructing tools for Gauss-like actions. Such actions can therefore only be expected in those groups that have non-trivial triangular forms. For brevity, only block upper triangular forms are given in the third column. Lower triangular forms are constructed analogously. Finally, 2×2 forms are useful in many ways such as in designing tools for Givens-like actions and scaling actions. These are given in the last column; the derivation of these parameterizations is included in Appendix A.

3.2.2. $\mathbb{G}$-reflectors. Following Householder [26], we define an *elementary transformation* to be a linear map $G\colon \mathbb{K}^n \to \mathbb{K}^n$ of the form $G = I + uv^*$ for some nonzero $u, v \in \mathbb{K}^n$. It is not hard to see that G has an $(n-1)$-dimensional fixed point subspace $\mathcal{H}$, i.e., a hyperplane $\mathcal{H}$ on which it acts as the identity. In [34], Mackey, Mackey and Tisseur consider elementary transformations G in automorphism groups $\mathbb{G}$ and refer to such maps as *generalized $\mathbb{G}$-reflectors*, or $\mathbb{G}$-reflectors for short. If $\mathbb{G} = O(n, \mathbb{R})$, then any $\mathbb{G}$-reflector is expressible in the form $G = I - 2uu^T$ with $u^Tu = 1$. The elementary transformation G is precisely a perpendicular reflection through the hyperplane $\mathcal{H} = \{v \in \mathbb{R}^n : \langle u, v\rangle = 0\}$ and is referred to as a reflector [38] or Householder transformation [23].

We state three main results about $\mathbb{G}$-reflectors and refer to [34] for the proofs. The first result gives a characterization of $\mathbb{G}$-reflectors for general automorphism groups.

Theorem 3.2. *Any $\mathbb{G}$-reflector G is expressible in the form*

$$G = \begin{cases} I + \beta uu^TM & \text{if } \langle\cdot,\cdot\rangle_{\mathrm{M}} \text{ is bilinear,} \\[2ex] I + \beta uu^*M & \text{if } \langle\cdot,\cdot\rangle_{\mathrm{M}} \text{ is sesquilinear,} \end{cases} \tag{3.6}$$

for some $\beta \in \mathbb{K} \setminus \{0\}$ and $u \in \mathbb{K}^n \setminus \{0\}$. Not every G given by (3.6) is in $\mathbb{G}$; the parameters β and u must satisfy an additional relation:

$$\text{For bilinear forms:} \qquad G \in \mathbb{G} \quad \Leftrightarrow \quad \big(M + (1 + \beta q_{\mathrm{M}}(u))M^T\big)u = 0.$$

$$\text{For sesquilinear forms:} \quad G \in \mathbb{G} \quad \Leftrightarrow \quad \big(\beta M + (\overline{\beta} + |\beta|^2 q_{\mathrm{M}}(u))M^*\big)u = 0.$$

The characterization of $\mathbb{G}$-reflectors in Theorem 3.2 can be refined if one assumes additional properties of the matrices M associated with the underlying scalar product.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

THEOREM 3.3 ($\mathbb{G}$-reflectors for specific classes of scalar product).

- *Symmetric bilinear forms* ($M^T = M$ *and* $q_{\mathrm{M}}(u) \in \mathbb{K}$)
  $G = I + \beta uu^T M \in \mathbb{G}$ *if and only if* u *is non-isotropic, and* $\beta = -2/q_{\mathrm{M}}(u)$.

- *Skew-symmetric bilinear forms* ($M^T = -M$ *and* $q_{\mathrm{M}}(u) \equiv 0$)
  $G = I + \beta uu^T M \in \mathbb{G}$ *for any* $u \in \mathbb{K}^{2n}$ *and any* $\beta \in \mathbb{K}$.

- *Hermitian sesquilinear forms* ($M^* = M$ *and* $q_{\mathrm{M}}(u) \in \mathbb{R}$)
  $G = I + \beta uu^* M \in \mathbb{G}$ *if and only if* u *is isotropic and* $\beta \in i\mathbb{R}$, *or* u *is non-isotropic and* $\beta \in \mathbb{C}$ *is on the circle*
  $$|\beta - r| = |r|, \quad \text{where} \quad r \stackrel{\text{def}}{=} -\frac{1}{q_{\mathrm{M}}(u)} \in \mathbb{R}.$$
- *Skew-Hermitian sesquilinear forms* ($M^* = -M$ *and* $q_{\mathrm{M}}(u) \in i\mathbb{R}$)
  $G = I + \beta uu^* M \in \mathbb{G}$ *if and only if* u *is isotropic and* $\beta \in \mathbb{R}$, *or* u *is non-isotropic and* $\beta \in \mathbb{C}$ *is on the circle*
  $$|\beta - r| = |r|, \quad \text{where} \quad r \stackrel{\text{def}}{=} -\frac{1}{q_{\mathrm{M}}(u)} \in i\mathbb{R}.$$

The next theorem gives necessary and sufficient conditions for the existence of a $\mathbb{G}$-reflector G such that $Gx = y$.

THEOREM 3.4 ($\mathbb{G}$-reflector mapping theorem). *Suppose* $\mathbb{K}^n$ *has a scalar product* $\langle\cdot,\cdot\rangle_{\mathrm{M}}$ *that is either symmetric bilinear, skew-symmetric bilinear, Hermitian sesquilinear, or skew-Hermitian sesquilinear. Then for distinct nonzero vectors* $x, y \in \mathbb{K}^n$, *there exists a* $\mathbb{G}$*-reflector* G *such that* $Gx = y$ *if and only if* $q_{\mathrm{M}}(x) = q_{\mathrm{M}}(y)$ *and* $\langle y - x, x\rangle_{\mathrm{M}} \neq 0$. *Furthermore, whenever* G *exists, it is unique and can be specified by taking* $u = y - x$ *and* $\beta = 1/\langle u, x\rangle_{\mathrm{M}}$ *in* (3.6). *Equivalently,* G *may be specified by taking* $u = x - y$ *and* $\beta = -1/\langle u, x\rangle_{\mathrm{M}}$ *in* (3.6).

It follows from Theorem 3.4 that $\mathbb{G}$-reflectors can be used to simultaneously introduce up to $n-1$ zeros into an n-vector, and therefore will play an important role when deriving Householder-like actions.

3.2.3. $\mathbb{G}$-orthogonal and $\mathbb{G}$-unitary forms. We describe the intersection of each automorphism group listed in Table 3.1 with the orthogonal or unitary group, as appropriate:

$$\begin{aligned}
O(n,\mathbb{C}) \cap U(n) &= O(n,\mathbb{R}),\\
O(p,q,\mathbb{R}) \cap O(n,\mathbb{R}) &= \left\{E \oplus H,\ E \in O(p,\mathbb{R}),\ H \in O(q,\mathbb{R})\right\},\\
O(p,q,\mathbb{C}) \cap U(n) &= \left\{ \begin{bmatrix} E & G \\ F & H \end{bmatrix} \in U(n), \quad \begin{matrix} E \in \mathbb{R}^{p\times p},\ H \in \mathbb{R}^{q\times q}, \\ F \in i\mathbb{R}^{q\times p},\ G \in i\mathbb{R}^{p\times q} \end{matrix} \right\},\\
\mathcal{P}(n) \cap O(n,\mathbb{R}) &= \{A \in O(n,\mathbb{R})\colon\ A = (a_{ij}) \text{ is centrosymmetric, i.e.,}\\
&\qquad a_{i,j} = a_{n-i+1,n-j+1} \text{ for } 1 \le i,j \le n\},\\
Sp(2n,\mathbb{R}) \cap O(2n,\mathbb{R}) &= \left\{ \begin{bmatrix} E & G \\ -G & E \end{bmatrix} \in O(2n,\mathbb{R}),\ E, G \in \mathbb{R}^{n\times n} \right\},
\end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

$$\begin{aligned} Sp(2n,\mathbb{C})\cap U(2n) &= \left\{ A = \begin{bmatrix} E & G \\ -\overline{G} & \overline{E} \end{bmatrix} \in U(2n),\ E,G\in\mathbb{C}^{n\times n} \right\}, \\ U(p,q)\cap U(n) &= \{E\oplus H,\ E\in U(p),\ H\in U(q)\}, \\ Sp^*(2n,\mathbb{C})\cap U(2n) &= \left\{ \begin{bmatrix} E & G \\ -G & E \end{bmatrix} \in U(2n),\ E,G\in\mathbb{C}^{n\times n} \right\}. \end{aligned}$$

Matrices with these double structures are likely to have good numerical properties. They also preserve the double structure of the matrices in the intersection of the corresponding Lie or Jordan algebras. For example Hamiltonian or skew-Hamiltonian structures that are also symmetric or skew-symmetric are preserved under similarity transformations with symplectic orthogonal matrices [20].

Two particular results concerning 2×2 and 4×4 real symplectic and perplectic matrices will be needed in section 4 when deriving Givens-like actions for these groups.

(i) *Symplectic orthogonals*: The set of 2×2 symplectic orthogonals is the same as $SO(2)$, the group of all 2×2 rotations. The 4×4 symplectic orthogonals can all be expressed as products

$$\begin{bmatrix} p_0 & -p_1 & -p_2 & -p_3 \\ p_1 & p_0 & -p_3 & p_2 \\ p_2 & p_3 & p_0 & -p_1 \\ p_3 & -p_2 & p_1 & p_0 \end{bmatrix} \begin{bmatrix} q_0 & 0 & q_2 & 0 \\ 0 & q_0 & 0 & q_2 \\ -q_2 & 0 & q_0 & 0 \\ 0 & -q_2 & 0 & q_0 \end{bmatrix}, \tag{3.7}$$

where $p_0^2+p_1^2+p_2^2+p_3^2=1$ and $q_0^2+q_2^2=1$ (see [20]).

(ii) *Perplectic orthogonals*: There are only four 2×2 perplectic orthogonals: $\pm I_2$ and $\pm R_2$. Any 4×4 perplectic rotation can be expressed either as a product of the form

$$\begin{bmatrix} p_0 & 0 & -p_1 & 0 \\ 0 & p_0 & 0 & p_1 \\ p_1 & 0 & p_0 & 0 \\ 0 & -p_1 & 0 & p_0 \end{bmatrix} \begin{bmatrix} q_0 & q_1 & 0 & 0 \\ -q_1 & q_0 & 0 & 0 \\ 0 & 0 & q_0 & -q_1 \\ 0 & 0 & q_1 & q_0 \end{bmatrix} \tag{3.8}$$

or of the form

$$\begin{bmatrix} 0 & -p_0 & 0 & -p_1 \\ p_0 & 0 & -p_1 & 0 \\ 0 & p_1 & 0 & -p_0 \\ p_1 & 0 & p_0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & q_0 & q_1 \\ 0 & 0 & -q_1 & q_0 \\ -q_0 & q_1 & 0 & 0 \\ -q_1 & -q_0 & 0 & 0 \end{bmatrix}, \tag{3.9}$$

where $p_0^2+p_1^2=1=q_0^2+q_1^2$ (see [33]). Additional details about the full group of 4×4 perplectic orthogonals, as well as an explicit parameterization of the group of 3×3 perplectic orthogonals can be found in [33].

4. Structured tools. For each of the eight automorphism groups listed in Table 3.1, we now describe structured matrices for performing the zeroing and scaling actions discussed in section 3.1.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

The scope and form of many of the tools is closely tied to the quadratic functional $q_{\mathrm{M}}(x) = \langle x, x\rangle_{\mathrm{M}}$ associated with the group. In large part this is due to the preservation of q-values by automorphisms, i.e., $q_{\mathrm{M}}(Gx) = q_{\mathrm{M}}(x)$ for any $G \in \mathbb{G}$, $x \in \mathbb{K}^n$ (see the proof of Proposition 3.1). We devote each of the following subsections to a particular group, and begin with a specification of its associated scalar product and quadratic functional. To simplify the formulas and emphasize the similarities and differences between the groups, we will abbreviate $q_{\mathrm{M}}(x)$ to $q(x)$.

It will become apparent that the distinction between isotropic and non-isotropic vectors is often crucial in determining the actions that can or cannot be realized in a given group. Indeed, we will see that some tools work only on non-isotropic vectors, while others work only on isotropic ones. Though there are many instances when isotropic vectors are not generic, and form a set of measure zero in $\mathbb{K}^n$, it is still important to have structured tools capable of working on them because they may of necessity be present in structured matrices. Two of the automorphism groups treated in this paper are worth mentioning in this regard. If A is an automorphism in $Sp^*(2n, \mathbb{C})$ or $\mathcal{P}(2n)$, then every column of A is isotropic; if $A \in \mathcal{P}(2n+1)$ then every column except the $(n+1)$th column is isotropic.

4.1. Complex orthogonals: $\mathbf{O}(\mathbf{n}, \mathbb{C})$.

$$\langle x, y\rangle = x^T y \in \mathbb{C}, \qquad q(x) = x^T x \in \mathbb{C}.$$

Complex orthogonal transformations have numerical uses when solving complex symmetric eigenproblems, since similarities with them preserve the complex symmetry in the problem. Such eigenproblems arise in quantum physics in the solution of differential equations such as the Schrödinger equation.

4.1.1. Givens-like action. A Givens-like action can be effected by complex orthogonal matrices of the form

$$G = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}, \quad \alpha^2 + \beta^2 = 1, \quad \alpha, \beta \in \mathbb{C}. \tag{4.1}$$

Let $x = [x_1, x_2]^T \in \mathbb{C}^2$ be non-isotropic, that is, $q(x) = x^T x \neq 0$, or equivalently, $x_1 \neq \pm i x_2$. Then choosing

$$(\alpha, \beta) = \frac{1}{\sqrt{q(x)}}(x_1, x_2), \tag{4.2}$$

gives $Gx = \sqrt{q(x)}\, e_1$. By suitable choice of the complex square root, $\sqrt{q(x)}$ can always be taken to be in the upper half-plane. We remark that there are only two choices for α, β, unlike the continuum of choices offered by (3.2) for unitary Givens. Note that in general, G will not be unitary.

By embedding G as a principal submatrix of I_n, a Givens-like action can be effected on any pair of coordinates of $x \in \mathbb{C}^n$ that do not form an isotropic 2-vector. The matrices G were used, for example, by Cullum and Willoughby [18] in their derivation of a QL procedure to compute all the eigenvalues of a complex symmetric tridiagonal matrix.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

4.1.2. Householder-like action. We list two ways of constructing matrices in $O(n,\mathbb{C})$ that perform Householder-like actions.

1. **$\mathbb{G}$-reflector:** By Theorem 3.4, when $q(x) = q(y) \neq \langle y, x\rangle$, the $\mathbb{G}$-reflector

$$H = I + \frac{(y-x)(y-x)^T}{(y-x)^T x}$$

can be used to map x to y. Isotropic x cannot be aligned with any e_j, since e_j is non-isotropic for all j. If x is non-isotropic, then with $\alpha^2 = q(x) \in \mathbb{C}$, the $\mathbb{G}$-reflector

$$H = I + \frac{(x-\alpha e_j)(x-\alpha e_j)^T}{\alpha(x_j - \alpha)} \tag{4.3}$$

has the property that $Hx = \alpha e_j$. The sign of α is chosen to ensure that $x_j - \alpha \neq 0$, or more generally, to avoid cancellation in the computation of this quantity. The formula for H can also be expressed as

$$H = I - 2\frac{uu^T}{q(u)}, \quad u = x - \alpha e_j. \tag{4.4}$$

Note that H is complex symmetric, and in general, neither Hermitian nor unitary.

2. **Composite Householder-Givens:** A non-isotropic vector $x = x_{\mathrm{R}} + i x_{\mathrm{I}}$, $x_{\mathrm{R}}, x_{\mathrm{I}} \in \mathbb{R}^n$, can be sent to $\pm\sqrt{q(x)}\, e_1$ by a product of two real orthogonal Householder reflectors (3.4) followed by a complex orthogonal G of the form (4.1)–(4.2) as follows. Let the Householder matrix H_{I} be such that $H_{\mathrm{I}} x_{\mathrm{I}} = \pm\sqrt{q(x_{\mathrm{I}})}\, e_1$ and let $H_{\mathrm{I}} x_{\mathrm{R}} = [\tilde{x}_1, \tilde{x}_{\mathrm{R}}^T]^T$, $\tilde{x}_{\mathrm{R}} \in \mathbb{R}^{n-1}$. If H_{R} is the Householder matrix sending $\tilde{x}_{\mathrm{R}}$ to $\pm\sqrt{q(\tilde{x}_{\mathrm{R}})}\, e_1$, then

$$y = (1 \oplus H_{\mathrm{R}}) H_{\mathrm{I}} x = \left[\tilde{x}_1 \pm i\sqrt{q(x_{\mathrm{I}})}, \pm\sqrt{q(\tilde{x}_{\mathrm{R}})}, 0, \ldots, 0\right]^T = [y_1, y_2, 0, \ldots, 0]^T,$$

and $[y_1, y_2]^T$ can be mapped to $\pm\sqrt{q(x)}\, e_1$ by $G = \frac{1}{\sqrt{q(y)}}\left[\begin{smallmatrix} y_1 & y_2 \\ -y_2 & y_1 \end{smallmatrix}\right]$. Hence,

$$(G \oplus I_{n-2})(1 \oplus H_{\mathrm{R}}) H_{\mathrm{I}} x = \pm\sqrt{q(x)}\, e_1. \tag{4.5}$$

This composite transformation is likely to have better numerical properties than the complex orthogonal $\mathbb{G}$-reflector in (4.4). Transformations such as (4.5) have been used by Bar-On and Ryaboy [4] and Bar-On and Paprzycki [3] to reduce a complex symmetric matrix to complex symmetric tridiagonal form.

4.1.3. Gauss-like action. From Table 3.1 we have that the only triangular matrices are diag$\{\pm 1\}$, and hence no Gauss-like action can be performed using a complex orthogonal matrix.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

4.1.4. Scaling. Since the only diagonal matrices in $O(n,\mathbb{C})$ are diag$\{\pm 1\}$ (see Table 3.1) there is no non-trivial scaling action by diagonal matrices on vectors in $\mathbb{C}^n$. On the other hand, isotropic vectors in $\mathbb{C}^2$ can be arbitrary scaled by suitably chosen non-diagonal matrices in $O(2,\mathbb{C})$. If $x \in \mathbb{C}^2$ is isotropic, then x is a complex scalar multiple of $v = [1, i]^T$ or $w = [1, -i]^T$. Since $\mathrm{diag}(1,-1)w = v$ and $\mathrm{diag}(1,-1) \in O(2,\mathbb{C})$, we may assume, without loss of generality, that our isotropic vector is a multiple of v. From Table 3.1, $A = \left[\begin{smallmatrix} \alpha & \beta \\ -\beta & \alpha \end{smallmatrix}\right] \in O(2,\mathbb{C})$ whenever $\alpha^2 + \beta^2 = 1$, with $\alpha, \beta \in \mathbb{C}$. Then an easy calculation shows $Av = (\alpha + i\beta)v \stackrel{\text{def}}{=} \lambda v$, and hence by Proposition A.4 our isotropic vector can be scaled by any desired $\lambda \in \mathbb{C} \setminus \{0\}$. The complex parameters α, β that determine the matrix A can be directly calculated from the desired scaling factor λ by the equations $\alpha = \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)$ and $\beta = \frac{1}{2i}\left(\lambda - \frac{1}{\lambda}\right)$. Thus any isotropic vector in $\mathbb{C}^2$ can be mapped to $v = [1, i]^T$ by constructing an element of $O(2,\mathbb{C})$ as described.

4.2. Real pseudo-orthogonals: $\mathbf{O(p,q,\mathbb{R})}$.

$$\langle x, y\rangle_{\Sigma_{p,q}} = x^T \Sigma_{p,q} y \in \mathbb{R}, \qquad q(x) = x^T \Sigma_{p,q} x \in \mathbb{R}.$$

Pseudo-orthogonal matrices are used in the Cholesky downdating problem [8], [44], [45] and when solving the indefinite least squares problem [10], to cite but two applications. They also play a fundamental role in the study of J-contractive matrices [39]. We refer to Higham [25] for properties of pseudo-orthogonal matrices and an algorithm for generating pseudo-orthogonal matrices with specified condition number.

Being too restrictive for some applications such as the HR factorization [14], the set of pseudo-orthogonal matrices is often extended to the set of (Σ_1, Σ_2)-orthogonal matrices. A is (Σ_1, Σ_2)-orthogonal if it satisfies $A^T \Sigma_1 A = \Sigma_2$, where Σ_1, Σ_2 are diagonal matrices with p diagonal elements equal to 1 and q diagonal elements equal to -1, and where the ordering of the diagonal elements is arbitrary. Matrices in this set do not generally belong to $O(p,q,\mathbb{R})$ and are therefore outside the scope of this paper. For details on these (Σ_1, Σ_2)-orthogonal matrices see, for example, Bojanczyk, Qiao and Steinhardt [11] and the references therein.

4.2.1. Givens-like action. A Givens-like action on non-isotropic vectors in $\mathbb{R}^2$ can be effected by matrices in $O(1,1,\mathbb{R})$ of the form

$$G = \begin{bmatrix} c & s \\ s & c \end{bmatrix}, \qquad c^2 - s^2 = 1. \tag{4.6}$$

When $c \geq 1$, then we may write $c = \cosh(\theta)$ and $s = \sinh(\theta)$, and hence these matrices have been called "hyperbolic rotations" [23].

Let $x = [x_1, x_2]^T \in \mathbb{R}^2$ be non-isotropic, that is, $q(x) = x^T \Sigma_{1,1} x = x_1^2 - x_2^2 \neq 0$. Then choosing

$$(c, s) = \begin{cases} \dfrac{1}{\sqrt{q(x)}}\,(x_1, -x_2) & \text{if } q(x) > 0, \\[2ex] \dfrac{1}{\sqrt{-q(x)}}\,(x_2, -x_1) & \text{if } q(x) < 0 \end{cases} \tag{4.7}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

gives $Gx = \sqrt{q(x)}\, e_1$ in the first case, and $Gx = \sqrt{-q(x)}\, e_2$ in the second.

Embedding G given in (4.6) and (4.7) as a principal submatrix of $I_n = I_p \oplus I_q$ in rows and columns j, k, where $1 \le j \le p < k \le n$ gives a matrix in $O(p, q, \mathbb{R})$ that zeros out either x_j or x_k of $x \in \mathbb{R}^n$, provided the vector $[x_j, x_k]^T \in \mathbb{R}^2$ is not isotropic. If $1 \le j < k \le p$, then embed the orthogonal matrix G given by equations (3.1) and (3.2) into I_p; if $p < j < k \le n$ then the embedding of the orthogonal G should be in I_q.

Bojanczyk, Brent and Van Dooren [9] noticed that the manner in which hyperbolic rotations are applied to a vector is crucial to the stability of the computation; see [11] for details on how to implement them.

4.2.2. Householder-like action. We list three ways of constructing matrices in $O(p, q, \mathbb{R})$ that perform Householder-like actions.

1. **Double Householder:** Let H_1, H_2 be $k \times k$ and $m \times m$ real orthogonal Householder matrices as in (3.4), with $k \le p$, and $m \le q$. Partitioning I_n as $I_p \oplus I_q$ and independently embedding H_1 into I_p and H_2 into I_q as principal submatrices yields an element H of $O(p, q, \mathbb{R})$. H performs independent Householder actions on k of the first p coordinates of $x \in \mathbb{R}^n$ and m of the second q coordinates of x.
2. $\mathbb{G}-$**reflector**: By Theorem 3.4, whenever $q(x) = q(y) \ne \langle y, x\rangle_{\Sigma_{p,q}}$, the $\mathbb{G}$-reflector
$$H = I + \frac{(y - x)(y - x)^T \Sigma_{p,q}}{(y - x)^T \Sigma_{p,q} x} \tag{4.8}$$
can be used to map x to y.
An isotropic vector x cannot be aligned with any e_j, since e_j is non-isotropic for all j. If x is non-isotropic, then x can be aligned with e_j for $1 \le j \le p$ if and only if $q(x) > 0$, and with e_j for $p + 1 \le j \le n$ if and only if $q(x) < 0$. Choose $\alpha \in \mathbb{R}$ so that $q(\alpha e_j) = q(x)$, or equivalently so that $\alpha^2 = q(e_j)q(x)$. Clearly this can only be done if $q(e_j)q(x) > 0$. Then with any $y = \alpha e_j$ in (4.8) such that $\operatorname{sign}(q(e_j)) = \operatorname{sign}(q(x))$ and $\alpha^2 = q(e_j)q(x)$, we have a $\mathbb{G}$-reflector
$$H = I + \frac{q(e_j)}{\alpha(x_j - \alpha)}(x - \alpha e_j)(x - \alpha e_j)^T \Sigma_{p,q} \tag{4.9}$$
with the property that $Hx = \alpha e_j$. The choice among the two possible α's is made to ensure that $x_j - \alpha \ne 0$, or more generally to avoid cancellation in the computation of $x_j - \alpha$. The formula for H can also be expressed as
$$H = I - 2\frac{uu^T \Sigma_{p,q}}{q(u)}, \quad u = x - \alpha e_j. \tag{4.10}$$
Note that H is in general neither symmetric nor orthogonal, but it is always pseudosymmetric. Putting $u = \Sigma_{p,q} v$ in the above equation yields a variation that was used by Stewart and Stewart [45]
$$H = I - 2\frac{\Sigma_{p,q} v v^T}{q(v)}. \tag{4.11}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

If v is normalized such that $v^T v = 2$, this form has the property that $\mathrm{exc}(H) = I - uu^T$ is an orthogonal Householder matrix; here $\mathrm{exc}(\cdot)$ denotes the exchange operator (see Higham [25] and references therein). On the other hand, Rader and Steinhardt [40], [41] used the non-$\mathbb{G}$-reflector but symmetric form obtained by premultiplying (4.11) by $\Sigma_{p,q}$:

$$\widetilde{H} = \Sigma_{p,q} - 2\frac{vv^T}{q(v)}, \tag{4.12}$$

which they call "hyperbolic Householder" matrices.

3. **Composite Householder-Givens:** A non-isotropic vector $x \in \mathbb{R}^n$ can be sent to a multiple of e_1 or e_{p+1} by a double Householder $H_1 \oplus H_2$ followed by a hyperbolic rotation G. Such pseudo-orthogonal transformations have been used by Bojanczyk, Higham and Patel [10] for hyperbolic QR factorizations of rectangular matrices. Tisseur [46] shows that the condition number of the transformation $G(H_1 \oplus H_2)$ is always less than or equal to the condition number of the $\mathbb{G}$-reflector (4.11) or the hyperbolic Householder matrix (4.12) performing the same action.

4.2.3. Gauss-like action. From Table 3.1 we see that upper or lower triangular matrices must be diagonal, and hence no Gauss-like actions can be performed by real pseudo-orthogonal matrices.

4.2.4. Scaling. Since the only diagonal matrices in $O(p, q, \mathbb{R})$ are $\mathrm{diag}\{\pm 1\}$ (see Table 3.1) there is no non-trivial scaling action by diagonal matrices on vectors in $\mathbb{R}^n$. On the other hand, isotropic vectors in $\mathbb{R}^2$ can be arbitrary scaled by suitably chosen non-diagonal matrices in $O(1, 1, \mathbb{R})$.

If $x \in \mathbb{R}^2$ is isotropic, then x is a real scalar multiple of $v = [1, 1]^T$ or $w = [1, -1]^T$. Since $D = \mathrm{diag}(1, -1) \in O(1, 1, \mathbb{R})$ and $Dw = v$, we may assume, without loss of generality, that our isotropic vector is a multiple of v. From Table 3.1 we see that every $A_\theta = \left[\begin{smallmatrix} c & s \\ s & c \end{smallmatrix}\right]$ with $c = \cosh\theta$, $s = \sinh\theta$ and $\theta \in \mathbb{R}$ is an element of $O(1, 1, \mathbb{R})$. Then an easy calculation shows that $A_\theta v = e^\theta v$, and hence our isotropic vector can be scaled by any desired positive real scalar $\lambda = e^\theta$. The entries of the matrix A_θ can be directly calculated from λ by the equations $c = \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)$ and $s = \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)$. Finally, since $-A_\theta \in O(1, 1, \mathbb{R})$, we can scale our isotropic vector by any positive or negative scalar, using $\pm A_\theta$ as appropriate. In summary, any isotropic vector in $\mathbb{R}^2$ can be mapped to $v = [1, 1]^T$ by constructing an element of $O(1, 1, \mathbb{R})$ as described.

4.3. Complex pseudo-orthogonals: $\mathbf{O}(\mathbf{p}, \mathbf{q}, \mathbb{C})$.

$$\langle x, y\rangle_{\Sigma_{p,q}} = x^T \Sigma_{p,q} y \in \mathbb{C}, \qquad q(x) = x^T \Sigma_{p,q} x \in \mathbb{C}.$$

4.3.1. Givens-like actions. A Givens-like action on non-isotropic vectors in $\mathbb{C}^2$ can be effected by matrices in $O(1, 1, \mathbb{C})$ of the form

$$G = \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix}, \qquad \alpha^2 - \beta^2 = 1, \quad \alpha, \beta \in \mathbb{C}. \tag{4.13}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

Let $x = [x_1, x_2]^T \in \mathbb{C}^2$ be non-isotropic, that is, $q(x) = x^T \Sigma_{1,1} x = x_1^2 - x_2^2 \neq 0$. Then choosing

$$(\alpha, \beta) = \frac{1}{\sqrt{q(x)}}(x_1, -x_2) \tag{4.14}$$

gives $Gx = \sqrt{q(x)}\, e_1$. By suitable choice of the complex square root, $\sqrt{q(x)}$ can always be taken to be in the upper half-plane. Note that G is not unitary in general.

Embedding G given in (4.13) and (4.14) as a principal submatrix of $I_n = I_p \oplus I_q$ in rows and columns j, k where $1 \leq j \leq p < k \leq n$ gives a matrix in $O(p, q, \mathbb{C})$ that zeros out either x_j or x_k of $x \in \mathbb{C}^n$, provided the vector $[x_j, x_k]^T \in \mathbb{C}^2$ is not isotropic. If $1 \leq j < k \leq p$, then embed the complex orthogonal G given by equations (4.1) and (4.2) into I_p; if $p < j < k \leq n$ then the embedding should be in I_q.

4.3.2. Householder-like action. We list three ways of constructing matrices in $O(p, q, \mathbb{C})$ that perform Householder-like actions.

1. **Double Householder**: Let H_1, H_2 be $k \times k$ and $m \times m$ complex orthogonal Householder matrices as in (4.4), with $k \leq p$, and $m \leq q$. Partitioning I_n as $I_p \oplus I_q$ and independently embedding H_1 into I_p and H_2 into I_q as principal submatrices yields an element H of $O(p, q, \mathbb{C})$. H performs independent Householder actions on k of the first p coordinates of $x \in \mathbb{C}^n$ and m of the second q coordinates of x.
2. $\mathbb{G}$**-reflector**: By Theorem 3.4, whenever $q(x) = q(y) \neq \langle y, x \rangle_{\Sigma_{p,q}}$, the $\mathbb{G}$-reflector
$$H = I + \frac{(y - x)(y - x)^T \Sigma_{p,q}}{(y - x)^T \Sigma_{p,q} x}$$
can be used to map x to y.
An isotropic vector x cannot be aligned with any e_j, since e_j is non-isotropic for all j. If x is non-isotropic, then x can be aligned with any e_j, by contrast with the real pseudo-orthogonal and complex pseudo-unitary cases, where x can only be aligned with an e_j such that $q(e_j)$ and $q(x)$ have the same (real) sign. With $\alpha \in \mathbb{C}$ chosen so that $q(\alpha e_j) = q(x)$, or equivalently so that $\alpha^2 = q(e_j) q(x)$, the $\mathbb{G}$-reflector
$$H = I + \frac{q(e_j)}{\alpha(x_j - \alpha)}(x - \alpha e_j)(x - \alpha e_j)^T \Sigma_{p,q} \tag{4.15}$$
has the property that $Hx = \alpha e_j$. The choice among the two possible α's is made to ensure that $x_j - \alpha \neq 0$, or more generally to avoid cancellation in the computation of $x_j - \alpha$.
3. **Composite Householder-Givens:** Let $x = [x_p^T, x_q^T]^T \in \mathbb{C}^n$, be non-isotropic, where $x_p \in \mathbb{C}^p$ and $x_q \in \mathbb{C}^q$. Then as long as $x_p^T x_p \neq 0$ and $x_q^T x_q \neq 0$, x can be sent to αe_1 with $\alpha^2 = q(x)$ by a double complex Householder followed by a complex pseudo-orthogonal Givens (4.13)–(4.14), in a manner similar to that described in the real pseudo-orthogonal case.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

4.3.3. Gauss-like action. From Table 3.1 we see that upper or lower triangular matrices must be diagonal and hence no Gauss-like actions can be performed by complex pseudo-orthogonal matrices.

4.3.4. Scaling. Since the only diagonal matrices in $O(p,q,\mathbb{C})$ are $\mathrm{diag}\{\pm 1\}$ (see Table 3.1) there is no non-trivial scaling action by diagonal matrices on vectors in $\mathbb{C}^n$. On the other hand, isotropic vectors in $\mathbb{C}^2$ can be arbitrary scaled by suitably chosen non-diagonal matrices in $O(1,1,\mathbb{C})$.

If $x \in \mathbb{C}^2$ is isotropic, then x is a complex scalar multiple of $v = [1,1]^T$ or $w = [1,-1]^T$. Since $D = \mathrm{diag}(1,-1) \in O(1,1,\mathbb{C})$ and $Dw = v$, we may assume, without loss of generality, that our isotropic vector is a multiple of v. From Table 3.1 we see that $A = \left[\begin{smallmatrix} \alpha & \beta \\ \beta & \alpha \end{smallmatrix}\right]$ with any $\alpha, \beta \in \mathbb{C}$ such that $\alpha^2 - \beta^2 = 1$ is an element of $O(1,1,\mathbb{C})$. Then an easy calculation shows $Av = (\alpha+\beta)v \stackrel{\text{def}}{=} \lambda v$, and hence by Proposition A.4 our isotropic vector can be scaled by any desired $\lambda \in \mathbb{C} \setminus \{0\}$. The complex parameters α, β that determine the matrix A can be directly calculated from λ by the equations $\alpha = \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)$ and $\beta = \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)$. Thus any isotropic vector in $\mathbb{C}^2$ can be mapped to $v = [1,1]^T$ by constructing an element of $O(1,1,\mathbb{C})$ as described.

4.4. Real perplectics: $\mathbf{P(n)}$.

$$\langle x, y \rangle_{\mathrm{R}} = x^T R y \in \mathbb{R}, \qquad q(x) = x^T R x \in \mathbb{R}.$$

The following definition will be useful.

DEFINITION 4.1. *A principal submatrix P of a $n \times n$ matrix A is said to be* centrosymmetrically embedded *in A if $a_{ii} \in P \Leftrightarrow a_{n-i+1,n-i+1} \in P$.*

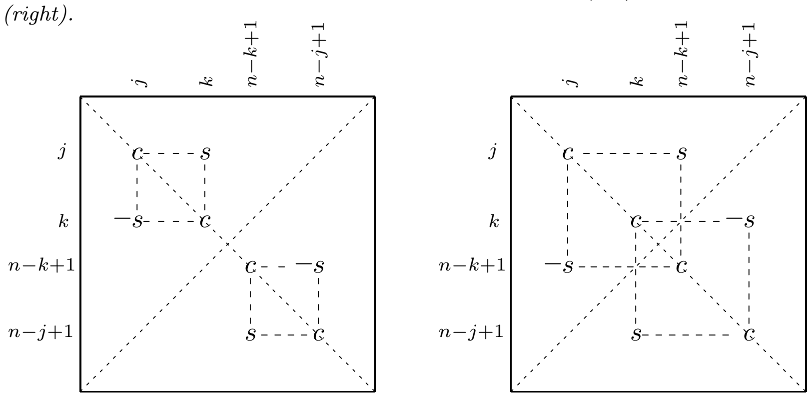
4.4.1. Givens-like action. We describe several ways of performing this action with matrices that are perplectic and orthogonal. Note that from Table 3.1 and section 3.2.3 there are no non-trivial 2×2 perplectic orthogonal matrices.

It will be convenient to use the "flip" operation [42], which transposes a matrix across its antidiagonal: $A^F \stackrel{\text{def}}{=} RA^TR$.

1. **Double Givens**: Let G denote a real 2×2 rotation. Even though G is not perplectic (other than the trivial case when $G = \pm I_2$), we can use G to build perplectic orthogonal matrices that have a Givens-like action. This is done by embedding G and G^{-F} in I_n as principal submatrices, in rows and columns $j < k < n-k+1 < n-j+1$, where $1 \le j,k \le \lfloor \frac{n}{2} \rfloor$. Depending on the action desired, there are two ways to do this. Both methods will in general zero out only one among the four affected coordinates of $x \in \mathbb{R}^n$. We note that $G^{-F} = G^T$ for any 2×2 rotation, so G^{-F} is also a rotation.
   - (i) *Direct sum embeddings* are used when a 2×2 Givens action is desired on a target pair of coordinates that are freely chosen from among either the first $m = \lfloor \frac{n}{2} \rfloor$ or the last m coordinates of $x \in \mathbb{R}^n$. As illustrated in Fig. 4.1, G is embedded in rows and columns j,k, while G^{-F} is embedded in rows and columns $n-k+1, n-j+1$.
   - (ii) *Interleaved embeddings* are used when one of the target pair of coordinates is to be chosen from among the first m and the other from among

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

FIG. 4.1. *Double Givens: perplectic direct sum embedding (left), perplectic interleaved embedding (right).*



the last m coordinates of $x \in \mathbb{R}^n$. G is embedded in rows and columns $j, n-k+1$, while G^{-F} is embedded in rows and columns $k, n-j+1$ (see Fig. 4.1).

Concentric embeddings (see Fig. 4.2) can not produce any nontrivial perplectic orthogonal matrices.

2. $\mathbf{3 \times 3}$: When $n = 2m+1$ is odd, the middle, i.e. the $(m+1)$th coordinate of $x \in \mathbb{R}^n$ cannot be reached by either of the double Givens described above in part (1). The 3×3 real perplectic orthogonal [33]

$$G = \frac{1}{2}\begin{bmatrix} 1+c & \sqrt{2}s & -1+c \\ -\sqrt{2}s & 2c & -\sqrt{2}s \\ -1+c & \sqrt{2}s & 1+c \end{bmatrix}, \quad c^2+s^2=1 \tag{4.16}$$

serves this purpose when centrosymmetrically embedded in I_n (see Definition 4.1). If $x \in \mathbb{R}^3$ with $x_2 \neq 0$, then choosing

$$c = \frac{x_1+x_3}{\sqrt{(x_1+x_3)^2+2x_2^2}}, \quad s = \frac{\sqrt{2}x_2}{\sqrt{(x_1+x_3)^2+2x_2^2}} \tag{4.17}$$

gives $y_2 = 0$ when $y = Gx$, regardless of whether x is isotropic or non-isotropic.

4.4.2. Householder-like action. We list two ways of constructing perplectic matrices that perform Householder-like actions.

1. **Double Householder**: For $k \leq m = \lfloor \frac{n}{2} \rfloor$ and $0 \neq u \in \mathbb{R}^k$, let $H(u)$ be the $k \times k$ Householder matrix given in (3.4). A centrosymmetric embedding of $H(u) \oplus H(u)^F$ into I_n (see Definition 4.1) yields a perplectic orthogonal matrix. The vector u is chosen to map k coordinates from among the first m (alternatively, from among the last m) coordinates of $x \in \mathbb{R}^n$ to a specific vector in $\mathbb{R}^k$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

2. **$\mathbb{G}$-reflector**: By Theorem 3.4, when $q(x) = q(y) \neq \langle y, x\rangle_{\mathrm{R}}$, the $\mathbb{G}-$reflector

$$G = I + \frac{(y-x)(y-x)^T R}{(y-x)^T Rx} \tag{4.18}$$

can be used to map x to y.
If n is even, all the coordinate vectors e_j are isotropic, and hence non-isotropic vectors $x \in \mathbb{R}^n$ cannot be mapped to e_j by any $G \in \mathbb{G}$. However, if x is isotropic, then taking $y = e_j$ in (4.18) gives the $\mathbb{G}-$reflector

$$G = I + \frac{(x-e_j)(x-e_j)^T R}{x_{n-j+1}} \tag{4.19}$$

with the property that $Gx = e_j$, whenever $x^T Rx = 0 \neq x_{n+j-1}$.
If $n = 2m+1$ is odd, all the coordinate vectors except e_{m+1} are isotropic. Thus if x is isotropic, then x can be mapped to e_j for any $j \neq m+1$ by the map G specified in (4.19), as long as $x_{n-j+1} \neq 0$. If x is non-isotropic, then a necessary condition for mapping x to αe_{m+1} is that $q(x) > 0$, since $q(x)$ must equal $q(\alpha e_{m+1}) = \alpha^2$. Then putting $y = \alpha e_{m+1}$ in (4.18) gives

$$G = I + \frac{(x-\alpha e_{m+1})(x-\alpha e_{m+1})^T R}{\alpha(x_{m+1}-\alpha)} \tag{4.20}$$

with the property $Gx = \alpha e_{m+1}$, whenever $\alpha^2 = x^T Rx > 0$. The sign of α is chosen so that $x_{m+1} \neq \alpha$, or more generally to avoid cancellation in the computation of $x_{m+1} - \alpha$.

4.4.3. Gauss-like action. Gauss-like actions on $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{R}^{2m}$ can be performed using *perplectic shears* $\left[\begin{smallmatrix} I & 0 \\ Z & I \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} I & Y \\ 0 & I \end{smallmatrix}\right]$, where $Y, Z \in \mathbb{R}^{m\times m}$ are perskew-symmetric (see Table 3.1). For nonzero $x, y \in \mathbb{R}^m$ we have

$$\begin{bmatrix} I & 0 \\ Z & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ Zx+y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix} \tag{4.21}$$

whenever $Zx = -y$. Because $\left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right]$ is isotropic for any $x \in \mathbb{R}^m$ and q-values must be preserved by automorphisms, this can only be achieved if $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ is also isotropic.

Suppose $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ is isotropic, i.e., $y^T Rx = 0 = x^T Ry$. Then for any k such that $x_k \neq 0$, let $w_k = e_k/x_k$ and define the $m \times m$ perskew-symmetric matrix

$$Z_k = -yw_k^T + (yw_k^T)^F = -yw_k^T + \left(\frac{e_{m-k+1}}{x_k}\right) y^T R. \tag{4.22}$$

Then $Z_k x = -y$, and (4.21) is satisfied using Z_k in place of Z. Alternatively, $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ can be mapped to $\left[\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right]$ using the upper triangular perplectic shear $\left[\begin{smallmatrix} I & Y_k \\ 0 & I \end{smallmatrix}\right]$, where the $m \times m$ perskew-symmetric matrix Y_k is given by

$$Y_k = -xv_k^T + (xv_k^T)^F = -xv_k^T + \left(\frac{e_{m-k+1}}{y_k}\right) x^T R, \tag{4.23}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

with $v_k = e_k/y_k$ and $y_k \neq 0$.

However, the typical vector $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ is not isotropic, so that $y^T Rx = \alpha \neq 0$. In this case,

$$\begin{bmatrix} I & 0 \\ Z_k & I \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ \left(\frac{\alpha}{x_k}\right) e_{m-k+1} \end{bmatrix}, \qquad \begin{bmatrix} I & Y_k \\ 0 & I \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \left(\frac{\alpha}{y_k}\right) e_{m-k+1} \\ y \end{bmatrix}.$$

Thus all but one among the first or the last m coordinates of $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ can always be zeroed out.

More generally, we may zero out a selected subset of coordinates of x or y, but at the price of a "side effect" in one coordinate. For example, suppose we wish to zero out coordinates y_j of y for all j in some index set $\mathcal{S} \subseteq \{1, 2, \dots, m\}$. Let $\widetilde{y}_{\mathcal{S}} = \sum_{j \in \mathcal{S}} y_j e_j$, and for any k such that $x_k \neq 0$, define the $m \times m$ perskew-symmetric matrix

$$W_k = -\widetilde{y}_{\mathcal{S}} w_k^T + (\widetilde{y}_{\mathcal{S}} w_k^T)^F = -\widetilde{y}_{\mathcal{S}} w_k^T + \left(\frac{e_{m-k+1}}{x_k}\right) \widetilde{y}_{\mathcal{S}}^T R, \tag{4.24}$$

where $w_k = e_k/x_k \in \mathbb{R}^m$. Then $W_k x = -\widetilde{y}_{\mathcal{S}} + \left(\frac{\widetilde{y}_{\mathcal{S}}^T Rx}{x_k}\right) e_{m-k+1}$, so that the perplectic shear $\left[\begin{smallmatrix} I & 0 \\ W_k & I \end{smallmatrix}\right]$ has the effect

$$\begin{bmatrix} I & 0 \\ W_k & I \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y - \widetilde{y}_{\mathcal{S}} + \left(\frac{\widetilde{y}_{\mathcal{S}}^T Rx}{x_k}\right) e_{m-k+1} \end{bmatrix},$$

zeroing out all the coordinates of y with indices in $\mathcal{S}$, and then altering the $(m-k+1)$st coordinate of y as a side effect. (Note that $k \in \mathcal{S}$ and $k \notin \mathcal{S}$ are both permitted here.) Care must be taken in choosing k so that this side effect is harmless in the given application. In a similar manner one may zero out a selected subset of coordinates of x using an upper triangular perplectic shear.

If $n = 2m+1$, we leave the middle coordinate invariant and then simply use the results for the even case. Note that the middle coordinate can be set to zero by a Givens-like action using the 3×3 matrix in (4.16)–(4.17).

4.4.4. Scaling. Arbitrary nonzero scaling factors d_i may be chosen to act on the first m components of a vector $x = \left[\begin{smallmatrix} y \\ z \end{smallmatrix}\right]$ where $y, z \in \mathbb{R}^m$. Then with $D = \mathrm{diag}(d_i)$, the diagonal matrix $\widetilde{D} = \mathrm{diag}(D, D^{-F})$ is perplectic, and $\widetilde{D}x = \left[\begin{smallmatrix} Dy \\ D^{-F} z \end{smallmatrix}\right]$. For example, if $x = \alpha e_1 \neq 0$ and $D = \mathrm{diag}(\alpha^{-1}, 1, \dots, 1)$ then $\widetilde{D}x = e_1$. Alternatively, the scaling factors may be chosen to act as desired on the second m components. Finally, when $x \in \mathbb{R}^{2m+1}$, $\widetilde{d}_{m+1} = \pm 1$, whereas the other d_i's for $1 \leq i \leq m$ can be chosen as desired.

4.5. Real symplectics: $\mathbf{Sp(2n}, \mathbb{R})$.

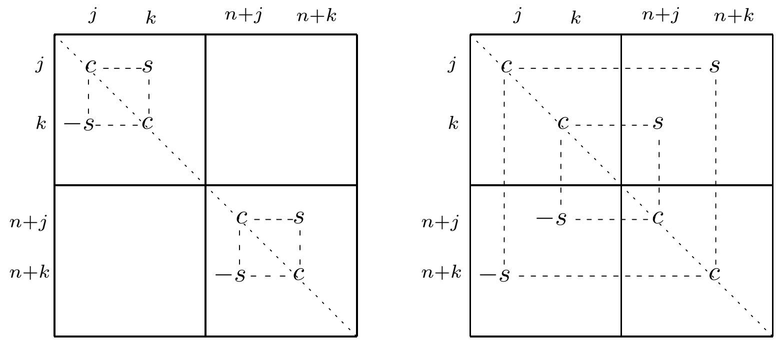
$$\langle x, y \rangle_{\mathrm{J}} = x^T J y \in \mathbb{R}, \qquad q(x) = x^T J x \equiv 0.$$

Symplectic matrices arise in a variety of scientific applications including control theory, in particular; see Faßbender [19] and references therein. The following definition will be useful.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

FIG. 4.2. *Double Givens: symplectic direct sum embedding (left), symplectic concentric embedding (right).*



DEFINITION 4.2. *A principal submatrix P of a $2n \times 2n$ matrix A is said to be* symplectically embedded *in A if $a_{ii} \in P \Leftrightarrow a_{n+i,n+i} \in P$.*

4.5.1. Givens-like action. We list three ways of constructing real symplectic orthogonal matrices that perform Givens-like actions. For brevity, let $G = \left[\begin{smallmatrix} c & s \\ -s & c \end{smallmatrix}\right]$ denote a real 2×2 Givens rotation.

1. **$\mathbf{2} \times \mathbf{2}$**: Such an action can only be performed on a restricted pair of components of $x \in \mathbb{R}^{2n}$, by symplectically embedding G into rows and columns $j, n+j$ of I_{2n}, where $1 \le j \le n$ [37]. G is chosen to zero either the jth or the $(n+j)$th component of x.
2. **Double Givens**: A coupled pair of plane rotations can be embedded as principal submatrices in rows and columns $j, k, n+j, n+k$ of I_{2n}, where $1 \le j < k \le n$. There are two ways to do this, depending on the action desired. Both methods yield symplectic orthogonal matrices and in general will zero out only one among the four affected coordinates of $x \in \mathbb{R}^{2n}$.
   (i) *Direct sum embeddings* symplectically embed $G \oplus G$ into rows and columns $j, k, n+j, n+k$ (see Fig. 4.2); they are used when a 2×2 Givens action is desired on a target pair of coordinates that are freely chosen from among either the first n or the last n coordinates of $x \in \mathbb{R}^{2n}$ [37]. These matrices are frequently used in symplectic and Hamiltonian eigenvalue problems.
   (ii) *Concentric embeddings* can be used when one of the target pair of coordinates is to be chosen from among the first n and the other independently chosen from among the last n coordinates of $x \in \mathbb{R}^{2n}$. One copy of G is embedded in rows and columns $j, n+k$, while the other copy is embedded in rows and columns $k, n+j$ (see Fig. 4.2). The concentric embedding does not seem to be as well-known as the direct sum embedding.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

3. **4 × 4**: Given $0 \neq x \in \mathbb{R}^4$, the matrix

$$G_4 = \frac{1}{\sqrt{x^T x}} \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2 & x_1 & x_4 & -x_3 \\ -x_3 & -x_4 & x_1 & x_2 \\ -x_4 & x_3 & -x_2 & x_1 \end{bmatrix} \tag{4.25}$$

is symplectic orthogonal since it is of the form (3.7) with $[q_0, q_2] = [1, 0]$ and $[p_0, p_1, p_2, p_3] = [x_1, -x_2, -x_3, -x_4]/\sqrt{x^T x}$. The transformation G_4 acts as a four-dimensional Givens rotation [20], [31], that is, if $y = G_4 x$, then $y_2 = y_3 = y_4 = 0$ and $y_1 = \sqrt{x^T x}$. Thus a symplectic embedding of G_4 into I_{2n} simultaneously zeroes out three out of the four affected components of $x \in \mathbb{R}^{2n}$. These doubly structured matrices have been used by Faßbender, Mackey and Mackey [20] when deriving Jacobi-like algorithms for doubly structured Hamiltonian eigenproblems. We refer to Tisseur [47] for a backward stable implementation of (4.25).
Symplectically embedding $G_4 \oplus G_4$ into I_{2n} (see Definition 4.2) yields a symplectic double Givens which, in general, zeroes out three of the eight affected coordinates. Various analogs of the concentric embedding described in 2(ii) can also be used.

4.5.2. Householder-like action. We list two ways of constructing symplectic matrices that perform Householder-like actions.

1. **Double Householder**: For $k \leq n$ and $0 \neq u \in \mathbb{R}^k$, let $H(u)$ be the real $k \times k$ Householder matrix given in (3.4). Symplectically embedding $H(u) \oplus H(u)$ into I_{2n} (see Definition 4.2) yields a symplectic orthogonal matrix that is usually called a *symplectic Householder* [37]. The vector u is chosen to map k coordinates from among the first n (alternatively, from among the last n) coordinates of $x \in \mathbb{R}^{2n}$ to a specific vector in $\mathbb{R}^k$. Such matrices are frequently used in Hamiltonian eigenproblems.
2. **$\mathbb{G}$-reflector**: By Theorem 3.4, the $\mathbb{G}$-reflector

$$G = I + \frac{(y-x)(y-x)^T J}{y^T J x} \tag{4.26}$$

can be used to map x to y, whenever $y^T J x \neq 0$. Taking $y = e_j$ in (4.26) gives

$$G = I + \frac{(x-e_j)(x-e_j)^T J}{\alpha}, \quad \alpha = \begin{cases} x_{n+j}, & \text{if } 1 \leq j \leq n \\ -x_{j-n}, & \text{if } n+1 \leq j \leq 2n, \end{cases} \tag{4.27}$$

with the property that $Gx = e_j$, as long as $\alpha \neq 0$. Note that these G's are real symplectic, but not orthogonal. On the other hand, they can introduce up to $2n-1$ zeros in $x \in \mathbb{R}^{2n}$ whereas symplectic Householder matrices of the form $H(u) \oplus H(u)$ generally zero out less than n components of a $2n$-vector. The transformation given in (4.26)was also described by Mehrmann in [36].

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

4.5.3. Gauss-like action. A *symplectic shear* $\left[\begin{smallmatrix} I & 0 \\ Z & I \end{smallmatrix}\right]$ with $Z^T = Z$ (see Table 3.1) can be used to zero out the last n coordinates of a vector in $\mathbb{R}^{2n}$. For nonzero $x, y \in \mathbb{R}^n$, we have

$$\begin{bmatrix} I & 0 \\ Z & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ Zx + y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix} \tag{4.28}$$

whenever $Zx = -y$. For any k such that $x_k \neq 0$, let $w_k = e_k/x_k$ and define the symmetric matrix

$$Z_k = -y w_k^T - w_k y^T + (y^T x) w_k w_k^T. \tag{4.29}$$

Then $Z_k x = -y$, so taking $Z = Z_k$ in (4.28) gives a real symplectic matrix with $Zx = -y$ as desired. In a similar fashion, one can send $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right]$ by using the symmetric matrix

$$Y_k = -x v_k^T - v_k x^T + (x^T y) v_k v_k^T, \quad \text{where} \quad v_k = e_k/y_k, \quad y_k \neq 0. \tag{4.30}$$

in the upper triangular symplectic shear $\left[\begin{smallmatrix} I & Y_k \\ 0 & I \end{smallmatrix}\right]$.

Symplectic shears can also be used to annihilate a selected subset of the components of x or y, or even just a single particular component. For example, suppose we want to zero out the components x_j for all $j \in \mathcal{S}$, where $\mathcal{S} \subseteq \{1, 2, \ldots, n\}$. Let $\widetilde{x}_{\mathcal{S}} = \sum_{j \in \mathcal{S}} x_j e_j$. Then for any k such that $y_k \neq 0$, this zeroing action can be achieved by using the symmetric matrix

$$W_k = -\widetilde{x}_{\mathcal{S}} v_k^T - v_k \widetilde{x}_{\mathcal{S}}^T + (\widetilde{x}_{\mathcal{S}}^T y) v_k v_k^T, \quad \text{where} \quad v_k = e_k/y_k \tag{4.31}$$

in the upper triangular symplectic shear $\left[\begin{smallmatrix} I & W_k \\ 0 & I \end{smallmatrix}\right]$, since $W_k y = -\widetilde{x}_{\mathcal{S}}$.

We remark that (4.30) and (4.31) are particular instances of a more general fact: an upper triangular symplectic shear can be designed to transform $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} x+u \\ y \end{smallmatrix}\right]$ for any $u \in \mathbb{R}^n$ (to get $Y_k y = u$, use $-u$ in place of x in (4.30)), thus altering x in any way we desire. This result in turn is a special case of the structured mapping theorem for Jordan algebras [35]. Similarly we may arbitrarily alter y by using lower triangular symplectic shears.

Certain special cases of the symplectic shear $\left[\begin{smallmatrix} I & W_k \\ 0 & I \end{smallmatrix}\right]$ using (4.31) have been referred to as "symplectic Gauss" transformations [19], and used to reduce a Hamiltonian matrix to J-Hessenberg form in [13]. The more general symplectic shears defined by (4.29), (4.30) and (4.31) appear to be new.

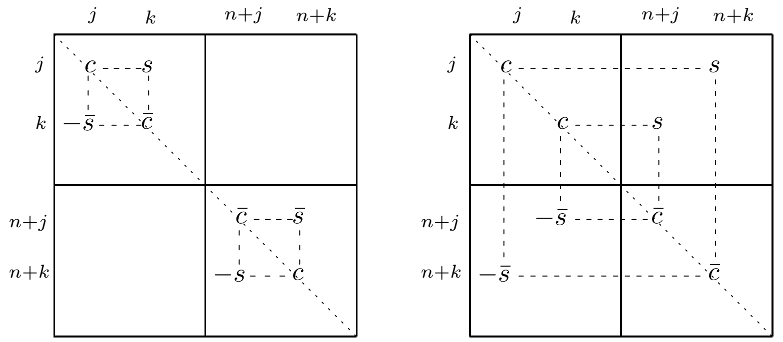
Finally, we note that symplectic shears were used for a different type of action in [21] when block-diagonalizing skew-Hamiltonian matrices by symplectic similarities.

4.5.4. Scaling. Arbitrary nonzero scaling factors d_i may be chosen to act on the first n components of a vector $x = \left[\begin{smallmatrix} y \\ z \end{smallmatrix}\right]$ where $y, z \in \mathbb{R}^n$. Then with $D = \mathrm{diag}(d_i)$, the diagonal matrix $\widetilde{D} = \mathrm{diag}(D, D^{-1})$ is symplectic, and $\widetilde{D}x = \left[\begin{smallmatrix} Dy \\ D^{-1}z \end{smallmatrix}\right]$. For example, if $x = \alpha e_1 \neq 0$ and $D = \mathrm{diag}(\alpha^{-1}, 1, \ldots, 1)$ then $\widetilde{D}x = e_1$. Alternatively, the scaling factors may be chosen to act as desired on the second n components.

Such scaling has been used by Benner [6] and Benner et al. [5] when implementing Van Loan's square reduced method for Hamiltonian matrices.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

FIG. 4.3. *Double Givens: complex symplectic direct sum embedding (left), symplectic concentric embedding (right).*



4.6. Complex symplectics: $\mathbf{Sp(2n},\mathbb{C})$.

$$\langle x, y\rangle_{\mathrm{J}} = x^T J y \in \mathbb{C}, \qquad q(x) = x^T J x \equiv 0.$$

4.6.1. Givens-like action. Direct analogy with the real symplectic case gives us the following ways of performing this action with matrices that are both complex symplectic and unitary. For brevity, let G denote a unitary 2×2 Givens as specified in (3.1)–(3.2).

1. $\mathbf{2 \times 2}$: Such an action can only be performed on a restricted pair of components of $x \in \mathbb{C}^{2n}$, by symplectically embedding G into rows and columns $j, n+j$ of I_{2n}, where $1 \le j \le n$. G is chosen to zero out either the jth or the $(n+j)$th component of x.
2. **Double Givens**: A coupled pair of Givens rotations can be embedded in two ways depending on the action desired. Both methods yield complex symplectic unitary matrices and in general will zero out only one among the four affected coordinates of $x \in \mathbb{C}^{2n}$. Here $1 \le j < k \le n$.
   (i) *Direct sum embeddings* symplectically embed $G \oplus \overline{G}$ into rows and columns $j, k, n+j, n+k$ of I_{2n} (see Fig. 4.3); they are used when a 2×2 Givens action is desired on a target pair of coordinates that are freely chosen from among either the first n or the last n coordinates of $x \in \mathbb{C}^{2n}$.
   (ii) *Concentric embeddings* can be used when one of the target pair of coordinates is to be chosen from among the first n and the other independently chosen from among the last n coordinates of $x \in \mathbb{C}^{2n}$. One copy of G is embedded in I_{2n} in rows and columns $j, n+k$, and a second copy of G (not $\overline{G}$) in rows and columns $k, n+j$ (see Fig. 4.3).
3. $\mathbf{4 \times 4}$: Unfortunately, there seems to be no 4×4 analog of (4.25) that is both complex symplectic and unitary and has Givens-like action for all $0 \neq x = [x_1, x_2, x_3, x_4]^T \in \mathbb{C}^4$. However, if the imaginary parts of $x_1\bar{x}_4$ and $x_2\bar{x}_3$ are

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

equal, then

$$G_4 = \frac{1}{\sqrt{x^* x}} \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \bar{x}_3 & \bar{x}_4 \\ -x_2 & x_1 & x_4 & -x_3 \\ -x_3 & -x_4 & x_1 & x_2 \\ -\bar{x}_4 & \bar{x}_3 & -\bar{x}_2 & \bar{x}_1 \end{bmatrix} \tag{4.32}$$

will be complex symplectic and unitary, and have a Givens-like action; that is, if $y = G_4 x$, then $y_2 = y_3 = y_4 = 0$ and $y_1 = \sqrt{x^* x}$, whenever $\mathrm{Im}(x_1 \bar{x}_4) = \mathrm{Im}(x_2 \bar{x}_3)$ and $x \neq 0$.

4.6.2. Householder-like action. We list two ways of constructing complex symplectic matrices that perform Householder-like actions.

1. **Double Householder**: For $k \leq n$ and $0 \neq u \in \mathbb{C}^k$, let $H(u)$ be the $k \times k$ Householder matrix given in (3.4). Symplectically embedding $H(u) \oplus \overline{H(u)}$ into I_{2n} (see Definition 4.2) yields a complex symplectic matrix that is also unitary. The vector u is chosen to map k coordinates from among the first n (alternatively, from among the last n) coordinates of $x \in \mathbb{C}^{2n}$ to a specific vector in $\mathbb{C}^k$.
2. **$\mathbb{G}$-reflector**: The $\mathbb{G}$-reflectors specified in (4.26) and (4.27) will map $x \in \mathbb{C}^{2n}$ to $y \in \mathbb{C}^{2n}$ or to e_j, with the same restrictions. Note that these matrices will be complex symplectic but not unitary.

4.6.3. Gauss-like action. The matrices Z_k, Y_k given in (4.29) and (4.30) are now complex symmetric, yielding complex symplectic shears $\left[\begin{smallmatrix} I & 0 \\ Z_k & I \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} I & Y_k \\ 0 & I \end{smallmatrix}\right]$ which can be used for Gauss-like zeroing actions on $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{C}^{2n}$ as described for the real symplectic case. To annihilate a selection of coordinates of x, (4.31) can be used in the complex symplectic shear $\left[\begin{smallmatrix} I & W_k \\ 0 & I \end{smallmatrix}\right]$. An analogous lower symplectic shear can be used for a similar zeroing effect on y.

4.6.4. Scaling. The scaling action for complex symplectics is similar to the real symplectic case, the only difference being that the scaling factors are complex.

4.7. Pseudo-unitaries: $\mathbf{U(p,q)}$.

$$\langle x, y \rangle_{\Sigma_{p,q}} = x^* \Sigma_{p,q} y \in \mathbb{C}, \qquad q(x) = x^* \Sigma_{p,q} x \in \mathbb{R}.$$

4.7.1. Givens-like actions. A Givens-like action on non-isotropic vectors in $\mathbb{C}^2$ can be effected by matrices in $U(1,1)$ of the form

$$G = \begin{bmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{bmatrix}, \quad |\alpha|^2 - |\beta|^2 = 1, \quad \alpha, \beta \in \mathbb{C}. \tag{4.33}$$

Let $x = [x_1, x_2]^T \in \mathbb{C}^2$ be non-isotropic, that is, $q(x) = x^* \Sigma_{1,1} x = |x_1|^2 - |x_2|^2 \neq 0$. Then choosing

$$(\alpha, \beta) = \begin{cases} \dfrac{1}{\sqrt{q(x)}} (\bar{x}_1, -\bar{x}_2) & \text{if } q(x) > 0, \\ \dfrac{1}{\sqrt{-q(x)}} (x_2, -x_1) & \text{if } q(x) < 0 \end{cases} \tag{4.34}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

gives $Gx = \sqrt{q(x)}\, e_1$ in the first case, and $Gx = \sqrt{-q(x)}\, e_2$ in the second. Note that in general, G is not unitary.

Embedding G given in (4.33)–(4.34) as a principal submatrix of $I_n = I_p \oplus I_q$ in rows and columns j, k where $1 \le j \le p < k \le n$ gives a matrix in $U(p,q)$ that zeros out either x_j or x_k, provided the vector $\left[\begin{smallmatrix} x_j \\ x_k \end{smallmatrix}\right] \in \mathbb{C}^2$ is not isotropic. If $1 \le j < k \le p$, then embed the unitary matrix G given by equations (3.1)–(3.2) into I_p in order to zero out x_j or x_k; if $p < j < k \le n$ then the embedding of the unitary G should be in I_q.

4.7.2. Householder-like action. We list three ways of constructing matrices in $U(p,q)$ that perform Householder-like actions.

1. **Double Householder**: Let H_1, H_2 be $k \times k$ and $m \times m$ unitary Householder matrices as in (3.4), with $k \le p$, and $m \le q$. Partitioning I_n as $I_p \oplus I_q$ and independently embedding H_1 into I_p and H_2 into I_q as principal submatrices yields an element H of $U(p,q)$. H performs independent Householder actions on k of the first p coordinates of $x \in \mathbb{C}^n$ and m of the second q coordinates of x.
2. $\mathbb{G}$**-reflector**: By Theorem 3.4 whenever $q(x) = q(y) \ne \langle y, x\rangle_{\Sigma_{p,q}}$, the $\mathbb{G}$-reflector

$$H = I + \frac{(y-x)(y-x)^*\Sigma_{p,q}}{(y-x)^*\Sigma_{p,q}x} \tag{4.35}$$

can be used to map x to y.
Recall that $q(x) \in \mathbb{R}$ for all $x \in \mathbb{C}^n$. An isotropic vector x cannot be aligned with any e_j, since e_j is non-isotropic for all j. If x is non-isotropic, then x can be aligned with e_j for $1 \le j \le p$ if and only if $q(x) > 0$, and with e_j for $p+1 \le j \le n$ if and only if $q(x) < 0$. Choose $\alpha \in \mathbb{C}$ so that $q(\alpha e_j) = q(x)$, or equivalently so that $|\alpha|^2 = q(e_j)q(x)$. Clearly this can only be done if $q(e_j)q(x) > 0$. Then with any $y = \alpha e_j$ in (4.35) such that $\operatorname{sign}(q(e_j)) = \operatorname{sign}(q(x))$ and $|\alpha|^2 = q(e_j)q(x)$, we have a $\mathbb{G}$-reflector

$$H = I + \frac{q(e_j)}{\overline{\alpha}(x_j - \alpha)}(x - \alpha e_j)(x - \alpha e_j)^*\Sigma_{p,q} \tag{4.36}$$

with the property that $Hx = \alpha e_j$. Usually every choice of α on the circle of radius $\sqrt{q(e_j)q(x)} = \sqrt{|q(x)|}$ in the complex plane will yield a $\mathbb{G}$-reflector. The only exception is when x_j also lies on this circle, since $\alpha = x_j$ is prohibited by (4.36). This freedom in choosing the polar angle of α can be used, for example, to ensure that the quantity $x_j - \alpha$ is not small, or to choose $\alpha \in \mathbb{R}$ thereby mapping x to a real vector, or to make the transformation (4.36) pseudo-Hermitian, or even a combination of these properties. We remark that mapping x to a real vector is of interest when tridiagonalizing a pseudo-Hermitian matrix, since the resulting tridiagonal matrix is pseudo-symmetric and the real HR algorithm [12], [14] can be used to compute its eigenvalues. To make H pseudo-Hermitian, one must choose α such that $\operatorname{sign}(\alpha) = \pm \operatorname{sign}(x_j)$. At least one of these choices (marked by $\times$ in Figure 4.4) will

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

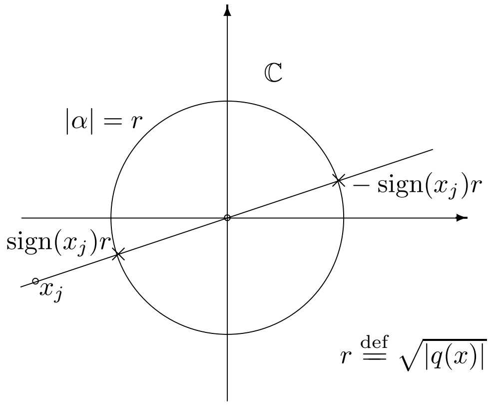


FIG. 4.4. *Circle of α's corresponding to pseudo-unitary $\mathbb{G}$-reflectors that align x with e_j.*

always be available, even when x_j lies on the circle. The formula for this pseudo-Hermitian H can be expressed as

$$H = I - 2\frac{uu^*\Sigma_{p,q}}{q(u)}, \quad u = x - \pm\operatorname{sign}(x_j)\sqrt{|q(x)|}\, e_j. \tag{4.37}$$

Finally, we remark that Rader and Steinhardt [40], [41] used the non-$\mathbb{G}$-reflector but Hermitian form obtained by post-multiplying (4.37) by $\Sigma_{p,q}$:

$$\widetilde{H} = \Sigma_{p,q} - 2\frac{uu^*}{q(u)}. \tag{4.38}$$

3. **Composite Householder-Givens:** A non-isotropic vector $x \in \mathbb{C}^n$ can be sent to a multiple of e_j with $1 \le j \le p$ if $q(x) > 0$, or $p+1 \le j \le n$ if $q(x) < 0$, by a direct sum of two independently chosen unitary Householder matrices $H_1 \oplus H_2$ with $H_1 \in U(p)$ and $H_2 \in U(q)$ followed by a 2×2 pseudo-unitary G of the form (4.33)–(4.34) appropriately embedded in I_n.

4.7.3. Gauss-like action. From Table 3.1 we see that block upper or lower triangular matrices must have the form $\left[\begin{smallmatrix} E & 0 \\ 0 & F \end{smallmatrix}\right]$, with $E \in U(p)$ and $F \in U(q)$. If E, F are also triangular, this forces them to be diagonal. Thus no Gauss-like actions can be performed by pseudo-unitary matrices.

4.7.4. Scaling. From Table 3.1 we see that diagonal pseudo-unitary matrices D have diagonal entries d_k of unit modulus. So given $z \in \mathbb{C}^n$, one may obtain $Dz = |z| \in \mathbb{R}^n$ by choosing $d_k = e^{-i\theta_k}$, where $z_k = |z_k|e^{i\theta_k}, \theta_k \in \mathbb{R}$.

For the special case when $z = [z_1, z_2]^T \in \mathbb{C}^2$ is isotropic, z can be scaled to $|z_1|[1,1]^T$, since $|z_1| = |z_2|$ in this case.

4.8. Conjugate symplectics: $\mathbf{Sp}^*(\mathbf{2n}, \mathbb{C})$.

$$\langle x, y\rangle_{\mathrm{J}} = x^*Jy \in \mathbb{C}, \qquad q(x) = x^*Jx \in i\mathbb{R}, \quad x, y \in \mathbb{C}^{2n}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

In contrast with the real and complex symplectic cases where $q \equiv 0$ so that all vectors are isotropic, for $x, y \in \mathbb{C}^n$ we now have

$$q\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) = x^*y - y^*x = 2i\,\mathrm{Im}(x^*y). \tag{4.39}$$

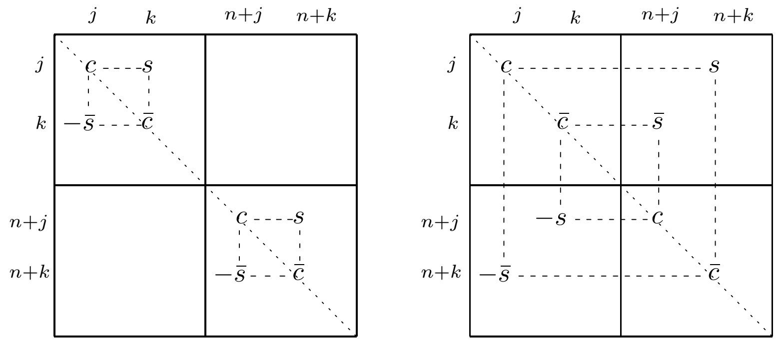
Thus $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ is isotropic if and only if $x^*y = y^*x \in \mathbb{R}$, and hence most vectors in $\mathbb{C}^{2n}$ are non-isotropic, while the coordinate vectors e_j are all isotropic. Consequently tools for zeroing actions built for $Sp(2n, \mathbb{R})$ and $Sp(2n, \mathbb{C})$ may not necessarily be directly transferable to this group.

4.8.1. Givens-like action. We list three ways of constructing conjugate symplectic unitary matrices that perform Givens-like actions.

1. $\mathbf{2 \times 2}$: Since most vectors in $\mathbb{C}^2$ are non-isotropic, while e_1, e_2 are isotropic, it is usually not possible to zero out a coordinate of $z \in \mathbb{C}^2$ using a 2×2 conjugate symplectic matrix, regardless of whether the matrix is unitary or not.
   The following construction (a special case of the intertwined Householder discussed in section 4.8.2) yields the simplest vector to which a general vector $z = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{C}^2$ can be mapped by a unitary conjugate symplectic matrix. Let $a = \frac{\overline{x+iy}}{|x+iy|}$ and $b = \frac{\overline{x-iy}}{|x-iy|}$. Then
   $$B = \frac{1}{2}\begin{bmatrix} a+b & i(a-b) \\ i(b-a) & a+b \end{bmatrix} \tag{4.40}$$
   is unitary and conjugate symplectic (see section 3.2.3), with the property that $Bz = \left[\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right]$, where
   $$\alpha = \frac{1}{2}\left(|x+iy| + |x-iy|\right) \in \mathbb{R}^+ \text{ and } \beta = \frac{i}{2}\left(|x-iy| - |x+iy|\right) = \frac{q(z)}{2\alpha} \in i\mathbb{R}$$
   .
   Observe that $Bz = \alpha e_1$ if and only if $q(z) = 0$, so that Givens-like action is indeed achieved by (4.40) on isotropic vectors. If $x,\ y \in \mathbb{R}$, then (4.40) simplifies to the ordinary real orthogonal Givens in (3.1)–(3.2).
   By symplectically embedding B into rows and columns j, $n+j$ of I_{2n}, where $1 \le j \le n$, the action defined by B can be performed on a restricted pair of coordinates of $z \in \mathbb{C}^{2n}$.
2. **Double Givens**: There are two ways to do this, depending on the action desired. Both methods yield conjugate symplectic unitary matrices and in general will zero out only one among the four affected coordinates of $x \in \mathbb{C}^{2n}$. Here $1 \le j < k \le n$. For brevity, let G denote a unitary 2×2 Givens as specified in (3.1)–(3.2).
   (i) *Direct sum embeddings* symplectically embed $G \oplus G$ in rows and columns $j, k, n+j, n+k$ (see Fig. 4.5); they are used when a 2×2 Givens action is desired on a target pair of coordinates that are freely chosen from among either the first n or the last n coordinates of $x \in \mathbb{C}^{2n}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

FIG. 4.5. *Double Givens: conjugate symplectic direct sum embedding (left), symplectic concentric embedding (right).*



(ii) *Concentric embeddings* can be used when one of the target pair of coordinates is to be chosen from among the first n and the other independently chosen from among the last n coordinates of $x \in \mathbb{C}^{2n}$. After embedding G in rows and columns $j, n+k$ of I_{2n}, embed $\overline{G}$ (not G) in rows and columns $k, n+j$ (see Fig. 4.5).

3. $\mathbf{4 \times 4}$: Unfortunately, there seems to be no 4×4 analog of (4.25) that is both conjugate symplectic and unitary, and has Givens-like action for all $0 \neq x = [x_1, x_2, x_3, x_4]^T \in \mathbb{C}^4$. However, if the imaginary parts of $x_1\bar{x}_3$ and $\bar{x}_2 x_4$ are equal, then

$$G_4 = \frac{1}{\sqrt{x^*x}} \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \bar{x}_3 & \bar{x}_4 \\ -x_2 & x_1 & x_4 & -x_3 \\ -\bar{x}_3 & -\bar{x}_4 & \bar{x}_1 & \bar{x}_2 \\ -x_4 & x_3 & -x_2 & x_1 \end{bmatrix} \tag{4.41}$$

will be conjugate symplectic and unitary, and have a Givens-like action. That is, if $y = G_4 x$, then $y_2 = y_3 = y_4 = 0$ and $y_1 = \sqrt{x^*x}$, whenever $\mathrm{Im}(x_1\bar{x}_3) = \mathrm{Im}(\bar{x}_2 x_4)$ and $x \neq 0$.

4.8.2. Householder-like action. We list three ways of constructing conjugate symplectic matrices that perform Householder-like actions.

1. **Double Householder**: For $k \leq n$ and $0 \neq v \in \mathbb{C}^k$, let $H(v)$ be the $k \times k$ Householder matrix given in (3.4). Symplectically embedding $H(v) \oplus H(v)$ into I_{2n} (see Definition 4.2) yields a conjugate symplectic matrix that is also unitary. The vector v is chosen to map k coordinates from among the first n (alternatively, from among the last n) coordinates of $x \in \mathbb{C}^{2n}$ to a specific vector in $\mathbb{C}^k$.
2. $\mathbb{G}$**-reflector**: By Theorem 3.4, whenever $q(x) = q(y) \neq \langle y, x \rangle_{\mathrm{J}}$, the $\mathbb{G}$-reflector

$$G = I + \frac{(y-x)(y-x)^*J}{(y-x)^*Jx} \tag{4.42}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

can be used to map x to y. Since all the coordinate vectors e_j are isotropic, a non-isotropic vector $x \in \mathbb{C}^{2n}$ cannot be mapped to e_j by any $G \in \mathbb{G}$. However if x is isotropic, then taking $y = e_j$ in (4.42) gives the $\mathbb{G}$-reflector

$$G = I + \frac{(x - e_j)(x - e_j)^* J}{\alpha}, \quad \alpha = \begin{cases} x_{n+j}, & \text{if } 1 \le j \le n \\ -x_{j-n}, & \text{if } n+1 \le j \le 2n, \end{cases} \tag{4.43}$$

with the property that $Gx = e_j$, as long as $\alpha \neq 0$.

3. **Intertwined Householder**: In [15] it is shown that any $z \in \mathbb{C}^{2n}$ may be mapped to $\left[\begin{smallmatrix} \alpha e_1 \\ \beta e_1 \end{smallmatrix}\right]$ by a unitary conjugate symplectic matrix, where $e_1 \in \mathbb{R}^n$, $\alpha \in \mathbb{R}^+$, $\beta \in i\mathbb{R}$, and $\alpha \ge |\beta|$. The following is a modified version of the presentation in [15].
Let $z = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ with $x, y \in \mathbb{C}^n$. Define $v_1 = x + iy$ and $v_2 = x - iy$, and find $n \times n$ unitary Householder matrices H_1 and H_2 as in (3.4) so that $H_1 v_1 = \gamma e_1$ and $H_2 v_2 = \delta e_1$, where $\gamma, \delta \in \mathbb{R}^+$. Note that H_1 and H_2 will usually not be Hermitian. Then

$$B = \frac{1}{2} \begin{bmatrix} H_1 + H_2 & i(H_1 - H_2) \\ i(H_2 - H_1) & H_1 + H_2 \end{bmatrix} \tag{4.44}$$

is unitary and conjugate symplectic, and $Bz = \left[\begin{smallmatrix} \alpha e_1 \\ \beta e_1 \end{smallmatrix}\right]$, where $\alpha = \frac{1}{2}(\delta + \gamma) \in \mathbb{R}^+$ and $\beta = \frac{1}{2} i(\delta - \gamma) = \frac{q(z)}{2\alpha} \in i\mathbb{R}$. If instead, H_1 and H_2 are chosen so that $H_1 v_1 = \gamma e_j$ and $H_2 v_2 = \delta e_j$, then the resulting B will send z to $\left[\begin{smallmatrix} \alpha e_j \\ \beta e_j \end{smallmatrix}\right]$. Whenever z is isotropic, we get $\beta = 0$ and hence $Bz = \left[\begin{smallmatrix} \alpha e_j \\ 0 \end{smallmatrix}\right]$.

4.8.3. Gauss-like action. There are important differences between the Gauss-like actions that are possible using conjugate symplectic matrices and the Gauss-like actions that have been described for real and complex symplectic matrices. These differences all stem from two sources, the contrasting nature of the q functionals described in section 4.8, and the form of shear matrices. Shears $\left[\begin{smallmatrix} I & Y \\ 0 & I \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} I & 0 \\ Z & I \end{smallmatrix}\right]$ are conjugate symplectic if and only if the $n \times n$ blocks Y and Z are Hermitian (see Table 3.1), rather than symmetric as was the case for real and complex symplectic shears.

A consequence of the presence of isotropic and non-isotropic vectors is that the zeroing actions $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right]$ as in (4.28)–(4.30) will not always be possible using conjugate symplectic shears, since both $\left[\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right]$ are isotropic, while $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ may not be.

Suppose we imitate (4.28) and (4.29): for any k such that $x_k \neq 0$, let $w_k = e_k / \overline{x}_k$ and define

$$Z_k = -y w_k^* - w_k y^* + (y^* x) w_k w_k^*, \tag{4.45}$$

with the aim of using $\left[\begin{smallmatrix} I & 0 \\ Z_k & I \end{smallmatrix}\right]$ to map $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right]$, since $Z_k x = -y$. Since Z_k will usually not be Hermitian, $\left[\begin{smallmatrix} I & 0 \\ Z_k & I \end{smallmatrix}\right]$ will unfortunately not always be conjugate symplectic. Observe that Z_k is Hermitian if and only if $y^* x \in \mathbb{R}$, that is, if and only if $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ is isotropic. A similar restriction applies to the analog of (4.30). We can map $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right]$ if and only if $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ is isotropic, that is, $x^* y \in \mathbb{R}$, using the upper triangular

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

conjugate symplectic shear $\left[\begin{smallmatrix} I & Y_k \\ 0 & I \end{smallmatrix}\right]$, where the Hermitian Y_k is given by

$$Y_k = -xv_k^* - v_k x^* + (x^*y)v_k v_k^*, \quad \text{with} \quad v_k = e_k/\overline{y}_k, \quad y_k \neq 0. \tag{4.46}$$

Since the typical vector $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{C}^{2n}$ is not isotropic, we usually cannot zero out all the coordinates of x or y using a conjugate symplectic shear. However, by using the Hermitian matrices

$$\widehat{Z}_k = -yw_k^* - w_k y^*, \qquad \widehat{Y}_k = -xv_k^* - v_k x^* \tag{4.47}$$

in place of Z_k and Y_k, respectively, one can zero out all but one among the first n or the last n coordinates of $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$, as is shown by the calculations

$$\begin{bmatrix} I & 0 \\ \widehat{Z}_k & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ (-y^*x)e_k/\overline{x}_k \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} I & \widehat{Y}_k \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (-x^*y)e_k/\overline{y}_k \\ y \end{bmatrix}.$$

More generally, we may zero out a selected subset of coordinates of x or y, but at the price of a "side-effect" in one coordinate. For example, suppose we wish to zero out coordinates x_j of x for all $j \in \mathcal{S}$, where $\mathcal{S} \subseteq \{1, 2, \ldots, n\}$. Let $\widetilde{x}_{\mathcal{S}} = \sum_{j\in\mathcal{S}} x_j e_j$, and for any k such that $y_k \neq 0$, define the Hermitian matrix

$$\widehat{W}_k = -\widetilde{x}_{\mathcal{S}} v_k^* - v_k \widetilde{x}_{\mathcal{S}}^*, \quad \text{with} \quad v_k = e_k/\overline{y}_k. \tag{4.48}$$

Then $\widehat{W}_k y = -\widetilde{x}_{\mathcal{S}} - (\widetilde{x}_{\mathcal{S}}^* y)e_k/\overline{y}_k$, so that the conjugate symplectic shear $\left[\begin{smallmatrix} I & \widehat{W}_k \\ 0 & I \end{smallmatrix}\right]$ has the effect

$$\begin{bmatrix} I & \widehat{W}_k \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - \widetilde{x}_{\mathcal{S}} - (\widetilde{x}_{\mathcal{S}}^* y)e_k/\overline{y}_k \\ y \end{bmatrix},$$

zeroing out all the coordinates of x with indices in $\mathcal{S}$, and then altering the kth coordinate of x as a side effect. (Note that both $k \in \mathcal{S}$ and $k \notin \mathcal{S}$ are permitted here.) Clearly one must exercise care in choosing k so that this side effect is harmless.

If it should happen that $\widetilde{x}_{\mathcal{S}}^* y \in \mathbb{R}$, then one may zero out all the $\mathcal{S}$-coordinates of x without any side effect by using the Hermitian matrix

$$W_k = -\widetilde{x}_{\mathcal{S}} v_k^* - v_k \widetilde{x}_{\mathcal{S}}^* + (\widetilde{x}_{\mathcal{S}}^* y)v_k v_k^* \tag{4.49}$$

in place of $\widehat{W}_k$.

In a similar manner one may zero out a selected subset of coordinates of y using a lower triangular conjugate symplectic shear, with a side effect on one coordinate.

4.8.4. Scaling. Arbitrary nonzero scaling factors d_i may be chosen to act on the first n components of a vector $x = \left[\begin{smallmatrix} y \\ z \end{smallmatrix}\right]$ where $y, z \in \mathbb{C}^n$. Then with $D = \operatorname{diag}(d_i)$, the diagonal matrix $\widetilde{D} = \operatorname{diag}(D, D^{-*})$ is conjugate symplectic, and $\widetilde{D}x = \left[\begin{smallmatrix} Dy \\ D^{-*}z \end{smallmatrix}\right]$. For example, if $x = \alpha e_1 \neq 0$ and $D = \operatorname{diag}(\alpha^{-1}, 1, \ldots, 1)$ then $\widetilde{D}x = e_1$. Alternatively, the scaling factors may be chosen to act as desired on the last n components of $x \in \mathbb{C}^{2n}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

5. Concluding summary. We have presented an extensive collection of structure-preserving transformations that we believe will be useful both in theory and in practice: in deriving structure-preserving factorizations and canonical forms, and in designing new algorithms or improving existing ones to compute them. The transformations in this paper perform the basic actions on which the majority of the algorithms of numerical linear algebra rely—introducing zeros into a vector, and scaling a vector—but they do so with the added constraint of preserving structure.

Structured tools for the three prototypical ways of introducing zeros into a vector have been provided: à la Givens, Householder and Gauss. In each case, both the scope and the restrictions on the use of each individual tool are delineated, so as to make its availability or appropriateness for the desired action more transparent. Concentric and interleaved embeddings of coupled 2×2 rotations yield new tools for Givens-like actions. The new theory of $\mathbb{G}$-reflectors developed in [34] enables a unified treatment of Householder-like actions. The repertoire of structured tools for Gauss-like actions is also significantly enhanced for the various automorphism groups.

By expressing the formulas in terms of the quadratic functional $q_{\mathrm{M}}(x) = \langle x, x\rangle_{\mathrm{M}}$, the tools for each action are presented in a parallel manner. This not only brings out the resemblances and differences between the groups, but also makes the a priori limitations on certain actions more evident. While a certain family likeness between the tools for related groups (e.g., the symplectic groups) may be expected, there can also be subtle differences between them. As an example, the transformations developed for the conjugate symplectic group are not always an automatic or obvious extension of the tools developed for the real symplectic group.

An important factor influencing the design of the toolkit provided for each automorphism group is the presence of isotropic vectors. Even when such vectors are non-generic, they cannot be ignored. For example, every column of a conjugate symplectic matrix is isotropic, even though a generic vector in this scalar product space is non-isotropic. Tools for performing zeroing and scaling actions on isotropic vectors are included in all automorphism groups in which they are applicable.

A combination of the tools described in this paper are used in [35] for mapping any vector to its "vector canonical form". An analysis of the numerical behavior of the new tools introduced here will be the subject of future work.

Appendix A. 2×2 **Forms.**

The primary aim of this appendix is to give explicit characterizations of the matrices in various 2×2 automorphism groups. To that end we first introduce some special sets of easily constructible 2×2 matrices that will be the building blocks of these characterizations:

$$
\begin{aligned}
&\mathcal{D} \stackrel{\text{def}}{=} \{\operatorname{diag}(a,d) : a = \pm 1, d = \pm 1\}, && \mathcal{D}_+ \stackrel{\text{def}}{=} \{\operatorname{diag}(1,d) : d = \pm 1\},\\
&\mathcal{D}^* \stackrel{\text{def}}{=} \{\operatorname{diag}(\alpha,\delta) : \alpha,\delta \in \mathbb{C}, |\alpha| = |\delta| = 1\}, && \mathcal{D}_+^* \stackrel{\text{def}}{=} \{\operatorname{diag}(1,\delta) : \delta \in \mathbb{C}, |\delta| = 1\},\\
&\mathcal{R}_{\mathbb{R}} \stackrel{\text{def}}{=} \left\{ \begin{bmatrix} c & s \\ -s & c \end{bmatrix} : \begin{array}{l} c = \cos\theta \\ s = \sin\theta \end{array}, \theta \in \mathbb{R} \right\}, && \mathcal{H}_{\mathbb{R}} \stackrel{\text{def}}{=} \left\{ \begin{bmatrix} c & s \\ s & c \end{bmatrix} : \begin{array}{l} c = \cosh\theta \\ s = \sinh\theta \end{array}, \theta \in \mathbb{R} \right\},\\
&\mathcal{R}_{\mathbb{C}} \stackrel{\text{def}}{=} \left\{ \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} : \begin{array}{l} \alpha,\beta \in \mathbb{C}, \\ \alpha^2 + \beta^2 = 1 \end{array} \right\}, && \mathcal{H}_{\mathbb{C}} \stackrel{\text{def}}{=} \left\{ \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} : \begin{array}{l} \alpha,\beta \in \mathbb{C}, \\ \alpha^2 - \beta^2 = 1 \end{array} \right\},
\end{aligned}
$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

$$\mathcal{R}^*_{\mathbb{C}} \overset{\text{def}}{=} \left\{ \begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{bmatrix} : \begin{array}{l} \alpha, \beta \in \mathbb{C}, \\ |\alpha|^2 + |\beta|^2 = 1 \end{array} \right\}, \quad \mathcal{H}^*_{\mathbb{C}} \overset{\text{def}}{=} \left\{ \begin{bmatrix} \alpha & \beta \\ \overline{\beta} & \overline{\alpha} \end{bmatrix} : \begin{array}{l} \alpha, \beta \in \mathbb{C}, \\ |\alpha|^2 - |\beta|^2 = 1 \end{array} \right\},$$

and $SL(2, \mathbb{K}) \overset{\text{def}}{=} \{A \in \mathbb{K}^{2\times 2} : \det A = +1\}$, where $K = \mathbb{R}, \mathbb{C}$. It is easy to check that each of the above sets of matrices forms a group, and that the following containment relations hold:

$$\begin{array}{rll} \mathcal{D}_+,\ \mathcal{D},\text{ and } \mathcal{R}_{\mathbb{R}} & \text{are subgroups of} & O(2, \mathbb{R}), \\ \mathcal{D}_+,\ \mathcal{D},\ \mathcal{R}_{\mathbb{R}},\ \mathcal{R}_{\mathbb{C}}\ ,\text{ and } O(2, \mathbb{R}) & \text{are subgroups of} & O(2, \mathbb{C}), \\ \mathcal{D}^*_+,\ \mathcal{D}^*,\text{ and } \mathcal{R}^*_{\mathbb{C}} & \text{are subgroups of} & U(2), \\ \mathcal{D}_+\ ,\ \mathcal{D},\text{ and } \mathcal{H}_{\mathbb{R}} & \text{are subgroups of} & O(1, 1, \mathbb{R}), \\ \mathcal{D}_+\ ,\ \mathcal{D}\ ,\ \mathcal{H}_{\mathbb{R}}\ ,\ \mathcal{H}_{\mathbb{C}}\ ,\text{ and } O(1, 1, \mathbb{R}) & \text{are subgroups of} & O(1, 1, \mathbb{C}), \\ \mathcal{D}^*_+\ ,\ \mathcal{D}^*,\text{ and } \mathcal{H}^*_{\mathbb{C}} & \text{are subgroups of} & U(1, 1). \end{array}$$

The first proposition gives characterizations of all 2×2 real orthogonal, complex orthogonal, and complex unitary matrices. The well-known descriptions of real orthogonals and complex unitaries are included in order to provide a context in which to view the less familiar automorphism groups.

Proposition A.1 (Orthogonals and Unitaries). *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$.*

(i) The only upper (lower) triangular matrices in $O(2, \mathbb{K})$ are the four matrices in $\mathcal{D}$. The only upper (lower) triangular matrices in $U(2)$ are the matrices in $\mathcal{D}^*$.

(ii) The real and complex 2×2 orthogonals and unitaries are characterized by

$$\begin{aligned} O(2, \mathbb{R}) &= \{RD : R \in \mathcal{R}_{\mathbb{R}}, D \in \mathcal{D}_+\}\,, \\ O(2, \mathbb{C}) &= \{RD : R \in \mathcal{R}_{\mathbb{C}}, D \in \mathcal{D}_+\}\,, \\ U(2) &= \left\{RD : R \in \mathcal{R}^*_{\mathbb{C}}, D \in \mathcal{D}^*_+\right\}. \end{aligned}$$

The results also hold if RD is replaced by DR in each of these equations.

Proof. (i): Let $A = \left[\begin{smallmatrix} \alpha & \beta \\ 0 & \delta \end{smallmatrix}\right]$ be in $O(2, \mathbb{K})$ so that $A^T A = I$. Then we have $A^T A = \left[\begin{smallmatrix} \alpha^2 & \alpha\beta \\ \alpha\beta & \beta^2+\delta^2 \end{smallmatrix}\right] = I$ if and only if $\alpha = \pm 1$, $\beta = 0$, and $\delta = \pm 1$. On the other hand if $A = \left[\begin{smallmatrix} \alpha & \beta \\ 0 & \delta \end{smallmatrix}\right] \in U(2)$, then $A^* A = \left[\begin{smallmatrix} \alpha\overline{\alpha} & \overline{\alpha}\beta \\ \alpha\overline{\beta} & \beta\overline{\beta}+\delta\overline{\delta} \end{smallmatrix}\right] = I$ if and only if $\alpha\overline{\alpha} = 1$, $\beta = 0$, and $\delta\overline{\delta} = 1$. The arguments for lower triangular matrices in $O(2, \mathbb{K})$ or $U(2)$ are almost identical.

(ii) $O(2, \mathbb{K})$: The proof is the same for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Suppose $A = \left[\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right] \in O(2, \mathbb{K})$. From $A^T A = I$, we get $\alpha^2 + \gamma^2 = 1$, so that $B \overset{\text{def}}{=} \left[\begin{smallmatrix} \alpha & \gamma \\ -\gamma & \alpha \end{smallmatrix}\right] \in \mathcal{R}_{\mathbb{K}} \subseteq O(2, \mathbb{K})$. But then the product $BA = \left[\begin{smallmatrix} 1 & \times \\ 0 & \times \end{smallmatrix}\right]$ is an upper triangular matrix in $O(2, \mathbb{K})$, so necessarily $BA = \left[\begin{smallmatrix} 1 & 0 \\ 0 & \pm 1 \end{smallmatrix}\right]$. Thus $A = B^{-1} \left[\begin{smallmatrix} 1 & 0 \\ 0 & \pm 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} \alpha & -\gamma \\ \gamma & \alpha \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 & 0 \\ 0 & \pm 1 \end{smallmatrix}\right]$.

(ii) $U(2)$: Suppose $A = \left[\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right] \in U(2)$. From $A^* A = I$, we get $|\alpha|^2 + |\gamma|^2 = 1$, so that $B \overset{\text{def}}{=} \left[\begin{smallmatrix} \bar{\alpha} & \bar{\gamma} \\ -\gamma & \alpha \end{smallmatrix}\right] \in \mathcal{R}^*_{\mathbb{C}} \subseteq U(2)$. But then the product $BA = \left[\begin{smallmatrix} 1 & \times \\ 0 & \times \end{smallmatrix}\right]$ is an upper triangular matrix in $U(2)$, so necessarily $BA = \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{smallmatrix}\right]$. Thus $A = B^{-1} \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{smallmatrix}\right] = \left[\begin{smallmatrix} \alpha & -\overline{\gamma} \\ \gamma & \overline{\alpha} \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{smallmatrix}\right]$. ∎

Proposition A.2 (Pseudo-orthogonals and Pseudo-unitaries).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

(i) The only 2×2 *upper (lower) triangular matrices in* $O(1,1,\mathbb{K})$ *where* $\mathbb{K}=\mathbb{R},\mathbb{C}$ *are the four matrices in* $\mathcal{D}$*. The only upper (lower) triangular matrices in* $U(1,1)$ *are the matrices in* $\mathcal{D}^*$*.*

(ii) The real and complex 2×2 *pseudo-orthogonals and pseudo-unitaries are characterized by*

$$\begin{aligned} O(1,1,\mathbb{R}) &= \{HD : H\in\mathcal{H}_{\mathbb{R}}, D\in\mathcal{D}\}, \\ O(1,1,\mathbb{C}) &= \{HD : H\in\mathcal{H}_{\mathbb{C}}, D\in\mathcal{D}_+\}, \\ U(1,1) &= \{HD : H\in\mathcal{H}^*_{\mathbb{C}}, D\in\mathcal{D}^*_+\}. \end{aligned}$$

The results also hold if HD *is replaced by* DH *in each of these equations.*

Proof. (i): Let $A=\left[\begin{smallmatrix}\alpha & \beta\\ 0 & \delta\end{smallmatrix}\right]\in O(1,1,\mathbb{C})$. Then

$$A^T\Sigma_{1,1}A=\begin{bmatrix}\alpha^2 & \alpha\beta\\ \alpha\beta & \beta^2-\delta^2\end{bmatrix}=\Sigma_{1,1}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix},$$

which implies that $\alpha=\pm1$, $\beta=0$ and $\delta=\pm1$. A similar argument gives the same result for lower triangular $A\in O(1,1,\mathbb{C})$. Since $O(1,1,\mathbb{R})\subseteq O(1,1,\mathbb{C})$, the same result holds for triangular matrices in $O(1,1,\mathbb{R})$. By contrast, if $A=\left[\begin{smallmatrix}\alpha & \beta\\ 0 & \delta\end{smallmatrix}\right]\in U(1,1)$, then

$$A^*\Sigma_{1,1}A=\begin{bmatrix}\alpha\overline{\alpha} & \overline{\alpha}\beta\\ \alpha\overline{\beta} & \beta\overline{\beta}-\delta\overline{\delta}\end{bmatrix}=\Sigma_{1,1}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix},$$

so that $\alpha\overline{\alpha}=1,\beta=0$, and $\delta\overline{\delta}=1$. A similar argument gives the same result for lower triangular $A\in U(1,1)$.

(ii) $O(1,1,\mathbb{R})$: Suppose that $A=\left[\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right]\in O(1,1,\mathbb{R})$. Then

$$\langle\left[\begin{smallmatrix}\alpha\\ \gamma\end{smallmatrix}\right],\left[\begin{smallmatrix}\alpha\\ \gamma\end{smallmatrix}\right]\rangle_{\Sigma_{1,1}}=\langle Ae_1,Ae_1\rangle_{\Sigma_{1,1}}=\langle e_1,e_1\rangle_{\Sigma_{1,1}}=1, \tag{A.1}$$

so that we must have $\alpha^2-\gamma^2=1$. Letting $p\stackrel{\text{def}}{=}\operatorname{sign}(\alpha), c\stackrel{\text{def}}{=}p\alpha$, and $s\stackrel{\text{def}}{=}-p\gamma$, it follows that $c^2-s^2=1$, and $c\geq 1$. Hence, there exists $\theta\in\mathbb{R}$ such that $c=\cosh\theta$ and $s=\sinh\theta$, so that $B\stackrel{\text{def}}{=}\left[\begin{smallmatrix}c & s\\ s & c\end{smallmatrix}\right]\in\mathcal{H}_{\mathbb{R}}\subseteq O(1,1,\mathbb{R})$. Then

$$B\begin{bmatrix}\alpha\\ \gamma\end{bmatrix}=\begin{bmatrix}c & s\\ s & c\end{bmatrix}\begin{bmatrix}\alpha\\ \gamma\end{bmatrix}=\begin{bmatrix}c\alpha+s\gamma\\ s\alpha+c\gamma\end{bmatrix}=\begin{bmatrix}p(\alpha^2-\gamma^2)\\ 0\end{bmatrix}=\begin{bmatrix}\pm1\\ 0\end{bmatrix}.$$

Thus $BA=\left[\begin{smallmatrix}\pm1 & \times\\ 0 & \times\end{smallmatrix}\right]$ is an upper triangular element of $O(1,1,\mathbb{R})$, so that $BA=\left[\begin{smallmatrix}\pm1 & 0\\ 0 & \pm1\end{smallmatrix}\right]$. Hence $A=B^{-1}\left[\begin{smallmatrix}\pm1 & 0\\ 0 & \pm1\end{smallmatrix}\right]=\left[\begin{smallmatrix}c & -s\\ -s & c\end{smallmatrix}\right]\left[\begin{smallmatrix}\pm1 & 0\\ 0 & \pm1\end{smallmatrix}\right]$.

(ii) $O(1,1,\mathbb{C})$: A simpler version of the previous proof works for this case. Suppose that $A=\left[\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right]\in O(1,1,\mathbb{C})$. Then $\alpha^2-\gamma^2=1$, by the computation in (A.1). But now we can immediately say that $B\stackrel{\text{def}}{=}\left[\begin{smallmatrix}\alpha & -\gamma\\ -\gamma & \alpha\end{smallmatrix}\right]\in\mathcal{H}_{\mathbb{C}}\subseteq O(1,1,\mathbb{C})$, and $BA=\left[\begin{smallmatrix}1 & \times\\ 0 & \times\end{smallmatrix}\right]$ is an upper triangular element of $O(1,1,\mathbb{C})$. Thus $BA=\left[\begin{smallmatrix}1 & 0\\ 0 & \pm1\end{smallmatrix}\right]$, so $A=B^{-1}\left[\begin{smallmatrix}1 & 0\\ 0 & \pm1\end{smallmatrix}\right]=\left[\begin{smallmatrix}\alpha & \gamma\\ \gamma & \alpha\end{smallmatrix}\right]\left[\begin{smallmatrix}1 & 0\\ 0 & \pm1\end{smallmatrix}\right]$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

(ii) $U(1,1)$: Suppose that $A = \left[\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right] \in U(1,1)$. Then the computation in (A.1) now implies that $|\alpha|^2 - |\gamma|^2 = 1$. Thus $B \stackrel{\text{def}}{=} \left[\begin{smallmatrix} \overline{\alpha} & -\overline{\gamma} \\ -\gamma & \alpha \end{smallmatrix}\right]$ is in $\mathcal{H}^*_{\mathbb{C}} \subseteq U(1,1)$. But then $BA = \left[\begin{smallmatrix} 1 & \times \\ 0 & \times \end{smallmatrix}\right]$ is an upper triangular matrix in $U(1,1)$, so that $BA = \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{smallmatrix}\right]$. Thus $A = B^{-1} \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{smallmatrix}\right] = \left[\begin{smallmatrix} \alpha & \bar{\gamma} \\ \gamma & \bar{\alpha} \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{smallmatrix}\right]$. ◻

Proposition A.3 (Real, complex, and conjugate symplectics).

$$\begin{aligned} Sp(2,\mathbb{R}) &= SL(2,\mathbb{R}), \\ Sp(2,\mathbb{C}) &= SL(2,\mathbb{C}), \\ Sp^*(2,\mathbb{C}) &= \{e^{i\theta}B : \theta \in \mathbb{R}, B \in SL(2,\mathbb{R})\}. \end{aligned}$$

Proof. $Sp(2,\mathbb{K})$: For $A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ we have $A^\star = \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix}$. Then

$$\begin{aligned} A \in Sp(2,\mathbb{K}) &\Longleftrightarrow AA^\star = I \\ &\Longleftrightarrow \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} = \begin{bmatrix} \alpha\delta - \beta\gamma & 0 \\ 0 & \alpha\delta - \beta\gamma \end{bmatrix} = I \\ &\Longleftrightarrow \det(A) = +1. \end{aligned}$$

$Sp^*(2,\mathbb{C})$: Suppose $A = e^{i\theta}B$, where $B = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in SL(2,\mathbb{R})$. Then $A^\star = \overline{e^{i\theta}} B^\star = e^{-i\theta} \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right]$, so that $A^\star A = \left[\begin{smallmatrix} ad-bc & 0 \\ 0 & ad-bc \end{smallmatrix}\right] = I$. Thus $\{e^{i\theta}B : \theta \in \mathbb{R}, B \in SL(2,\mathbb{R})\} \subset Sp^*(2,\mathbb{C})$. Conversely, suppose $A = \left[\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right] \in Sp^*(2,\mathbb{C})$. Since $|\det A| = 1$, at least two of the entries of A are nonzero. Suppose α is one of them. (The proof using some other entry is similar.) Then we can write $A = \alpha \left[\begin{smallmatrix} 1 & e \\ f & g \end{smallmatrix}\right]$, so that $A^\star = \overline{\alpha} \left[\begin{smallmatrix} \bar{g} & -\bar{e} \\ -\bar{f} & 1 \end{smallmatrix}\right]$. From the $(2,1)$ and $(2,2)$ entries of the equation $A^\star A = I$, together with the $(1,2)$ entry of $AA^\star = I$, we see that e, f, g are real, and $|\alpha|^2(g - ef) = 1$. Thus

$$A = \alpha \begin{bmatrix} 1 & e \\ f & g \end{bmatrix} = (\alpha/|\alpha|) \begin{bmatrix} |\alpha| & e|\alpha| \\ f|\alpha| & g|\alpha| \end{bmatrix} = e^{i\theta}B,$$

with $B \in SL(2,\mathbb{R})$. Thus $Sp^*(2,\mathbb{C}) \subset \{e^{i\theta}B : \theta \in \mathbb{R}, B \in SL(2,\mathbb{R})\}$. ◻

The following proposition characterizes two sets of complex scalars relevant to the "structured scaling" of certain isotropic (null) vectors.

Proposition A.4.

$$\begin{aligned} \mathcal{K}_1 &\stackrel{\text{def}}{=} \{\alpha + i\beta : \alpha, \beta \in \mathbb{C} \quad \textit{and} \quad \alpha^2 + \beta^2 = 1\} = \mathbb{C} \setminus \{0\}. \\ \mathcal{K}_2 &\stackrel{\text{def}}{=} \{\alpha + \beta : \alpha, \beta \in \mathbb{C} \quad \textit{and} \quad \alpha^2 - \beta^2 = 1\} = \mathbb{C} \setminus \{0\}. \end{aligned}$$

Proof. Let $\mathcal{E}_1$ be the range of the entire function $z \mapsto (\cos z + i \sin z) = e^{iz}$. Clearly $0 \notin \mathcal{E}_1 \subseteq \mathcal{K}_1$. But by the Little Picard Theorem [17], the range of a non-constant entire function contains all of $\mathbb{C}$ with at most one exception. Hence $\mathcal{E}_1 =$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

$\mathbb{C}\setminus\{0\} \subseteq \mathcal{K}_1$. To see that $0 \notin \mathcal{K}_1$, observe that $\alpha+i\beta = 0 \Rightarrow \alpha = -i\beta \Rightarrow \alpha^2+\beta^2 = 0$, which contradicts the definition of $\mathcal{K}_1$. Thus $\mathcal{K}_1 = \mathbb{C} \setminus \{0\}$.

Letting $\mathcal{E}_2$ be the range of the entire function $z \mapsto (\cosh z + \sinh z) = e^z$, a similar argument shows that $\mathcal{K}_2 = \mathbb{C} \setminus \{0\}$. $\square$

Acknowledgement. It is a pleasure to thank Nick Higham for helpful comments on an earlier draft of this paper.

REFERENCES

[1] Gregory Ammar, Christian Mehl, and Volker Mehrmann. Schur-like forms for matrix Lie groups, Lie algebras and Jordan algebras. *Linear Algebra Appl.*, 287(1-3), 1999.

[2] E. Anderson, Z. Bai, C. H. Bischof, S. Blackford, J. W. Demmel, J. J. Dongarra, J. J. Du Croz, A. Greenbaum, S. J. Hammarling, A. McKenney, and D. C. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, third edition, 1999.

[3] Ilan Bar-On and Marcin Paprzycki. High performance solution of the complex symmetric eigenproblem. *Numerical Algorithms*, 18(2):195–208, 1998.

[4] Ilan Bar-On and Victor Ryaboy. Fast diagonalization of large and dense complex symmetric matrices, with applications to quantum reaction dynamics. *SIAM J. Sci. Comput.*, 18(5):1412–1435, 1997.

[5] P. Benner, R. Byers, and E. Barth. Algorithm 800: Fortran 77 subroutines for computing the eigenvalues of Hamiltonian matrices I: The square-reduced method. *ACM Trans. Math. Software*, 26(1):49–77, 2000.

[6] Peter Benner. Symplectic balancing of Hamiltonian matrices. *SIAM J. Sci. Comput.*, 22(5):1885–1904, 2001.

[7] David Bindel, James Demmel, William Kahan, and Osni Marques. On computing Givens rotations reliably and efficiently. *ACM Trans. Math. Software*, 28(2):206–238, 2002.

[8] Åke Björck. *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1996.

[9] A. Bojanczyk, R. P. Brent, P. Van Dooren, and F. R. de Hoog. A note on downdating the Cholesky factorization. *SIAM J. Sci. Statist. Comput.*, 8(3):210–221, 1987.

[10] Adam Bojanczyk, Nicholas J. Higham, and Harikrishna Patel. Solving the indefinite least squares problem by hyperbolic QR factorization. Numerical Analysis Report No. 397, Manchester Centre for Computational Mathematics, Manchester, England, January 2002. To appear in SIAM J. Matrix Anal. Appl.

[11] Adam W. Bojanczyk, Sanzheng Qiao, and Allan O. Steinhardt. Unifying unitary and hyperbolic transformations. *Linear Algebra Appl.*, 316:183–197, 2000.

[12] M. A. Brebner and J. Grad. Eigenvalues of $Ax = \lambda Bx$ for real symmetric matrices A and B computed by reduction to a pseudosymmetric form and the HR process. *Linear Algebra Appl.*, 43:99–118, 1982.

[13] A. Bunse-Gerstner and V. Mehrmann. A symplectic QR like algorithm for the solution of the real algebraic Riccati equation. *IEEE Trans. Automat. Control*, 31:1104–1113, 1986.

[14] Angelika Bunse-Gerstner. An analysis of the HR algorithm for computing the eigenvalues of a matrix. *Linear Algebra Appl.*, 35:155–173, 1981.

[15] Angelika Bunse-Gerstner. Matrix factorizations for symplectic QR-like methods. *Linear Algebra Appl.*, 83:49–77, 1986.

[16] Angelika Bunse-Gerstner, Ralph Byers, and Volker Mehrmann. A chart of numerical methods for structured eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 13(2):419–453, 1992.

[17] John B. Conway. *Functions of One Complex Variable*. Springer Verlag, New York, 2nd edition, 1978.

[18] Jane K. Cullum and Ralph A. Willoughby. A QL procedure for computing the eigenvalues of complex symmetric tridiagonal matrices. *SIAM J. Matrix Anal. Appl.*, 17(1):83–109, 1996.

[19] Heike Faßbender. *Symplectic Methods for the Symplectic Eigenproblem*. Kluwer Academic Publishers, New York, 2000.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

ELA
http://math.technion.ac.il/iic/ela

[20] Heike Faßbender, D. Steven Mackey, and Niloufer Mackey. Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian problems. *Linear Algebra Appl.*, 332-334:37–80, 2001.

[21] Heike Faßbender, D. Steven Mackey, Niloufer Mackey, and Hongguo Xu. Hamiltonian square roots of skew-Hamiltonian matrices. *Linear Algebra Appl.*, 287:125–159, 1999.

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

[23] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, USA, 3rd edition, 1996.

[24] Roe Goodman and Nolan R. Wallach. *Representations and Invariants of the Classical Groups*. Cambridge University Press, Cambridge., 1998.

[25] Nicholas J. Higham. J-orthogonal matrices: Properties and generation. Numerical Analysis Report No. 408, Manchester Centre for Computational Mathematics, Manchester, England, September 2002.

[26] Alston S. Householder. *The Theory of Matrices in Numerical Analysis*. Blaisdell, New York, 1964. Reprinted by Dover, New York, 1975.

[27] Nathan Jacobson. *Basic Algebra I*. W. H. Freeman and Company, San Francisco, 1974.

[28] Peter Lancaster and Leiba Rodman. *Algebraic Riccati Equations*. Oxford University Press, 1995.

[29] Serge Lang. *Algebra*. Addison-Wesley, Reading, MA, USA, 3rd edition, 1993.

[30] R. B. Lehoucq. The computation of elementary unitary matrices. *ACM Trans. Math. Software*, 22(4):393–400, 1996.

[31] D. Steven Mackey and Niloufer Mackey. On n-dimensional Givens rotations. Numerical Analysis Report No. 423, Manchester Centre for Computational Mathematics, Manchester, England, 2003.

[32] D. Steven Mackey and Niloufer Mackey. On the determinant of symplectic matrices. Numerical Analysis Report No. 422, Manchester Centre for Computational Mathematics, Manchester, England, 2003.

[33] D. Steven Mackey, Niloufer Mackey, and Daniel M. Dunlavy. A Perplectic family of eigenproblems and algorithms for their solution. Numerical Analysis Report No. 427, Manchester Centre for Computational Mathematics, Manchester, England, 2003.

[34] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. $\mathbb{G}$-reflectors: analogues of Householder transformations in scalar product spaces. Numerical Analysis Report No. 420, Manchester Centre for Computational Mathematics, Manchester, England, February 2003. To appear in *Linear Algebra Appl.*

[35] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured mapping problems for automorphism groups, Lie algebras and Jordan algebras associated with scalar products Numerical Analysis Report No. 421, Manchester Centre for Computational Mathematics, Manchester, England, 2003.

[36] Volker Mehrmann. Der SR-algorithmus zur bestimmung der eigenwerte einer matrix. Diplomarbeit, Universität Bielefeld, 1979.

[37] C. C. Paige and Charles F. Van Loan. A Schur decomposition for Hamiltonian matrices. *Linear Algebra Appl.*, 41:11–32, 1981.

[38] Beresford N. Parlett. *The Symmetric Eigenvalue Problem*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998. Unabridged, amended version of book first published by Prentice-Hall in 1980.

[39] V. P. Potapov. The multiplicative structure of J-contractive matrix functions. *American Mathematical Society Translations, Series 2*, 15:131–243, 1960.

[40] Charles M. Rader and Allan O. Steinhardt. Hyperbolic Householder transformations. *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-34(6):1589–1602, 1986.

[41] Charles M. Rader and Allan O. Steinhardt. Hyperbolic Householder transforms. *SIAM J. Matrix Anal. Appl.*, 9(2):269–290, 1988.

[42] Russell M. Reid. Some eigenvalue properties of persymmetric matrices. *SIAM Rev.*, 39(2):313–316, 1997.

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

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 106-145, May 2003

[44] G. W. Stewart. The effects of rounding error on an algorithm for downdating a Cholesky factorization. *J. Inst. Maths. Applics.*, 23:203–213, 1979.

[45] Michael Stewart and G. W. Stewart. On hyperbolic triangularization: Stability and pivoting. *SIAM J. Matrix Anal. Appl.*, 19(4):847–860, 1998.

[46] Françoise Tisseur. Tridiagonal-diagonal reduction of symmetric indefinite pairs. Numerical Analysis Report No. 409, Manchester Centre for Computational Mathematics, Manchester, England, 2002.

[47] Françoise Tisseur. Stability of structured Hamiltonian eigensolvers. *SIAM J. Matrix Anal. Appl.*, 23(1):103–125, 2001.

[48] Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1997.