On Elementary Unitary and Φ-unitary Transformations

Xiaobai Sun*

Duke CS Preprint CS-1995-27 and PRISM Working Note #27
Nov.30, 1995

Abstract. We introduce a simple uniform structure of elementary Φ-unitary matrices. Elementary Φ-unitary transformations are the extentions of Householder transformations in real and complex spaces with a definite or an indefinite inner product induced by a nonsingular signature matrix Φ. The extentions are found useful in the computations of complex matrices or special structured matrices. For two (complex) vectors a and b, we present the sufficient and necessary conditions on the existence of an elementary Φ-unitary transformation U such that $Ua = b$. And we discuss some impacts of the theory on matrix computations involving Φ-unitary transformations.

1 Introduction

Householder transformations have been widely recognized as one of the fundamental and powerful transformations in matrix theory and computations. Given a vector, a Householder transformation is applied to introduce certain zero elements in the transformed vector. A sequence of Householder transformations is most often used in algorithms such as the QR factorization and reductions of dense matrices to various condensed forms [1]. Mathematical and numerical properties of both the real and complex Householder matrices and transformations are well understood. Given any two vectors $a, b \in \mathcal{R}^n$, $a^{\mathrm{T}}a = b^{\mathrm{T}}b$, one can always construct a Householder matrix H such that $Ha = b$, or symmetrically, $Hb = a$. It is also known all along that this does not necessarily extend to the case of $a, b \in \mathcal{C}^n$ [2]. Nevertheless, there is always an elementary unitary transformation U such that $Ua = b$ and $U^{\mathrm{H}}b = a$. A matrix is elementary if it is a rank-1 modification to the identity matrix [3]. Elementary non-Hermitian unitary transformations

*Duke University, Department of Computer Science. This work was supported by the Applied and Computational Mathematics Program, Advanced Research Projects Agency, under contract DM28E04120.

have been given in different and particular forms in terms of given a and b [4,5,6,7], partially due to the lack of recognizing the general structure of elementary unitary matrices. We show that a nontrivial elementary matrix U is unitary if and only if it can be simply expressed as

$$U = I - \frac{2}{1+i\gamma} y\, y^{\mathrm{H}}, \quad \text{with} \quad y^{\mathrm{H}} y = 1, \quad \gamma \in \mathcal{R}. \tag{1}$$

The matrix U is Hermitian and hence of Householder if and only if $\gamma = 0$. To construct a transformation U such that $Ua = b$ is then to determine the vector y and the scalar γ from a and b.

Let Φ be a nonsingular signature matrix, namely, Φ is a diagonal matrix with ± 1 diagonal elements.* A matrix U is unitary with respect to Φ, or simply Φ-unitary, if $U^{\mathrm{H}}\Phi U = \Phi$ [8]. Unitary matrices are the special case of $\Phi = \pm I$. Hyperbolic rotations are well-known 2×2 Φ-unitary transformations with $\Phi = \pm\mathrm{diag}(1, -1)$, they have been used to efficiently implement down-dating of various matrix factorizations [9,10,11,12,13]. In comparison, hyperbolic Householder transformations were first used in linear algebra algorithms by Bunse-Gerstner [14] for a theoretical interest in generalizing the QR algorithm for eigenvalue problems. They were further studied and proposed by Rader and Steinhardt [8,15] to use in practical computing such as in solving signal processing problems with least-squares-based adaptive methods. The hyperbolic Householder matrices used are of the Hermitian form $\tilde{H} = \Phi - 2yy^{\mathrm{H}}$, $y^{\mathrm{H}}\Phi y = 1$, cf., for instance, [14] and [8]. As for elementary unitary transformations, the Hermitian property needs to be relaxed as well for elementary Φ-unitary transformations in a complex space. In addition, there are elementary Φ-unitary matrices of the form $\Phi - \beta yy^{\mathrm{H}}$ with $y \neq 0$ but $y^{\mathrm{H}}\Phi y = 0$, when $\Phi \neq \pm I$. In this paper we explore first the general structure of elementary Φ-unitary matrices, which turns out to have the following uniform quasi-Householder structure

$$U = \Phi - \alpha yy^{\mathrm{H}} \tag{2}$$

for some y and nonzero α, see Section 3 for details.

The paper is outlined as follows. In Section 2, we focus on elementary unitary matrices and point it out that the extra degree of freedom in γ of (1) guarantees the existence of an elementary unitary transformation U between any two vectors of equal Euclidean length. In Section 3, we first

*In the literature of matrix computations J is often used to denote the signature matrix. We use Φ for the symmetric and sparse structure and perhaps the indefiniteness indicated in the letter.

treat Φ-unitary matrices of (2) in two separate cases, the degenerate case of $y^{\mathtt{H}}\Phi y = 0$ with $y \neq 0$ and the non-degenerate case. We then present a uniform necessary and sufficient condition on the existence of an elementary Φ-unitary transformation between two given vectors with Φ being any nonsingular signature matrix. We also discuss the impacts of the theory on matrix computations with Φ-unitary transformations. In Section 4, we describe certain relations between general and elementary Φ-unitary matrices and give a block representation for products or compositions of elementary Φ-unitary matrices. Section 5 summaries the results presented in this paper and offers some comments on related computational issues.

2 Elementary Unitary Transformations

In this section we present first a simple uniform structure of elementary unitary matrices and then describe a general approach to constructing an elementary unitary transformation U such that $Ua = b$ for given vectors a and b with $\|a\|_2 = \|b\|_2$. We show that elementary unitary transformations are the completion of Householder transformations in complex spaces. For a nonzero vector y, we use $\mathcal{R}(y)$ to denote the subspace spanned by y, and $\mathcal{R}(y)^{\perp}$, the orthogonal complementary subspace.

Let $H = I - 2yy^{\mathtt{H}}$ with $y^{\mathtt{H}}y = 1$ be a Householder matrix with the normalized Householder vector y. For any $x \in \mathcal{C}^n$, one can decompose x into $x = \alpha y + z$, where $z \in \mathcal{R}(y)^{\perp}$ and α is a scalar. We have $Hx = -\alpha y + z$. That is, the component of x along y is reflected with respect to $\mathcal{R}(y)^{\perp}$, and its component along $\mathcal{R}(y)^{\perp}$ is not altered. For two real vectors a and b with $\|a\|_2 = \|b\|_2$, there is always a Householder matrix H such that $Ha = b$, or symmetrically, $Hb = a$, cf. [3,2,16]. If a and b are complex, there exists no such a Householder transformation if $a^{\mathtt{H}}b$ is not real [2]. Nevertheless, we will show that there always exists an elementary unitary transformation U such that $Ua = b$. We will need the following lemma several times. The proof of the lemma is elementary and is hence omitted.

Lemma 1 *If $uv^{\mathtt{H}} + vu^{\mathtt{H}} = 0$, $u, v \in \mathcal{C}^n$, $u \neq 0$, then*

$$v = i\gamma u, \quad \textit{for some} \quad \gamma \in \mathcal{R}.$$

It is easily seen that if both u and v are real, then $\gamma = 0$ and $v = 0$.

Theorem 2 *A nontrivial elementary matrix U is unitary if and only if it can be represented as*

$$U = I - \frac{2}{1 + i\gamma} yy^{\mathtt{H}}, \quad \textit{with} \quad y^{\mathtt{H}}y = 1 \quad \textit{and} \quad \gamma \in \mathcal{R}. \tag{3}$$

Moreover, U is a Householder matrix if and only if $\gamma = 0$.

Proof. First, it can be checked easily that the matrix U of (3) is an elementary unitary matrix and that U is Hermitian if and only if $\gamma = 0$. Now let U be a non-trivial elementary matrix, i.e., $U = I - 2xy^{\mathtt{H}}$ for some non-zero vectors x and y. Without loss of generality we assume $y^{\mathtt{H}}y = 1$. If U is unitary, then from $UU^{\mathtt{H}} = I$ we have

$$x(y-x)^{\mathtt{H}} + (y-x)x^{\mathtt{H}} = 0. \tag{4}$$

By Lemma 1, $y - x = i\gamma x$ for some $\gamma \in \mathcal{R}$. Thus, $x = (1+i\gamma)^{-1}y$ and U can be expressed as in (3). ∎

An equivalent form of (3) is

$$U = I - \frac{2}{1+i\gamma}\frac{y\,y^{\mathtt{H}}}{y^{\mathtt{H}}\,y}, \quad y \neq 0, \quad \gamma \in \mathcal{R}. \tag{5}$$

Let U be a unitary matrix of (3). For any $x \in \mathcal{C}^n$, we have a unique decomposition of $x = \alpha y + z$, $z \in \mathcal{R}(y)^{\perp}$, and we have

$$Ux = e^{i(\pi - 2\arctan(\gamma))}\alpha y + z. \tag{6}$$

We depict the transformation in Figure 1. An elementary unitary transformation can be called a *circular* transformation not only for its unitary property that the trasformed vector is the same Euclidean length and hence on the same circle or hyper-circle, but also for its elementary property that the component of a vector along the *basis vector* y [17] is altered by a unit or 'circular' factor, $e^{i(\pi-2\arctan(\gamma))}$, while its component in $\mathcal{R}(y)^{\perp}$ remains the same. Since the unit factor is completely determined by γ, we may call γ the *circular factor.* Geometrically, a circular matrix is a reflector with respect to $\mathcal{R}(y)^{\perp}$ if and only if $\gamma = 0$.

Theorem 2 reveals the Householder-like structure for all elmentary Φ-unitary matrices, and (6) characterize the the geometric relation between x and Ux. Moreover, with a fixed basis vector y, U of (3) converges to a Householder matrix H at one end as $\gamma \to 0$ and to the identity at the other end as $\gamma \to \infty$. In between, there are infinite many non-Hermitian circular matrices. The additional degree of freedom promoted by γ underpins the following theorem on transformations with circular matrices.

Theorem 3 *Given two nonzero vectors a and b with $a^{\mathtt{H}}a = b^{\mathtt{H}}b$, if $a \neq b$, there exists a unique elementary circular unitary transformation U of (5)*

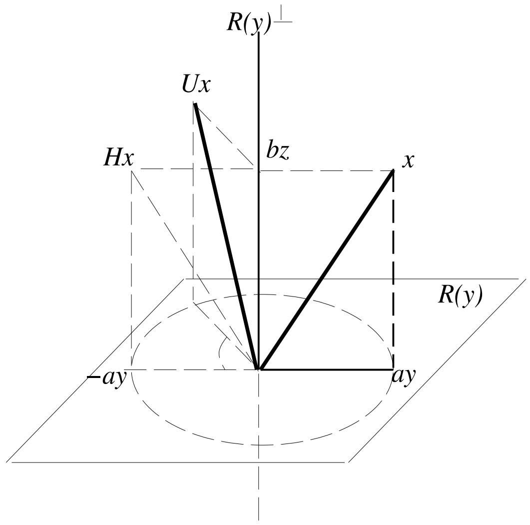

Figure 1: Elementary Unitary (Circular) Transformations

such that $Ua = b$ *with*

$$
\begin{aligned}
y &= a - b, \quad \textit{or,} \quad y = \alpha(a-b), \quad \alpha \neq 0, \\
\gamma &= \frac{\operatorname{Im}(a^{\mathtt{H}} b)}{a^{\mathtt{H}} a - \operatorname{Re}(a^{\mathtt{H}} b)}.
\end{aligned} \tag{7}
$$

Proof. From $0 < (b-a)^{\mathtt{H}}(b-a) = 2a^{\mathtt{H}}a - 2\operatorname{Re}(b^{\mathtt{H}}a)$, we know that

$$\operatorname{Re}(b^{\mathtt{H}}a) < a^{\mathtt{H}}a, \quad \text{and} \quad (b-a)^{\mathtt{H}}a \neq 0.$$

Thus, the transformation U given by (5) and (7) is well defined and it can be verified directly that $Ua = b$. Now, let U be an elementary unitary matrix such that $Ua = b$. Then, by (5),

$$a - b = \frac{2}{1 + i\gamma} \frac{y^{\mathtt{H}} a}{y^{\mathtt{H}} y} y,$$

and hence $y \in \mathcal{R}(a - b)$. Since γ is invariant to scaling in y, we may let $y = a - b$. Then,

$$1 + i\gamma = \frac{2(a-b)^{\mathtt{H}} a}{2a^{\mathtt{H}} a - 2\operatorname{Re}(a^{\mathtt{H}} b)} = \frac{a^{\mathtt{H}} a - \operatorname{Re}(b^{\mathtt{H}} a) + i \operatorname{Im}(a^{\mathtt{H}} b)}{a^{\mathtt{H}} a - \operatorname{Re}(a^{\mathtt{H}} b)},$$

and we therefore have γ as shown in (7). ■

By Theorem 3, whether the transformation matrix U is Hermitian or not depends on if $a^{\mathtt{H}}b$ is real or not. In matrix computation practice, b is often chosen a multiple of the unit vector e_1. The multiplier can be any scalar, it can be chosen for an easily obtained better numerical property [16] or for a more convenient algorithm arrangement in complex matrix computation [5,4,18]. For a fixed multiplier, a transformation does not necessarily suffer numerical stability, the reader is referred to Parlett's computational techniques in [19].

3 Elementary Φ-Unitary Transformations

In this section we extend the results in Section 2 for spaces with Euclidean metric to spaces with a pseudo-metric, in particular, a so-called 'hyperbolic-metric' induced by the following hyperbolic inner product with an indefinite signature matrix:

$$\langle a, b\rangle_\Phi \stackrel{\text{def}}{=} a^{\mathtt{H}}\Phi b, \tag{8}$$

where $a, b \in \mathcal{R}^n$ or $a, b \in \mathcal{C}^{n\dagger}$. Note that a pseudo-metric $\langle\cdot,\cdot\rangle_\Phi$ allows $\langle a, a\rangle_\Phi = 0$ for a nonzero vector a. Such extensions are called forth in computations of certain special structured matrices [8, 13, 20]. We assume for the remaining of the paper that a nonsingular signature matrix Φ is given once and for all. Unless specified otherwise, we use the simplified notation $\langle a, b\rangle$ for $\langle a, b\rangle_\Phi$.

We define that a matrix U is elementary with respect to Φ, or simply Φ-elementary, if it is a rank-1 modification to Φ :

$$U = \Phi - xy^{\mathtt{H}},$$

for some vectors x and y. When $xy^{\mathtt{H}} \neq 0$, U is said non-trivial. Recall that a matrix U is unitary with respect to Φ if $U^{\mathtt{H}}\Phi U = \Phi$. We may call such a matrix *hyperbolic-unitary* when we emphasize the indefinite property of the underlying signature matrix. We say U is elementary Φ-unitary matrix if it is both Φ-elementary and Φ-unitary. A Hermitian-like structure for all elementary Φ-unitary matrices, including $\Phi = \pm I$, is given below.

Theorem 4 *A nontrivial Φ-elementary matrix U is unitary with respect to Φ if and only if U can be represented as*

$$U = \Phi - \frac{2}{1 + i\gamma} yy^{\mathtt{H}}, \quad \textit{with} \quad \langle y, y\rangle = 1, \quad \gamma \in \mathcal{R}, \tag{9}$$

†For convenience, the inner product $\langle a, b\rangle$ defined here is linear with respect to the second argument instead of the first one.

or,

$$U = \Phi - \frac{2}{i\gamma}\, yy^{\mathtt{H}}, \quad \textit{with} \quad \langle y, y\rangle = 0, \quad \gamma \in \mathcal{R} - \{0\}. \tag{10}$$

Proof. It can be verified directly that the elementary matrix of (9) or (10) is unitary with respect to Φ. Now, let U be a non-trivial Φ-elementary matrix, i.e., $U = \Phi - 2xy^{\mathtt{H}}$ for some nonzero vectors x and y. If U is unitary with respect to Φ, we have from $\Phi = U\Phi U^{\mathtt{H}}$ that

$$x(y - (y^{\mathtt{H}}\Phi y)x)^{\mathtt{H}} + (y - (y^{\mathtt{H}}\Phi y)x)x^{\mathtt{H}} = 0.$$

For the case $y^{\mathtt{H}}\Phi y = 0$, we have $x = i\alpha y$ for some $\alpha \in \mathcal{R}$ by Lemma 1, and $\alpha \neq 0$ for $x \neq 0$. For the case $y^{\mathtt{H}}\Phi y \neq 0$, we may assume that $y^{\mathtt{H}}\Phi y = 1$. By Lemma 1, $y - \langle y, y\rangle x = i\gamma x$ for some $\gamma \in \mathcal{R}$ and hence $x = y/(1 + i\gamma)$. ■

For convenience we take the following representations equivalent to (9) and (10), respectively,

$$\begin{aligned} U &= \Phi - \frac{1}{2 + i\gamma}\, \Phi y\, y^{\mathtt{H}}\Phi, \qquad \langle y, y\rangle = 1, \quad \gamma \in \mathcal{R}; \\ U &= \Phi - \frac{2}{i\gamma}\, \Phi y\, y^{\mathtt{H}}\Phi, \qquad \langle y, y\rangle = 0, \quad \gamma \in \mathcal{R} - \{0\}. \end{aligned} \tag{11}$$

In the case $\langle y, y\rangle = 0$ and $y \neq 0$, U is non-Hermitian and we have $Uy = \Phi y$. When $\langle y, y\rangle \neq 0$, U is Hermitian if and only if $\gamma = 0$. In such case, the subspace

$$\mathcal{R}(y)^{\perp}_{\Phi} \stackrel{\text{def}}{=} \{z | \langle z, y\rangle = 0\}$$

is disjoint with $\mathcal{R}(y)$. Hence for any x, there is a unique decomposition of $x = \alpha y + z$ with $z \in \mathcal{R}(y)^{\perp}_{\Phi}$. And,

$$Ux = \Phi(e^{i(\pi - 2\arctan(\gamma))}\alpha y + z), \tag{12}$$

which is an extension of (6). The matrix U with a non-normalized vector y for such case can be represented by

$$U = \Phi - \frac{2}{1 + i\gamma}\, \frac{\Phi y\, y^{\mathtt{H}}\Phi}{\langle y, y\rangle}, \quad \langle y, y\rangle \neq 0. \tag{13}$$

Given two vectors a and b with $\langle a, a\rangle = \langle b, b\rangle$, we consider the transformation problem of finding an elementary Φ-unitary matrix U such that

$$Ua = \Phi b. \tag{14}$$

We have seen the convenience of such form from the geometrical viewpoints discussed above. This form of (14) is symmetrical in the sense that $U^{\mathtt{H}}b = \Phi a$ and $U^{\mathtt{H}}$ is also an elementary Φ-unitary matrix. For the resolution of problem (14), we need first to distinguish two cases according to whether the 'distance' between a and b is zero or not.

Theorem 5 *Given two vectors a and b with $\langle a, a\rangle = \langle b, b\rangle$. If $\langle a-b, a-b\rangle \neq 0$, there exists a unique elementary Φ-unitary transformation U of (13) such that $Ua = \Phi b$ with*

$$\begin{aligned} y &= a - b, \quad \textit{or,} \quad y = \alpha(a-b) \quad \textit{for some } \alpha \neq 0, \\ \gamma &= \frac{\mathrm{Im}(\langle a, b\rangle)}{\langle a, a\rangle - \mathrm{Re}(\langle a, b\rangle)}. \end{aligned} \tag{15}$$

Proof. Since $\langle y, y\rangle = |\alpha|^2(a-b)^{\mathtt{H}}\Phi(a-b) = 2|\alpha|^2(a^{\mathtt{H}}\Phi a - \mathrm{Re}(a^{\mathtt{H}}\Phi b) \neq 0$, the matrix U of (13) and (15) is well defined. It is a straightforward verification that $Ua = \Phi b$. On the other hand, let V be an elementary Φ-unitary matrix such that $Va = \Phi b$. According to Theorem 4, $V = \Phi - \beta z z^{\mathtt{H}}$ for some z and β. From $Va = \Phi b$, we have $z = \mathcal{R}(\Phi(a-b))$ and hence $y = \mathcal{R}(a-b)$ by our convention in (11). It can then be shown easily that $V = U$. ∎

Theorem 5 is an extension of Theorem 3 to the case of non-zero 'distance' between a and b. We now turn to the degenerate case $\langle a-b, a-b\rangle = 0$. If $a = b$, U can be the trivial matrix Φ. If $a \neq b$, the transformation, if exists, can not be trivial. By Theorem 4 and our convention in (11), the solution to the problem (14) must meet the requirement that $y = \alpha(a-b)$ for some nonzero α and that $0 \neq \gamma = 2\langle y, a\rangle = 2\alpha\langle a-b, a\rangle$. Since $0 = \langle a-b, a-b\rangle = 2(\langle a, a\rangle - \mathrm{Re}(\langle a, b\rangle)$, $\langle a-b, a\rangle = -\mathrm{Im}(\langle b, a\rangle) = \mathrm{Im}(\langle a, b\rangle)$. Thus, $\gamma \neq 0$ if and only if $\mathrm{Im}(\langle a, b\rangle) \neq 0$. We have proved the following theorem.

Theorem 6 *Given two distinctive vectors a and b with $\langle a, a\rangle = \langle b, b\rangle$. If $\langle a-b, a-b\rangle = 0$, then there exists an elementary Φ-unitary transformation U such that $Ua = \Phi b$ if and only if $\mathrm{Im}(\langle a, b\rangle) \neq 0$. Moreover, when $\mathrm{Im}(\langle a, b\rangle) \neq 0$, the transformation U is of the form (10) with*

$$y = a - b, \quad \textit{and} \quad \gamma = \frac{\mathrm{Im}(\langle a, b\rangle)}{2}, \tag{16}$$

or more generally,

$$y = \alpha(a-b), \quad \textit{and} \quad \gamma = \frac{\mathrm{Im}(\langle a, b\rangle)}{2\alpha^2}, \quad \textit{with} \quad \alpha \neq 0,$$

and U is unique.

The above two theorems can be summarized in the following corollary.

Corollary 7 *Given two distinctive vectors a and b with $\langle a, a\rangle = \langle b, b\rangle$. There exists an elementary Φ-unitary transformation U such that $Ua = \Phi b$ if and only if*

$$\langle a - b, a\rangle = -\langle b, a - b\rangle \neq 0. \tag{17}$$

Proof. Part 1, the condition (17) is necessary. In the case $\langle a-b, a-b\rangle \neq 0$, if $\langle a - b, a\rangle = 0$, then $\langle b, a\rangle \in \mathcal{R}$ and $\langle a - b, a - b\rangle = 2\langle a - b, a\rangle = 0$, a contradiction. In the case $\langle a-b, a-b\rangle = 0$, we have $\mathrm{Im}(\langle a, b\rangle) = \langle a-b, a\rangle$. If U exists, $\langle a-b, a\rangle \neq 0$ by Theorem 6. Part 2, the condition (17) is sufficient. Assume $\langle a - b, a\rangle \neq 0$. If $\mathrm{Re}(\langle a - b, a\rangle) \neq 0$, then $\langle a - b, a - b\rangle \neq 0$, and U exists as described in Theorem 5; otherwise, $\langle a - b, a\rangle = \mathrm{Im}(\langle a - b, a\rangle)$, U exists as described in Theorem 6. ■

The condition $\langle a - b, a\rangle \neq 0$ is recognized as a necessary condition on the existence of U for the problem equivalent to (14) in [6,7]. By exploring the general structure of elementary Φ-unitary matrices, we have shown that the uniform condition (17) is both necessary and sufficient.

Example 8 *Let $\Phi = diag(1, 1, -1)$, $a = (1, 1, 1)^{\mathtt{H}}$, and $b = (1, 0, 0)^{\mathtt{H}}$. Then $\langle a, a\rangle = \langle b, b\rangle = 1$ and $\langle b, a - b\rangle = 0$. According to Corollary 7, there is no elementary Φ-unitary matrix U such that $Ua = \Phi b$. However, for $\tilde{b} = e^{i\theta} b$ with $e^{i\theta} \neq 1$, $\langle \tilde{b}, \tilde{b}\rangle = \langle b, b\rangle = \langle a, a\rangle$, and $\langle a - \tilde{b}, \tilde{b}\rangle = e^{i\theta} - 1 \neq 0$.*

We see from Example 8 that the necessary and sufficient condition in Corollary 7 actually suggests an approach to turning a failing case to a successful case. Once the condition is satisfied, one can construct the transformation matrix according to either Theorem 5 or Theorem 6.

As mentioned earlier, b is most often a multiple of the unit vector e_1, $b = \beta e_1$ for some nonzero β. With such b, the condition of $\langle a - b, b\rangle \neq 0$ in (17) is easier to check, that is, one can tell if the condition holds from one single element of $a-b$. Furthermore, for any a, one can always pick up a vector $b = e^{i\theta} e_1$ such that $\langle a - b, b\rangle \neq 0$, as illustrated in Example 8. On the other side, since the condition $\langle a, a\rangle = \langle b, b\rangle$ must be met, $b = \beta e_1$ with $\beta \neq 0$ requires that $\langle a, a\rangle \neq 0$. In summary, for any a, $\langle a, a\rangle \neq 0$, there always exist a vector b as a multiple of e_1 and an elementary Φ-unitary transformation U such that $Ua = \Phi b$, or simply, $Ua = \pm\beta e_1$ for some nonzero multiplier β. Notice that the multiplier β can always be chosen so that the transformation is also Hermitian. The Hermitian property would not be brought back if the multiplier is not allowed to be flexible. For the case $\langle a, a\rangle = 0$ for a nonzero a, it is clear that b must be nonzero and that b

must be chosen different from a multiple of e_1, which imply either a fatal break-down in a factorization algorithm and/or invoke an essential change in algorithm design and development.

4 General and Block Φ-Unitary Transformations

In this section we discribe some relations among Φ-unitary matrices for their impacts on matrix computations. We start with a few basic facts, their proofs are simple and hence are omitted.

a. A product of Φ-unitary matrices is Φ-unitary.
b. A block triangular Φ-unitary matrix is block diagonal. In particular, a triangular Φ-unitary matrix is diagonal.
c. A diagonal unitary matrix is Φ-unitary with any signature matrix Φ.

Theorem 9 *A Φ-unitary matrix of order n can be expressed as a product of at most n elementary Φ-unitary matrices with a diagonal unitary matrix.*

Proof. Let F be an $n \times n$ Φ-unitary matrix. We prove the theorem by induction on the order n. When $n = 1$, F is a scalar with $|F| = 1$, and hence a diagonal unitary matrix. Let $n > 1$. Assume that the theorem is true for matrices of order $n-1$ with respect to $\Phi_{n-1} = \Phi(2:n, 2:n)$. Let $a = Fe_1$. If a is a multiple of e_1, $a = e^{i\theta_1}e_1$, then $F = e^{i\theta_1} \oplus F_{n-1}$ (2×2 block diagonal) and F is a product of elementary Φ-unitary matrices by the induction hypothesis on F_{n-1}. Otherwise, $a^{\mathtt{H}}\Phi a = \phi_{11} \neq 0$. From the discuss in Section 3, there is an elementary Φ-unitary matrix U_1 and a unit scalar $e^{i\theta_1}$ such that $U_1 a = e^{i\theta_1}e_1$. Thus, $U_1 F$ is a product of elementary Φ-unitary matrices and so is F. ■

Theorem 9 tells us that a general Φ-unitary transformation can be equivalently and effectively carried out by a sequence of elementary Φ-unitary transformations. On the other hand, for a sequence of elementary transformations

$$U_j a_{j-1} = \Phi a_j, \quad \text{or,} \quad V_j \stackrel{\text{def}}{=} \Phi U_j a_{j-1} = a_j, \quad 1 \leq j \leq k,$$

one may be also interested in effcient computations by employing a block representation for the composition of the transformations [21]. Let $a = a_0$, $b = a_k$ and

$$G \stackrel{\text{def}}{=} \prod_{j=1:k} \Phi U_j = \prod_{j=1:k} V_j.$$

Then $Ga = b$. Let $U = \Phi G$, then $Ua = \Phi b$, which is of the same form as (14). Here we give a block representation of ΦG, based solely on the structure of elementary Φ-unitary matrices.

Theorem 10 *Assume $k < n$. Let $G = \prod_{j=1:k} \Phi U_j$, where U_j are elementary Φ-unitary matrices of order n as in (11). Then,*

$$\Phi G = \Phi - \Phi Y R^{-1} Y^{\mathtt{H}} \Phi, \quad R = D + \mathrm{triu}(Y^{\mathtt{H}} \Phi Y, 1), \tag{18}$$

where $Y = (y_1, \ldots, y_k)$, $\mathrm{triu}(A, 1)$ is the strict upper triangular part of A, and $D = \mathrm{diag}(\delta_j)$ with

$$\delta_j = \left\{ \begin{array}{ll} i\gamma_j/2, & \textit{if } \langle y_j, y_j \rangle = 0, \\ (1 + i\gamma_j)/2, & \textit{if } \langle y_j, y_j \rangle = 1, \end{array} \right\} \neq 0, \quad j = 1 : k. \tag{19}$$

Proof. We prove the result by induction on k. When $k = 1$, $Y = y_1$ and $R = \delta_1$ and it is true that $G = U_1 = \Phi - \Phi Y R^{-1} Y \Phi$. Let $k > 1$ and $G_{k-1} = \prod_{j=1:k-1} \Phi U_j$. Assume that the theorem is true for $\Phi G_{k-1} = \Phi - \Phi Y_{k-1} R_{k-1}^{-1} Y_{k-1} \Phi$. Then,

$$\Phi G = \Phi G_{k-1} \Phi U_k = (\Phi - \Phi Y_{k-1} R_{k-1}^{-1} Y_{k-1} \Phi) \Phi (\Phi - \Phi y_k y_k^{\mathtt{H}} \Phi / \delta_k)$$

and hence

$$\begin{aligned} \Phi G &= \Phi - \Phi (Y_{k-1}, y_k) \begin{pmatrix} R_{k-1}^{-1} & -\delta_k^{-1} R_{k-1}^{-1} Y_{k-1}^{\mathtt{H}} \Phi y_k \\ 0 & \delta_k^{-1} \end{pmatrix} \begin{pmatrix} Y_{k-1}^{\mathtt{H}} \\ y_k^{\mathtt{H}} \end{pmatrix} \Phi \\ &= \Phi - \Phi (Y_{k-1}, y_k) \begin{pmatrix} R_{k-1} & Y_{k-1}^{\mathtt{H}} \Phi y_k \\ 0 & \delta_k \end{pmatrix}^{-1} \begin{pmatrix} Y_{k-1}^{\mathtt{H}} \\ y_k^{\mathtt{H}} \end{pmatrix} \Phi \\ &= \Phi - \Phi (Y_{k-1}, y_k) R_k^{-1} \begin{pmatrix} Y_{k-1}^{\mathtt{H}} \\ y_k^{\mathtt{H}} \end{pmatrix} \Phi. \end{aligned}$$

From the illustrated relation between R_k and R_{k-1} and by the induction hypothesis, the theorem is true for G_k as well. ■

The above theorem is a generalization of the compact WY representation for products of real Householder matrices [22,23,24]. If all transformations are Hermitian, then $D = I/2$, and ΦG can be determined by Y only; otherwise, $\{\gamma_j\}_{j=1:k}$ are needed. The theory here can also lead to an extension of the basis-kernel representation for orthogonal matrices [17].

We are now in a position to promote an alternative and equivalent uniform form for Φ-unitary matrices. Notice that the matrices $V_j \stackrel{\mathrm{def}}{=} \Phi U_j$ defined above are of the form

$$V \stackrel{\mathrm{def}}{=} \Phi U = I - \delta\, y\, y^{\mathtt{H}} \Phi \quad \text{for some} \quad \delta \neq 0. \tag{20}$$

We call it the *V-form* for elementary Φ-unitary matrices, in comparison to the *U-form* we have used so far. Here V is elementary with respect to

I. Though not symmetrical nor traditional, the V-form offers more convenience in expressing relations among matrices or transformed vectors as listed below.

(i) The inverse $V^{-1} = \Phi U^{\mathtt{H}} \Phi = I - \delta^{\mathtt{H}} y\, y^{\mathtt{H}} \Phi$ is of the same form as V except that the scalar δ is replaced by its conjugate.

(ii) The problem of (14) with a U-form transformation is equivalent to the problem of $Va = b$ with a V-form transformation. The symmetrical positions between a and b are preserved by the common structure of V and V^{-1}.

(iii) Let V of (20) be defined with $\langle y, y \rangle = 1$ and $\delta = 2/(1+i\gamma)$, cf. (11) and (20). Let $x = \alpha y + z$ with $z \in \mathcal{R}(z)^{\perp}_{\Phi}$. Then, $Vx = e^{i(\pi - 2\arctan(\gamma))}\alpha y + z$, cf. (6) and (12).

(iv) Let G be defined as above. Then, $G = I - YR^{-1}Y^{\mathtt{H}}\Phi$ and $G^{-1} = \Phi G^{\mathtt{H}} \Phi = I - YR^{-\mathtt{H}}Y^{\mathtt{H}}\Phi$ where R is as shown in Theorem 10.

5 Concluding Remarks

We have presented a simple U-form structure for all elementary Φ-unitary matrices, whether such a matrix is Hermitian or not depend on its scalar in the U-form is real or complex. The V-form is the composition of the U-form with the signature matrix Φ, and it is convenient in expressing relationships among matrices and transformed vectors. The analysis of the general structure has given more insights into matrix computations with such transformations. In particular, the necessary and sufficient conditions on the existence of a desired elementary Φ-unitary transformation tell us when a fatal breakdown occurs and when the success of a transformation or a sequence of transformations can be ensured.

Our results are helpful to the analysis of some numerical issues and the design of implementation details. For example, even with unitary transformations between complex vectors, one should be careful about the merge of the 'circular' scalar $2/(1+i\gamma)$ and the the scaling factor $y^{\mathtt{H}}y$ of (5). Since γ can be arbitrarily large by Theorem 3, a merging scalar well-bounded from below and above does not imply a well-scaled vector y. With an indefinite signature matrix Φ, elementary Φ-unitary matrices of the case $\langle y, y \rangle = 0$ may not be well distinguished in numerical computations from that of the non-degenerate case as in exact arithmetic computations, due to roundoffs and affected by schemes for scaling y. It seems necessary to have a uniform scheme for scaling y, independent of Φ. However, a small relative error in a well-scaled y does not necessarily imply a small relative error in

$y^{\mathrm{H}}\Phi y$, a Φ-unitary transformation is not backward stable in general. On the other hand, things are not as pessimistic as the worst case may suggest. Given the fact that the transformed vector is often in matrix computations a multiple of e_1, one may use a strategy in choosing the multiplier so that $\langle y, y\rangle = 0$ is bounded away from zero if possible. Moreover, for a sequence of transformations, one may still get reasonable error bounds on well behaved matrices with carefully designed algorithms. The reader is referred to Stewart's work [25] for the concept and method of relational stability analysis on a sequence of transformations.

6 Acknowledgments

Many sincere thanks to Weimin Han for his careful proofreading and his useful comments on every section of the manuscript.

References

[1] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, 2nd. edition, 1989.

[2] J. H. Wilkinson. *The Algebraic Eigenvalue Problem.* Claredon Press, Oxford, England, 1965.

[3] A. S. Householder. *The Theory of Matrices in Numerical Analysis.* Dover Publications, New York, 1964.

[4] Numerical Algorithms Group Ltd. *NAG Fortran Library Manual, Mark 16.* Oxford, 1993.

[5] E. Anderson and et. al. *LAPACK User's Guide.* SIAM pub., Philadephia, 2nd. edition, 1995.

[6] V. Ch. Venkaiah, V. Krishna, and A. Paulraj. Householder transform in C^m. *Digital Signal Processing*, 3:226–227, 1993.

[7] X. G. Xia and B.W. Suter. On the Householder transform in C^m. *Digital Signal Processing*, 5:116–117, 1995.

[8] C. M. Rader and A. O. Steinhardt. Hyperbolic Householder transformations. *IEEE Tras. on A.S.S.P.*, 34:1589–1602, 1986.

[9] G. H. Golub. Matrix decompositions and statistical computation. In R. C. Milton and J. A. Nelder, editors, *Statistical Computation*, pages 365–397. Academic Press, New York, 1969.

[10] P. E. Gill, G. H. Golub, W. Murray, and M. A. Saunders. Methods for modeling matrix factorizations. *Math. Comp.*, 28:505–535, 1974.

[11] J.-M. Delosme and I. Ipsen. Parallel solution of symmetric positive definite systems with hyperbolic rotations. *Lin. Alg. Appl.*, 77:75–111, 1986.

[12] A. W. Bojanczyk, R. P. Brent, P. Van Dooren, and F. R. de Hoog. A note on downdating the cholesky factorization. *SIAM J. Sci. Stat. Comp.*, 8:210–221, 1987.

[13] S. T. Alexander, C.-T. Pan, and R. J. Plemmons. Analysis of a recursive least squares hyperbolic rotation algorithm for signal processing. *Lin. Alg. Appl.*, 98:3–40, 1988.

[14] A. Bunse-Gerstner. An analysis of the HR algorithms for computing the eigenvalues of a matrix. *Lin. Alg. & Appl.*, 35:155–173, 1981.

[15] C. M. Rader and A. O. Steinhardt. Hyperbolic Householder transforms. *SIAM J. Matrix Anal. Appl.*, 9:269–290, 1988.

[16] G. W. Stewart. *Introduction to Matrix Computations.* Academic Press, Inc., 1973.

[17] X. Sun and C. H. Bischof. A basis-kernel representation of orthogonal matrices. *SIAM J. Matrix Anal. Appl.*, 16(4), Oct. 1995.

[18] R. B. Lehoucq. The computation of elementary unitary matrices. Draft, Department of Computer Science, Rice University, July 1994.

[19] B. N. Parlett. Analysis of algorithms for reflectors in bisectors. *SIAM Review*, 13(2):197–208, April 1971.

[20] G. Cybenko and M. Berry. Hyperbolic Householder algorithms for factoring structured matrices. *SIAM J. Matrix Anal. Appl.*, 11:499–520, 1990.

[21] W.-M. Yan and K.-L. Chung. A block representation for hyperbolic householder transform. Draft, National Taiwan University, Department of Computer Science and Information Engineering, Nov. 1995.

[22] C. H. Bischof and C. Van Loan. The WY representation for products of householder matrices. *SIAM J. Sci. Stat. Comp.*, 8:12–13, 1987.

[23] R. Schreiber and C. Van Loan. A storage efficient WY representation for products of householder transformations. *SIAM J. Sci. Stat. Comp.*, 10(1):53–57, 1989.

[24] C. Puglisi. Modification of the Householder method based on the compact WY representation. *SIAM J. Matrix Anal. Appl.*, 3(3):723–726, 1992.

[25] G. W. Stewart. On the stability of sequential updates and downdates. Technical Report CS-TR-3263, University of Maryland, Department of Computer Science, 1993.