Representing Unitary Matrices by Independent Parameters - Rev 0

Article · October 2016

John J Polcari
Oak Ridge National Laboratory

<u>Representing Unitary Matrices by Independent Parameters</u>

<u>Working Paper</u>

Dr. John Polcari

Revision 0
October 10, 2016

Abstract

A representation of any orthogonal or unitary matrix in terms of independently selectable degrees of freedom (DOF) is derived. The representation includes closed form results for both analysis (matrix to DOF) and synthesis (DOF to matrix) mappings. The synthesis mapping takes on the convenient form of a factorization of the unitary matrix into a series of component matrices, each of which is also unitary. As the mappings in both directions are one-to-one, the representation is necessarily unique.

1. Introduction

Central to the understanding and larger exploitation of the structure of a unitary matrix is an orderly representation of such a matrix in terms by a set of independently selectable parameters. This paper demonstrates that any $N \; x \; N$ unitary matrix may be uniquely defined in terms of $N(N-1)/2$ complex DOF plus an additional N real DOF, all independently selectable. Both the analysis mapping, for computing the DOF associated with a specified matrix, and the synthesis mapping, for recovering the matrix from specified DOF values, are developed. While the representation provides a useful factorization of a unitary matrix into unitary components, it does not appear to be well known, as [1] makes no mention of it, even though it is of commensurate utility to many of the results provided there.

Denoting the composite DOF set as $\left\{ \underline{\mathbf{w}}_j, \varphi_j \middle| j = 1 \to N \right\}$, with each complex vector $\underline{\mathbf{w}}_j$ being of size $j-1 \; x \; 1$ and the N real phase values φ_j being scalars, and requiring only that each $\underline{\mathbf{w}}_j$ not exceed unit length, an equivalent unitary matrix $\underline{\overline{\mathbf{U}}}$ may be synthesized as

$$\left\{ \underline{\mathbf{w}}_j, \varphi_j \middle| j = 1 \to N \right\} \to \underline{\overline{\mathbf{U}}}: \quad \underline{\overline{\mathbf{U}}} = \left(\prod_{j=N}^{1} \underline{\overline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j \right) \right) \begin{bmatrix} e^{j\varphi_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{j\varphi_N} \end{bmatrix} , \tag{1}$$

where the unitary component matrices are

$$\underline{\overline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_i \right) = \begin{bmatrix} \underline{\overline{\mathbf{I}}}_{[i-1]} - \dfrac{\underline{\mathbf{w}}_i \underline{\mathbf{w}}_i^+}{1 + \sqrt{1 - \underline{\mathbf{w}}_i^+ \underline{\mathbf{w}}_i}} & \underline{\mathbf{w}}_i & \underline{\overline{\mathbf{0}}} \\ -\underline{\mathbf{w}}_i^+ & \sqrt{1 - \underline{\mathbf{w}}_i^+ \underline{\mathbf{w}}_i} & \underline{\mathbf{0}}^+ \\ \underline{\overline{\mathbf{0}}} & \underline{\mathbf{0}} & \underline{\overline{\mathbf{I}}}_{[N-i]} \end{bmatrix} . \tag{2}$$

Conversely, given $\underline{\overline{\mathbf{U}}}$ with columns $\left\{ \underline{\mathbf{u}}_j \middle| j = 1 \to N \right\}$, the associated DOF may be found from

$$\underline{\overline{\mathbf{U}}} \to \left\{ \underline{\mathbf{w}}_j, \varphi_j \middle| j = 1 \to N \right\}: \quad \begin{aligned} \varphi_i &= \angle \left(\underline{\mathbf{e}}_i^+ \left(\prod_{j=i+1}^{N} \underline{\overline{\mathbf{\Psi}}}^+ \left(\underline{\mathbf{w}}_j \right) \right) \underline{\mathbf{u}}_i \right) \\ \underline{\mathbf{w}}_i &= e^{-j\varphi_i} \underline{\overline{\mathbf{P}}}_{N(i-1)}^+ \left(\prod_{j=i+1}^{N} \underline{\overline{\mathbf{\Psi}}}^+ \left(\underline{\mathbf{w}}_j \right) \right) \underline{\mathbf{u}}_i \end{aligned} . \tag{3}$$

As the mappings are one-to-one in both directions, the representation is necessarily unique. Equivalent results for real orthogonal matrices can be found in Theorem 2 of Section 6.

2. Notational Preliminaries

The following notational conventions are employed in this paper:

1. Scalars are denoted by italics (x), vectors by a bolded underscore ($\underline{\mathbf{x}}$), and matrices by bolded capitals that are doubly scored ($\overline{\underline{\mathbf{X}}}$). By convention, all vectors are column vectors.

2. Whenever possible, explicit notation of vector/matrix size is avoided in favor of notational conciseness. In those cases where mathematical precision dictates explicit annotation of size, the bracket subscript

$$\underline{\mathbf{x}} = \underline{\mathbf{x}}_{[i]} \quad \text{and} \quad \overline{\underline{\mathbf{X}}} = \overline{\underline{\mathbf{X}}}_{[k]} \tag{4}$$

respectively denotes an $i \ x\ 1$ vector and a $k \ x\ k$ square matrix.

3. Except for the partial identity matrices defined below, non-square matrices arise only in the context of a block matrix form, so that the size of non-square matrices may be readily inferred from the size of the diagonal blocks; for example

$$\overline{\underline{\mathbf{X}}}_{[i+k]} = \begin{bmatrix} \overline{\underline{\mathbf{Y}}}_{[i]} & \overline{\underline{\mathbf{U}}} \\ \overline{\underline{\mathbf{V}}} & \overline{\underline{\mathbf{Z}}}_{[k]} \end{bmatrix} \tag{5}$$

implies that $\overline{\underline{\mathbf{U}}}$ is $i \ x\ k$ and $\overline{\underline{\mathbf{V}}}$ is $k \ x\ i$.

4. Columns and elements of a matrix are denoted as vectors and scalers of the same letter, that is

$$\text{and} \quad \overline{\underline{\mathbf{X}}}_{[i]} = \begin{bmatrix} \underline{\mathbf{x}}_{1[i]} & \cdots & \underline{\mathbf{x}}_{i[i]} \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{1i} \\ \vdots & \ddots & \vdots \\ x_{i1} & \cdots & x_{ii} \end{bmatrix} \tag{6}$$

5. The $i \ x\ 1$ k^{th} unit vector (i.e., column k of the identity matrix) is written

$$\underline{\mathbf{e}}_{k[i]} = \begin{bmatrix} \underline{\mathbf{0}}_{[k-1]} \\ 1 \\ \underline{\mathbf{0}}_{[i-k]} \end{bmatrix} \quad \Rightarrow \quad \overline{\underline{\mathbf{I}}}_{[i]} = \begin{bmatrix} \underline{\mathbf{e}}_{1[i]} & \cdots & \underline{\mathbf{e}}_{i[i]} \end{bmatrix} \quad . \tag{7}$$

6. The $i \ x\ k$ leading partial identity matrix (identity matrix columns $1 \rightarrow k$) is written

$$\overline{\underline{\mathbf{P}}}_{ik} = \begin{bmatrix} \underline{\mathbf{e}}_{1[i]} & \cdots & \underline{\mathbf{e}}_{k[i]} \end{bmatrix} \quad . \tag{8}$$

3. Identification of Independent Degrees of Freedom (DOF) in Unitary Matrices

Central to the understanding and larger exploitation of the structure of a unitary matrix is an orderly representation of such a matrix in terms by a set of independently selectable parameters. In this section, development of such a representation begins by identifying of the number of degrees of freedom (DOF) that are independently selectable in the creation of an arbitrary unitary matrix.

Consider a unitary matrix $\underline{\overline{\mathbf{U}}}$, so that

$$\underline{\overline{\mathbf{U}}}_{[N]} = \left[\underline{\mathbf{u}}_1 \quad \cdots \quad \underline{\mathbf{u}}_N \right] \quad \text{with} \quad \underline{\overline{\mathbf{U}}}^+ \underline{\overline{\mathbf{U}}} = \underline{\overline{\mathbf{U}}}\,\underline{\overline{\mathbf{U}}}^+ = \underline{\overline{\mathbf{I}}} \quad . \tag{9}$$

It is obvious that $\underline{\overline{\mathbf{U}}}$ requires specification of N^2 complex parameters (or, equivalently, $2N^2$ real parameters, in terms of either real and imaginary parts or amplitudes and phases) to be fully defined. However, unlike a fully arbitrary matrix, not all these elements may be independently selected, as the requirements on the right hand side of (9) constrain many of the choices. Consider the final column of $\underline{\overline{\mathbf{U}}}$, which may be conveniently written as

$$\underline{\mathbf{u}}_N = \begin{bmatrix} \underline{\mathbf{w}}_N \\ |u_{NN}| \end{bmatrix} e^{j\varphi_N} \quad , \tag{10}$$

where the only constraint on the selection of the $N-1\,x\,1$ vector $\underline{\mathbf{w}}_N$ is that it not exceed unit length. However, the final element u_{NN} cannot be arbitrarily chosen; while φ_N is independently selectable, the requirement that $\underline{\mathbf{u}}_N$ be of unit length implies that

$$|\underline{\mathbf{u}}_N| = 1 \quad \Rightarrow \quad |u_{NN}| = \sqrt{1 - \underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N} \quad \Rightarrow \quad \underline{\mathbf{u}}_N = \begin{bmatrix} \underline{\mathbf{w}}_N \\ \sqrt{1 - \underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N} \end{bmatrix} e^{j\varphi_N} \quad . \tag{11}$$

Of the N complex elements in $\underline{\mathbf{u}}_N$, only $N-1$, together with the real phase parameter φ_N, are independent DOF. The latter may be usefully considered as one half of a complex DOF, for a total of $N - 1/2$ DOF.

Now consider $\underline{\mathbf{u}}_{N-1}$, the next to last column of $\underline{\overline{\mathbf{U}}}$. In addition to being of unit length, this column must necessarily be orthogonal to the final column. If one writes

$$\underline{\mathbf{u}}_{N-1} = \begin{bmatrix} \begin{bmatrix} u_{1(N-1)} \\ \vdots \\ u_{(N-1)(N-1)} \end{bmatrix} \\ u_{N(N-1)} \end{bmatrix} \quad , \tag{12}$$

this additional requirement can be seen to dictate the final element in $\underline{\mathbf{u}}_{N-1}$, as

$$\underline{\mathbf{u}}_N^+ \underline{\mathbf{u}}_{N-1} = 0 \quad \Rightarrow \quad u_{N(N-1)} = -\frac{1}{\sqrt{1-\underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N}} \left(\underline{\mathbf{w}}_N^+ \begin{bmatrix} u_{1(N-1)} \\ \vdots \\ u_{(N-1)(N-1)} \end{bmatrix} \right) \quad ; \tag{13}$$

Then the earlier procedure to enforce the unit length requirement dictates the magnitude (but not the phase) of $u_{(N-1)(N-1)}$, so that that this column represents $N-3/2$ independent DOF.

Continuing in this fashion, each earlier column of $\overline{\underline{\mathbf{U}}}$ possesses one less DOF, since there is one more column to which it must be orthogonal. Hence, the number of DOF comprising column i is

$$DOF_i = N - i + \frac{1}{2} \quad \Rightarrow \quad DOF = \sum_{i=1}^{N} DOF_i = \frac{N^2}{2} \tag{14}$$

Note that this result implies that the first column possesses only $1/2$ DOF, which is intuitively obvious; given all the remaining columns, the first is fully determined to within a phase factor.

In the case of a real orthogonal matrix, the complex DOF count reduces to a real DOF count, with the partial real DOF representing a (binary) choice of sign rather than a phase.

4. Unitary Component Matrices

Temporarily ignoring the phase parameter, let the remaining $N-1$ independently selectable DOF comprising $\underline{\mathbf{u}}_N$ be represented by the $N-1\ x\ 1$ vector $\underline{\mathbf{w}}_N$. It is useful to define the N^{th} unitary component matrix as follows.

DEFINITION 1 (Unitary Component Matrix): Let

$$\overline{\underline{\boldsymbol{\Psi}}}_{[N]}\left(\underline{\mathbf{w}}_N\right)=\begin{bmatrix}\overline{\underline{\mathbf{Q}}}\left(\underline{\mathbf{w}}_N\right) & \dfrac{\underline{\mathbf{w}}_N}{\sqrt{1-\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N}} \\ -\underline{\mathbf{w}}_N^{+} & \sqrt{1-\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N}\end{bmatrix} \quad \text{with} \quad \overline{\underline{\mathbf{Q}}}_{[N-1]}\left(\underline{\mathbf{w}}_N\right)=\overline{\underline{\mathbf{I}}}-\frac{\underline{\mathbf{w}}_N\underline{\mathbf{w}}_N^{+}}{1+\sqrt{1-\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N}} \quad . \tag{15}$$

This matrix and its partitions have several important properties.

PROPERTY 1:

$$\det\left(\overline{\underline{\mathbf{Q}}}\left(\underline{\mathbf{w}}_N\right)\right)=\sqrt{1-\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N} \quad . \tag{16}$$

and

$$\overline{\underline{\mathbf{Q}}}^{-1}\left(\underline{\mathbf{w}}_N\right)=\overline{\underline{\mathbf{I}}}+\frac{\underline{\mathbf{w}}_N\underline{\mathbf{w}}_N^{+}}{\sqrt{1-\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N}\left(1+\sqrt{1-\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N}\right)} \quad , \tag{17}$$

so that

$$\overline{\underline{\mathbf{Q}}}^{-1}\left(\underline{\mathbf{w}}_N\right)\underline{\mathbf{w}}_N=\left(\frac{1}{\sqrt{1-\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N}}\right)\underline{\mathbf{w}}_N \quad \text{and} \quad \underline{\mathbf{w}}_N^{+}\overline{\underline{\mathbf{Q}}}^{-1}\left(\underline{\mathbf{w}}_N\right)\underline{\mathbf{w}}_N=\frac{\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N}{\sqrt{1-\underline{\mathbf{w}}_N^{+}\underline{\mathbf{w}}_N}} \quad . \tag{18}$$

Proof: The determinant results from the Matrix Determinant Lemma (MDL) [2], while the inverse is a consequence of the Woodbury Matrix Identity (WMI) [3]; the final two results follow directly.

PROPERTY 2: The matrix $\overline{\underline{\boldsymbol{\Psi}}}\left(\underline{\mathbf{w}}_N\right)$ is unitary, so that

$$\overline{\underline{\boldsymbol{\Psi}}}^{+}\left(\underline{\mathbf{w}}_N\right)\overline{\underline{\boldsymbol{\Psi}}}\left(\underline{\mathbf{w}}_N\right)=\overline{\underline{\boldsymbol{\Psi}}}\left(\underline{\mathbf{w}}_N\right)\overline{\underline{\boldsymbol{\Psi}}}^{+}\left(\underline{\mathbf{w}}_N\right)=\overline{\underline{\mathbf{I}}} \quad \text{with} \quad \det\left(\overline{\underline{\boldsymbol{\Psi}}}\left(\underline{\mathbf{w}}_N\right)\right)=1 \quad . \tag{19}$$

Proof:

$$\overline{\underline{\boldsymbol{\Psi}}}^{+}\left(\underline{\mathbf{w}}_N\right)\overline{\underline{\boldsymbol{\Psi}}}\left(\underline{\mathbf{w}}_N\right)=\begin{bmatrix}\overline{\underline{\mathbf{A}}} & \underline{\mathbf{b}} \\ \underline{\mathbf{b}}^{+} & c\end{bmatrix} \quad , \tag{20}$$

where

$$\underline{\overline{\mathbf{A}}} = \left(\underline{\overline{\mathbf{I}}} - \frac{\underline{\mathbf{w}}_N \underline{\mathbf{w}}_N^+}{1+\sqrt{1-\underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N}} \right)^2 + \underline{\mathbf{w}}_N \underline{\mathbf{w}}_N^+ = \underline{\overline{\mathbf{I}}}$$

$$\underline{\mathbf{b}} = \left(\underline{\overline{\mathbf{I}}} - \frac{\underline{\mathbf{w}}_N \underline{\mathbf{w}}_N^+}{1+\sqrt{1-\underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N}} \right) \underline{\mathbf{w}}_N - \left(\sqrt{1-\underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N} \right) \underline{\mathbf{w}}_N = \underline{\mathbf{0}} \quad . \tag{21}$$

$$c = \underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N + \left(\sqrt{1-\underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N} \right)^2 = 1$$

Similarly, $\underline{\overline{\mathbf{\Psi}}}(\underline{\mathbf{w}}_N)\underline{\overline{\mathbf{\Psi}}}^+(\underline{\mathbf{w}}_N) = \underline{\overline{\mathbf{I}}}$. Using the rules for the determinant of a block matrix [4], together with Property 1, yields

$$\det\left(\underline{\overline{\mathbf{\Psi}}}(\underline{\mathbf{w}}_N)\right) = \det\left(\underline{\overline{\mathbf{Q}}}(\underline{\mathbf{w}}_N)\right)\left(\sqrt{1-\underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N} + \underline{\mathbf{w}}_N^+ \underline{\overline{\mathbf{Q}}}^{-1}(\underline{\mathbf{w}}_N)\underline{\mathbf{w}}_N\right) = 1 \quad . \tag{22}$$

For DOF arising from earlier columns, the component matrix is defined as the embedded form

$$\underline{\overline{\mathbf{\Psi}}}_{[N]}(\underline{\mathbf{w}}_i) = \begin{bmatrix} \underline{\overline{\mathbf{\Psi}}}_{[i]}(\underline{\mathbf{w}}_i) & \underline{\overline{\mathbf{0}}} \\ \underline{\overline{\mathbf{0}}} & \underline{\overline{\mathbf{I}}}_{[N-i]} \end{bmatrix} = \begin{bmatrix} \underline{\overline{\mathbf{Q}}}(\underline{\mathbf{w}}_i) & \underline{\mathbf{w}}_i & \underline{\overline{\mathbf{0}}} \\ -\underline{\mathbf{w}}_i^+ & \sqrt{1-\underline{\mathbf{w}}_i^+ \underline{\mathbf{w}}_i} & \underline{\mathbf{0}}^+ \\ \underline{\overline{\mathbf{0}}} & \underline{\mathbf{0}} & \underline{\overline{\mathbf{I}}} \end{bmatrix} \quad , \tag{23}$$

where the definition of $\underline{\overline{\mathbf{Q}}}(\underline{\mathbf{w}}_i)$ from (15) continues to hold. Note that the size of $\underline{\overline{\mathbf{\Psi}}}(\underline{\mathbf{w}}_i)$ remains constant while the size of $\underline{\overline{\mathbf{Q}}}(\underline{\mathbf{w}}_i)$ is implicitly defined by the size of the argument. Typically, all component matrices are the same size as the unitary matrix under consideration; when necessary to differentiate component matrices of varying size (as in (23)), explicit size notation is used.

Obviously, Properties 1 and 2 continue to hold for the embedded form found in (23). Additionally, while relatively trivial, one last property of unitary component matrices is subsequently important.

PROPERTY 3: For $k > i$

$$\underline{\overline{\mathbf{\Psi}}}(\underline{\mathbf{w}}_i)\underline{\mathbf{e}}_k = \underline{\mathbf{e}}_k \quad \text{and} \quad \underline{\overline{\mathbf{\Psi}}}^+(\underline{\mathbf{w}}_i)\underline{\mathbf{e}}_k = \underline{\mathbf{e}}_k \quad . \tag{24}$$

5. Representation of Unitary Matrix Columns

In this section, a quantitative mapping between any particular column $\underline{\mathbf{u}}_i$ of a unitary matrix and the DOF $\{\underline{\mathbf{w}}_i, \varphi_i\}$ associated with that column is developed. While other definitions of the associated DOF are possible, this particular choice is convenient for two reasons. First, as shown next, the mapping is one-to-one; that is, knowledge of either $\left\{\underline{\mathbf{u}}_j \,\middle|\, j = i \to N\right\}$ or $\left\{\underline{\mathbf{w}}_j, \varphi_j \,\middle|\, j = i \to N\right\}$ [1] allows calculation of the other in a reversible manner. Second, as shown in Section 6, it leads to a simple factorization of the full unitary matrix in terms of the previously defined component matrices.

From the final result in (11), it is reasonable to adopt the following mapping between the final column $\underline{\mathbf{u}}_N$ and an arbitrarily selectable $\{\underline{\mathbf{w}}_N, \varphi_N\}$

$$\underline{\mathbf{u}}_N = \begin{bmatrix} \begin{bmatrix} u_{1N} \\ \vdots \\ u_{(N-1)N} \end{bmatrix} \\ u_{NN} \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{w}}_N \\ \sqrt{1 - \underline{\mathbf{w}}_N^+ \underline{\mathbf{w}}_N} \end{bmatrix} e^{j\varphi_N} \qquad \begin{array}{l} \varphi_N = \angle u_{NN} = \angle \underline{\mathbf{e}}_N^+ \underline{\mathbf{u}}_N \\ \underline{\mathbf{w}}_N = e^{-j\varphi_N} \begin{bmatrix} u_{1N} \\ \vdots \\ u_{(N-1)N} \end{bmatrix} = e^{-j\varphi_N} \overline{\underline{\mathbf{P}}}_{N(N-1)}^+ \underline{\mathbf{u}}_N \end{array} \quad , \tag{25}$$

subject only to the constraint that $\left|\underline{\mathbf{w}}_N\right| \le 1$. Notice that this choice also implies that

$$\underline{\mathbf{u}}_N = e^{j\varphi_N} \overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right) \underline{\mathbf{e}}_N \quad \Rightarrow \quad \overline{\underline{\mathbf{\Psi}}}^+\left(\underline{\mathbf{w}}_N\right) \underline{\mathbf{u}}_N = e^{j\varphi_N} \underline{\mathbf{e}}_N \quad . \tag{26}$$

A useful (although not immediately obvious) way of relating $\underline{\mathbf{u}}_{N-1}$ to $\{\underline{\mathbf{w}}_{N-1}, \varphi_{N-1}\}$ is

$$\underline{\mathbf{u}}_{N-1} = e^{j\varphi_{N-1}} \overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right) \begin{bmatrix} \underline{\mathbf{w}}_{(N-1)} \\ \sqrt{1 - \underline{\mathbf{w}}_{N-1}^+ \underline{\mathbf{w}}_{N-1}} \\ 0 \end{bmatrix} = e^{j\varphi_{N-1}} \overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right) \overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_{N-1}\right) \underline{\mathbf{e}}_{N-1} \quad , \tag{27}$$

subject to the additional constraint that $\left|\underline{\mathbf{w}}_{N-1}\right| \le 1$. Since

$$\underline{\mathbf{u}}_N^+ \underline{\mathbf{u}}_{N-1} = e^{j(\varphi_{N-1} - \varphi_N)} \left(\underline{\mathbf{e}}_N^+ \overline{\underline{\mathbf{\Psi}}}^+\left(\underline{\mathbf{w}}_N\right) \right) \left(\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right) \begin{bmatrix} \underline{\mathbf{w}}_{(N-1)} \\ \sqrt{1 - \underline{\mathbf{w}}_{N-1}^+ \underline{\mathbf{w}}_{N-1}} \\ 0 \end{bmatrix} \right) = 0 \quad , \tag{28}$$

---

[1] From Section 3, each $\underline{\mathbf{w}}_j$ must necessarily have $j-1$ elements, further implying that $\underline{\mathbf{w}}_1 = \varnothing$.

this choice naturally enforces orthogonality, while also ensuring unit length, as

$$\underline{\mathbf{u}}_{N-1}^{+}\underline{\mathbf{u}}_{N-1}=\left(\begin{bmatrix}\underline{\mathbf{w}}_{N-1}\\ \sqrt{1-\underline{\mathbf{w}}_{N-1}^{+}\underline{\mathbf{w}}_{N-1}}\\ 0\end{bmatrix}^{+}\overline{\underline{\mathbf{\Psi}}}^{+}\left(\underline{\mathbf{w}}_{N}\right)\right)\left(\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_{N}\right)\begin{bmatrix}\underline{\mathbf{w}}_{N-1}\\ \sqrt{1-\underline{\mathbf{w}}_{N-1}^{+}\underline{\mathbf{w}}_{N-1}}\\ 0\end{bmatrix}\right)=1 \quad . \tag{29}$$

In general, $\underline{\mathbf{u}}_i$ must be orthogonal to all later columns, suggesting the forward mapping

$$\left\{\underline{\mathbf{w}}_i,\varphi_i\right\}\rightarrow\underline{\mathbf{u}}_i:\quad \underline{\mathbf{u}}_i=e^{j\varphi_i}\left(\prod_{j=N}^{i+1}\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right)\begin{bmatrix}\underline{\mathbf{w}}_{i[i-1]}\\ \sqrt{1-\underline{\mathbf{w}}_i^{+}\underline{\mathbf{w}}_i}\\ \underline{\mathbf{0}}_{[N-i]}\end{bmatrix}=e^{j\varphi_i}\left(\prod_{j=N}^{i}\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right)\underline{\mathbf{e}}_i \quad , \tag{30}$$

subject to the cumulative constraint that $\left\{\left|\underline{\mathbf{w}}_j\right|\leq 1\,\middle|\, j=i\rightarrow N\right\}$.

The following lemma proves that this form enforces the required orthonormal structure.

LEMMA 1: Given $\left\{\underline{\mathbf{u}}_k\,\middle|\,k=i\rightarrow N\right\}$ constructed per (30), then for any choice of k

$$\underline{\mathbf{u}}_k^{+}\underline{\mathbf{u}}_i=\begin{cases}1 & k=i\\ 0 & k>i\end{cases} \quad . \tag{31}$$

Proof: For $k=i$

$$\underline{\mathbf{u}}_i^{+}\underline{\mathbf{u}}_i=\left(\begin{bmatrix}\underline{\mathbf{w}}_{i[i-1]}\\ \sqrt{1-\underline{\mathbf{w}}_i^{+}\underline{\mathbf{w}}_i}\\ \underline{\mathbf{0}}_{[N-i]}\end{bmatrix}^{+}\left(\prod_{j=i+1}^{N}\overline{\underline{\mathbf{\Psi}}}^{+}\left(\underline{\mathbf{w}}_j\right)\right)\right)\left(\left(\prod_{j=N}^{i+1}\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right)\begin{bmatrix}\underline{\mathbf{w}}_{i[i-1]}\\ \sqrt{1-\underline{\mathbf{w}}_i^{+}\underline{\mathbf{w}}_i}\\ \underline{\mathbf{0}}_{[N-i]}\end{bmatrix}\right)=1 \quad . \tag{32}$$

For $k>i$, using Property 3

$$\begin{aligned}\underline{\mathbf{u}}_k^{+}\underline{\mathbf{u}}_i&=e^{j(\varphi_i-\varphi_k)}\left(\underline{\mathbf{e}}_k^{+}\left(\prod_{j=k}^{N}\overline{\underline{\mathbf{\Psi}}}^{+}\left(\underline{\mathbf{w}}_j\right)\right)\right)\left(\left(\prod_{j=N}^{i}\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right)\underline{\mathbf{e}}_i\right)\\ &=e^{j(\varphi_i-\varphi_k)}\underline{\mathbf{e}}_k^{+}\left(\prod_{j=k-1}^{i}\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right)\underline{\mathbf{e}}_i=e^{j(\varphi_i-\varphi_k)}\underline{\mathbf{e}}_k^{+}\underline{\mathbf{e}}_i=0\end{aligned} \quad . \tag{33}$$

The inverse mapping to that in (30) is now developed, implicitly proving that $\{\underline{\mathbf{w}}_i, \varphi_i\}$ must exist.

LEMMA 2: Given $\{\underline{\mathbf{u}}_k \mid k = i \to N\}$, drawn from the final $N-i+1$ columns of $\overline{\underline{\mathbf{U}}}$ so that

$$\underline{\mathbf{u}}_k^+ \underline{\mathbf{u}}_i = \begin{cases} 1 & k = i \\ 0 & k > i \end{cases} , \tag{34}$$

then

$$\underline{\mathbf{u}}_i \to \{\underline{\mathbf{w}}_i, \varphi_i\} : \begin{array}{l} \varphi_i = \angle\left(\underline{\mathbf{e}}_i^+ \left(\prod_{j=i+1}^{N} \overline{\underline{\mathbf{\Psi}}}^+ \left(\underline{\mathbf{w}}_j\right)\right) \underline{\mathbf{u}}_i\right) \\ \underline{\mathbf{w}}_i = e^{-j\varphi_i} \overline{\underline{\mathbf{P}}}_{N(i-1)}^+ \left(\prod_{j=i+1}^{N} \overline{\underline{\mathbf{\Psi}}}^+ \left(\underline{\mathbf{w}}_j\right)\right) \underline{\mathbf{u}}_i \end{array} . \tag{35}$$

<u>Proof:</u> (By construction). Adopting the inverse mapping for the final column from (25)

$$\varphi_N = \angle \underline{\mathbf{e}}_N^+ \underline{\mathbf{u}}_N \quad \text{and} \quad \underline{\mathbf{w}}_N = e^{-j\varphi_N} \overline{\underline{\mathbf{P}}}_{N(N-1)}^+ \underline{\mathbf{u}}_N \quad , \tag{36}$$

where

$$\left|\underline{\mathbf{u}}_N\right| = 1 \quad \Rightarrow \quad \left|\underline{\mathbf{w}}_N\right| \le 1 \quad ; \tag{37}$$

hence, the matrix $\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right)$ may be constructed using (15), and

$$\underline{\mathbf{u}}_N = e^{j\varphi_N} \overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right) \underline{\mathbf{e}}_N \quad . \tag{38}$$

Now, for the remaining columns, let

$$\underline{\mathbf{u}}_k^{\{N-1\}} = \overline{\underline{\mathbf{\Psi}}}^+ \left(\underline{\mathbf{w}}_N\right) \underline{\mathbf{u}}_k \quad \forall k = i \to N-1 \quad ; \tag{39}$$

since $\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right)$ is unitary, this is a generalized coordinate rotation which preserves the orthonormal structure of $\left\{\underline{\mathbf{u}}_k^{\{N-1\}} \middle| k = 1 \to N-1\right\}$. However, using (26), $\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right)$ may be written in alternate block form as

$$\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_N\right) = \left[\overline{\underline{\mathbf{X}}} \quad e^{-j\varphi_N} \underline{\mathbf{u}}_N\right] \quad , \tag{40}$$

where $\underline{\overline{\mathbf{X}}}$ is an $N \; x \; (N-1)$ matrix, so that

$$\underline{\mathbf{u}}_j^{\{N-1\}} = \underline{\overline{\mathbf{\Psi}}}^+ (\underline{\mathbf{w}}_N) \underline{\mathbf{u}}_j = \begin{bmatrix} \underline{\overline{\mathbf{X}}}^+ \underline{\mathbf{u}}_j \\ e^{-j\varphi_N} \underline{\mathbf{u}}_N^+ \underline{\mathbf{u}}_j \end{bmatrix} = \begin{bmatrix} \underline{\overline{\mathbf{X}}}^+ \underline{\mathbf{u}}_j \\ 0 \end{bmatrix} \quad . \tag{41}$$

In essence, the effect of the rotation in (39) is to zero the final element in each of the rotated columns. Then, for column $N-1$, one may adopt the inverse mapping

$$\varphi_{N-1} = \angle \underline{\mathbf{e}}_{N-1}^+ \underline{\mathbf{u}}_{N-1}^{\{N-1\}} \quad \text{and} \quad \underline{\mathbf{w}}_{N-1} = e^{-j\varphi_{N-1}} \underline{\overline{\mathbf{P}}}_{N(N-2)}^+ \underline{\mathbf{u}}_{N-1}^{\{N-1\}} \quad , \tag{42}$$

where again

$$\left| \underline{\mathbf{u}}_{N-1}^{\{N-1\}} \right| = 1 \quad \Rightarrow \quad \left| \underline{\mathbf{w}}_{N-1} \right| \leq 1 \quad , \tag{43}$$

so that

$$\underline{\mathbf{u}}_{N-1}^{\{N-1\}} = \underline{\overline{\mathbf{\Psi}}} (\underline{\mathbf{w}}_{N-1}) \underline{\mathbf{e}}_{N-1} \quad ; \tag{44}$$

now, the matrix $\underline{\overline{\mathbf{\Psi}}}(\underline{\mathbf{w}}_{N-1})$ may be constructed using (23) and

$$\underline{\mathbf{u}}_{N-1}^{\{N-1\}} = e^{j\varphi_{N-1}} \underline{\overline{\mathbf{\Psi}}} (\underline{\mathbf{w}}_{N-1}) \underline{\mathbf{e}}_{N-1} \quad \Rightarrow \quad \underline{\mathbf{u}}_{N-1} = e^{j\varphi_{N-1}} \left(\prod_{j=N}^{N-1} \underline{\overline{\mathbf{\Psi}}} (\underline{\mathbf{w}}_j) \right) \underline{\mathbf{e}}_{N-1} \quad . \tag{45}$$

Repeating this process for the remaining columns, one arrives at the result

$$\varphi_i = \angle \underline{\mathbf{e}}_i^+ \underline{\mathbf{u}}_i^{\{i\}} \quad \text{and} \quad \underline{\mathbf{w}}_i = e^{-j\varphi_i} \underline{\overline{\mathbf{P}}}_{N(i-1)}^+ \underline{\mathbf{u}}_i^{\{i\}} \quad , \tag{46}$$

which, upon recovering $\underline{\mathbf{u}}_i$ using

$$\underline{\mathbf{u}}_i^{\{i\}} = \underline{\overline{\mathbf{\Psi}}}^+ (\underline{\mathbf{w}}_{i+1}) \underline{\mathbf{u}}_k^{\{i+1\}} \quad \Rightarrow \quad \underline{\mathbf{u}}_i^{\{i\}} = \left(\prod_{j=i+1}^{N} \underline{\overline{\mathbf{\Psi}}}^+ (\underline{\mathbf{w}}_j) \right) \underline{\mathbf{u}}_k^{\{N\}} = \left(\prod_{j=i+1}^{N} \underline{\overline{\mathbf{\Psi}}}^+ (\underline{\mathbf{w}}_j) \right) \underline{\mathbf{u}}_k \quad , \tag{47}$$

results in (35).

COROLLARY 1: The mapping between $\{\underline{\mathbf{w}}_i, \varphi_i\}$ and $\underline{\mathbf{u}}_i$ is unique.

<u>Proof:</u> The previously demonstrated bidirectional one-to-one structure guarantees uniqueness.

6. Representation of a Unitary Matrix by Independent DOF

The previously developed ability to uniquely represent each column by independent DOF implies the ability to uniquely factor any unitary matrix into a sequence of unitary component matrices, providing a unique representation of the matrix by the composite set of DOF.

THEOREM 1 (Unitary Matrix Representation Theorem): Any unitary matrix may be written as

$$\overline{\underline{\mathbf{U}}} = \overline{\underline{\mathbf{U}}}\left(\underline{\mathbf{w}}_1, \cdots \underline{\mathbf{w}}_N, \varphi_1, \cdots, \varphi_N\right) = \left(\prod_{j=N}^{1} \overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right) \overline{\underline{\mathbf{E}}}\left(\varphi_1, \cdots, \varphi_N\right) \quad , \tag{48}$$

where $\overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)$ is given by (23) and

$$\overline{\underline{\mathbf{E}}}\left(\varphi_1, \cdots, \varphi_N\right) = \begin{bmatrix} e^{j\varphi_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{j\varphi_N} \end{bmatrix} \quad , \tag{49}$$

with $\left\{\underline{\mathbf{w}}_j, \varphi_j \middle| j = i \to N\right\}$ specified by Lemma 2.

Proof: Lemma 2 guarantees that any column i of $\overline{\underline{\mathbf{U}}}$ may be written in the form of (30), that is

$$\underline{\mathbf{u}}_i = e^{j\varphi_i}\left(\prod_{j=N}^{i} \overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right)\underline{\mathbf{e}}_i \quad \Rightarrow \quad \underline{\mathbf{u}}_i = e^{j\varphi_i}\left(\prod_{j=N}^{1} \overline{\underline{\mathbf{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right)\underline{\mathbf{e}}_i \quad , \tag{50}$$

where the second result is a consequence of Property 3. Then

$$\overline{\underline{\mathbf{U}}} = \left[\underline{\mathbf{u}}_1 \quad \cdots \quad \underline{\mathbf{u}}_N\right] = \left(\prod_{j=N}^{1}\left(\underline{\mathbf{w}}_j\right)\right)\left[\underline{\mathbf{e}}_1 e^{j\varphi_1} \quad \cdots \quad \underline{\mathbf{e}}_N e^{j\varphi_N}\right] \quad , \tag{51}$$

leading to (48).

COROLLARY 2: The mapping between $\left\{\underline{\mathbf{w}}_j, \varphi_j \middle| j = i \to N\right\}$ and $\overline{\underline{\mathbf{U}}}$ is unique.

Proof: Since each of the individual column mappings is unique, the composite mapping is also.

The reduction to the real case is obvious.

THEOREM 2 (Orthogonal Matrix Representation Theorem): Any orthogonal matrix may be uniquely written as

$$\underline{\overline{\mathbf{U}}} = \underline{\overline{\mathbf{U}}}\left(\underline{\mathbf{w}}_1, \cdots \underline{\mathbf{w}}_N, s_1, \cdots, s_N\right) = \left(\prod_{j=N}^{1} \underline{\overline{\boldsymbol{\Psi}}}\left(\underline{\mathbf{w}}_j\right)\right) \underline{\overline{\mathbf{S}}}\left(s_1, \cdots, s_N\right) \quad , \tag{52}$$

where

$$\underline{\overline{\mathbf{S}}}\left(s_1, \cdots, \mathrm{s}_N\right) = \begin{bmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_N \end{bmatrix} \quad \text{with} \quad s_j = \pm 1 \quad \forall j = 1 \rightarrow N \quad . \tag{53}$$

References

1. Wikipedia. *Matrix decomposition.* [Online], 2016. Available: https://en.wikipedia.org/wiki/Matrix_decomposition.

2. Wikipedia. *Matrix determinant lemma.* [Online], 2015. Available: https://en.wikipedia.org/wiki/Matrix_determinant_lemma.

3. Wikipedia. *Woodbury matrix identity.* [Online], 2016. Available: https://en.wikipedia.org/wiki/Woodbury_matrix_identity.

4. Wikipedia. *Determinant.* [Online], 2016. Available: https://en.wikipedia.org/wiki/Determinant.