1. INTRODUCTION

The purpose of this chapter is to introduce mathematical tools that will be used in later chapters. In Sections 2–6 we define proper random variables and processes, study some of their properties, and derive the probability density function of proper Gaussian random vectors. Proper random variables and processes are complex-valued. We will see that the use complex-valued random variables simplifies the mathematics. More importantly, proper Gaussian random vectors will be the tools that we need to model the noise (which we have neglected so far) as well as the filter-coefficients of baseband-equivalent channel models. In subsequent chapters we will learn how to use eigenvectors to simplify our channel model.

We now give a few background definitions and facts that will be useful throughout this chapter.

**Definition 1.1.** A matrix $U \in \mathbb{C}^{n \times n}$ is said to be unitary if $U^\dagger U = I$. If, in addition, $U \in \mathbb{R}^{n \times n}$, $U$ is said to be orthogonal.

The following theorem lists a number of handy facts about unitary matrices. Most of them are straightforward.

**Theorem 1.1.** If $U \in \mathbb{C}^{n \times n}$, the following are equivalent:

(a) $U$ is unitary;

(b) $U$ is nonsingular and $U^\dagger = U^{-1}$;

(c) $UU^\dagger = I$;

(d) $U^\dagger$ is unitary;

(e) The columns of $U$ form an orthonormal set;

(f) The rows of $U$ form an orthonormal set; and

(g) For all $x \in \mathbb{C}^n$ the Euclidean length of $y = Ux$ is the same as that of $x$; that is, $y^\dagger y = x^\dagger x$.

**Lemma 1.1.** (Schur) For any square matrix $A \in \mathbb{C}^{n \times n}$ there exists a unitary $V$ and upper triangular $R$ such that

$$A = VRV^\dagger.$$  

**Proof.** See Homework.

**Definition 1.2.** A matrix $A \in \mathbb{C}^{n \times n}$ is said to be Hermitian if $A = A^\dagger$. It is said to be Skew-Hermitian if $A = -A^\dagger$. 

Recall that an $n \times n$ matrix has exactly $n$ eigenvalues in $\mathbb{C}$.

**Lemma 1.2.** Let $H \in \mathbb{C}^{n \times n}$ be Hermitian. Then

(i) All $n$ eigenvalues of $H$ are real.

(ii) $H$ has a set of eigenvectors $\{u_i : i = 1, \ldots, n\}$ that form an orthonormal basis of $\mathbb{C}^n$.

**Proof.** Using the Schur lemma there exists a unitary $V$ and upper triangular $R$ such that $H = VRV^\dagger$. Since $H$ is Hermitian,

$$VRV^\dagger = (VRV^\dagger)^\dagger = VR^\dagger V^\dagger$$

from which we obtain $R = R^\dagger$. But since $R$ is upper triangular, this implies that $R$ is real and diagonal. It is now easy to verify that the columns of $V$ are the eigenvectors, with the $i$th column of $V$ being an eigenvector with eigenvalue $R_{ii}$. \qed

Notice that all covariance matrices are Hermitian.

**Exercise 1.1.** Show that if $A \in \mathbb{C}^{n \times n}$ is Hermitian, then $u^\dagger Au$ is real for all $u \in \mathbb{C}^n$.

A class of Hermitian matrices with a special positivity property arises naturally in many applications, including in communication theory. They provide one generalization to matrices of the notion of positive numbers.

**Definition 1.3.** An Hermitian matrix $A \in \mathbb{C}^{n \times n}$ is said to be **positive definite** if

$$u^\dagger Au > 0$$

for all non-zero $u \in \mathbb{C}^n$.

If the above strict inequality is weakened to $u^\dagger Au \geq 0$, then $A$ is said to be **positive semidefinite**. Implicit in these defining inequalities is the observation that if $A$ is Hermitian, the left hand side is always a real number.

## 2. Complex-Valued Random Variables

A complex-valued random variable $U$ (hereafter simply called complex random variable) is defined as a random variable of the form

$$U = U_R + jU_I, \quad j = \sqrt{-1},$$

where $U_R$ and $U_I$ are real-valued random variables.

The statistical properties of $U = U_R + jU_I$ are determined by the joint distribution $P_{U_RU_I}(u_R, u_I)$ of $U_R$ and $U_I$.

The expectation of a real random vector $\mathbf{x}$ is naturally generalized to the complex case

$$E[\mathbf{U}] = E[\mathbf{U}_R] + jE[\mathbf{U}_I].$$

Recall that the covariance matrix of two real-valued random vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$K_{\mathbf{x}\mathbf{y}} = \text{cov}[\mathbf{x}, \mathbf{y}] \triangleq E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{y} - E[\mathbf{y}])^T].$$

To specify the “covariance” of the two complex random vectors $\mathbf{U} = \mathbf{U}_R + j\mathbf{U}_I$ and $\mathbf{V} = \mathbf{V}_R + j\mathbf{V}_I$ the four covariance matrices

$$\begin{align*}
K_{\mathbf{U}_R\mathbf{V}_R} &= \text{cov}[\mathbf{U}_R, \mathbf{V}_R] \\
K_{\mathbf{U}_I\mathbf{V}_R} &= \text{cov}[\mathbf{U}_I, \mathbf{V}_R] \\
K_{\mathbf{U}_R\mathbf{V}_I} &= \text{cov}[\mathbf{U}_R, \mathbf{V}_I] \\
K_{\mathbf{U}_I\mathbf{V}_I} &= \text{cov}[\mathbf{U}_I, \mathbf{V}_I]
\end{align*}$$

(2)
are needed. These four real-valued matrices are equivalent to the following two complex-valued matrices, each of which is a natural generalization of (1)

\[ K_{UV} \triangleq E[(U - E[U])(V - E[V])^\dagger] \]
\[ J_{UV} \triangleq E[(U - E[U])(V - E[V])^T] \]

The reader is encouraged to verify that the following (straightforward) relationships hold:

\[ K_{UV} = K_{U_RV_R} + K_{U_IV_I} + j(K_{U_IV_R} - K_{U_RV_I}) \]
\[ J_{UV} = K_{U_RV_R} - K_{U_IV_I} + j(K_{U_IV_R} + K_{U_RV_I}) \]

This system may be solved for \( K_{U_RV_R}, K_{U_IV_I}, K_{U_IV_R}, \) and \( K_{U_RV_I} \) to obtain

\[ K_{U_RV_R} = \frac{1}{2} \Re\{K_{UV} + J_{UV}\} \]
\[ K_{U_IV_I} = \frac{1}{2} \Re\{K_{UV} - J_{UV}\} \]
\[ K_{U_IV_R} = \frac{1}{2} \Im\{K_{UV} + J_{UV}\} \]
\[ K_{U_RV_I} = \frac{1}{2} \Im\{-K_{UV} + J_{UV}\} \]

proving that indeed the four real-valued covariance matrices in (2) are in one-to-one relationship with the two complex-valued covariance matrices in (3).

In the literature \( K_{UV} \) is widely used and it is called covariance matrix (of the complex random vectors \( U \) and \( V \)). Hereafter \( J_{UV} \) will be called the pseudo-covariance matrix (of \( U \) and \( V \)). For notational simplicity we will write \( K_U \) instead of \( K_{UU} \) and \( J_U \) instead of \( J_{UU} \).

**Definition 2.1.** \( U \) and \( V \) are said to be **uncorrelated** if all four covariances in (2) vanish.

From (3), we now obtain the following.

**Lemma 2.1.** The complex random vectors \( U \) and \( V \) are uncorrelated iff \( K_{UV} = J_{UV} = 0 \).

**Proof.** The “if” part follows from (5) and the “only if” part from (4). \( \square \)

3. **Complex-Valued Random Processes**

We focus on discrete-time random processes since corresponding results for continuous-time random processes follow in a straightforward fashion.

A discrete-time complex random process is defined as a random process of the form

\[ U[n] = U_R[n] + jU_I[n] \]

where \( U_R[n] \) and \( U_I[n] \) are a pair of real discrete-time random processes.

**Definition 3.1.** A complex random process is **wide-sense stationary** (w.s.s.) if its real and imaginary parts are jointly w.s.s.

**Definition 3.2.** We define

\[ r_U[m,n] \triangleq E[U[n+m]U^*[n]] \]
\[ s_U[m,n] \triangleq E[U[n+m]U[n]] \]

as the autocorrelation and pseudo-autocorrelation functions of \( U[n] \).

**Lemma 3.1.** A complex random process \( U[n] \) is w.s.s. if and only if \( E[U[n]], r_U[m,n], \) and \( s_U[m,n] \) are independent of \( n \).

**Proof.** The proof is left as an exercise. \( \square \)
4. Proper Complex Random Variables

Proper random variables are of interest to us since they arise in practical applications and since they are mathematically easier to deal with than their non-proper counterparts. We will also see that proper Gaussian random vectors maximize entropy among all random variables of a given covariance matrix. This last property tells us that proper Gaussian random vectors are fundamentally important.

**Definition 4.1.** A complex random vector $\mathbf{U}$ is called *proper* if its pseudo-covariance $J_{\mathbf{U}}$ vanishes. The complex random vectors $\mathbf{U}_1$ and $\mathbf{U}_2$ are called *jointly proper* if the composite random vector $\begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix}$ is proper.

**Lemma 4.1.** Two jointly proper, complex random vectors $\mathbf{U}$ and $\mathbf{V}$ are uncorrelated, if and only if their covariance matrix $K_{\mathbf{UV}}$ vanishes.

*Proof.* The proof follows immediately from the definition of joint properness and Lemma 2.1. $\square$

Note that any subvector of a proper random vector is also proper. By this we mean that if $\begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix}$ is proper, then $\mathbf{U}_1$ and $\mathbf{U}_2$ are proper. However, two individual proper random vectors are not necessarily jointly proper.

Using the fact that (by definition) $K_{\mathbf{U}_R\mathbf{U}_I} = K_{\mathbf{U}_I\mathbf{U}_R}^T$, the pseudo-covariance matrix $J_{\mathbf{U}}$ may be written as

$$J_{\mathbf{U}} = (K_{\mathbf{U}_R} - K_{\mathbf{U}_I}) + j(K_{\mathbf{U}_I \mathbf{U}_R} + K_{\mathbf{U}_R}^T).$$

Thus:

**Lemma 4.2.** A complex random vector $\mathbf{U}$ is proper iff

$$K_{\mathbf{U}_R} = K_{\mathbf{U}_I}, \quad \text{and} \quad K_{\mathbf{U}_I \mathbf{U}_R} = -K_{\mathbf{U}_R}^T,$$

i.e. $J_{\mathbf{U}}$ vanishes, iff $\mathbf{U}_R$ and $\mathbf{U}_I$ have identical auto-covariance matrices and their cross-covariance matrix is skew-symmetric.\(^1\)

Notice that the skew-symmetry of $K_{\mathbf{U}_I \mathbf{U}_R}$ implies that $K_{\mathbf{U}_I \mathbf{U}_R}$ has a zero main diagonal, which means that the real and imaginary part of each component $U_k$ of $\mathbf{U}$ are uncorrelated. The vanishing of $J_{\mathbf{U}}$ does not, however, imply that the real part of $U_k$ and the imaginary part of $U_l$ are uncorrelated for $k \neq l$.

Notice that a *real* random vector is a proper complex random vector, if and only if it is constant (with probability 1), since $K_{\mathbf{U}_I} = 0$ and Lemma 4.2 imply $K_{\mathbf{U}_R} = 0$.

**Lemma 4.3 (Closure Under Affine Transformations).** Let $\mathbf{U}$ be a proper $n$-dimensional random vector, i.e., $J_{\mathbf{U}} = 0$. Then any vector obtained from $\mathbf{U}$ by an affine transformation, i.e. any vector $\mathbf{V}$ of the form $\mathbf{V} = A\mathbf{U} + \mathbf{b}$, where $A \in \mathbb{C}^{m \times n}$ and $\mathbf{b} \in \mathbb{C}^m$ are constant, is also proper.

*Proof.* From

$$E[\mathbf{V}] = AE[\mathbf{U}] + \mathbf{b}$$

it follows

$$\mathbf{V} - E[\mathbf{V}] = A(\mathbf{U} - E[\mathbf{U}]).$$

\(^1\)A matrix $A$ is skew-symmetric if $A^T = -A$. 

4
Hence we have

\[
J_V = E[(V - E[V])(V - E[V])^T]
= E\{A(U - E[U])(U - E[U])^T A^T\}
= AJ_U A^T = 0
\]

**Corollary 4.1.** Let \( U \) and \( V \) be as in the previous Lemma. Then \( U \) and \( V \) are jointly proper.

**Proof.** The vector having \( U \) and \( V \) as subvectors is obtained by the affine transformation

\[
\begin{bmatrix}
U \\
V
\end{bmatrix} = \begin{bmatrix}
I_n \\
A
\end{bmatrix} U + \begin{bmatrix}
0 \\
b
\end{bmatrix}.
\]

The claim now follows from Lemma 4.3.

**Lemma 4.4.** Let \( U \) and \( V \) be two independent complex random vectors and let \( U \) be proper. Then the linear combination \( W = a_1 U + a_2 V, a_1, a_2 \in \mathbb{C}, a_2 \neq 0 \), is proper iff \( V \) is also proper.

**Proof.** The independence of \( U \) and \( V \) and the properness of \( U \) imply

\[
J_W = a_1^2 J_U + a_2^2 J_V = a_2^2 J_V.
\]

Thus \( J_W \) vanishes iff \( J_V \) vanishes.

5. **Relationship Between Real-Valued and Complex-Valued Operations**

Consider now an arbitrary vector \( u \in \mathbb{C}^n \) (not necessarily a random vector), let \( A \in \mathbb{C}^{m \times n} \), and suppose that we would like to implement the operation that maps \( u \) to \( v = Au \). Suppose also that we implement this operation on a DSP which is programmed at a level at which we can’t rely on routines that handle complex-valued operations. A natural question is: how do we implement \( v = Au \) using real-valued operations? More generally, what is the relationship between complex-valued variables and operations with respect to their real-valued counterparts? We need this knowledge in the next section to derive the probability density function of proper Gaussian random vectors.

A natural approach is to define the operation that maps a general complex vector \( u \) into a real vector \( \hat{u} \) according to

\[
\hat{u} = \begin{bmatrix} u_R \\ u_I \end{bmatrix} \triangleq \begin{bmatrix} \Re[u] \\ \Im[u] \end{bmatrix}
\]

and hope for the existence of a real-valued matrix \( \hat{A} \) such that

\[
\hat{v} = \hat{A}\hat{u}.
\]

From \( \hat{v} \) we can then immediately obtain \( v \). Fortunately such a matrix exists and it is straightforward to verify that

\[
\hat{A} = \begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix} \triangleq \begin{bmatrix} \Re[A] & -\Im[A] \\ \Im[A] & \Re[A] \end{bmatrix}.
\]

A set of operations on complex-valued vectors and matrices and the corresponding real-valued operations are described in the following Lemma.
Lemma 5.1. The following properties hold:

\[ \hat{A}\hat{B} = \hat{A}\hat{B} \]
\[ \hat{A} + \hat{B} = \hat{A} + \hat{B} \]
\[ \hat{A}^\dagger = \hat{A}^\dagger \]
\[ \hat{A}^{-1} = \hat{A}^{-1} \]
\[ \det(\hat{A}) = |\det(A)|^2 = \det(AA^\dagger) \]
\[ \hat{u} + \hat{v} = \hat{u} + \hat{v} \]
\[ \hat{A}\hat{u} = \hat{A}\hat{u} \]
\[ \Re\{\hat{u}^\dagger\hat{v}\} = \hat{u}^\dagger\hat{v} \]

Proof. The properties (8a), (8b) and (8c) are immediate. For instance, property (8a) is verified as follows:

\[
\hat{A}\hat{B} = \begin{bmatrix} (AB)_R & -(AB)_I \\ (AB)_I & (AB)_R \end{bmatrix} = \begin{bmatrix} A_RB_R - A_I B_I & -A_RB_I - A_I B_R \\ A_RB_I + A_I B_R & A_RB_R - A_I B_I \end{bmatrix} = \hat{A}\hat{B}
\]

Property (8d) follows from (8a) and the fact that \( \hat{I}_n = I_{2n} \). To prove (8e) we use the fact that the determinant of a product is the product of the determinant and the determinant of a block triangular matrix is the product of the determinants of the diagonal blocks. Hence:

\[
\det(\hat{A}) = \det\left( \begin{bmatrix} I & iI \\ 0 & I \end{bmatrix} \hat{A} \begin{bmatrix} I & -iI \\ 0 & I \end{bmatrix} \right) = \det\left( \begin{bmatrix} A & 0 \\ \Im(A) & A^* \end{bmatrix} \right) = \det(A)\det(A^*).
\]

Properties (8f), (8g) and (8h) are immediate.

Corollary 5.1. If \( A \in \mathbb{C}^{n \times n} \) is unitary then \( \hat{A} \in \mathbb{R}^{2n \times 2n} \) is orthonormal.

Proof. \( U^\dagger U = I_n \Longleftrightarrow (\hat{U})^\dagger \hat{U} = \hat{I}_n = I_{2n}. \)

Corollary 5.2. If \( Q \in \mathbb{C}^{n \times n} \) is non-negative definite, then so is \( \hat{Q} \in \mathbb{R}^{2n \times 2n} \). Moreover, \( \hat{u}^\dagger \hat{Q}\hat{u} = \hat{u}^\dagger \hat{Q}\hat{u} \).

Proof. Assume that \( Q \) is non-negative definite. Then \( \hat{u}^\dagger \hat{Q}\hat{u} \) is a non-negative real-valued number for all \( \hat{u} \in \mathbb{C}^n \). Hence,

\[
\hat{u}^\dagger \hat{Q}\hat{u} = \Re\{\hat{u}^\dagger (\hat{Q}\hat{u})\} = \hat{u}^\dagger (\hat{Q}\hat{u}) = \hat{u}^\dagger \hat{Q}\hat{u}
\]

where in the last two equalities we used (8h) and (8g), respectively.

Exercise 5.1. A random vector \( U \) is proper iff \( 2K\hat{U} = \hat{K}_U \).
6. Complex-Valued Gaussian Random Variables

A complex-valued Gaussian random vector \( U \) is defined as a vector with jointly Gaussian real and imaginary parts. Following Feller\(^2\), we consider Gaussian distributions to include degenerate distributions concentrated on a lower-dimensional manifold, i.e., when the \( 2n \times 2n \)-covariance matrix

\[
\text{cov} \left( \begin{bmatrix} U_R \\ U_I \end{bmatrix}, \begin{bmatrix} U_R \\ U_I \end{bmatrix} \right) = \begin{bmatrix} K_{U_R} & K_{U_I} U_R \\ K_{U_I} U_R & K_{U_I} \end{bmatrix}
\]

is singular and the pdf does not exist unless one admits generalized functions.

Hence, by definition, a complex-valued random vector \( U \in \mathbb{C}^n \) with nonsingular covariance matrix \( \hat{K}_U \) is Gaussian iff

\[
f_U(u) = f_{\hat{U}}(\hat{u}) = \frac{1}{\sqrt{\det(2\pi \hat{K}_U)}} e^{-\frac{1}{2} (\hat{u} - \hat{m})^T \hat{K}_U^{-1} (\hat{u} - \hat{m})}. \tag{9}
\]

**Theorem 6.1.** Let \( U \in \mathbb{C}^n \) be a proper Gaussian random variable with mean \( \text{m} \) and nonsingular covariance matrix \( K_U \). Then the pdf of \( U \) is given by

\[
f_U(u) = f_{\hat{U}}(\hat{u}) = \frac{1}{\sqrt{\det(\pi K_U)}} e^{-(u - \text{m})^T K_U^{-1} (u - \text{m})}. \tag{10}
\]

Conversely, let the pdf of a random \( U \) be given by (10) where \( K_U \) is some Hermitian and positive definite matrix. Then \( U \) is proper and Gaussian with covariance matrix \( K_U \) and mean \( \text{m} \).

**Proof.** If \( U \) is proper then by Exercise 5.1

\[
\sqrt{\det 2\pi K_U} = \sqrt{\det \pi K_U} = | \det \pi K_U | = \det \pi K_U,
\]

where the last equality holds since the determinant of an Hermitian matrix is always real. Moreover, letting \( \hat{v} = \hat{u} - \hat{m} \), again by Exercise 5.1

\[
\hat{v}^T (2K_U)^{-1} \hat{v} = \hat{v}^T (K_U)^{-1} \hat{v} = v^T (K_U)^{-1} v
\]

where for last equality we used Corollary 5.2 and the fact that if a matrix is positive definite, so is its inverse. Using the last two relationships in (9) yields the direct part of the theorem. The converse follows similarly. \( \square \)

**Exercise 6.1.** Show that if \( U \) is a proper Gaussian random vector with zero mean, then the pdf of \( U \) is circularly symmetric, i.e., \( p_U(e^{j\theta}u) \) does not depend on \( \theta \). For this reason some authors use circular symmetric as a synonym of proper in the Gaussian case.

Notice that two jointly proper Gaussian random vectors \( U \) and \( V \) are independent, iff \( K_{UV} = 0 \), which follows from Lemma 4.1 and the fact that uncorrelatedness and independence are equivalent for Gaussian random variables.

7. Linear Transformations

The Fourier transform is a useful tool in dealing with linear time-invariant (LTI) systems. This is so since the input/output relationship if a LTI system is easily described in the Fourier domain. In this section we learn that this is just a special case of a more general principle that applies to linear transformations (not necessarily time-invariant). Key ingredients are the eigenvectors.

7.1. Linear Transformations, Toepliz, and Circulant Matrices

A linear transformation from $\mathbb{C}^n$ to $\mathbb{C}^n$ can be described by an $n \times n$ matrix $H$. If the matrix is Toepliz, meaning that $H_{ij} = h_{i-j}$, then the transformation which sends $u \in \mathbb{C}^n$ to $v = Hu$ can be described by the convolution sum

$$v_i = \sum_k h_{i-k}u_k.$$

A Toepliz matrix is a matrix which is constant along its diagonals.

In this section we focus attention on Toepliz matrices of a special kind called circulant. A matrix $H$ is circulant if $H_{ij} = h_{i-j}$ where here and hereafter the operator $[\cdot]$ applied to an index denotes the index taken modulo $n$. When $H$ is circulant, the operation that maps $u$ to $v = Hu$ may be described by the circulant convolution

$$v_i = \sum_k h_{i-k}u_k.$$

Example 7.1. 

$$H = \begin{bmatrix} 3 & 1 & 5 \\ 5 & 3 & 1 \\ 1 & 5 & 3 \end{bmatrix}$$

is a circulant matrix.

A circulant matrix $H$ is completely described by its first column $h$ (or any column or row for that matter).

7.2. The DFT

The discrete Fourier transform of a vector $u \in \mathbb{C}^n$ is the vector $U \in \mathbb{C}^n$ defined by

$$U = F^\dagger u$$

$$F = (f_1, f_2, \ldots, f_n)$$

$$f_i = \frac{1}{\sqrt{n}} \begin{bmatrix} \beta^0 \\ \beta^1 \\ \vdots \\ \beta^{(n-1)} \end{bmatrix} \quad i = 1, 2, \ldots, n,$$

(11)

where $\beta = e^{2\pi i/n}$ is the primitive $n$-th root of unity in $\mathbb{C}$. Notice that $f_1, f_2, \ldots, f_n$ is an orthonormal basis for $\mathbb{C}^n$.

Usually, the DFT is defined without the $\sqrt{n}$ in (11) and with a factor $1/n$ (instead of $1/\sqrt{n}$) in the inverse transform. The resulting transformation is not orthonormal, and a factor $n$ must be inserted in Parseval’s identity when it is applied to the DFT. In this class we call $F^\dagger u$ the DFT of $u$.

7.3. Eigenvectors of Circulant Matrices

Lemma 7.1. Any circulant matrix $H \in \mathbb{C}^{n \times n}$ has exactly $n$ (normalized) eigenvectors which may be taken as $f_1, \ldots, f_n$. Moreover, the vector of eigenvalues $(\lambda_1, \ldots, \lambda_n)^T$ equals $\sqrt{n}$ times the DFT of the first column of $H$, namely $\sqrt{n}F^\dagger h$. 

8
Example 7.2. Consider the matrix

\[
H = \begin{bmatrix}
h_0 & h_1 \\
h_1 & h_0 \\
\end{bmatrix} \in \mathbb{C}^{2 \times 2}.
\]

This is a circulant matrix. Hence

\[
f_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad f_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\]

are eigenvectors and the eigenvalues are

\[
\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \sqrt{2} F^\dagger h = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} h_0 - h_1 \\ h_0 + h_1 \end{bmatrix}
\]

indeed

\[
H f_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} h_0 - h_1 \\ h_1 - h_0 \end{bmatrix} = \frac{h_0 - h_1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \lambda_1 f_1
\]

and

\[
H f_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} h_0 + h_1 \\ h_1 + h_0 \end{bmatrix} = \frac{h_0 + h_1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda_2 f_2
\]

Proof.

\[
(H f_i)_k = \frac{1}{\sqrt{n}} \sum_{e=0}^{n-1} h_{k-e\beta^{ie}}
\]

\[
= \left( \sum_{m=0}^{n-1} h_m \beta^{-im} \right) \frac{1}{\sqrt{n}} \beta^{ik}
\]

\[
= \sqrt{n} f_i^\dagger h \frac{1}{\sqrt{n}} \beta^{ik} = \lambda_i \frac{1}{\sqrt{n}} \beta^{ik},
\]

where \( \lambda_i = \sqrt{n} f_i^\dagger h \). Going to vector notation we obtain \( H f_i = \lambda_i f_i \).

7.4. Eigenvectors to Describe Linear Transformations

When the eigenvectors of a transformation \( H \in \mathbb{C}^{n \times n} \) (not necessarily Toepliz) span \( \mathbb{C}^n \), both the vectors and the transformation can be represented with respect to a basis of eigenvectors. In that new basis the transformation takes the form \( H' = \text{diag}(\lambda_1, \ldots, \lambda_n) \), where \( \text{diag}(\cdot) \) denotes a matrix with the arguments on the main diagonal and 0s elsewhere, and \( \lambda_i \) is the eigenvalue of the \( i \)-th eigenvector. In the new basis the input/output relationship is

\[
v' = H' u'
\]

or equivalently, \( v'_i = \lambda_i u'_i, \ i = 1, 2, \ldots, n \). To see this, let \( \varphi_i, i = 1 \ldots n \), be \( n \) eigenvectors of \( H \) spanning \( \mathbb{C}^n \). Letting \( u = \sum_i \varphi_i u'_i \) and \( v = \sum_i \varphi_i v'_i \) and plugging into \( Hu \) we obtain

\[
Hu = H \left( \sum_i \varphi_i u'_i \right) = \sum_i H \varphi_i u'_i = \sum_i \varphi_i \lambda_i u'_i
\]

showing that \( v'_i = \lambda_i u'_i \).

Notice that the key aspects in the proof are the linearity of the transformation and the fact that \( \varphi_i u'_i \) is sent to \( \varphi_i \lambda_i u'_i \), as shown in Fig. 1.
It is often convenient to use matrix notation. To see how the proof goes with matrix notation we define $\Phi = (\varphi_1, \ldots, \varphi_n)$ as the matrix whose columns span $\mathbb{C}^n$. Then $u = \Phi u'$ and the above proof in matrix notation is

$$v = H u = H \Phi u' = \Phi H' u',$$

showing that $v' = H' u'$.

For the case where $H$ is circulant, $u = F u'$ and $v = F v'$. Hence $u' = F^\dagger u$ and $v' = F^\dagger v$ are the DFT of $u$ and $v$, respectively. Similarly, the diagonal elements of $H'$ are $\sqrt{n}$ times the DFT of the first column of $H$. Hence the above representation via the new basis says (the well-know result) that a circular convolution corresponds to a multiplication in the DFT domain.

8. Karhunen-Loève Expansion

We have seen that the eigenvectors of a linear transformation $H$ can be used as a basis and in the new basis the linear transformation of interest becomes a componentwise multiplication.

A similar idea can be used to describe a random vector $u$ as a linear combination of deterministic vectors with orthogonal random coefficients. Now the eigenvectors are those of the correlation matrix $r_u$. The procedure, that we now describe, is the Karhunen-Loève expansion.

Let $\varphi_1, \ldots, \varphi_n$ be a set of eigenvectors of $r_u$ that form an orthonormal basis of $\mathbb{C}^n$. Such a set exists since $r_u$ is Hermitian. Hence

$$\lambda_i \varphi_i = r_u \varphi_i, i = 1, 2, \ldots, n$$

or, in matrix notation,

$$\Phi \Lambda = r_u \Phi$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and $\Phi = [\varphi_1, \ldots, \varphi_n]$ is the matrix whose columns are the eigenvectors. Since the eigenvectors are orthonormal, $\Phi$ is unitary (i.e. $\Phi^\dagger \Phi = I$). Solving for $\Lambda$ we obtain

$$\Lambda = \Phi^\dagger r_u \Phi.$$

Notice that if we solve for $r_u$ we obtain $r_u = \Phi \Lambda \Phi^\dagger$ which is the well known result that an Hermitian matrix can be diagonalized.

Since $\Phi$ forms a basis of $\mathbb{C}^n$ we can write

$$u = \Phi u'$$

for some vector of coefficient $u'$ with correlation matrix

$$r_{u'} = E[u'(u')^\dagger] = \Phi^\dagger E[u u^\dagger] \Phi = \Phi^\dagger r_u \Phi = \Lambda$$
Hence (12) expresses $u$ as a linear combination of deterministic vectors $\varphi_1, \ldots, \varphi_n$ with orthogonal random coefficients $u'_1, \ldots, u'_n$. This is the Karhunen-Loève expansion of $u$.

If $r_u$ is circulant, then $\Phi = F$ and $u' = \Phi^* u$ is the DFT of $u$.

**Remark 8.1.** $\|u\|^2 = \|u'\|^2 = \sum |u'_i|^2$. Also $E\|u\|^2 = \sum \lambda_i$.

### 9. Circularly Wide-Sense Stationary Random Vectors

Now we have all the background material for the next result. We consider random vectors in $\mathbb{C}^n$. We will continue using the notation that $u$ and $U$ denote DFT pairs. Observe that if $U$ is random then $u$ is also random. This forces us to abandon the convention that we use capital letters for random variables.

The following definitions are natural.

**Definition 9.1.** A random vector $u \in \mathbb{C}^n$ is **circularly wide sense stationary (c.w.s.s.)** if

\[
m_u \triangleq E[u] \text{ is a constant vector} \\
r_u \triangleq E[uu^\dagger] \text{ is a circulant matrix} \\
s_u \triangleq E[uu^T] \text{ is a circulant matrix}
\]

**Definition 9.2.** A random vector $u$ is uncorrelated if $K_u$ and $J_u$ are diagonal.

We will call $r_u$ and $s_u$ the **circular correlation matrix** and **circular pseudo-correlation matrix**, respectively.

The following Theorem is a key tool for finding the capacity of a complex discrete-time Gaussian channel with memory.

**Theorem 9.1.** Let $u \in \mathbb{C}^n$ be a zero-mean proper random vector and $U = F^\dagger u$ be its DFT. Then $u$ is c.w.s.s. iff $U$ is uncorrelated. Moreover,

\[r_u = \text{circ}(a)\] (13)

if and only if

\[r_U = \sqrt{n} \text{ diag}(A)\] (14)

for some $a$ and its DFT $A$.

**Proof.** Let $u$ be a zero-mean proper random vector. If $u$ is c.w.s.s. then we can write $r_u = \text{circ}(a)$ for some vector $a$. Then, using Lemma 7.1,

\[
r_U \triangleq E[F^\dagger uu^\dagger F] = F^\dagger r_u F \\
= F^\dagger \sqrt{n} F \text{ diag}(F^\dagger a) \\
= \sqrt{n} \text{ diag}(A),
\]

proving (14). Moreover, $m_U = 0$ since $m_u = 0$ and therefore $s_U = J_U$. But $J_U = 0$ since the properness of $u$ and Lemma 4.3 imply the properness of $U$. Conversely, let $r_U = \text{diag}(A)$. Then

\[r_u = E[uu^\dagger] = Fr_U F^\dagger.\]
Due to the diagonality of $r_U$, the element $(k, l)$ of $r_u$ is

$$\sqrt{n} \sum_m F_{k,m} A_m (F^\dagger)_m = \sum_m F_{k,m} F^*_{l,m} A_m \sqrt{n}$$

$$= \frac{1}{\sqrt{n}} \sum_m A_m e^{\frac{2\pi i m(k-l)}{n}}$$

$$= a_{k-l}$$