LECTURE 9

DISCRETE SIGNALS
(or processing of signals that were discretized and became vectors in a high dimensional space)

9.1. From this point on we shall deal with signals/images/data that are vectors in some N-dimensional space. We, by now, understand that these vectors are representations of some continuous entities that have been digitized/digitized and now are just vectors of values in $\mathbb{R}$ or $\mathbb{C}$. We shall see that the continuous domain results on operators map nicely to matrix/vector operations in the discrete domain.
A discrete signal $\overline{\phi}$ will be denoted by a vector

$$\overline{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix}$$

and we shall write $\overline{\phi}$ as follows:

$$\overline{\phi} = \sum_{k=1}^{N} \phi_k \overline{\beta}_k^S = \phi_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \phi_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \phi_N \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

and we call this the representation of $\overline{\phi}$ in the standard basis for $\mathbb{R}^N$, $\{\overline{\beta}_k^S\}_{k=1}^{N}$

where $\overline{\beta}_k^S = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ at $k^{th}$ location.

Clearly the standard basis $\{\overline{\beta}_k^S\}_{k=2}^{N}$ obeys

$$<\overline{\beta}_k^S, \overline{\beta}_l^S> = \overline{\beta}_k^S \overline{\beta}_l^s = \delta_{kl} = \begin{cases} 1 & \text{if } k=l \\ 0 & \text{if } k \neq l \end{cases}$$
Note that the standard representation can be written as
\[ \overline{\varphi} = I_{nxn} \, \overline{\varphi} \]
where \( \overline{\varphi} \) here becomes the vector of coefficients that multiply the columns of \( I_{nxn} \) that are the standard basis for vectors in \( \mathbb{R}^n \).

9.2. Given any unitary matrix \( U \), whose columns are clearly an alternative basis for \( \mathbb{T}^n \) or \( \mathbb{C}^n \) we have
\[ UU^* = U^* U = I_{nxn} \]
hence we can readily write:
\[ \overline{\varphi} = UU^* \overline{\varphi} = U \begin{bmatrix} \langle \overline{u}_1, \overline{\varphi} \rangle \\ \langle \overline{u}_2, \overline{\varphi} \rangle \\ \vdots \\ \langle \overline{u}_n, \overline{\varphi} \rangle \end{bmatrix} \]
and we have from this the representation of \( \overline{\varphi} \) in the basis formed by the columns of the matrix \( U \).
\[ \Phi = \sum_{k=1}^{N} \langle \bar{u}_k, \Phi \rangle \bar{u}_k \]

if \( \Phi = \begin{bmatrix} \bar{u}_1 & \bar{u}_2 & \ldots & \bar{u}_n \end{bmatrix} \).

As we have seen in the continuous case, we can use for \( \Phi \), the Hadamard matrices or the Haar matrices and using them we have representations of the vector \( \Phi \) in various bases providing the benefits of better \( K \)-term approximations, as discussed before.

In the discrete case we have, clearly, that all these representations, i.e., the coefficient vectors that correspond to the columns of \( \Phi \), are zero error descriptions of the original \( \Phi \). Indeed
We have clearly

\[ \overline{\varphi}^* \overline{\varphi} = \langle \overline{\varphi}, \overline{\varphi} \rangle = \sum_{k=1}^{N} \overline{\phi_k}^2 = \]

\[ = \overline{\varphi}^* \mathbf{U} \mathbf{U}^* \overline{\varphi} = \sum_{k=1}^{N} \langle \overline{\varphi}, \overline{u_k} \rangle^2 = \sum_{k=1}^{N} (\phi_k^m)^2 \]

and the total energy (Squared length) of the vector \( \overline{\varphi} \) equals the Energy of the vector with the new representation coefficients.

However, if we consider partial sums of the form

\[ \overline{\varphi}(K) = \sum_{k=1}^{K} \langle \overline{u}_k, \overline{\varphi} \rangle \overline{u}_k \]

we get approximations that improve as \( K \) goes from 1 to \( N \), eventually reaching a MSE of zero for \( K = N \). The way to the perfect description is however very
important (as we have seen in the continuos case too!)

Clearly the error vector is

\[ \overline{E}(K) = \overline{\Phi} - \hat{\overline{\Phi}}(K) = \]

\[ = \sum_{k=1}^{N} \langle \overline{\Phi}, \overline{U}_k \rangle \overline{U}_k - \sum_{k=1}^{K} \langle \overline{\Phi}, \overline{U}_k \rangle \overline{U}_k = \]

\[ = \sum_{k=K+1}^{N} \langle \overline{\Phi}, \overline{U}_k \rangle \overline{U}_k = \sum_{k=K+1}^{N} \phi_k \overline{U}_k \]

and the (mean) squared error in this approximation is the squared length of \( \overline{E}(K) \)

i.e.

\[ \overline{E}(K)^T \overline{E}(K) = \left( \sum_{k=K+1}^{N} \phi_k \overline{U}_k \right)^T \left( \sum_{k=K+1}^{N} \phi_k \overline{U}_k \right) = \]

\[ = \sum_{k=K+1}^{N} (\phi_k \overline{U}_k)^2 = \sum_{k=K+1}^{N} \phi_k^2 \overline{U}_k^2 = \sum_{k=K+1}^{N} \phi_k^2 \]

Using orthonormality of the vectors of \( \overline{U}_k \), \( k = 1, 2, \ldots, N \)
The various unitary matrices that we have encountered, and many others that were proposed by researchers in the field, differ in the rate at which the squared approximation error decreases to zero at $K=N$. For various classes of signals, i.e. sets of vectors $\bar{P}$ denoted by

$$\bar{P}_w = \{ \text{a set of vectors } \bar{P} \text{ indexed by } w \} \quad \text{(with } w \in \mathbb{R})$$

one can often determine the best $U$ matrix in the sense of yielding the fastest rate of decrease with $K$ in the $K$-approximation process we discussed. We shall see this theory in the sequel.
9.3. OPERATIONS ON DISCRETE SIGNALS

As in the continuous case one can devise various operators acting on vectors $\vec{p}$ yielding some "output" signal vectors $\vec{p}_{out}$. We denote such operators by $X$. Like in the continuous case and write

$$\vec{p}_{out} = X \vec{p}_{in}$$

or schematically we draw:

\[
\begin{array}{c}
\vec{p}_{in} \\
\downarrow \\
X \\
\downarrow \\
\vec{p}_{out}
\end{array}
\]

In fact all signal and image processing algorithms take a vector input $\vec{p}_{in}$ and produce an "improved" vector $\vec{p}_{out}$.
As in the continuous case, there are many possible operators $\mathcal{X}$ that one may design. The important operators for us here will be the linear ones.

An operator $\mathcal{X}$ is linear if

$$\mathcal{X}\{a \bar{\psi}_1 + b \bar{\psi}_2\} = a \mathcal{X}\{\bar{\psi}_1\} + b \mathcal{X}\{\bar{\psi}_2\}$$

for $a, b$ scalars in $\mathbb{R}$ or $\mathbb{C}$.

Clearly, this shows that

$$\mathcal{X}\{\bar{\psi}_j\} = \mathcal{X}\{\sum_{k=1}^{N} \phi_k \bar{\psi}_k\} =$$

$$= \sum_{k=1}^{N} \phi_k \mathcal{X}\{\bar{\psi}_k\} =$$

$$= \begin{bmatrix}
\mathcal{X}\{\bar{\psi}_1\} & \mathcal{X}\{\bar{\psi}_2\} & \cdots & \mathcal{X}\{\bar{\psi}_N\}
\end{bmatrix}
\begin{bmatrix}
\phi_1 \\
\phi_2 \\
\vdots \\
\phi_N
\end{bmatrix} =\begin{bmatrix}
\psi_1 \\
\psi_2 \\
\vdots \\
\psi_N
\end{bmatrix} = \bar{\phi}_{\text{out}}$$
We conclude that any linear operator acting on a signal vector $\vec{v}$ is simply a matrix multiplying the vector $\vec{v}$. The matrix, which we shall denote by $H$, has columns which denote the "impulse responses" of the operator $H$, i.e., its output to the set of vectors $\vec{p}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\vec{p}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, ..., $\vec{p}_k = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$.

Denoting by $h_k$ the response to $\vec{p}_k$, i.e.,

$\vec{h}_k = H \vec{p}_k$ we get

$\vec{v}_{\text{out}} = \begin{bmatrix} h_1 & h_2 & \cdots & h_n \end{bmatrix} \vec{v}_{\text{in}}$

$H$ - the linear system matrix
9.4. Shift Invariance

A linear system will be called shift invariant if its response to a cyclically shifted input vector is the cyclically shifted response to the original input.

A cyclic shift of a vector $\mathbf{\phi}$ is the following operator $J^k$:

$$J^k\left\{ \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix} \right\} = \begin{bmatrix} \phi_N \\ \phi_1 \\ \vdots \\ \phi_{N-1} \end{bmatrix}$$

The operator $J^k$ is clearly linear, hence it is described by a matrix, whose columns are the impulse responses of $J^k$.

$$I = \begin{bmatrix} J^k \{0\} & J^k \{0\} & \cdots & J^k \{0\} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ J^k \{0\} & J^k \{0\} & \cdots & J^k \{0\} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$
An operator is shift invariant if it commutes with the shift operator \( T \) i.e. if:

\[ X \{ T \{ \bar{\phi} \} \} = T \{ X \{ \bar{\phi} \} \} \quad \text{for all} \quad \bar{\phi}. \]

Written in matrix notation we have:

\[
\begin{align*}
H \cdot \bar{\phi} &= \mathbf{T} \cdot \bar{\phi} \\
\text{the process by which} \quad \mathbf{H} \mathbf{x} &= \mathbf{y} \\
\text{yields} \quad \mathbf{H}, \quad \text{we obtain that} \\
\end{align*}
\]

\[
\begin{align*}
H &= \begin{bmatrix}
\mathbf{H}_{LSE} \left[ \begin{array}{c} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \end{array} \right]
\end{bmatrix} \\
&= \begin{bmatrix}
h_0 & h_{N-1} & h_{N-2} & \cdots & h_1 \\
h_1 & h_0 & h_{N-1} & \cdots & h_2 \\
h_2 & h_1 & h_0 & \cdots & h_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
h_{N-1} & h_{N-2} & h_{N-3} & \cdots & h_0
\end{bmatrix}
\end{align*}
\]

Note that we used indices from 0 to \( N-1 \) here, for future convenience!
We see that $H_{LSI}$ has a very interesting structure: the values on all the diagonals are the same – this makes $H$ into a *Toeplitz* matrix and the first row is cyclically shifted and produces all the subsequent rows as its shifts by one, two, \ldots and $(n-1)$ places – this makes $H_{LSI}$ a *circulant matrix*.

**Toeplitz Circulant Matrices have a series of wonderful properties**, as we shall see!

Clearly we must have:

(a) $H_{LSI}^T = TH_{LSI}$

Note that $T$ is also a *Toeplitz circulant matrix*, i.e. the cyclic shift operator is a linear and shift invariant operator too!
In general we shall see that:

\[(b) \quad H_{LSI}^1 \cdot H_{LSI}^2 = H_{LSI}^2 \cdot H_{LSI}^1 \triangleq H_{LSI}^{1+2}\]

The product of two circulant matrices commutes and yields a circulant matrix.

\[(c) \quad \text{Circulant Matrices have eigenvectors that are independent of their entries!}\]

This property states that there is a set of orthonormal vectors \(\beta_1^c, \beta_2^c, \ldots, \beta_n^c\), so that for any \(H_{LSI}\)

We have

\[H_{LSI}^k \beta_k^c = \lambda_k^c (H_{LSI}) \beta_k^c\]

This means that we can write:

\[H_{LSI} \begin{bmatrix} \beta_1^c & \beta_2^c & \cdots & \beta_n^c \end{bmatrix} = \begin{bmatrix} \beta_1^c & \beta_2^c & \cdots & \beta_n^c \end{bmatrix} \begin{bmatrix} \lambda_1^H & \lambda_2^H & \cdots & \lambda_n^H \end{bmatrix}\]

or

\[H_{LSI} = \beta \Lambda \beta^* \quad \beta - \text{unitary!}\]
We shall prove these properties by considering the third one first.

Indeed given property (c) the others are immediate consequences since we have that

\[ H_{LSI}^{(0)} = \beta \lambda \beta^* \quad H_{LSI}^{(2)} = \beta \lambda \beta^* \quad \text{and} \quad T = \beta \lambda \beta^* \]

hence

\[ H_{LSI}^{(0)} + H_{LSI}^{(2)} = \beta \lambda \beta^* + \beta \lambda \beta^* = \beta \lambda \lambda \beta^* = I_{\lambda \lambda} \]

\[ = \beta \lambda \beta^* = H_{LSI}^{(2)} \cdot H_{LSI}^{(0)} \quad \text{(6)} \]

and the same is true for \( H^{(2)} = T \) proving property (c).

The amazing property of circulants is even more amazing when we realize that the \( \beta \)-matrix of its eigenvectors is the \([DFT]^*\) matrix.

This is what we prove next!
IMPORTANT RESULT

Circulant Matrices have eigenvectors given by the columns of the DFT matrix.

Proof: (a straightforward calculation)

\[
\begin{bmatrix}
    h_0 & h_{N-1} & \cdots & h_1 \\
    h_1 & h_0 & \cdots & h_2 \\
    \vdots & \vdots & \ddots & \vdots \\
    h_{N-1} & h_{N-2} & \cdots & h_0
\end{bmatrix}
\begin{bmatrix}
    W^{k_0} \\
    W^{k_1} \\
    \vdots \\
    W^{k_{N-1}}
\end{bmatrix}
= 
\begin{bmatrix}
    h_0 W^{k_0} + h_{1} W^{k_1} + \cdots + h_{l-1} W^{k_{l-1} \mod N} W^{k_{N-1}} \\
    h_1 W^{k_0} + h_0 W^{k_1} + \cdots + h_{l-2} W^{k_{l-2} \mod N} W^{k_{N-1}} \\
    \vdots \\
    h_{N-1} W^{k_0} + h_{N-2} W^{k_1} + \cdots + h_0 W^{k_{N-1}}
\end{bmatrix}
\]

Let us consider the \( l \)-th entry of the vector:

\[
h_0 W^{k_0} + h_{1} W^{k_1} + \cdots + h_{l-1} W^{k_{l-1} \mod N} W^{k_{N-1}} + h_0 W^{k_{l+1}} + h_{N-1} W^{k_{N-1} - (N-l) \mod N} = 
\begin{bmatrix}
1 & 2 & \cdots & (N-l+1) \\
(N-1) & 0 & 1 & 2 & \cdots & (N-l+1) - l
\end{bmatrix}
\]

\( 1 \text{st}, 2 \text{nd}, 3 \text{rd}, \ldots, (N-l+1) \text{th} \)
\[\begin{align*}
&= h_e^k W + h_{e-1}^k W + \ldots + h_0^k W + \\
&\quad + h_{N-1}^k W + h_{N-2}^k W + \ldots + h_{(N-1)-(N-(l+1))-1}^k W = \\
&\quad \left(\begin{array}{c}
\begin{array}{cccc}
k(0) & k(1) & k(2) & k(N-1) \\
0 & 1 & 2 & 1
\end{array} \\
\begin{array}{cc}
h_0 W & h_1 W \\
W & h_{N-1} W + h_{N-2} W + \ldots + h_{(N-1)-1} W \\
& + h_{(N-1)-2} W + \ldots + h_0 W
\end{array}
\end{array}\right) \\
&\quad \left(\begin{array}{c}
\begin{array}{cccc}
k(-l) & k(-l+1) & k(-l+2) & k(N-l) \\
-l & 1 & 1 & 1
\end{array} \\
\begin{array}{cc}
h_0 W & h_1 W \\
W & h_{N-1} W + h_{N-2} W + \ldots + h_{(N-1)-1} W \\
& + h_{(N-1)-2} W + \ldots + h_0 W
\end{array}
\end{array}\right) \\
&\quad \left(\begin{array}{c}
\begin{array}{cc}
k(N-l+1) \\
1
\end{array} \\
\begin{array}{cc}
h_0 W & h_1 W \\
W & h_{N-1} W + h_{N-2} W + \ldots + h_{(N-1)-1} W \\
& + h_{(N-1)-2} W + \ldots + h_0 W
\end{array}
\end{array}\right) \\
&\quad = W^{kl} \left[h_0, h_{N-1}, h_{N-2}, \ldots, h_1, h_0, h_{N-1}, h_{N-2}, \ldots, h_1\right] \left[h_0, W, W, \ldots, W\right]
\end{align*}\]
Therefore we obtain:

\[
\begin{bmatrix}
W^0 \\
W^1 \\
\vdots \\
W^{k(N-1)} \\
\end{bmatrix}
\begin{bmatrix}
H_{151}^{(k)} \\
\vdots \\
H_{151}^{(k(N-1))} \\
\end{bmatrix}
= \langle \text{first row of } H_{151}^{(k(N-1))}, \begin{bmatrix}
W^0 \\
W^1 \\
\vdots \\
W^{k(N-1)} \\
\end{bmatrix} \rangle 
\]

which is the kth eigenvalue.

Hence we see that, indeed, the columns of \( D_1 \) are the eigenvectors of \( H_{151} \) (any circulant matrix) and the eigenvalues are given by

\[
\begin{align*}
\lambda_0 &= \begin{bmatrix} W^0 & W^1 & \cdots & W^{(N-1)} \end{bmatrix} \begin{bmatrix} h_0 \\ h_{-1} \\ h_{-2} \\ \vdots \end{bmatrix} \\
\lambda_1 &= \begin{bmatrix} W^0 & W^1 & \cdots & W^{(N-1)} \end{bmatrix} \begin{bmatrix} h_0 \\ h_{-1} \\ h_{-2} \\ \vdots \end{bmatrix} \\
\vdots & \vdots \\
\lambda_{(N-1)} &= \begin{bmatrix} W^0 & W^1 & \cdots & W^{(N-1)} \end{bmatrix} \begin{bmatrix} h_0 \\ h_{-1} \\ h_{-2} \\ \vdots \end{bmatrix}
\end{align*}
\]

or written in a matrix form.
\[
\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} = \text{DFT}^* \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_N \end{bmatrix}
\]

Hence the eigenvectors of $\mathbf{H}_{LSI}$ are the columns (or rows) of $\text{DFT}^*$ and the eigenvalues are given by the $[\text{DFT}^*]$ matrix applied to the first row of $\mathbf{H}_{LSI}$, i.e. the reverse ordered impulse response of the operator $x_{LSI}$.
In summary, a linear shift-invariant operator in the discrete case yields a circulant-Toeplitz matrix action on the input vectors $\mathbf{x}$ and

$$
\mathbf{x}_{\text{LSI}} \rightarrow \mathbf{H}_{\text{LSI}} \mathbf{x}_{\text{LSI}} \text{ obeys:}
$$

$$
\mathbf{H}_{\text{LSI}} (\text{DFT}^*) = (\text{DFT})^* \begin{bmatrix}
\lambda_0^H & \lambda_1^H & 0 \\
0 & \lambda_2^H & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_{N-1}^H
\end{bmatrix}
$$

Hence

$$
[\text{DFT}] \mathbf{H}_{\text{LSI}} [\text{DFT}]^* = \text{diag}[\lambda_0, \lambda_1, \ldots, \lambda_{N-1}]
$$

where

$$
\begin{bmatrix}
\lambda_0^H \\
\lambda_1^H \\
\vdots \\
\lambda_{N-1}^H
\end{bmatrix} = [\text{DFT}^*] 
\begin{bmatrix}
h_0^* \\
h_{N-1}^*
\end{bmatrix}
$$

and also, every $\mathbf{H}_{\text{LSI}}$ (i.e., every Toeplitz circulant) can be written as:

$$
\mathbf{H}_{\text{LSI}} = [\text{DFT}^*] 
\begin{bmatrix}
\lambda_0^H & \lambda_1^H & 0 \\
0 & \lambda_2^H & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_{N-1}^H
\end{bmatrix} [\text{DFT}]
$$
95. The action of a circulant matrix, i.e. a linear shift invariant operator on a signal vector $\vec{x}$ is a discrete version of the convolution process that we have seen in the continuous case.

Consider a circulant matrix's action on a vector $\vec{x}$ as follows:

$$
\begin{pmatrix}
    h_0 & h_{N-1} & h_{N-2} & h_1 \\
    h_1 & h_0 & h_{N-1} & h_2 \\
    h_2 & h_1 & h_0 & h_3 \\
    \vdots & \vdots & \vdots & \vdots \\
    h_{N-1} & h_{N-2} & h_{N-3} & h_0
\end{pmatrix}
\begin{pmatrix}
    x_0 \\
    x_1 \\
    x_2 \\
    \vdots \\
    x_{N-1}
\end{pmatrix} = 
\begin{pmatrix}
    y_0 \\
    y_1 \\
    y_2 \\
    \vdots \\
    y_{N-1}
\end{pmatrix}
$$

We have that

$$
y_l = h_l x_0 + h_{l-1} x_1 + \ldots + h_0 x_l + 
+ h_{N-1} x_{l+1} + \ldots + h_{N-1-l} x_{N-1-l}
$$

$$
\uparrow \downarrow
l+1 \quad [N-1-l]
$$
Hence we can write:

\[ y_l = \sum_{k=0}^{N-1} x_k h(l-k) \mod N \]

\[ (l+1 \to l-(l+1) = -1 \mod N = (N-1) \]

\[ l+2 \to l-(l+2) = -2 \mod N = (N-2) \]

\[ N-1 = l \quad [l+1], \quad l \quad [l+(N-1)-l] = l+1-N = l+1 \mod N \]

This process is called "cyclic convolution" and we write

\[ y = \overline{h} \circ \overline{x} \]

We have clearly that

\[ \overline{h} \circ \overline{x} = \overline{x} \circ \overline{h} \]
Indeed:

\[
\sum_{k=0}^{N-1} x_k h((l-k) \mod N) = \sum_{r=0}^{N-1} h_r x_r (-r) \mod N
\]

since defining \((l-k) \mod N = r\)

\[
\Rightarrow l - k = r + \alpha N
\]

\[
l - r = k + \alpha N
\]

\[
\Rightarrow k = (l - r) \mod N
\]

(clearly as \(k\) runs from 0 to \(N-1\), \((l-k) \mod N\) runs over all these)

Using the result on diagonalization of circulants we therefore have

\[
[DFT] \begin{bmatrix}
x_0^N & x_1^N & \cdots & x_{N-1}^N \\
0 & x_0^N & \cdots & x_{N-2}^N \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & x_0^N
\end{bmatrix} =
[DFT] \begin{bmatrix}
x_0 \\
x_1 \\
\vdots \\
x_{N-1}
\end{bmatrix} = \begin{bmatrix}
y_0 \\
y_1 \\
\vdots \\
y_{N-1}
\end{bmatrix}
\]
yielding
\[
\begin{bmatrix}
\lambda_0^* & \lambda_1^* & \cdots & 0 \\
0 & \lambda_0 & \cdots & \lambda_1^* \\
\cdots & \cdots & \cdots & \cdots \\
0 & \lambda_{N-1} & \cdots & \lambda_0^*
\end{bmatrix}
\begin{bmatrix}
x_0 \\
x_1 \\
\vdots \\
x_{N-1}
\end{bmatrix}
= [DFT]
\begin{bmatrix}
y_0 \\
y_1 \\
\vdots \\
y_{N-1}
\end{bmatrix}
\]

Now recall that
\[
\begin{bmatrix}
\lambda_0^* \\
\lambda_1^* \\
\cdots \\
\lambda_N^*
\end{bmatrix}
= [DFT]^* \begin{bmatrix}
h_0 \\
h_{N-1} \\
\vdots \\
h_1
\end{bmatrix}
= [DFT]^*
\]

The matrix: \( AC \) is "anti-circulant" as the columns are shifted up rather than down for the rows are shifted to the left rather than to the right. We clearly have that
\[ [AC][AC] = I \quad (AC \text{ is a permutation matrix!}) \]

but we also have that (amazingly)

\[ [DFT][DFT] = AC \quad \text{(also implying that } [DCT][DCT] = I \text{)} \]

as it is easy to verify by direct computation.

Hence readily we get that:

\[
\begin{bmatrix}
\lambda_0^* \\
\lambda_1^* \\
\vdots \\
\lambda_N^*
\end{bmatrix}
= \left[ DFT^* \right] \left[ DFT \right] \left[ DFT \right]
\begin{bmatrix}
h_0 \\
h_1 \\
\vdots \\
h_{N-1}
\end{bmatrix}
= [DFT]
\begin{bmatrix}
h_0 \\
h_1 \\
\vdots \\
h_{N-1}
\end{bmatrix}
\]
Hence we obtained the RESULT:

If \( y = x \circ h \)

Then \( \text{DFT}y = \text{DFT}x \circ \text{DFT}h \)

where \( \circ \) denotes the element-wise (or Shur) product of the vectors \( \text{DFT}x \) and \( \text{DFT}h \).

This very important result states that any Linear Shift-Invariant Operator acts on the input \( \tilde{\phi} \) by multiplying the entries of \( \text{DFT}\tilde{\phi} \) with the entries of the \( \text{DFT}h \), i.e., the Discrete Fourier Transform of the Impulse response of \( X_{\text{LTI}} \) acts as a "mask" multiplying the entries of the DFT of \( \tilde{\phi} \)!
9.6. **Signal Representations**

and how operators affect them

Suppose we have a signal $\bar{\mathbf{e}}$ and it is acted upon by a linear operator $\mathbf{H}_L$. We clearly have

$$\bar{\mathbf{e}} = \mathbf{I} \bar{\mathbf{e}} = \mathbf{W}^* \bar{\mathbf{e}} = \left[ \bar{\beta}_1 \bar{\beta}_2 \ldots \bar{\beta}_N \right] \mathbf{u}^T \bar{\mathbf{e}}$$

where $\bar{\beta}_1 = \mathbf{u}_1$, $\bar{\beta}_2 = \mathbf{u}_2$, $\ldots$, $\bar{\beta}_N = \mathbf{u}_N$ are a new basis for representing $\bar{\mathbf{e}}$ and $\mathbf{u}^T \bar{\mathbf{e}} = [\langle \mathbf{e}_1, \mathbf{e} \rangle \ldots \langle \mathbf{e}_N, \mathbf{e} \rangle]^T$ are the coefficients.

The output of $\mathbf{H}_L$ acting on $\mathbf{e}^* = \bar{\mathbf{e}}$ is clearly:

$$\bar{\mathbf{e}}_{\text{out}} = \mathbf{H}_L \left\{ \sum_{k=1}^{N} \langle \bar{\beta}_k, \bar{\mathbf{e}} \rangle \bar{\beta}_k \right\} = \sum_{k=1}^{N} \langle \bar{\beta}_k, \bar{\mathbf{e}} \rangle \mathbf{H}_L \bar{\beta}_k$$
Notice that $\bar{\mathbf{C}}_{\text{aut}}$ is expressed in terms of a new set of vectors $\{\mathbf{H}^k \tilde{\mathbf{B}}_k\}_{k=1,2,...N}$ and this set is not necessarily orthonormal! These vectors are given by

$$\{\mathbf{H}^k \tilde{\mathbf{B}}_k\}_{k=1,2,...N}$$

where $\mathbf{H}_k$ is the matrix describing the action of $\mathbf{A}_k$. We could express the vectors $\mathbf{H}_k \tilde{\mathbf{B}}_k$ in terms of the orthonormal set of vectors $\{\tilde{\mathbf{B}}_1, \tilde{\mathbf{B}}_2, ..., \tilde{\mathbf{B}}_N\}$ and obtain a representation of $\bar{\mathbf{C}}_{\text{aut}}$ in terms of them but the coefficient computation would be quite involved, since

$$\bar{\mathbf{C}}_{\text{aut}} = \mathbf{H} \tilde{\mathbf{C}}^m = \mathbf{H} \mathbf{U} \mathbf{U}^\dagger \tilde{\mathbf{C}}^m \Rightarrow$$

$$\mathbf{U}^\dagger \bar{\mathbf{C}}_{\text{aut}} = \mathbf{U}^\dagger \mathbf{H} \mathbf{U} \mathbf{U}^\dagger \tilde{\mathbf{C}}^m$$

yielding $\bar{\mathbf{C}}_{\text{aut}} = \left[\tilde{\mathbf{B}}_1, \tilde{\mathbf{B}}_2, ..., \tilde{\mathbf{B}}_N\right] \mathbf{U}^\dagger \bar{\mathbf{C}}_{\text{aut}}$. 
However if we had
\[ H_l \beta_k^* = \lambda_k \beta_k^* \]
for some set of orthonormal vectors \( \beta_1, \beta_2, \ldots, \beta_N \)
we'd obtain readily that:
\[
\phi_{\text{out}} = \sum_{k=1}^{n} \langle \beta_k^* \phi \rangle \lambda_k \beta_k^*
\]
hence the coefficient would be simply multiplied by the eigenvalues of \( H_l \)
i.e. \( \lambda_1, \lambda_2, \ldots, \lambda_N \) - respectively.

Therefore, in general if the operator \( \mathcal{X}_l \rightarrow H_l \) has a complete set of eigenvectors \( \beta_1^*, \beta_2^*, \ldots, \beta_N^* \) then representing the input in the basis \( \phi_k^* \sum_{k=1}^{n} \lambda_k \) achieves
the result that the action of $\mathcal{A}_L$ is mapped to a multiplication of each representation coefficient by a number $\mathcal{N}_k$ / i.e. a mask in the "transform domain" (representation coefficient domain!).

The results we had on circulants showed that if $\mathcal{A}_L$ is a shift invariant operator the DFT is the preferred representation for signals from this point of view.