

**DISCOVERING TRANSFORMS: A
TUTORIAL ON CIRCULANT MATRICES,
CIRCULAR CONVOLUTION, AND THE
DISCRETE FOURIER TRANSFORM**

Diagonalization of CMs and the DFT

- The DFT can be derived as the change of basis that simultaneously diagonalizes all circulant matrices.
- The simultaneous diagonalization of any class of linear operators or matrices is the ultimate way to understand their actions, by reducing the entire class to the simplest form of linear operations (diagonal matrices) simultaneously.

Circulant Matrices

- Given a vector $\mathbf{a} = (a_0, \dots, a_{n-1})$, the associated matrix C_a would have the first column made up of these numbers, and each subsequent column is obtained by a circular shift (shifting 1 unit and wrapping the spare numbers to the front) of the previous column.

$$C_a := \begin{bmatrix} a_0 & a_{n-1} & a_{n-2} & \cdots & a_1 \\ a_1 & a_0 & a_{n-1} & & a_2 \\ a_2 & a_1 & a_0 & & a_3 \\ \vdots & & \ddots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 \end{bmatrix}.$$

Diagonalization Properties

- Since diagonalizing transformations are made up of eigenvectors of a matrix, then a set of matrices is simultaneously diagonalizable iff they share a full set of eigenvectors
- An equivalent condition is that they each are diagonalizable, and they all mutually commute
- When a matrix M can be diagonalized with a similarity transformation (i.e. $\Lambda = V^{-1}MV$, where Λ is diagonal), then we have a change of basis (into the eigenbasis) in which the linear transformation has that simple diagonal matrix representation, and its properties can be easily understood.
- The decomposition originally decomposes the matrix into the diagonal scaling matrix with its eigenvalues and the eigenvectors and inverse eigenvectors matrices which correspond to shifting into the eigenbasis and back to the regular basis. In that example, the equation is simply solving for Λ , which is the eigenvalue diagonalized matrix.

Matrix set Diagonalization

- Often one has to work with a set of transformations rather than a single one, and usually with sums and products of elements of that set.
- If we require a different similarity transformation for each member of that set, then sums and products will each require finding their own diagonalizing transformation, which is a lot of work.
- It is natural to ask if there exists one basis in which all members of a set of transformations have diagonal forms. This is the simultaneous diagonalization problem. If such a basis exists, then the properties of the entire set, as well as all sums and products (i.e. the algebra generated by that set) can be easily deduced from their diagonal forms.
- **If a basis that is shared by all members of the transformation can be found then the entire set of transformations can be generalized into one matrix, by decomposing these transformations and multiplying their diagonal matrices.**

Simultaneous Diagonalization

- A set M of matrices is called simultaneously diagonalizable if there exists a single similarity transformation that diagonalizes all matrices in M .
- A set of diagonalizable matrices can be simultaneously diagonalized **iff** they all mutually commute (if the matrices commute then their scaling occurs in the same eigen basis, they share eigenvectors).
- If two matrices commute, they are also simultaneously diagonalizable.

$$A = V^{-1}\Lambda_a V \text{ and } B = V^{-1}\Lambda_b V$$

$$\begin{aligned} AB &= (V^{-1}\Lambda_a V)(V^{-1}\Lambda_b V) = V^{-1}\Lambda_a\Lambda_b V = V^{-1}\Lambda_b\Lambda_a V \\ &= (V^{-1}\Lambda_b V)(V^{-1}\Lambda_a V) = BA \end{aligned}$$

Diagonalization Multiplication Notation

- Diagonal matrix multiplication can be represented with the following notation.

$$Av_i = \lambda_i v_i, \quad i = 1, \dots, n, \quad \text{and } \lambda_i \neq \lambda_j \text{ if } i \neq j.$$

- This is simply saying that because of the way matrix multiplication works, multiplying a diagonal matrix A by a vector v corresponds to multiplying each of the eigenvalues by the corresponding vector value (from top to bottom). The index i ranges from 1 to the number of eigenvalues present.

Finding the simultaneous set diagonalization

- If a matrix A has simple (non-repeating) eigenvalues, then A and B are simultaneously diagonalizable iff they commute. In that case, the diagonalizing basis is made up of the eigenvectors of A .
- If we can find one matrix $A \in M$ with simple eigenvalues, then find its eigenvectors, those will yield the simultaneously diagonalizing transformation for the entire set.

Modular arithmetic

- In modular arithmetic, we say k equals l modulo n if $k - l$ is an integer multiple of n . This is shown formally in the following state

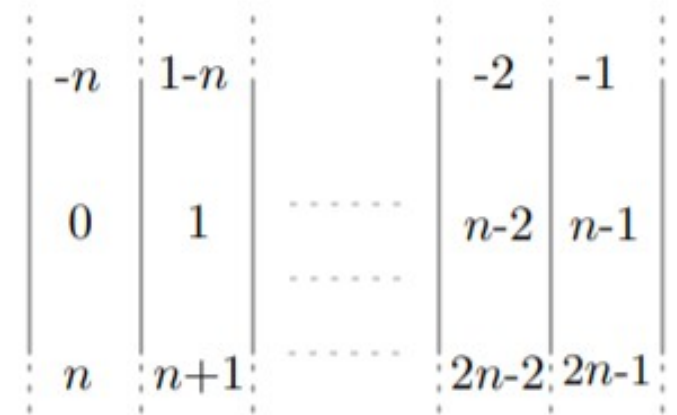
$$k = l \pmod{n}$$

or $k \equiv_n l$

- The statement says k is congruent to $l \pmod{n}$. This means that in the 'world' of modulus n , k and l will have the same value. So for instance, in modulus 2 world, 5 and 7 have the same value; they're both 1. Following the first statement, $7-5$ is equal to 2 (a multiple of 2), so we consider 5 and 7 to be congruent in the modulus 2 domain.

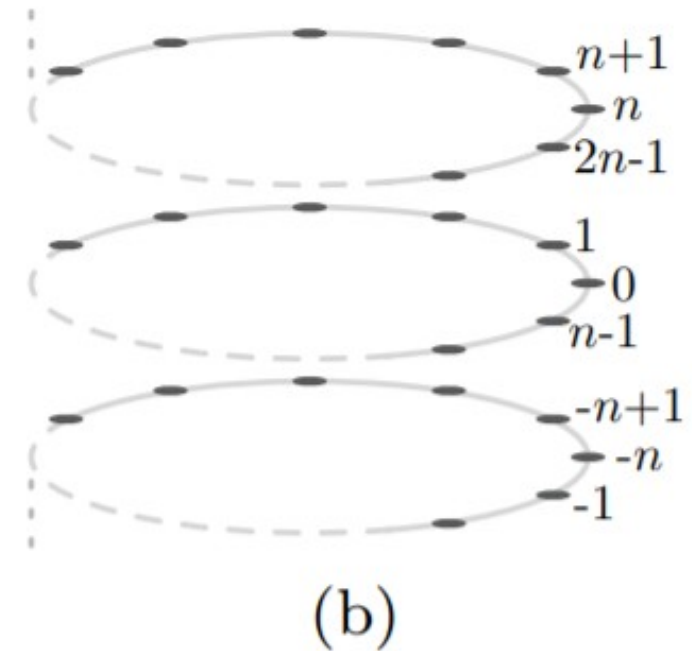
Visualizing modular arithmetic

- Definition of Z_n as the decomposition of the integers Z into equivalence classes each indicated as a “vertical bin”.
- Two integers in Z belong to the same equivalence class if they differ by an integer multiple of n .
- To gain a simple understanding of this plug in values for n . Each equivalence class has numbers in increments of n , so they are all congruent in the modulus n domain.
- The equivalence classes on the right are the same concept except they represent the end of the domain (size n), but the same relationship can be viewed by the simplification of the whole column being ‘shifted’ one unit up. If the -2 and the -1 were in the middle, the concept would be the same as the first two columns. The $2n-x$ simply represents the second iteration of an increment in n . X just corresponds to the ‘base’ of the equivalence class.



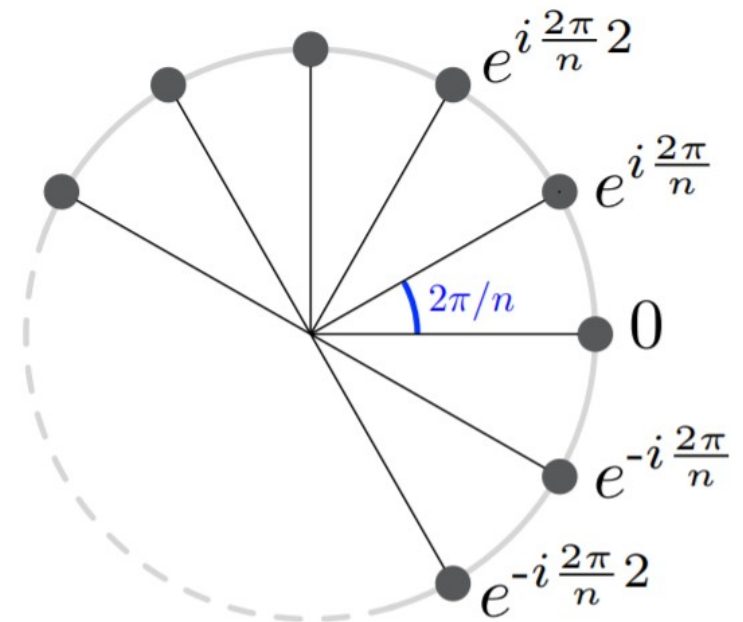
Visualizing modular arithmetic

- Another depiction of the decomposition where two integers that are vertically aligned in this figure belong to the same equivalence class.
- The values going around the circle just vary in the base of the domain, it is essentially shifting the entire domain by one unit, but if n remains the same the increment steps will be the same for all.



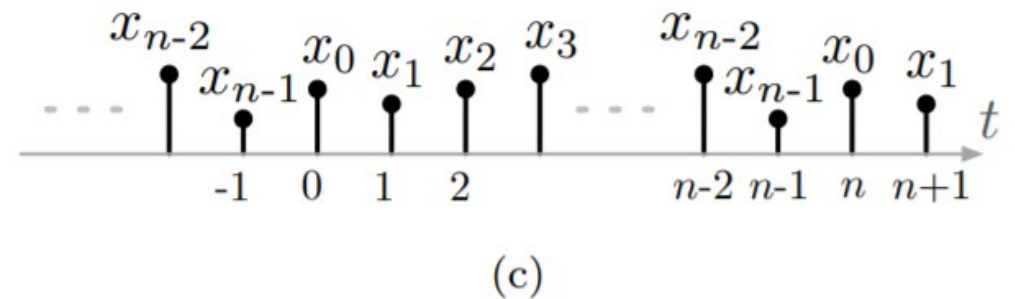
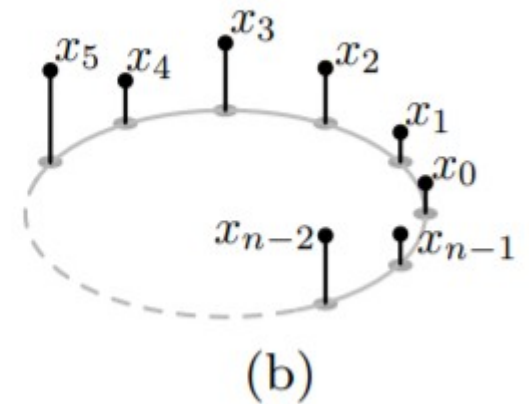
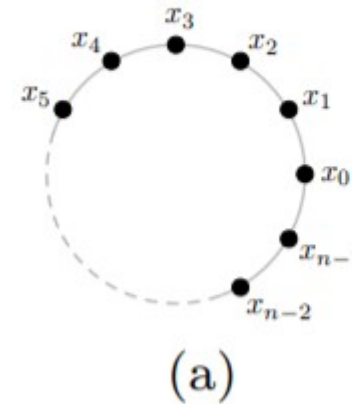
Visualizing modular arithmetic

- This circle represents the n th roots of unity $\rho_m = \exp(im)$ lying on the unit circle in the complex plane. Identifying ρ_m with $m \in \mathbb{Z}_n$ shows that complex multiplication on $\{\rho_m\}$ (which corresponds to angle addition) is equivalent to modular addition in \mathbb{Z}_n .
- This is simply saying that modular addition corresponds to the same concept applied in angle addition on an evenly sliced circle. If each angle is $\frac{2\pi}{n}$, then scaling the exponent by a factor of X corresponds to adding X to the angle $\frac{2\pi}{n}$ times.



Circular Visualization of Vectors

- A vector $x = (x_0, \dots, x_{n-1})$ visualized as (a) a set of numbers arranged counter-clockwise on a discrete circle.
- (b) is the same concept except it shows the amplitude of the vector x .
- (c) shows the same idea on a number line, where the vector wraps around itself in a modulus fashion, repeating itself every n values. (n is the size of the vector)



Symmetry Properties of Circulant Matrices

- Let S and its adjoint S^* be the circular shift operators defined by the following action on vectors

$$S(x_0, \dots, x_{n-2}, x_{n-1}) = (x_{n-1}, x_0, \dots, x_{n-2})$$

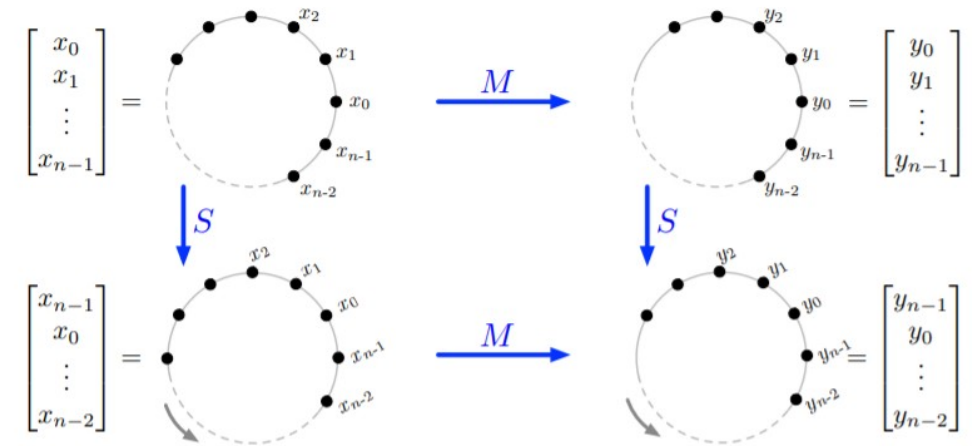
$$S^*(x_0, x_1, \dots, x_{n-1}) = (x_1, \dots, x_{n-1}, x_0).$$

- S is therefore called the circular right-shift operator while S^* is the circular leftshift operator. It is clear that S^* is the inverse of S , and it is easy to show that it is the

$$Sx = \begin{bmatrix} 0 & & & 1 \\ 1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 1 & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} x_{n-1} \\ x_0 \\ \vdots \\ x_{n-2} \end{bmatrix}, \quad S^*x = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_0 \end{bmatrix}$$

Circular Matrix shift-invariance property

- An important property of S is that it commutes with any circulant matrix. One way to see this is to observe that for any matrix M , left (right) multiplication by S amounts to row (column) circular permutation.
- This is just an illustration of how a circular shift can be applied before or after a circulant matrix transformation and the result is the same.



Circulant Matrix Commutativity

- A matrix M is circulant iff it commutes with the circular shift operator S , i.e. $SM = MS$.
- A simple corollary that a matrix is circulant iff it commutes with S^*

$$SM = MS \iff S^* SM S^* = S^* MS S^* \iff MS^* = S^* M,$$

- This is expected as S^* is simply the inverse of S .
- Commutation with S also expresses a shift invariance property, the transformation is the same regardless of shifts.
- $S(Mx) = M(Sx)$ means that rotating the result of the action of M on x is the same as rotating x first and then acting with M .

Circular Convolution

- The multiplication of a vector with the circulant matrix of another vector is equivalent to circular convolution. Every row of the circulant matrix gets multiplied by the vector, and each row is a circular shifted version of the last row. So, this operation is similar to convolution but it is circular in the way that each row in the circulant matrix is representative of the same vector just shifted circularly (there is no real shifting where values are left out, only

$$y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} = \begin{bmatrix} a_0 & a_{n-1} & \cdots & a_1 \\ a_1 & a_0 & & a_2 \\ \vdots & & \ddots & \vdots \\ a_{n-1} & a_{n-2} & \cdots & a_0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} = C_a x.$$

Associativity and Commutativity of CMs

- Since vector multiplication is associative and commutative, circulant matrices are simply shifting the vector circularly along its rows, and circulant matrices are shift invariant; it is expected that:
 - *Associativity:* for any three n -vectors a , b and c we have

$$a \star (b \star c) = (a \star b) \star c$$

- *Commutativity:* for any two n -vectors a and b

$$a \star b = b \star a$$

Circulant Matrix Commutivity

THEOREM 3.3. 1. *Circular convolution of any two vectors can be written as a matrix-vector product with a circulant matrix*

$$a \star x = C_a x = C_x a.$$

2. *The product of any two circulant matrices is another circulant matrix*

$$C_a C_b = C_{a \star b}.$$

3. *All circulant matrices mutually commute since for any two C_a and C_b*

$$C_a C_b = C_{a \star b} = C_{b \star a} = C_b C_a.$$

Deriving the DFT from Circulant Matrices

- Since all circulant matrices mutually commute, we can look for a circulant matrix that has simple (non repeating) eigenvalues. The eigenvectors of that matrix will then give the simultaneously diagonalizing transformation.
- The shift operator is in some sense the most fundamental circulant matrix, and is therefore a good candidate for an eigenvector/eigenvalue decomposition

Finding Eigenvalues

- The first equation assumes the existence of v , an eigenvector.
- The last step can also be thought of as moving the λIv term to the right and multiplying by v inverse on the right of both sides. Then moving back and taking the determinant, since we want values.

We start by finding the **eigenvalue**: we know this equation must be true:

$$Av = \lambda v$$

Now let us put in an identity matrix so we are dealing with matrix-vs-matrix:

$$Av = \lambda Iv$$

Bring all to left hand side:

$$Av - \lambda Iv = 0$$

If v is non-zero then we can solve for λ using just the determinant:

$$| A - \lambda I | = 0$$

Eigenvalues of S^*

- Let w be an eigenvector (with eigenvalue λ) of the shift operator S^*
- Raising the shift to a power results in raising the eigenvalue to the same power.
- This relation “repeats” for $l \geq n$ (it loops around). For $l = n$ we have for each index k the equation on the right.
- Since the vector w isn't 0 then dividing both sides by w_k yields that the eigenvalues are n th roots of unity.

$$\begin{array}{llll} S^* w & = & \lambda w & \iff w_{k+1} = \lambda w_k, \\ (S^*)^l w & = & \lambda^l w & \iff w_{k+l} = \lambda^l w_k, \end{array}$$

$$w_{k+n} = \lambda^n w_k \iff w_k = \lambda^n w_k$$

$$\lambda^n = 1 \iff \lambda = \rho_m := e^{i \frac{2\pi}{n} m}, \quad m \in \mathbb{Z}_n.$$

Left shift operator eigenvectors

- The eigenvectors of the circular left-shift operator correspond to the n th roots of unity whose exponential term is multiplied by some index m which ranges from 0 to $n-1$. N being the dimension of the left-shift operator.

LEMMA 4.1. *The circular left-shift operator S^* on \mathbb{R}^n has n distinct eigenvalues. They are the n th roots of unity $\rho_m := e^{i\frac{2\pi}{n}m} = \rho_1^m =: \rho^m$, $m \in \mathbb{Z}_n$. The corresponding eigenvectors are*

$$(4.4) \quad w^{(m)} = \left(1, \rho^m, \rho^{2m}, \dots, \rho^{m(n-1)}\right), \quad m = 0, \dots, n-1,$$

Expanding the circulant matrix eigenvector

- Recall that since any circulant matrix commutes with S^* , and S^* has distinct eigenvalues, then C_a has the same eigenvectors as those previously found for S^* .
- The circulant matrix of some vector a multiplied by the eigenvalues of the left shift operator equals λ_m (the eigenvalues of C_a) multiplied by the eigenvectors of S^* because C_a and the left shift operator share the same eigenvectors. Each row of the equation represent the

sa

$$(4.5) \quad C_a w^{(m)} = \lambda_m w^{(m)} \quad \Leftrightarrow \quad \begin{bmatrix} a_0 & a_{n-1} & \cdots & a_1 \\ a_1 & a_0 & & a_2 \\ \vdots & & \ddots & \vdots \\ a_{n-1} & a_{n-2} & \cdots & a_0 \end{bmatrix} \begin{bmatrix} 1 \\ \rho_m \\ \vdots \\ \rho_m^{n-1} \end{bmatrix} = \lambda_m \begin{bmatrix} 1 \\ \rho_m \\ \vdots \\ \rho_m^{n-1} \end{bmatrix}$$

The DFT

- To obtain the eigenvalues of the circulant matrix multiply each row (take the first as an example as the others are the same multiplied by some factor of ρ_m) by the eigenvalues of the left shift and what is obtained is the sum of each value in the matrix multiplied by the corresponding root of unity (which change powers with the l i i

$$\begin{aligned}
 \lambda_m &= a_0 + a_{n-1} \rho_m + \cdots + a_1 \rho_m^{n-1} \\
 &= a_0 + a_1 \rho_m^{-1} + \cdots + a_{n-1} \rho_m^{-(n-1)} \\
 &= \sum_{l=0}^{n-1} a_l \rho_m^{-l} = \sum_{l=0}^{n-1} a_l \rho^{-ml} = \boxed{\sum_{l=0}^{n-1} a_l e^{-i \frac{2\pi}{n} ml} =: \hat{a}_m,}
 \end{aligned}$$

Circulant Matrix multiplied by Shift Eigenvectors

- If we use the eigenvectors of S^* as columns of a matrix W , the n eigenvalue/eigenvector relationships $C_a w_{(m)} = \lambda_m w_{(m)}$ can be written as a single matrix equation.
- Since the circulant matrix and the left shift operator have the same eigenvectors they are commutative and the multiplication by the circulant matrix can be changed to multiplying by the diagonal matrix of its eigenvalues, which correspond to the ones calculated in the previous slide as \hat{a} (DFT of a).

$$C_a \begin{bmatrix} w^{(0)} & \cdots & w^{(n-1)} \end{bmatrix} = \begin{bmatrix} w^{(0)} & \cdots & w^{(n-1)} \end{bmatrix} \begin{bmatrix} \hat{a}_0 & & \\ & \ddots & \\ & & \hat{a}_{n-1} \end{bmatrix},$$

$$\iff C_a W = W \text{diag}(\hat{a}),$$

Conjugate of W

- The matrix W (whose columns are the eigenvectors of S^*) is symmetric, W^* is thus the matrix W with each entry replaced by its complex conjugate. Furthermore, since for each root of unity $\rho^{k*} = \rho^{-k}$, we can write the conjugate of W to be the following matrix.

$$W^* = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \rho^{-1} & \dots & \rho^{-(n-1)} \\ \vdots & \vdots & & \vdots \\ 1 & \rho^{-(n-1)} & \dots & \rho^{-(n-1)(n-1)} \end{bmatrix}.$$

DFT

- Multiplying a vector by W^* is exactly taking its DFT. The m 'th row of W^*x is:

$$\hat{x}_m = \begin{bmatrix} 1 & \rho^{-m} & \dots & \rho^{-m(n-1)} \end{bmatrix} \begin{bmatrix} x_0 \\ \vdots \\ x_{n-1} \end{bmatrix},$$

- These are the conjugates of the eigenvectors of the left shift matrix multiplied by some vector x .

Inverse Fourier Transform

- Multiplication by W is the same as taking the inverse DFT.
- W is a matrix made up of the eigenvectors of the left shift matrix and n is the size of the matrix.

$$x_l = \frac{1}{n} \sum_{k=0}^{n-1} \hat{x}_k \rho^{kl} = \frac{1}{n} \sum_{k=0}^{n-1} \hat{x}_k e^{i \frac{2\pi}{n} kl}.$$

Circular Convolution

- Take this previously seen equation of a circulant matrix with W , which is a matrix whose columns are the eigenvectors of C_a

$$C_a \begin{bmatrix} w^{(0)} & \cdots & w^{(n-1)} \end{bmatrix} = \begin{bmatrix} w^{(0)} & \cdots & w^{(n-1)} \end{bmatrix} \begin{bmatrix} \hat{a}_0 & & \\ & \ddots & \\ & & \hat{a}_{n-1} \end{bmatrix},$$

$$\iff C_a W = W \text{diag}(\hat{a}),$$

- Multiplying both sides by the inverse of W results in the following

$$C_a = W \text{diag}(\hat{a}) W^{-1} = W \text{diag}(\hat{a}) \left(\frac{1}{n} W^* \right)$$

$$= \left(\frac{1}{n} W \right) \text{diag}(\hat{a}) W^* = \left(\frac{1}{\sqrt{n}} W \right) \text{diag}(\hat{a}) \left(\frac{1}{\sqrt{n}} W^* \right).$$

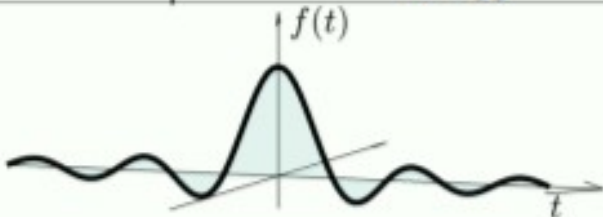
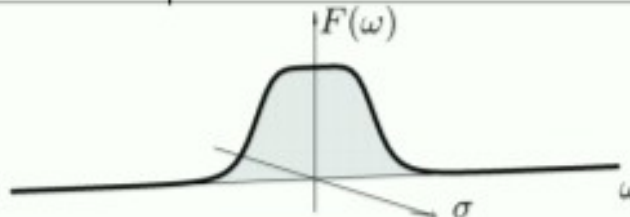
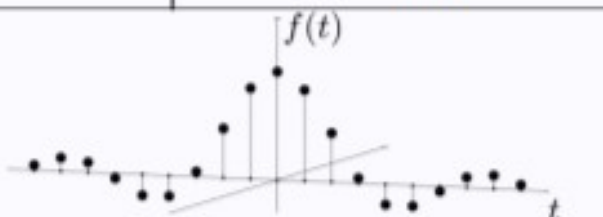
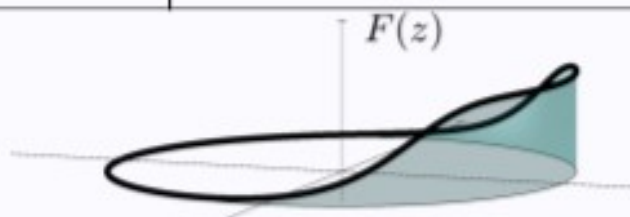
Circular Convolution

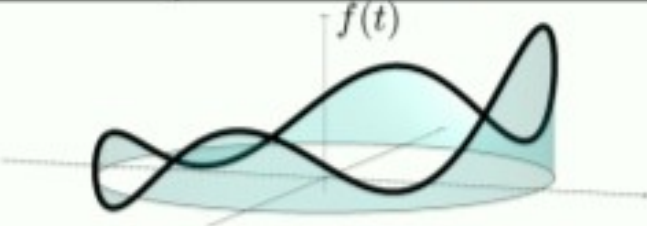
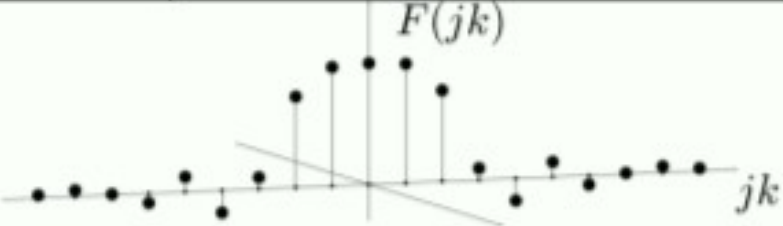
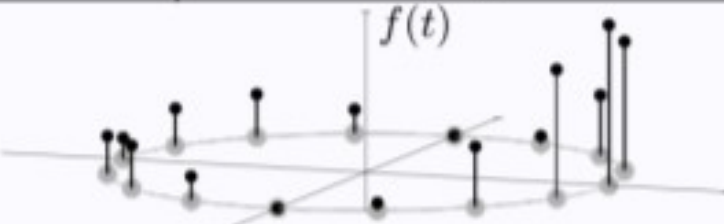
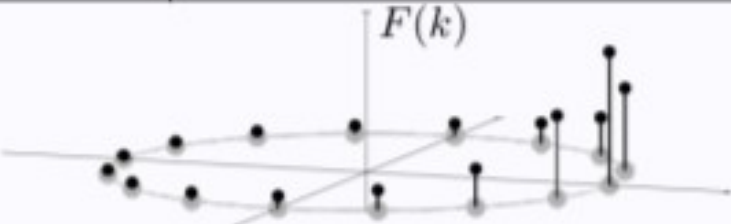
- The previously derived diagonalization can be interpreted as follows.

$$C_a x = \underbrace{\left(\frac{1}{n} W \right) \text{diag}(\hat{a}) \underbrace{W^* x}_{\text{DFT of } x}}_{\text{multiply by } \hat{a} \text{ entrywise}}_{\text{inverse DFT}}$$

- **Thus the action of C_a on x , or equivalently the circular convolution of a with x , can be performed by first taking the DFT of x , then multiplying the resulting vector component-wise by \hat{a} (the DFT of the vector a defining the matrix C_a), and then taking an inverse DFT. In other words, the diagonalization of a circulant matrix is equivalent to converting circular convolution to component-wise vector multiplication through the DFT.**

Fourier Transform

Time Axis	Transform	Frequency Axis (Frequency "Set")
$t \in \mathbb{R}$	Fourier Transform $F(\omega) := \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt$	$j\omega \in j\mathbb{R}$ <i>imaginary axis of \mathbb{C}</i>
		
$t \in \mathbb{Z}$	z-Transform (bilateral) $F(z) := \sum_{t \in \mathbb{Z}} f(t) z^{-t}$	$z = e^{j\theta} \in \mathbb{T}$ <i>unit circle of \mathbb{C}</i>
		

Time Axis	Transform	Frequency Axis (Frequency "Set")
$t \in \mathbb{T}$	Fourier Series $F(k) := \int_0^{2\pi} f(t) e^{-jkt} dt$	$jk \in j\mathbb{Z}$, <i>integers of imaginary axis of \mathbb{C}</i>
 		
$t \in \mathbb{Z}_n$	Discrete Fourier Transform (DFT) $F(k) := \sum_{t=0}^{n-1} f(t) e^{-j \frac{2\pi}{n} kt}$	$k \in \mathbb{Z}_n$ <i>n-roots of unity</i>
 		

Shift Invariance

- The time shift operation is defined in the following way. It takes a function on domain t and transfers it to domain T , that T is negated in the function which produces a flip and a delay by T , the function is then shifted along the domain.

$$(S_T f)(t) := f(t - T),$$

- We call an operator time invariant (or shift invariant) if it commutes with all possible time-shift operations, convolution is an example of this because it's a symmetrical operation due to the mirroring in the y axis seen above.