Discovering the Fourier Transform: A Tutorial on Circulant Matrices, Circular Convolution, and the DFT

Bassam Bamieh
Mechanical Engineering (bamieh@ucsb.edu)
University of California, Santa Barbara

May 10, 2018

Abstract

How could the Fourier transform be discovered if one didn’t know it? In the case of the Discrete Fourier Transform (DFT), we show how it arises naturally out of analysis of circulant matrices. In particular, the DFT can be derived as the change of basis that simultaneously diagonalizes all circulant matrices. Thus the DFT arises naturally from a linear algebra question. Rather than thinking of the DFT as a signal transform, it is more natural to think of it as a change of basis that renders a certain set of linear operations into a simple, diagonal form.

The Fourier transform in all its forms is ubiquitous. Its many useful properties are introduced early on in Mathematics, Science and Engineering curricula [1]. Typically, it is introduced as a transformation on functions and signals, and then its many useful and remarkable properties are derived. Those properties are then shown to be very effective in solving certain differential equations, or to analyze the action of time-invariant linear dynamical systems, amongst many other uses. To the student, the effectiveness of the Fourier transform in solving these problems may seem magical at first, before familiarity eventually suppresses that initial sense of wonder. In this tutorial, I’d like to step back to before one is shown the Fourier transform, and ask the following question: How would one naturally discover the Fourier transform rather than have it be postulated?

The above question is interesting for several reasons. First, it is more intellectually satisfying to introduce a new mathematical object from familiar and well-known objects rather than having it postulated “out of thin air”. In this tutorial we demonstrate how the DFT arises naturally from the problem of simultaneous diagonalization of all circulant matrices, which share symmetry properties that enable this diagonalization. It should be noted that 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. The same procedure can be applied to discover the other related transforms, namely the Fourier Transform, the z-Transform and Fourier Series. All can be arrived at by simultaneously diagonalizing a respective class of linear operators that obey their respective symmetry rules.

To make the point above, and to have a concrete discussion, we consider in this tutorial only the case of circulant matrices, and later on circulant operators which are the natural generalization to higher dimensions. This case is also particularly useful because it yields the DFT, which is the computational workhorse for all Fourier-type analysis.
Given an ordered set of \( n \) numbers (an \( n \)-tuple, or an \( n \)-vector) \( x := (x_0, \ldots, x_{n-1}) \), define the associated matrix \( C_x \) whose first column is made up of these numbers, and each subsequent column is obtained by a circular shift of the previous column

\[
C_x := \begin{bmatrix}
x_0 & x_{n-1} & x_{n-2} & \cdots & x_1 \\
x_1 & x_0 & x_{n-1} & \cdots & x_2 \\
x_2 & x_1 & x_0 & \cdots & x_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
x_{n-1} & x_{n-2} & x_{n-3} & \cdots & x_0
\end{bmatrix}.
\]  

(1)

Note that each row is also obtained from the previous row by a circular shift. Thus the entire matrix is completely determined by any one of its rows or columns. Such matrices are called circulant. They are a subclass of Toeplitz matrices, and have very special properties due to their intimate relation to the Discrete Fourier Transform (DFT) and circular convolution.

Given an \( n \)-vector \( x \) as above, its DFT \( \hat{x} \) is another \( n \)-vector defined by the expression

\[
\hat{x}_k := \sum_{l=0}^{n-1} x_l e^{-i \frac{2\pi kl}{n}}, \quad k = 0, 1, \ldots, n-1.
\]  

(2)

A remarkable fact is that given a circulant matrix \( C_x \), its eigenvalues are precisely the set of complex numbers \( \{\hat{x}_k\} \), i.e. the DFT of the vector \( x \) that defines the circulant matrix \( C_x \).

There are many ways to derive this conclusion and other beautiful properties of the DFT, and its connections with circulant matrices and circular convolutions. Most treatments start with the definition (2) of the DFT, from which many of its seemingly magical properties are easily derived. To restate the goal of this tutorial, the question we ask here is: what if we didn’t know what the DFT is? How can we arrive at it in a natural manner without having to guess its form?

There is a natural way to think about this problem. Given a class of matrices or operators, one asks if there is a transformation, a change of basis, in which their matrix representations all have the same structure such as diagonal, block diagonal, etc. The simplest such scenario is when a class of matrices can be simultaneously diagonalized with the same transformation. 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. Therefore given a mutually commuting set of matrices, by finding their shared eigenvectors, one finds that special transformation that simultaneously diagonalizes all of them. Thus, finding the “right transform” for a particular class of operators amounts to identifying the correct eigenvalue problem (which turns out to be relatively easy), and then calculating the eigenvectors, which then yield the transform.

An alternative but complementary view of the above procedure involves describing the class of operators using some underlying common symmetry. For example, circulant matrices such as (1) have a shift invariance property with respect to circular shifts of vectors. This can also be described as having a shift-invariant action on vectors over \( \mathbb{Z}_n \) (the integers modulo \( n \)), which is also equivalent to having a shift-invariant action on periodic functions (with period \( n \)). In more formal language, circulant matrices represent a class of mutually commuting operators that also commute with the action of the group \( \mathbb{Z}_n \). A basic shift operator generates that group, and the eigenvalue problem for that shift operator yields the DFT. This approach has the advantage of being generalizable to more complex symmetries that can be encoded in the action of other, possibly non-commutative, groups. These
techniques are part of the theory of group representations. However, we adopt here the approach described in the previous paragraph, which uses familiar Linear Algebra language and avoids the formalism of group representations. None the less, the two approaches are intimately linked. Perhaps the present approach can be thought of as a “gateway” treatment on a slippery slope to group representations [2, 3] if the reader is so inclined.

Finally, it should be noted that the same linear algebra language used in the present approach generalizes to the other related transforms, namely the Fourier transform, Fourier series, and the z-transform. They can all be discovered by understanding the class of shift-invariant operators involved, and then identifying the basic eigenvalue problem whose solutions yields the correct transform.

1 Modular Arithmetic, $\mathbb{Z}_n$, and Circular Shifts

To understand the symmetry properties of circulant matrices, it is useful to first study and establish some simple properties of the set $\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$ of integers modulo $n$. The arithmetic in $\mathbb{Z}_n$ is modular arithmetic, that is, we say $k$ equals $l$ modulo $n$ if $k - l$ is an integer multiple of $n$. The following notation can be used to describe this formally

$$k = l \pmod{n} \quad \text{or} \quad k \equiv_l n \quad \iff \quad \exists i \in \mathbb{Z}, \text{ s.t. } k - l = i n$$

Thus for example $n \equiv_n 0$, and $n + 1 \equiv_n 1$ and so on. There are two equivalent ways to define (and think) about $\mathbb{Z}_n$, one mathematically formal and the other graphical. The first is to consider the set of all integers $\mathbb{Z}$ and regard any two integers $k$ and $l$ such that $k - l$ is a multiple of $n$ as equivalent, or more precisely as members of the same equivalence class. The infinite set of integers $\mathbb{Z}$ becomes a finite set of equivalence classes with this equivalence relation. This is illustrated in Figure 1 where elements of $\mathbb{Z}_n$ are arranged in “vertical bins”

![Figure 1](image_url)

**Figure 1**: (a) Definition of $\mathbb{Z}_n$ as the decomposition of the integers $\mathbb{Z}$ into equivalence classes each indicated as a “vertical bin”. Two integers in $\mathbb{Z}$ belong to the same equivalence class (and represent the same element of $\mathbb{Z}_n$) if they differ by an integer multiple of $n$. (b) Another depiction of the decomposition where two integers that are vertically aligned in this figure belong to the same equivalence class. The arithmetic in $\mathbb{Z}_n$ is just angle addition in this diagram. For example, $(-1) + (n+1) \equiv_n 0 \equiv_n n$. (c) Using the set $\{0, 1, \ldots, n-1\}$ as $\mathbb{Z}_n$. A few equivalent members are shown, and the arithmetic of $\mathbb{Z}_n$ is just angle addition here. (d) The $n$th roots of unity $\rho_n := \exp(i \frac{2 \pi}{n} m)$ lying on the unit circle in the complex plane. Identifying $\rho_n$ with $m \in \mathbb{Z}_n$ shows that complex multiplication on $\{\rho_n\}$ (which corresponds to angle addition) is equivalent to modular addition in $\mathbb{Z}_n$

which are the equivalence classes. Each equivalence class can be identified with any of its members. One choice is to identify the first one with the element 0, the second one with 1, and so on up to the $n$th class identified with the integer $n - 1$. Figure 1 also shows how elements of $\mathbb{Z}_n$ can be arranged on a discrete circle so that the arithmetic in $\mathbb{Z}_n$ is identified
with angle addition. One more useful isomorphism is between $\mathbb{Z}_n$ and the $n$th roots of unity $\rho_m := e^{\frac{2\pi i}{n} m}$, $m = 0, \ldots, n - 1$. The complex numbers $\{\rho_m\}$ lie on the unit circle each at a corresponding angle of $\frac{2\pi}{n} m$ counter-clockwise from the real axis (Figure 1.d). Complex multiplication on $\{\rho_m\}$ corresponds to addition of their corresponding angles, and the mapping $\rho_m \rightarrow m$ is an isomorphism from complex multiplication on $\{\rho_m\}$ to modular arithmetic in $\mathbb{Z}_n$.

Using modular arithmetic, we can write down the definition of a circulant matrix (1) by specifying the $ij$'th entry $1$ of the matrix $C_x$ as

$$(C_x)_{ij} := x_{i-j}, \quad i, j \in \mathbb{Z}_n,$$

where we use (mod $n$) arithmetic for computing $i - j$. It is clear that with this definition, the first column of $C_x$ is just the sequence $x_0, x_1, \ldots, x_{n-1}$. The second column is given by the sequence $\{x_{i-1}\}$ and is thus $x_{-1}, x_0, \ldots, x_{n-2}$, which is exactly the sequence $x_{n-1}, x_0, \ldots, x_{n-2}$, i.e. a circular shift of the first column. Similarly each subsequent column is a circular shift of the column preceding it.

Finally, it is useful to visualize an $n$-vector $x := (x_0, \ldots, x_{n-1})$ as a set of numbers arranged at equidistant points along a circle, or equivalently as a function $x : \mathbb{Z}_n \rightarrow \mathbb{C}$ on the discrete circle. This is illustrated in figure 2. Note the difference between this figure and Figure 1, which depicts

![Figure 2: A vector $x := (x_0, \ldots, x_{n-1})$ visualized as (a) a set of numbers arranged counter-clockwise on a discrete circle, or equivalently (b) as a function $x : \mathbb{Z}_n \rightarrow \mathbb{C}$ on the discrete circle $\mathbb{Z}_n$. (c) An $n$-periodic function on the integers $\mathbb{Z}$ can equivalently be viewed as a function on $\mathbb{Z}_n$ as in (b).](image)

the elements of $\mathbb{Z}_n$ and modular arithmetic. Figure 2 instead depicts vectors as a set of numbers arranged in a discrete circle, or as functions on $\mathbb{Z}_n$. A function on $\mathbb{Z}_n$ can also be thought of as a periodic function (with period $n$) on the set of integers $\mathbb{Z}$ (Figure 2.c). In this case, periodicity of the function is expressed by the condition

$$x_{t+n} = x_t \quad \text{for all } t \in \mathbb{Z}. \quad (4)$$

It is however more natural to view periodic functions on $\mathbb{Z}$ as just functions on $\mathbb{Z}_n$. In this case, periodicity of the function is simply “encoded” in the modular arithmetic of $\mathbb{Z}_n$, and condition (4) does not need to be explicitly stated.

---

1Here, and in this entire appendix, matrix rows and columns are indexed from 0 to $n - 1$ rather than the more traditional 1 through $n$ indexing. This alternative indexing significantly simplifies notation, and corresponds more directly to modular arithmetic.
1.1 Symmetry Properties of Circulant Matrices

Amongst all circulant matrices, there is a special one. Let \( S \) and its adjoint \( S^* \) be the circular shift operators defined by the following action on vectors

\[
S \begin{bmatrix} x_0 & \cdots & x_{n-2} & x_{n-1} \end{bmatrix}^T = \begin{bmatrix} x_{n-1} & x_0 & \cdots & x_{n-2} \end{bmatrix}^T,
\]
\[
S^* \begin{bmatrix} x_0 & x_1 & \cdots & x_{n-1} \end{bmatrix}^T = \begin{bmatrix} x_1 & \cdots & x_{n-1} & x_0 \end{bmatrix}^T.
\]

\( S \) is therefore called the circular right-shift operator while \( S^* \) is the circular left-shift operator. It is clear that \( S^* \) is the inverse of \( S \), and it is easy to show that it is the adjoint of \( S \). The latter fact also becomes clear upon examining the matrix representations of \( S \) and \( S^* \)

\[
Sx = \begin{bmatrix} 0 & 1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} x_0 \\ 1 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \end{bmatrix} = \begin{bmatrix} x_{n-1} \\ x_0 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \end{bmatrix},
\]
\[
S^*x = \begin{bmatrix} 0 & 1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \end{bmatrix},
\]

which shows that \( S^* \) is indeed the transpose (and therefore the adjoint) of \( S \). Note that both matrix representations are circulant matrices since \( S = C_{(0,1,0,\ldots,0)} \) and \( S^* = C_{(0,0,\ldots,0,1)} \) in the notation of (1). The actions of \( S \) and \( S^* \) expressed in terms of vector indices are

\[
(Sx)_k := x_{k-1}, \quad (S^*x)_k := x_{k+1}, \quad k \in \mathbb{Z}_n, \tag{5}
\]

where modular arithmetic is used for computing vector indices. For example \((Sx)_0 = x_{n-1} \equiv x_{n-1}\).

An important property of \( S \) is that it commutes with any circulant matrix. To see this, note that the matrix representation of \( S \) implies its \( ij \)’th entry is given by \((S)_{ij} = \delta_{i-j-1}\). Now let \( C_x \) be any circulant matrix, and observe that

\[
(SC_x)_{ij} = \sum_l S_{il} (C_x)_{lj} = \sum_l \delta_{i-l-1} x_{l-j} = \sum_l \delta_{(i-1)-l} x_{l-j} = x_{i-1-j},
\]
\[
(C_xS)_{ij} = \sum_l (C_x)_{il} S_{lj} = \sum_l x_{i-l} \delta_{l-j-1} = \sum_l x_{i-l} \delta_{l-(j+1)} = x_{i-j-1},
\]

where (3) is used for the entries of \( C_x \). Thus \( S \) commutes with any circulant matrix. The converse is also true (see Exercise 1), and we state these conclusions in the next lemma.

Lemma 1 A matrix \( M \) is circulant iff it commutes with the circular shift operator \( S \), i.e. \( SM = MS \).

Note a simple corollary that a matrix is circulant iff it commutes with \( S^* \) since

\[
SM = MS \iff S^* SM S^* = S^* MS S^* \iff MS^* = S^* M,
\]

which could be an alternative statement of the Lemma. The fact that a circulant matrix commutes with \( S \) could have been used as a definition of a circulant matrix, with the structure in (1) derived as a consequence. Commutation with \( S \) also expresses a shift invariance property. If we think of an \( n \)-vector \( x \) as a functions on \( \mathbb{Z}_n \) (Figure 2.b), then \( SMx = MSx \) means that the action of \( M \) on \( x \) is shift invariant. Geometrically, \( Sx \) is a counter-clockwise rotation of the function \( x \) in Figure 2.b. \( S(Mx) = M(Sx) \) means that rotating the result of the action of \( M \) on \( x \) is the same as rotating \( x \) first and then acting with \( M \). This property is illustrated graphically in Figure 3.
2 Simultaneous Diagonalization of all Circulant Matrices Yields the DFT

In this section, we will derive the DFT as a byproduct of diagonalizing circulant matrices. Any two (diagonalizable) matrices that commute have the same eigenvectors, but possibly different eigenvalues. Since all circulant matrices mutually commute, they all have the same eigenvectors, and they only differ in their eigenvalues. Therefore, if we find the eigenvectors of any one particular circulant matrix, then we have found the eigenvectors of all circulant matrices. A diagonalizing transformation is made up of the eigenvectors of a matrix, we can then simultaneously diagonalize all circulant matrices with one transformation made up of those found eigenvectors.

The shift operator is in some sense the most fundamental circulant matrix, and is therefore a good candidate for an eigenvector/eigenvalue decomposition. The eigenvalue problem for \( S \) will turn out to be the simplest one. Its immediate that \( S \) is normal, meaning that it commutes with its adjoint (we in fact have a stronger statement \( S^* S = SS^* = I \)). Any normal matrix has a full set of mutually orthogonal eigenvectors. We will not need this fact here since we will explicitly construct all the eigenvectors. Once its eigenvectors are found, we will be able to diagonalize all circulant matrices using them. Note that we have two options. To find eigenvectors of \( S \) or alternatively of \( S^* \). We begin with \( S^* \) since this will end up yielding the classically defined DFT.

2.1 Construction of Eigenvectors/Eigenvalues of \( S^* \)

Let \( w \) be an eigenvector (with eigenvalue \( \lambda \)) of the shift operator \( S^* \). Note that it is also an eigenvector (with eigenvalue \( \lambda^l \)) of any power \( (S^*)^l \) of \( S^* \). Applying the definition (5) to

\[
\begin{bmatrix}
  x_0 \\
  x_1 \\
  \vdots \\
  x_{n-1}
\end{bmatrix}
= S
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_0
\end{bmatrix}
\]

\[
\begin{bmatrix}
  y_0 \\
  y_1 \\
  \vdots \\
  y_{n-1}
\end{bmatrix}
= M
\begin{bmatrix}
  y_0 \\
  y_1 \\
  \vdots \\
  y_{n-1}
\end{bmatrix}
\]

Figure 3: Illustration of the circular shift-invariance property of the matrix-vector product \( y = Mx \) in a commutative diagram. \( Sx \) is the circular shift of a vector, depicted also as a clockwise rotation of the vector components arranged on the discrete circle. A matrix \( M \) has the shift invariance property if \( SM = MS \). In this diagram, this means that the action of \( M \) on the rotated vector \( Sx \) is equal to acting \( y = Mx \) on \( x \) with \( M \) first, and then rotating the resulting vector to yield \( Sy \). A matrix \( M \) has this shift-invariance property iff it is circulant.

\[2\text{ To invoke this, we need to know that circulant matrices are diagonalizable. Indeed they are since the complex conjugate transpose (the adjoint) of a circulant matrix is also circulant, so any circulant matrix commutes with it’s adjoint, i.e. it is normal. Therefore, any circulant matrix has a full set of mutually orthogonal eigenvectors.}\]
the relation $Sw = \lambda w$ will reveal that an eigenvector $w$ has a very special structure

$$S^*w = \lambda w \iff w_{k+1} = \lambda w_k, \quad k \in \mathbb{Z}_n,$$

$$\quad (S^*)^l w = \lambda^l w \iff w_{k+l} = \lambda^l w_k, \quad k \in \mathbb{Z}_n, \quad l \in \mathbb{Z}.$$  \hfill (6)

These relations can be used to compute all eigenvectors/eigenvalues of $S^*$. First, observe that although (6) is valid for all $l \in \mathbb{Z}$, this relation “repeats” for $l \geq n$. In particular, for $l = n$ we have for each index $k$

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

since $k+n \equiv n k$. Now since the vector $w \neq 0$, then for at least one index $k$, $w_k \neq 0$, and the last equality implies that $\lambda^n = 1$, i.e. any eigenvalue of $S$ must be an $n$th root of unity

$$\lambda^n = 1 \iff \lambda = \rho_m := e^{i2\pi m/n}, \quad m \in \mathbb{Z}_n.$$  \hfill (7)

Thus we have discovered that the $n$ eigenvalues of $S$ are precisely the $n$ distinct $n$th roots of unity $\{\rho_m\}$.

To compute the corresponding eigenvectors, apply the last relation in (6) $w_{k+l} = \lambda^l w_k$, and use it to express the entries of an eigenvector $w$ in terms of the first entry

$$w_k = \lambda^l w_0.$$  \hfill (8)

Applying this with the discovered eigenvalues of roots of unity

$$w_l = \rho_m^l w_0 \iff w = w_0 \begin{bmatrix} 1 & \rho_m & \rho_m^2 & \cdots & \rho_m^{n-1} \end{bmatrix}^T$$

Note that $w_0$ is a scalar, and since eigenvectors are only unique up to multiplication by a scalar, we can set $w_0 = 1$ for a more compact expression for the eigenvector. We summarize the previous derivations in the following statement.

**Lemma 2** The circular left-shift operator $S^*$ on $\mathbb{R}^n$ has exactly $n$ eigenvalues, they are the $n$th roots of unity $\rho_m := e^{i2\pi m/n} = \rho_m^1 =: \rho^m$. The corresponding (mutually orthogonal) eigenvectors are

$$w^{(m)} = \begin{bmatrix} 1 & \rho^m & \rho^{2m} & \cdots & \rho^{m(n-1)} \end{bmatrix}^T, \quad m = 0, \ldots, n-1, \hfill (10)$$

Note that the eigenvectors $\{w^{(m)}\}$ are indexed with the same index as the eigenvalues $\{\lambda_m\}$. Once the form of the vector (10) is given, it is immediate to verify that it is an eigenvector

$$S^* \begin{bmatrix} 1 & \rho^m & \rho^{2m} & \cdots & \rho^{m(n-1)} \end{bmatrix}^T = \begin{bmatrix} \rho^m & \rho^{2m} & \cdots & \rho^{m(n-1)} & 1 \end{bmatrix}^T$$

$$= \rho_m \begin{bmatrix} 1 & \rho & \rho_m^2 & \cdots & \rho_m^{m(n-2)} & \rho^{-1} \end{bmatrix}^T$$

$$= \rho_m \begin{bmatrix} 1 & \rho & \rho_m^2 & \cdots & \rho_m^{m(n-2)} & \rho_m^{m(n-1)} \end{bmatrix}^T.$$

While this last calculation provides an easy and immediate verification of Lemma 2, it is not a satisfactory starting point for the development of the subject, as the form of the eigenvalues/eigenvectors would have to be “guessed”. Instead, as shown above, they can be constructed systematically from the basic definitions that result in (8), and no guessing is required.
2.2 Eigenvalues Calculation of a Circulant Matrix Yields the DFT

Now that we have calculated all the eigenvectors of the shift operator in Lemma 2, we can use them to find the eigenvalues of any circulant matrix $C_x$ from the relation

$$C_x w^{(m)} = \lambda_m w^{(m)} \iff \begin{bmatrix} x_0 & x_{n-1} & \cdots & x_1 \\ x_1 & x_0 & \cdots & x_2 \\ \vdots & \ddots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & \cdots & x_0 \end{bmatrix} \begin{bmatrix} 1 \\ \rho_m \\ \vdots \\ \rho_m^{n-1} \end{bmatrix} = \lambda_m \begin{bmatrix} 1 \\ \rho_m \\ \vdots \\ \rho_m^{n-1} \end{bmatrix}, \quad (12)$$

Each row of the above equation represent essentially the same equation (but multiplied by a power of $\rho_m$). The first row is the easiest equation to write down and gives

$$\lambda_m = x_0 + x_{n-1} \rho_m + \cdots + x_1 \rho_m^{n-1}$$

$$= x_0 + x_1 \rho_m^{-1} + \cdots + x_{n-1} \rho_m^{(n-1)}$$

$$= \sum_{l=0}^{n-1} x_l \rho_m^{-l} = \sum_{l=0}^{n-1} x_l \rho^{-ml} = \sum_{l=0}^{n-1} x_l e^{-i \frac{2\pi}{n} ml} =: \hat{x}_m, \quad (13)$$

which is precisely the classically-defined DFT (2) of the vector $x$.

Another way to state the previous conclusion is to define a complex function, the \"z\-transform\" of $x$, which is a polynomial in the variable $z^{-1}$ as follows

$$\hat{x}(z) := x_0 + x_1 z^{-1} + \cdots + x_{n-1} z^{-(n-1)}.$$  

Then our conclusion can be restated as the eigenvalues of $C_x$ being the values of the function $\hat{x}(z)$ evaluated at $z = \rho_m$, the $n$-roots of unity.

One might ask what the conclusion would have been if the eigenvectors of $S$ have been used instead of those of $S^*$. A repetition of the previous steps but now for the case of $S$ would yield that the eigenvalues of a circulant matrix $C_x$ are given by

$$\mu_k = \sum_{l=0}^{n-1} x_l e^{i \frac{2\pi}{n} kl}, \quad k = 0, 1, \ldots, n - 1. \quad (14)$$

While the expressions (13) and (14) may at first appear different, the sets of numbers $\{\lambda_m\}$ and $\{\mu_k\}$ are actually equal. So in fact, the expression (14) gives the same set of eigenvalues as (13) but arranged in a different order since $\mu_k = \lambda_{-k}$.

Along with the two choices of $S$ and $S^*$, there are also other possibilities. Let $p$ be any number that is coprime with $n$. It is easy to show (Exercise 2) that a $n \times n$ matrix is circulant iff it commutes with $S^p$. Therefore the eigenvectors of $S^p$ (rather than those of $S$) can be used to simultaneously diagonalize all circulant matrices. This would yield yet another transform distinct from the two transforms (13) or (14). However, the set of numbers produced from that transform will still be the same as those computed from the previous two transforms, but arranged in a different ordering.
3 Circular Convolution

We will start with examining the matrix-vector product when the matrix is circulant. By analyzing this product, we will obtain the circular convolution of two vectors. Let $C_x$ by some circulant matrix, and examine the action of a such a matrix on any vector $y = [y_0, y_1, \cdots, y_{n-1}]^T$. The matrix-vector multiplication $z = C_x y$ in detail reads

$$z = \begin{bmatrix} z_0 \\ z_1 \\ \vdots \\ z_{n-1} \end{bmatrix} = \begin{bmatrix} x_0 & x_{n-1} & \cdots & x_1 \\ x_1 & x_0 & \cdots & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & \cdots & x_0 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} = C_x y. \quad (15)$$

Using $(C_x)_{ij} = x_{i-j}$, the above matrix-vector multiplication can be rewritten as

$$z_i = \sum_{j=0}^{n-1} (C_x)_{ij} y_j = \sum_{j=0}^{n-1} x_{i-j} y_j. \quad (16)$$

Note that this can be viewed as an operation on the two vectors $x$ and $y$ to yield the vector $z$, and allows us to reinterpret the matrix-vector product of a circulant matrix as follows.

**Definition 1** Given two $n$-vectors $x$ and $y$, their circular convolution $z = x \star y$ is another $n$-vector defined by

$$z = x \star y \iff z_k = \sum_{l=0}^{n-1} x_l \cdot y_{k-l}, \quad (17)$$

where the indices in the sum are evaluated modulo $n$.

It follows immediately from (16) that multiplication by a circulant matrix $C_x$ is equivalent to convolving with its defining vector $x$, i.e.

$$z = C_x y = x \star y. \quad (18)$$

The sum in (17) defining circular convolution has a nice circular visualization due to modular arithmetic on $\mathbb{Z}_n$. This is illustrated in figure 4. The elements of $x$ are arranged in a discrete circle counter-clockwise, while the elements of $y$ are arranged in a circle clockwise (the reverse orientation is because elements of $x$ are indexed like $x_l$ while those of $y$ are indexed like $y_{-l}$ in the definition (17)). For each $k$, the array $y$ is rotated counter-clockwise by $k$ steps (figure 4 shows cases for three different values of $k$). The number $z_k$ in (17) is then obtained by multiplying the corresponding elements of the $x$ and rotated $y$ arrays, and then summing. This generates the $n$ numbers $z_0, \ldots, z_{n-1}$. From the definition, it is easy to show (see Exercise 3) that circular convolution is associative and commutative.

- **Associativity:** for any three $n$-vectors $x$, $y$ and $z$ we have
  $$x \star (y \star z) = (x \star y) \star z$$

- **Commutativity:** for any two $n$-vectors $x$ and $y$
  $$x \star y = y \star x$$
The above two facts have several interesting implications. First, since convolution is commutative, the above matrix-vector product can be written in two equivalent ways

\[ x \star y = C_x y = C_y x. \]

Applying this fact in succession to two circulant matrices

\[ C_x C_y z = C_x (y \star z) = x \star (y \star z) = (x \star y) \star z = C_{x \star y} z. \]

This means that the product of any two circulant matrices \( C_x \) and \( C_y \) is another circulant matrix \( C_{x \star y} \) whose defining vector is \( x \star y \), the circular convolution of the defining vectors for \( C_x \) and \( C_y \). We summarize this conclusion and an important corollary of it next.

**Fact 1**

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

\[ x \star y = C_x y = C_y x. \]

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

\[ C_x C_y = C_{x \star y}. \]

3. All circulant matrices mutually commute since for any two \( C_x \) and \( C_y \)

\[ C_x C_y = C_{x \star y} = C_{y \star x} = C_y C_x. \]

The set of all \( n \)-vectors forms a commutative algebra under the operation of circular convolution. The above shows that the set of \( n \times n \) circulant matrices is also a commutative algebra isomorphic to \( n \)-vectors with circular convolution.

### 4 The Big Picture

Let \( C_x \) be a circulant matrix made from a vector \( x \) as in (1). If we use the eigenvectors (10) of \( S^* \) as columns of a matrix \( W \), the \( n \) eigenvalue/eigenvector relationships (12) \( C_x w^{(m)} = \)}
\( \lambda_m w^{(m)} \) can be written as a single matrix equation as follows

\[
C_x \begin{bmatrix} w^{(0)} & \cdots & w^{(n-1)} \end{bmatrix} = \begin{bmatrix} w^{(0)} & \cdots & w^{(n-1)} \end{bmatrix} \begin{bmatrix} \hat{x}_0 & \cdots & \hat{x}_{n-1} \end{bmatrix},
\]

where we have used the fact (13) that the eigenvalues \( \{ \lambda_m \} \) of \( C_x \) are precisely \( \{ \hat{x}_m \} \), the elements of the DFT of the vector \( x \).

The columns of \( W \) are mutually orthogonal, and thus \( W \) is a unitary matrix (up to a rescaling) \( W^*W = WW^* = nI \), or equivalently \( W^{-1} = \frac{1}{n} W^* \). Since the matrix \( W \) is made up of the eigenvectors of \( S^* \), which in turn are made up of various powers of the roots of unity (10), it has some special structure which is worth examining

\[
W := \begin{bmatrix} w^{(0)} & \cdots & w^{(n-1)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \rho & \cdots & \rho^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \rho^{n-1} & \cdots & \rho^{(n-1)(n-1)} \end{bmatrix}.
\]

The matrix \( W^* \) 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 therefore write

\[
W^* = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \rho^{-1} & \cdots & \rho^{-(n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \rho^{-(n-1)} & \cdots & \rho^{-(n-1)(n-1)} \end{bmatrix}.
\]

Also observe that multiplying a vector by \( W^* \) is exactly taking its DFT. Indeed the \( m \)'th row of \( W^*y \) is

\[
\hat{y}_m = \begin{bmatrix} 1 & \rho^{-m} & \cdots & \rho^{-m(n-1)} \end{bmatrix} \begin{bmatrix} y_0 \\ \vdots \\ y_{n-1} \end{bmatrix},
\]

which is exactly the definition (2) of the DFT. Similarly, multiplication by \( \frac{1}{\sqrt{n}} W \) is taking the inverse DFT

\[
y_l = \frac{1}{n} \sum_{k=0}^{n-1} \hat{y}_k \rho^{kl} = \frac{1}{n} \sum_{k=0}^{n-1} \hat{y}_k e^{\frac{2\pi i kl}{n}}.
\]

Multiplying both sides of (19) from the right by \( W^{-1} \) gives the diagonalization of \( C_x \) which can be written in several equivalent forms

\[
C_x = W \text{ diag} \hat{x} W^{-1} = W \text{ diag} \hat{x} \left( \frac{1}{n} W^* \right) = \left( \frac{1}{n} W \right) \text{ diag} \hat{x} W^* \quad (20)
\]

\[
= \left( \frac{1}{\sqrt{n}} W \right) \text{ diag} \hat{x} \left( \frac{1}{\sqrt{n}} W^* \right). \quad (21)
\]

The diagonalization (20) can be interpreted as follows in terms of the action of a circulant matrix \( C_x \) on any vector \( y \)

\[
C_x y = \left( \frac{1}{n} W \right) \text{ diag} \hat{x} \underbrace{W^* y}_{\text{DFT of } y} \underbrace{\text{ multiply by } \hat{x} \text{ entrywise}}_{\text{inverse DFT}}.
\]
Thus the action of $C_x$ on $y$, or equivalently the circular convolution of $y$ with $x$, can be performed by first taking the DFT of $y$, then multiplying the resulting vector component-wise by $\hat{x}$ (the DFT of the vector $x$ defining the matrix $C_x$), 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. This is illustrated in Figure 5.

Figure 5: Illustration of the relationships between circulant matrices, circular convolution and the DFT. The matrix-vector multiplication $z = C_x y$ with the circulant matrix $C_x$ is equivalent to the circular convolution $z = x \ast y$. The DFT is a linear transformation $W$ on vectors with inverse $\frac{1}{n} W$. It converts multiplication by the circulant matrix $C_x$ into multiplication by the diagonal matrix $\text{diag}(\hat{x}_0, \ldots, \hat{x}_{n-1})$ whose entries are the DFT of the vector $x$ defining the matrix $C_x$.

Note that in the literature there is an alternative form for the DFT and its inverse

$$\hat{y}_k = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} y_l e^{-i \frac{2\pi}{n} kl}, \quad y_l = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \hat{y}_k e^{i \frac{2\pi}{n} kl},$$

which is sometimes preferred due to its symmetry (and is also truly unitary since with this definition $\|y\|_2 = \|\hat{y}\|_2$). This “unitary” DFT corresponds to the diagonalization given in (21) instead of (20). We do not adopt this unitary DFT definition here since it complicates the statement that the eigenvalues of $C_x$ are precisely entries of $\hat{x}$.

**Theorem 3**

1. The following sets are isomorphic commutative algebras
   
   (a) The set of $n$-vectors is closed under circular convolutions and is thus an algebra with the operations of addition and convolution.

   (b) The set of $n \times n$ circulant matrices is an algebra under the operations of addition and matrix multiplication.

   (c) The set of $n$-vectors is an algebra under the operations of additions and component-wise multiplication.

2. For any $n$-vector $x = (x_0, \ldots, x_{n-1})$, denote its associated circulant matrix (1) by $C_x$ and its DFT (2) by $(\hat{x}_0, \ldots, \hat{x}_{n-1})$.

   The set of $n$-vectors, their associated circulant matrices and their DFTs are isomorphic as algebras under the respective operations described above. Specifically,

   $$C_{x \ast y} = C_x C_y, \quad \hat{x} \circ \hat{y} = \hat{x} \otimes \hat{y},$$

   where $\circ$ denotes the element-by-element product of vectors.

---

3 If the unitary DFT is adopted, the equivalent statement would be that the eigenvalues of $C_x$ are the elements of the entries of $\sqrt{n} \hat{x}$. 

12
3. The eigenvalues of a circulant matrix $C_x$ are given by the values of the $z$-transform

$$\hat{x}(z) := x_0 + x_1 z^{-1} + \cdots + x_{n-1} z^{n-1}$$

evaluated at the roots of unity $\rho^k := e^{\frac{2\pi}{n} k}$, i.e.

$$\hat{x}(\rho^k) = \sum_{l=0}^{n-1} x_l (\rho^k)^{-l} = \sum_{l=0}^{n-1} x_l \rho^{-kl} = \sum_{l=0}^{n-1} x_l e^{-\frac{2\pi}{n} kl} =: \hat{x}_k,$$

which are precisely the entries of $\hat{x}$, the DFT of $x$.

5 The Multidimensional Case

In the one dimensional case covered earlier, one considers an $n$-vector as a real (or complex) valued function on the index set $\mathbb{Z}_n$, which is viewed as a discrete circle. In the multidimensional case, this set is replaced by $\mathbb{Z}_n^d$, the $d$-dimensional torus. Figure 6 illustrates its structure. It can be either thought of as a lattice with periodic boundaries, or equivalently as a torus embedded in a higher dimensional space. It is convenient to use multi-index notation for an element $k \in \mathbb{Z}_n^d$ where $k = (k_1, \ldots, k_d)$, with each $k_q \in \mathbb{Z}_n$ for $q = 1, \ldots, d$.

The proper generalization of circulant matrices is to operators that act on functions on $\mathbb{Z}_n^d$. Instead of $n$-vectors, we now have $d$-dimensional $n$-arrays which we equivalently consider as functions $x : \mathbb{Z}_n^d \rightarrow \mathbb{R}$. We simply refer to such objects as arrays when their dimensions are clear from context. Multi-index notation is also used for elements of such arrays, i.e. $x_k$ stands for $x_{(k_1,\ldots,k_d)}$ for any array $x : \mathbb{Z}_n^d \rightarrow \mathbb{R}$.

For functions defined on $\mathbb{Z}_n^d$, a general shift operator has the form

$$(S_l x)_{(k_1,\ldots,k_d)} := x_{(k_1-l_1,\ldots,k_d-l_d)},$$

and note how the operator is indexed by the multi-index $l = (l_1,\ldots,l_d)$. There are as many shift operators as the cardinality of $\mathbb{Z}_n^d$, but amongst all of them there are $q$ basic ones which we refer to as the coordinate-shift operators

$$(S_q x)_{(k_1,\ldots,k_d)} := x_{(k_1,\ldots,k_q-1,\ldots,k_d)},$$
i.e. \( S_q \) is a right circular-shift in only the \( q \) coordinate. There are \( d \) (the dimension of the Torus) such operators, and an important property they posses is that they generate all other shift operators by

\[
S_{(l_1, \ldots, l_d)} = S_{l_1}^{l_1} \cdots S_{l_d}^{l_d}.
\]

Observe that since all shift operators mutually commute, the order in which the coordinate shifts are performed does not matter.

The shift operators encode the basic symmetries of \( \mathbb{Z}_n^d \), and we therefore use them to define circulant operators.

**Definition 2** A linear operator \( M : \mathbb{R}^{\mathbb{Z}_n^d} \rightarrow \mathbb{R}^{\mathbb{Z}_n^d} \) is called circulant if it commutes with all the coordinate-shift operators

\[
S_q M = MS_q, \quad q = 1, \ldots, d.
\]

Since coordinate shifts generate all other shifts by (22), it then follows that a circulant operator also commutes with all possible shifts \( S_l, l \in \mathbb{Z}_n^d \).

An easy calculation will show that an operator is circulant iff it can be represented as circular convolution with an array. First define the multidimensional Kroenecker delta \( \delta \) and its shifted version \( \delta^{(l)} \)

\[
\delta_k := \begin{cases} 1 & \text{if } k = (0, \ldots, 0), \\ 0 & \text{if } k \neq (0, \ldots, 0), \end{cases} \quad \delta^{(l)}_k := \begin{cases} 1 & \text{if } k = (l_1, \ldots, l_d), \\ 0 & \text{if } k \neq (l_1, \ldots, l_d). \end{cases}
\]

i.e. \( \delta^{(l)} \) is the array that is 1 at index \( l \) and zero otherwise. Note that any shifted delta can be obtained from \( \delta \) by acting with a shift operator \( \delta^{(l)} = S_l \delta \), i.e. \( \delta^{(l)}_k = \delta_{k-l} \). We also note that any array \( y \) can be written as a weighted sum of shifted deltas

\[
y = \sum_{l \in \mathbb{Z}_n^d} y_l \delta^{(l)} = \sum_{l \in \mathbb{Z}_n^d} y_l S_l \delta.
\]

Let \( M \) be a circulant operator in the sense of Definition 2 and define the array \( x \) as the result of the action of \( M \) on the Kroenecker delta

\[
x := M \delta,
\]

i.e. \( x \) is the “impulse response” of \( M \). Now using the representation (23) and the commutativity of \( M \) with shifts, we can write the action of \( M \) on any array \( y \) with a formula that involves only the impulse response \( x \) and the array \( y \)

\[
z = My = M \left( \sum_{l \in \mathbb{Z}_n^d} y_l S_l \delta \right) = \sum_{l \in \mathbb{Z}_n^d} y_l S_l M \delta = \sum_{l \in \mathbb{Z}_n^d} y_l S_l x
\]

\[
\Leftrightarrow z_k = \sum_{l \in \mathbb{Z}_n^d} y_l x_{k-l},
\]

where we have used the commutativity \( MS_l = S_l M \) in the third equality, and the fact that \( M \delta \) is just the impulse response \( x \). We have therefore uncovered multidimensional circular convolution as the representation of circulant operators.
Lemma 4 A linear operator $M : y \mapsto z$ on $d$-dimensional $n$-arrays is circulant iff it can be written as a circular convolution

$$z_k = \sum_{l \in \mathbb{Z}_n^d} x_{k-1} \ y_l \iff z_{(k_1, \ldots, k_d)} = \sum_{l_1, \ldots, l_d \in \mathbb{Z}_n} x_{(k_1-l_1, \ldots, k_d-l_d)} \ y_{(l_1, \ldots, l_d)}$$

where $x$ is the “impulse response” $M \delta$ of $M$. In this case, we denote $M = C_x$, the circulant operator defined by the array $x$.

Thus a circulant operator $C_x$ is completely specified by the array $x$ (its “impulse response”) in the same way a circulant matrix is completely specified by its first column. In contrast to the one dimensional case, matrix-vector notation does not yield a simple visualization of multidimensional circulant operators. In other words, matrix representations of circulant operators do not have an easily discernible matrix structure in higher dimensions.

5.1 The Mutlidimensional DFT

In a manner similar to the one dimensional case, we would like to simultaneously diagonalize all circulant operators and discover the multidimensional DFT in the process. In the 1 dimensional case, there was a single shift operator, the commutation with which characterizes all circulant matrices. In the $d$ dimensional case, there are $d$ such basic operators that all mutually commute. It turns out that for each of those operators $S_q$ individually, their eigenvalues come with multiplicities, and there is not a unique way to select their eigenvectors. However, they do share one set of unique (up to scalar multiples) eigenvectors.

To find these shared eigenvectors, consider the following eigenvalue problem

$$S_q^* w = \lambda_q w, \quad \text{for } q = 1, \ldots, d,$$

i.e. look for an array $w$ that is an eigenvector of all coordinate-shift operators $S_q^*$ simultaneously, but possibly with different eigenvalues $\lambda_q$. This condition can be written out explicitly as

$$w_{(k_1, \ldots, k_{q+1}, \ldots, k_d)} = \lambda_q w_{(k_1, \ldots, k_d)}; \quad \text{for } q = 1, \ldots, d.$$

We also have that $(S_q^*)^l w = \lambda_q^l w$ which explicitly is

$$w_{(k_1, \ldots, k_{q+1}, \ldots, k_d)} = \lambda_q^l w_{(k_1, \ldots, k_d)}; \quad \text{for } q = 1, \ldots, d. \quad (24)$$

Since for each $q$, we have $k_{q+n} \equiv_n k_q$ we conclude that for each $q$

$$1 = \lambda_q^n,$$

i.e. each $\lambda_q$ is an $n$-root of unity, which means that $\lambda_q = \rho^{m_q} = e^{i \frac{2\pi m_q}{n}}$ for some $m_q \in \mathbb{Z}_n$. Now, since eigenvectors are defined only up to a scalar multiple, we set $w_{(0, \ldots, 0)} = 1$, and use (24) in succession to generate all components of $w$ from $w_{(0, \ldots, 0)} = 1$ by

$$w_{(l_1, 0, \ldots, 0)} = \rho^{l_1 m_1} w_{(0, \ldots, 0)} = \rho^{l_1 m_1}$$
$$w_{(l_1, l_2, 0, \ldots, 0)} = \rho^{l_1 m_2 + l_2 m_1} w_{(l_1, 0, \ldots, 0)} = \rho^{l_2 m_1 + l_1 m_1}$$
$$\vdots$$
$$w_{(l_1, \ldots, l_d)} = \rho^{l_d m_d \ldots + l_1 m_1} = \rho^{l_1 m_1 + \cdots + l_d m_d}$$

We will use the notation $m \cdot l := l_1 m_1 + \cdots + l_d m_d$ (a dot product of the indices). Note that in the above formula, the eigenvector $w$ is indexed by $m \in \mathbb{Z}_n^d$. There are as many of them as the cardinality of $\mathbb{Z}_n^d$. We summarize this conclusion next.
Lemma 5 The coordinate-shift operators $S_1^*, \ldots, S_d^*$ have a total of $n \times d$ common eigenvectors $w_{(m)}$, $m \in \mathbb{Z}_n$. The entries of the arrays $w_{(m)}$ are given by
\begin{equation}
w_{(m)} = w_{(m_1, \ldots, m_d)} = \rho^{l_1 m_1 + \cdots + l_d m_d} = \rho^{m^* l}, \quad l, m \in \mathbb{Z}_n,
\end{equation}
where $\rho := e^{i \frac{2\pi}{n}}$, the $n$-root of unity.

Now that we have discovered the common eigenvectors of the coordinate-shift operators, we can diagonalize any circulant operator. Indeed, let $C_x$ be a circulant operator generated by an array $x$. Since it is circulant, it commutes with all coordinate-shifts. It thus has the same eigenvectors as them. To find the eigenvalues of $C_x$, we simply apply it to the already computed eigenvectors (25)
\begin{equation}
(C_x w_{(m)})_k = \sum_{l \in \mathbb{Z}_n^d} x_{k-l} w_{(m)} = \sum_{l \in \mathbb{Z}_n^d} x_{k-l} \rho^{m^* l},
\end{equation}
where we have used Lemma 4 for the first equality, and (25) for the second equality. A reindexing of the sum (with $j := k - l$) uncovers the eigenvalues
\begin{equation}
(C_x w_{(m)})_k = \sum_{j \in \mathbb{Z}_n^d} x_j \rho^{m^*(k-j)} = \rho^{m^* k} \sum_{j \in \mathbb{Z}_n^d} x_j \rho^{-m^* j} = w_{(m)} \left( \sum_{j \in \mathbb{Z}_n^d} x_j \rho^{-m^* j} \right).
\end{equation}
Therefore, the eigenvalues of $C_x$ are the set of numbers
\begin{equation}
\sum_{j \in \mathbb{Z}_n^d} x_j \rho^{-m^* j} =: \hat{x}_m,
\end{equation}
which we will define to be the DFT of $x$. We summarize this in the next statement.

Theorem 6 Let $C_x$ be a circulant operator defined by the multidimensional array $x$ over $\mathbb{Z}_n^d$. The $n \times d$ eigenvalues of $C_x$ are given by the multidimensional DFT of the array $x$
\begin{equation}
\hat{x}_m = \hat{x}_{(m_1, \ldots, m_d)} = \sum_{l \in \mathbb{Z}_n^d} x_{(l_1, \ldots, l_d)} e^{-i \frac{2\pi}{n} (m_1 l_1 + \cdots + m_d l_d)} = \sum_{l \in \mathbb{Z}_n^d} x_l \rho^{-m^* l},
\end{equation}
where $\rho := e^{i \frac{2\pi}{n}}$ is the $n$-root of unity. Equivalently, the eigenvalues are given by evaluations of the multidimensional $z$-transform
\begin{equation}
\hat{x}(z) := \sum_{l \in \mathbb{Z}_n^d} x_{(l_1, \ldots, l_d)} z_1^{-l_1} \cdots z_d^{-l_d}
\end{equation}
with each coordinate evaluated at all possible roots of unity $z_q = \rho^m$, $q = 1, \ldots, d$, $m \in \mathbb{Z}_n$.

References

[1] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, “Signals and systems, vol. 2,” Prentice-Hall Englewood Cliffs, NJ, vol. 6, no. 7, p. 10, 1983.
[2] R. Plymen, “Noncommutative fourier analysis,” 2010.
[3] M. E. Taylor and J. Carmona, Noncommutative harmonic analysis. American Mathematical Soc., 1986, no. 22.
A Exercises

Exercise ch.1
Show that any matrix $M$ that commutes with the shift operator $S$ must be a circulant matrix, i.e. must have the structure shown in (1), or equivalently (3).

Solution ch.1
Starting from the relation $SM = MS$, and using the definition $(S)_{ij} = \delta_{i-j-1}$ compute

$$(SM)_{ij} = \sum_l S_{il} (M)_{lj} = \sum_l \delta_{i-l-1} (M)_{lj} = (M)_{i-1,j},$$

$$(MS)_{ij} = \sum_l (M)_{il} S_{lj} = \sum_l (M)_{il} \delta_{l-j-1} = (M)_{i,j+1}.$$ 

Note that since the indices $i-j-1$ of the Kroenecker delta are to be interpreted using modular arithmetic, then the indices $i-1$ and $j+1$ of $M$ above should also be interpreted with modular arithmetic. The statements

$$(M)_{i-1,j} = (M)_{i,j+1} \iff (M)_{i-1,j-1} = (M)_{i,j} \iff (M)_{i,j} = (M)_{i+1,j+1}$$

then mean that the $i$’th column is obtained from the previous $i-1$ column by circular right shift of it.

Alternatively, the last statement above implies that for any $k$, $(M)_{i,j} = (M)_{i+k,j+k}$, i.e. that entries of $M$ are constant along “diagonals”. Now take the first column of $M$ as $m_i := (M)_{i,0}$, then

$$(M)_{ij} = (M)_{i-j,j-j} = (M)_{i,j,0} = m_{i-j}.$$ 

Thus all entries of $M$ are obtained from the first column by circular shifts as in (3).

Exercise ch.2
Show that an $n \times n$ matrix $M$ is circulant iff it commutes with $S^p$ where $(p, n)$ are coprime.

Solution ch.2
If $M$ is circulant, then it commutes with $S$ and also commutes with any of its powers $S^p$. The other direction is more interesting.

The basic underlying fact for this conclusion has to do with modular arithmetic in $\mathbb{Z}_n$. If $(p, n)$ are coprime, then there are integers $a, b$ that satisfy the Bezout identity

$$ap + bn = 1,$$

which also implies that $ap$ is equivalent to 1 mod $n$ since $ap = 1 - bn$, i.e. it is equal to a multiple of $n$ plus 1. Therefore, there exists a power of $S^p$, namely $S^{ap}$ such that

$$S^{ap} = S.$$ 

Equation (26) has a nice geometric interpretation. $S^p$ is a rotation of the circle in Figure 3 by $p$ steps. If $p$ and $n$ were not coprime, then regardless of how many times the rotation $S^p$ is repeated, there will be some elements of the discrete circle that are not
reachable from the 0 element by these rotations. The condition $p$ and $n$ coprime insures that there is some repetition of the rotation $S^p$, namely $(S^p)^a$ which gives the basic rotation $S$. Repetitions of $S$ then of course generate all possible rotations on the discrete circle. In other words, $p$ and $n$ coprime insures that by repeating the rotation $S^p$, all elements of the discrete circle are eventually reachable from 0.

**Exercise ch.3**
Show that circular convolution (17) is commutative and associative.

**Solution ch.3**

*Commutativity:* Follows from

$$(a \ast b)_k = \sum_l a_l b_{k-l} = \sum_l a_{k-j} b_j = (b \ast a)_k$$

where we used the substitution $j = k - l$ (and consequently $l = k - j$).

*Associativity:* First note that $(b \ast c)_i = \sum_j b_j c_{i-j}$, and compare

$$(a \ast (b \ast c))_k = \sum_l a_l (b \ast c)_{k-l} = \sum_l a_l \left( \sum_j b_j c_{k-l-j} \right) = \sum_{l,j} a_l b_j c_{k-l-j}$$

$$(a \ast (b \ast c))_k = \sum_j (a \ast b)_j c_{k-j} = \sum_j \left( \sum_l a_l b_{j-l} \right) c_{k-j} = \sum_{l,j} a_l b_{j-l} c_{k-j}.$$  \[ 18 \]

Relabeling $j - l =: i$ (and therefore $j = l + i$) in the second sum makes it

$$\sum_{l,i} a_j b_i c_{k-(l+i)} = \sum_{l,i} a_j b_i c_{k-l-i},$$

Which is exactly the first sum, but with a different labeling of the indices.