Approximating the Spectrum
DTFT of a finite-length portion of the signal

- In sum, when applying DTFT for the samples in a finite portion (or segment) of the signal, the approximation is caused by

(1) frequency domain convolution with a sync function

(2) aliasing

- Although imperfect, this is a practically feasible way for evaluating the spectrum of a signal.

General form in analog domain: **Short-time Fourier Transform (STFT)**

- A general extension of spectrum is the short-time Fourier transform (STFT), where the window slides at every time sites.
- The signal is multiplied by a window function.

$$\mathbf{STFT}\{x(t)\}(\tau,\omega) \equiv X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j\omega t}\,dt$$

- The CFT of the resulting signal is taken as the window (usually of finite length) is slid along the time axis, resulting in a two-dimensional representation of the signal.

Approximating the Spectrum
DTFT of a finite-length portion of the signal

- **What can be done to further improve the performance?**
- Rectangular window is applied in the rest of this lecture.
- In general, the window can be replaced by other windowing functions (eg., Hann window, Hamming window, Kaiser window) that would distort the shape of spectrum less in general
  - **will be investigated in the future**

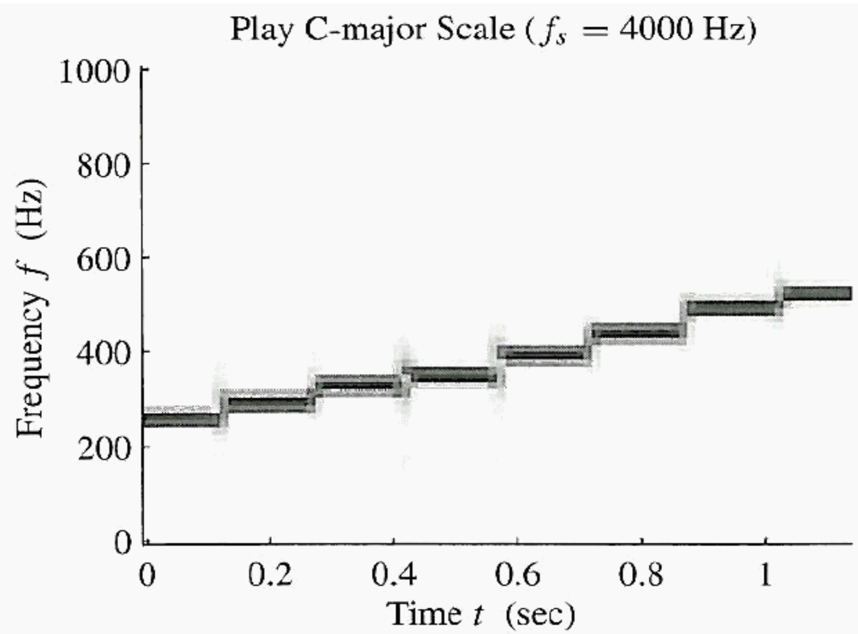
Segmenting the signal

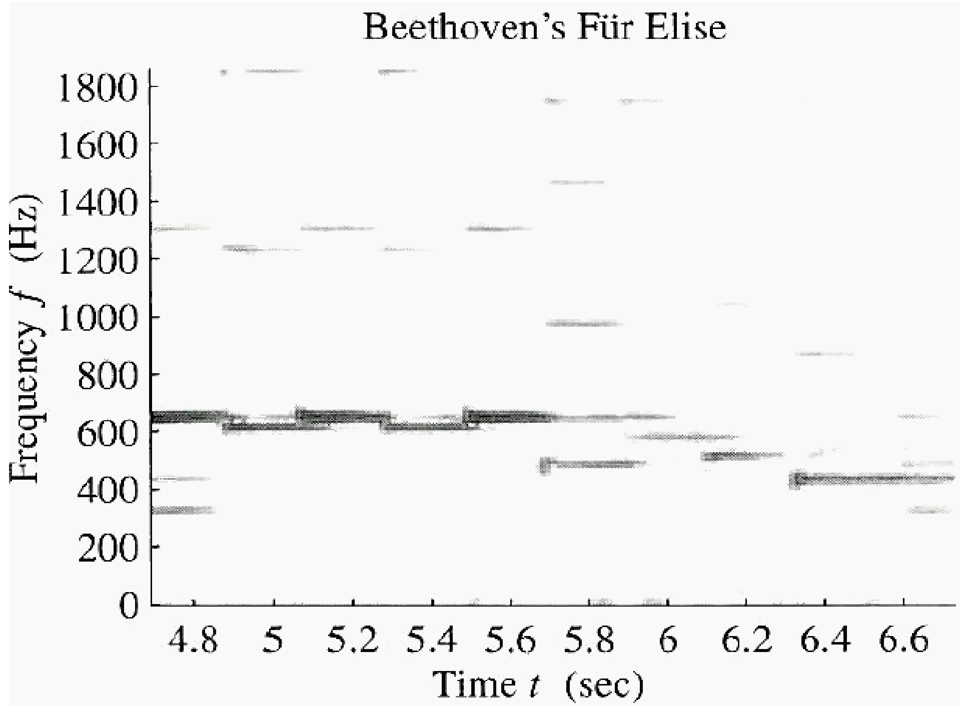
- In the above, we only use a single segment of the signal and compute its DTFT.
- It is natural to extend this idea by separating a signal into **multiple segments** in time, and compute the DTFT for each segment along the time axis.
- The segments can be either **non-overlapping** or **overlapping**.

Spectrogram

- In this way, we can obtain a two-dimensional map (like a 2D image)
  - the horizontal axis is the time interval
  - the vertical axis is the frequency in $[-\pi/T, \pi/T]$.
  - the gray level is the magnitude of DTFT of the associated segment.
- **This is called spectrogram, a common way for viewing the spectral domain response for a signal in practice.**
  - Sometimes the gray level is the magnitude square of DTFT, depending on how it is defined.

Spectrogram Examples





Synthesized notes, C,D,E,F,G,A,B,C (which are pure sinusoids)

Spectrogram of Fur Elise played on a piano

The blurs are caused by the following reason: As the time domain is only a finite portion of a sinusoid wave, the frequency domain is **convolved by a Sinc function**.

Spectrogram Examples

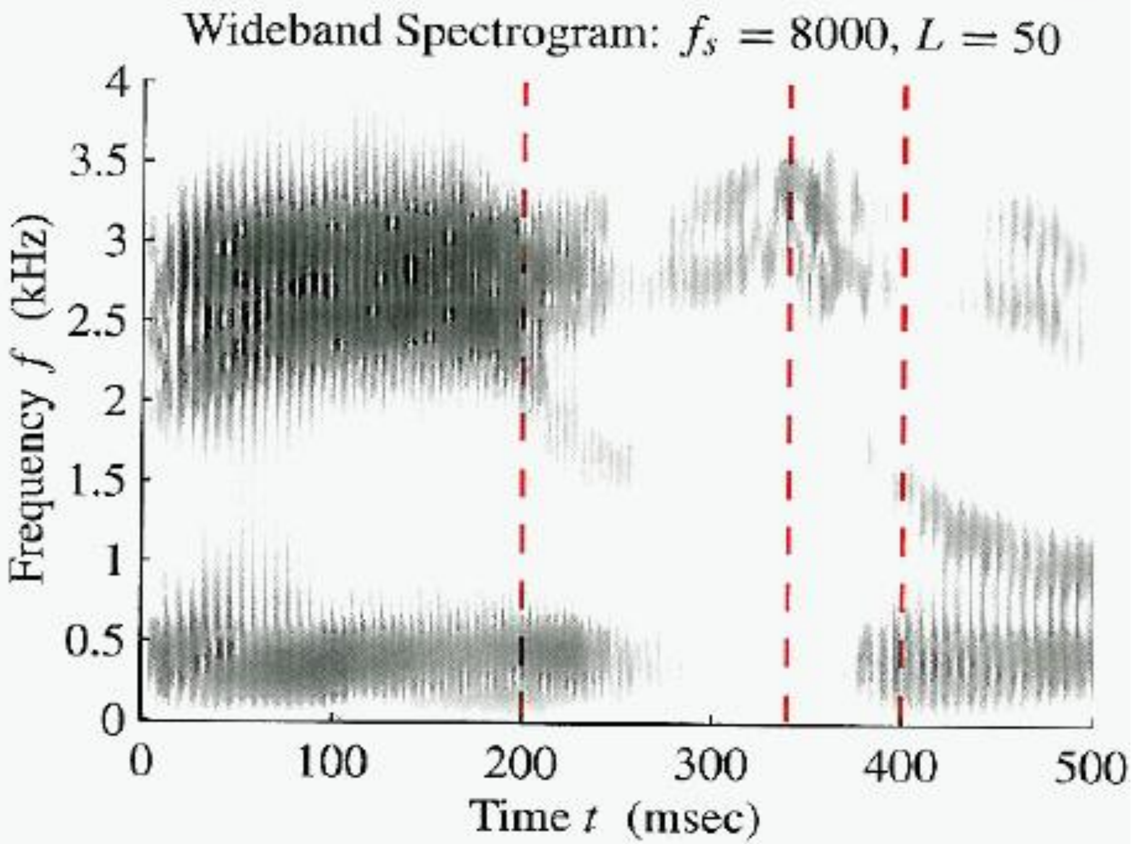


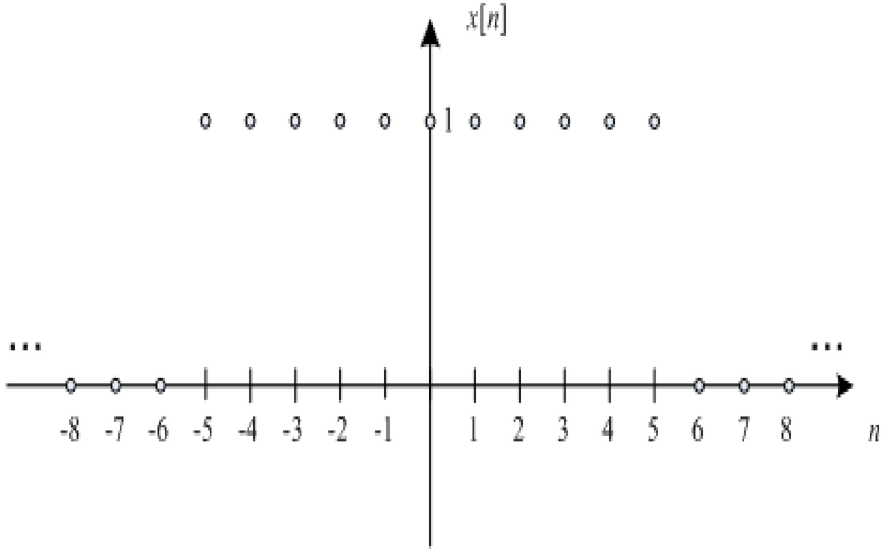
Figure 13-27: Spectrogram of speech signal; sampling frequency $f_s = 8000$ Hz, window length $L = 50$ (6.25 msec). Spectrum slices at times $nT_s = 200$, 340, and 400 msec are shown individually in Fig. 13-28.

- DTFT range: $[-\pi, \pi]$
- Associated CFT range: $[-\pi/T_s, \pi/T_s]$ in radians.
- That is, $[-\omega_s/2, \omega_s/2]$ in radians. (because $w_s = 2\pi/T_s$)
- This is equivalent to $[-f_s/2, f_s/2]$ in Hz.

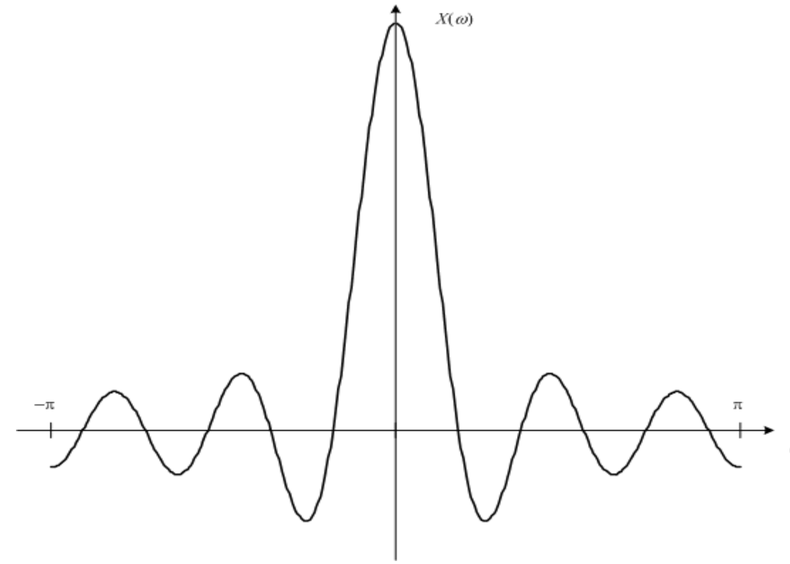
- In the example:
- $f_s = \mathbf{8000\ Hz}$
- $f_s/2 = 4000 = \mathbf{4\ kHz}$
- For real-valued signal, the magnitude is an even function. **Only show the positive side $[0, f_s/2]$** in the spectrogram to ease the viewing

A remark on DTFT viewing

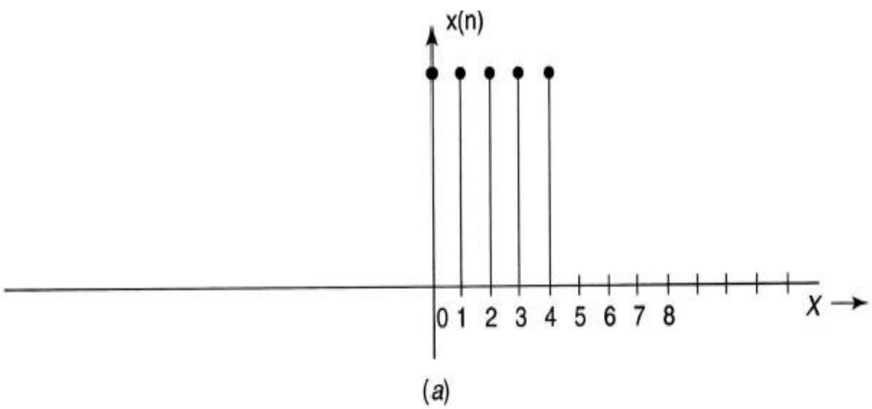
- Remember that DTFT is typically viewed in $[-\pi, \pi]$ (sometimes we draw all of the periods).
- As every period contains the same information, in some cases we draw $[0, 2\pi]$ for visualization.
- For real signal, the spectrum is an even function. Both cases $[-\pi, \pi]$ and $[0, 2\pi]$ can be reduced into $[0, \pi]$ due to symmetry.



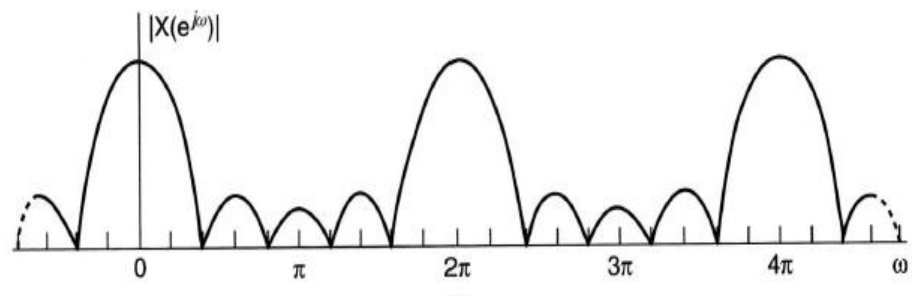
Eg, a time-domain *N*-point (*N*=11) discrete-time signal (an even function)



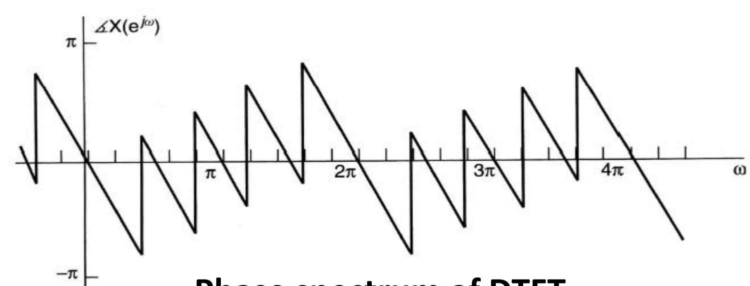
Its DTFT domain is continuous in [-π, π]



A finite-length signal (an odd function) of *N*=5



Magnitude spectrum of DTFT



Phase spectrum of DTFT

Further Improvement

- In the above, we compute the DTFT for a finite-length (i.e., finite-duration) signal (of length N).
- To record the frequency domain of DTFT, we have to record a continuous function in the range of $[-\pi, \pi]$.
- However, to record exactly a continuous function in a digital device is infeasible in practice.

Further Improvement

- There are only N points in time domain for a finite-duration signal (i.e., time domain is an N-dimensional vector).
- We know that time and frequency domains retain the same information for signal representation.
- So, for an N-points signal in time domain, it seems redundant if we need to use a continuous function (containing infinite points) in $[-\pi, \pi]$ or $[0, 2\pi]$ in the frequency domain to record the spectral information.

Further Improvement

- Can we use only a finite number points (eg., N points) to record the frequency domain of DTFT for a finite-duration signal?
- This leads to the development of the fourth type of Fourier transform, **Discrete Fourier Transform (DFT).**

Discrete Fourier Transform (DFT)

- Consider both the signal and the spectrum only within one period (N-point signals both in time and frequency domains)

DFT	$X[k]=\sum_{n=0}^{N-1} x[n]W_N^{kn},$	$k=0,1\ldots,N-1$
IDFT (inverse DFT)	$x[n]=\dfrac{1}{N}\sum_{k=0}^{N-1} X[k]W_N^{-kn},$	$n=0,1\ldots,N-1$

where $W_N = e^{-j(2\pi/N)}$, and W_N^n are the roots of the polynomial $W^N = 1$.

From Kuhn 2005

DFT in **matrix form**

Discrete Fourier Transform (DFT)

$$X_k = \sum_{i=0}^{n-1} x_i \cdot \mathrm{e}^{-2\pi \mathrm{j}\frac{ik}{n}} \qquad x_k = \frac{1}{n} \cdot \sum_{i=0}^{n-1} X_i \cdot \mathrm{e}^{2\pi \mathrm{j}\frac{ik}{n}}$$

The n-point DFT multiplies a vector with an $n \times n$ matrix

$$F_n = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \mathrm{e}^{-2\pi \mathrm{j}\frac{1}{n}} & \mathrm{e}^{-2\pi \mathrm{j}\frac{2}{n}} & \mathrm{e}^{-2\pi \mathrm{j}\frac{3}{n}} & \cdots & \mathrm{e}^{-2\pi \mathrm{j}\frac{n-1}{n}} \\ 1 & \mathrm{e}^{-2\pi \mathrm{j}\frac{2}{n}} & \mathrm{e}^{-2\pi \mathrm{j}\frac{4}{n}} & \mathrm{e}^{-2\pi \mathrm{j}\frac{6}{n}} & \cdots & \mathrm{e}^{-2\pi \mathrm{j}\frac{2(n-1)}{n}} \\ 1 & \mathrm{e}^{-2\pi \mathrm{j}\frac{3}{n}} & \mathrm{e}^{-2\pi \mathrm{j}\frac{6}{n}} & \mathrm{e}^{-2\pi \mathrm{j}\frac{9}{n}} & \cdots & \mathrm{e}^{-2\pi \mathrm{j}\frac{3(n-1)}{n}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \mathrm{e}^{-2\pi \mathrm{j}\frac{n-1}{n}} & \mathrm{e}^{-2\pi \mathrm{j}\frac{2(n-1)}{n}} & \mathrm{e}^{-2\pi \mathrm{j}\frac{3(n-1)}{n}} & \cdots & \mathrm{e}^{-2\pi \mathrm{j}\frac{(n-1)(n-1)}{n}} \end{pmatrix}$$

$$F_n \cdot \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \\ X_{n-1} \end{pmatrix}, \qquad \frac{1}{n} \cdot F_n^* \cdot \begin{pmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \\ X_{n-1} \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{pmatrix}$$

DFT pairs (more specifically)

DFT can be written in Matrix form

- Forward transform : $\mathbf{X} = \mathbf{D}_N \mathbf{x}$
- Inverse transform: $\mathbf{x} = \mathbf{D}_N^{-1} \mathbf{X}$

$$\mathbf{D}_N = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N & W_N^2 & \cdots & W_N^{(N-1)} \\ 1 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix}$$

where $W_N = e^{-j2\pi/N}$

$$\mathbf{D}_N^{-1} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N^{-1} & W_N^{-2} & \cdots & W_N^{-(N-1)} \\ 1 & W_N^{-2} & W_N^{-4} & \cdots & W_N^{-2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{-(N-1)} & W_N^{-2(N-1)} & \cdots & W_N^{-(N-1)(N-1)} \end{bmatrix} \frac{\mathbf{1}}{\boldsymbol{N}}$$

- The matrix of $\frac{1}{\sqrt{N}} D_N$ is a unitary matrix, i.e.,

$$\left(\frac{1}{\sqrt{N}} D_N\right)^* \left(\frac{1}{\sqrt{N}} D_N\right) = \left(\frac{1}{\sqrt{N}} D_N\right)\left(\frac{1}{\sqrt{N}} D_N\right)^* = I_N$$

where I_N is the $N \times N$ identity matrix.

- Hence, $\frac{1}{N} D_N D_N^* = I$, and thus $D_N^{-1} = \frac{1}{N} D_N^*$

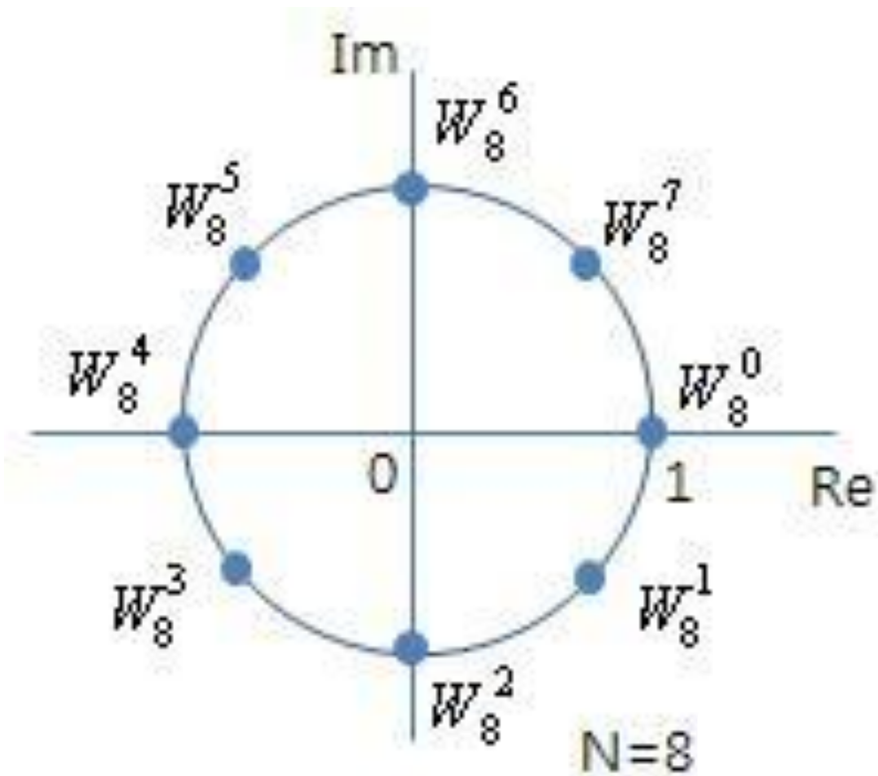
So

- The columns of D_N are orthogonal to each other.
- The rows of D_N are also orthogonal to each other.

$N = 8$

n-th Root of **1**

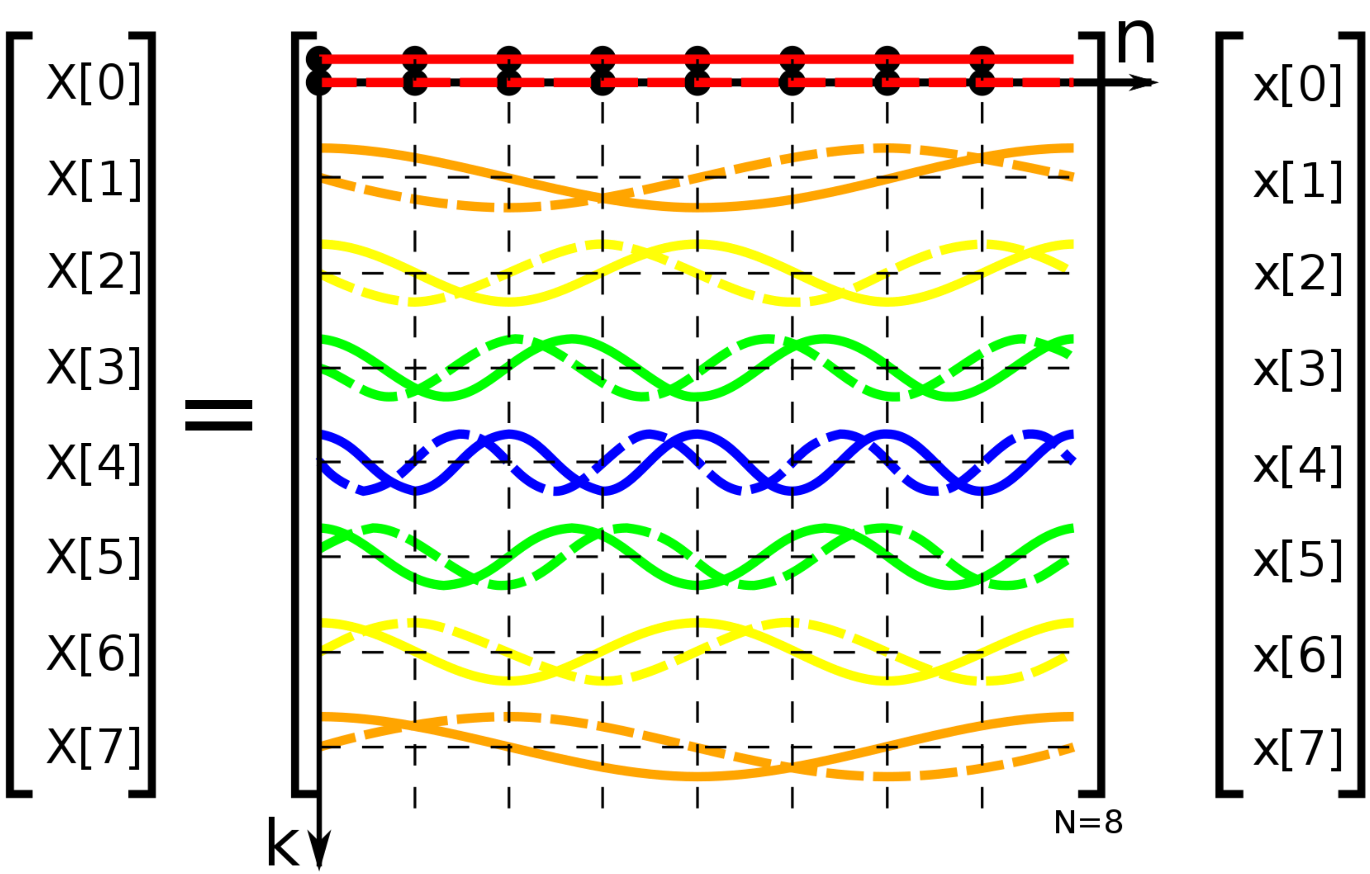
- W_N^n is the n-th root of the equation $\boldsymbol{W}^N = 1$.
- Eg., when n=8,



$N = 8$

Visualization of the DFT matrix ($N = 8$)

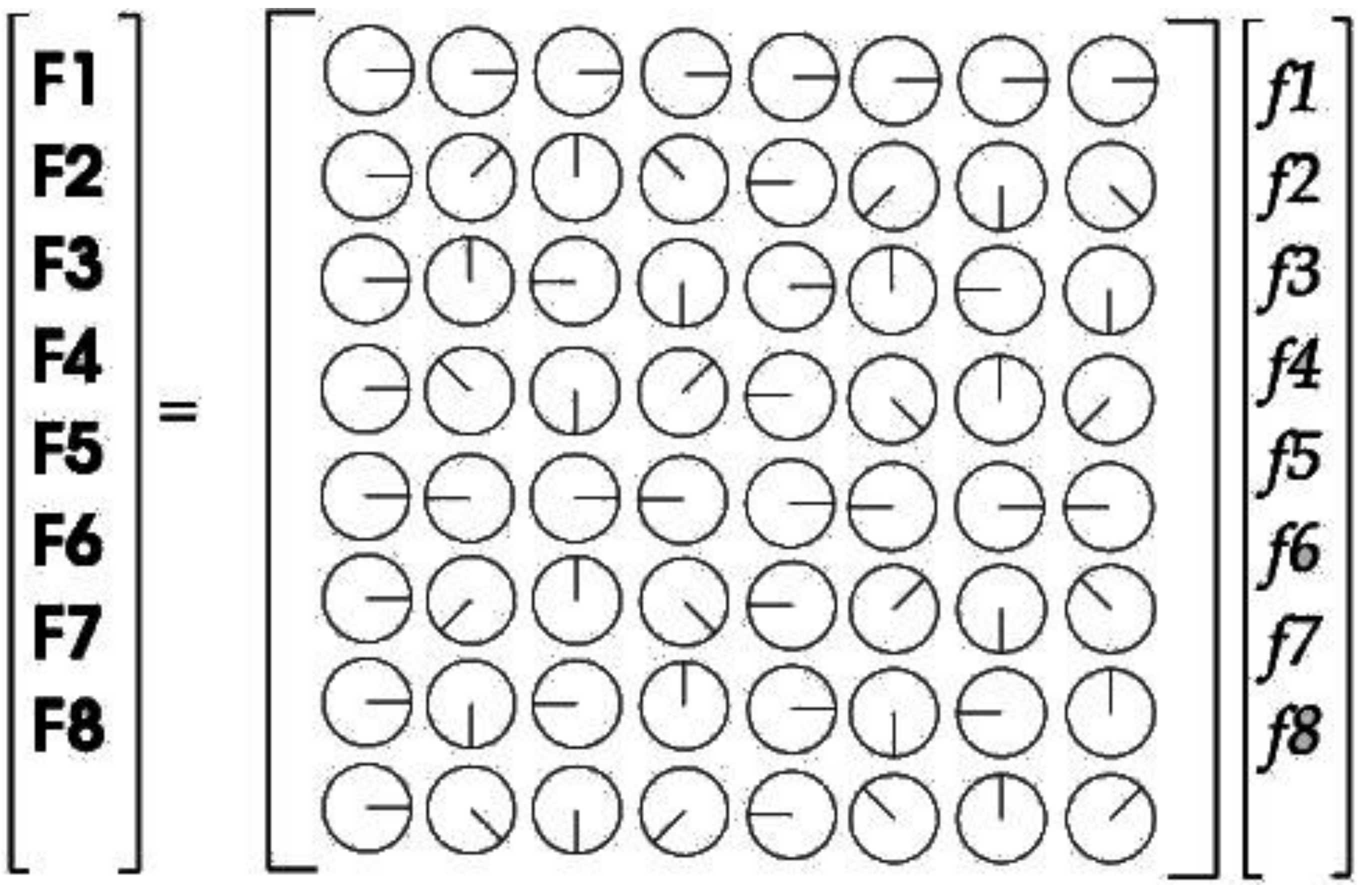
- $W_N = e^{-j(2\pi/N)} = cos\left(-\frac{2\pi}{N}\right) + jsin(-\frac{2\pi}{N})$



Because of the negative sign in the notation, the rotation in the polar-coordinate system is clockwise.

$N = 8$

Visualization of the DFT matrix ($N = 8$)



- In our definition, the rotation is clockwise.
- The illustration in the left figure is in a counterclockwise rotation due to several variations of the definition of DFT.

Relation between DFT and DTFT for finite-length signals

- Remember that the frequency domain of DTFT can be within $[0, 2\pi]$.

- **Key Property:** The frequency domain of an N-point DFT are the same as the N-point uniform samples of DTFT in $[0, 2\pi]$.

- IDFT (inverse DFT) gives us an algorithm to exactly reconstruct the N-point discrete-time signal.

Relation between DFT and DTFT for finite-length signals

- **Key Property:** The frequency domain of an N-point DFT are the same as the N-point uniform samples of DTFT in $[0, 2\pi]$.
  - IDFT gives us an algorithm to reconstruct the N-point discrete-time signal.

- Question: Can we reconstruct the DTFT spectrum (continuous in ω) from the N-point spectrum of DFT?

- $\rightarrow$ As the N-point signal in time domain can be exactly recovered from both the DFT and the DTFT spectra, we expect that the DTFT spectrum (continuous in $[0, 2\pi]$) can be exactly reconstructed by the DFT spectrum (which is discrete).

Reconstruct DTFT from DFT
(when the discrete-time signal is **finite-length**)

In the DTFT frequency domain, $X\left(e^{j\omega}\right)$, by substituting the inverse DFT into the $x(n)$, we have

$$X(e^{j\omega}) = \sum_{n=0}^{N-1} x(n)e^{-j\omega n} = \sum_{n=0}^{N-1} \underbrace{\left[\frac{1}{N}\sum_{k=0}^{N-1} X(k)e^{j(2\pi kn/N)}\right]}_{\text{Inverse DFT of } N \text{ points}} e^{-j\omega n}$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} X(k) \underbrace{\boxed{\sum_{n=0}^{N-1} e^{j(2\pi kn/N)} e^{-j\omega n}}}_{\text{a geometric sequence}}$$

Continuous DTFT spectrum

Inverse DFT of *N* points

a geometric sequence

By applying the geometric-sequence formula

$$\sum_{n=0}^{N-1} e^{j(2\pi kn/N)} e^{-j\omega n} = \frac{1 - e^{-j(\omega N - 2\pi k)}}{1 - e^{-j[\omega - (2\pi k/N)]}}$$

$$= \frac{e^{-j[(\omega N - 2\pi k)/2]}}{e^{-j[(\omega N - 2\pi k)/2N]}} \cdot \frac{\sin\left[\dfrac{\omega N - 2\pi k}{2}\right]}{\sin\left[\dfrac{\omega N - 2\pi k}{2N}\right]}$$

$$= \underbrace{\frac{\sin\left[\dfrac{\omega N - 2\pi k}{2}\right]}{\sin\left[\dfrac{\omega N - 2\pi k}{2N}\right]}}_{\text{magnitude}} e^{\underbrace{-j[\omega - (2\pi k/N)][(N-1)/2]}_{\text{phase}}}$$

So

Dirichlet Kernel

$$X(e^{j\omega}) = \frac{1}{N}\sum_{k=0}^{N-1} X(k) \underbrace{\boxed{\frac{\sin\left[\dfrac{\omega N - 2\pi k}{2}\right]}{\sin\left[\dfrac{\omega N - 2\pi k}{2N}\right]}} e^{-j[\omega-(2\pi k/N)][(N-1)/2]}}_{\text{The reconstruction formula}}$$

The reconstruction formula

From $X(k)$, we can reconstruct the spectrum of DTFT, $X(e^{jw})$ by the above interpolation.

- Hence, instead of computing the DTFT of a finite-duration signal of length N directly, we always use N-point DFT for the computation, and the DTFT spectrum can be exactly reconstructed for any $w \in [-\pi, \pi]$ or $[0, 2\pi]$.

Zero padding

- In the above, we have introduced how to reconstruct the DTFT spectrum of an length-N signal from the DFT spectrum of the same length-N signal.
- Another way to interpolate $X(e^{jw})$ from $X(k)$ is zero-padding that will be shown below.

Zero padding (cont.)

- Sometimes our purpose is for visualization or efficient maximum finding. The N samples of the DTFT spectrum is too few. We often like to see the M-point uniform samples of the DTFT of the length-N signal (with $M > N$).
- In practice, we can use a simple technique called zero-padding to achieve this purpose, which complement $M - N$ zeros in the end of the original length-N sequence and performing DFT.
- Then, the obtained M-point DFT spectrum is just the M-point uniform samples of the DTFT of the length-N signal.

Definition of DFT in Boaz Porat's Book

To explain zero-padding, we use other notations.

Remark

- There could be different notations from different articles and books.
- The Fourier transform is referred to as DTFT below.

Let the discrete-time signal $x[n]$ have finite duration, say in the range $0 \le n \le N-1$. The Fourier transform of this signal is

$$\boxed{X^{\mathrm{f}}(\theta) = \sum_{n=0}^{N-1} x[n]e^{-j\theta n}.} \qquad \textbf{DTFT} \qquad (4.1)$$

Let us sample the frequency axis using a total of N equally spaced samples in the range $[0, 2\pi)$, so the sampling interval is $2\pi/N$; in other words, we use the frequencies

$$\theta[k] = \frac{2\pi k}{N}, \quad 0 \le k \le N-1. \qquad (4.2)$$

The result is, by definition, the discrete Fourier transform. Mathematically,

$$\boxed{X^{\mathrm{d}}[k] = \{\mathcal{D}x\}[k] = \sum_{n=0}^{N-1} x[n] \exp\left(-\frac{j2\pi kn}{N}\right), \quad 0 \le k \le N-1.} \qquad \textbf{DFT} \qquad (4.3)$$

4.4 Zero Padding

The DFT of a length-N sequence is itself a length-N sequence, so it gives the frequency response of the signal at N points. Suppose we are interested in computing the frequency response at M equally spaced frequency points, where $M > N$. A simple device accomplishes this goal: We add $M - N$ zeros at the tail of the given sequence, thus forming a length-M sequence. The DFT of the new sequence has M frequency points. We prove that the values of the new DFT are indeed samples of the frequency response of the given signal at M equally spaced frequencies. Denote

$$\boxed{x_{\mathrm{a}}[n] = \begin{cases} x[n], & 0 \le n \le N-1, \\ 0, & N \le n \le M-1. \end{cases}} \tag{4.44}$$

The operation of adding zeros to the tail of a sequence is called *zero padding*. The DFT of the zero-padded sequence $x_{\mathrm{a}}[n]$ is given by

$$X_{\mathrm{a}}^{\mathrm{d}}[k] = \sum_{n=0}^{M-1} x_{\mathrm{a}}[n] \exp\left(-\frac{j2\pi kn}{M}\right) = \sum_{n=0}^{N-1} x[n] \exp\left(-\frac{j2\pi kn}{M}\right) = X^{\mathrm{f}}(\theta[k]), \tag{4.45}$$

where

$$\theta[k] = \frac{2\pi k}{M}, \quad 0 \le k \le M-1. \tag{4.46}$$

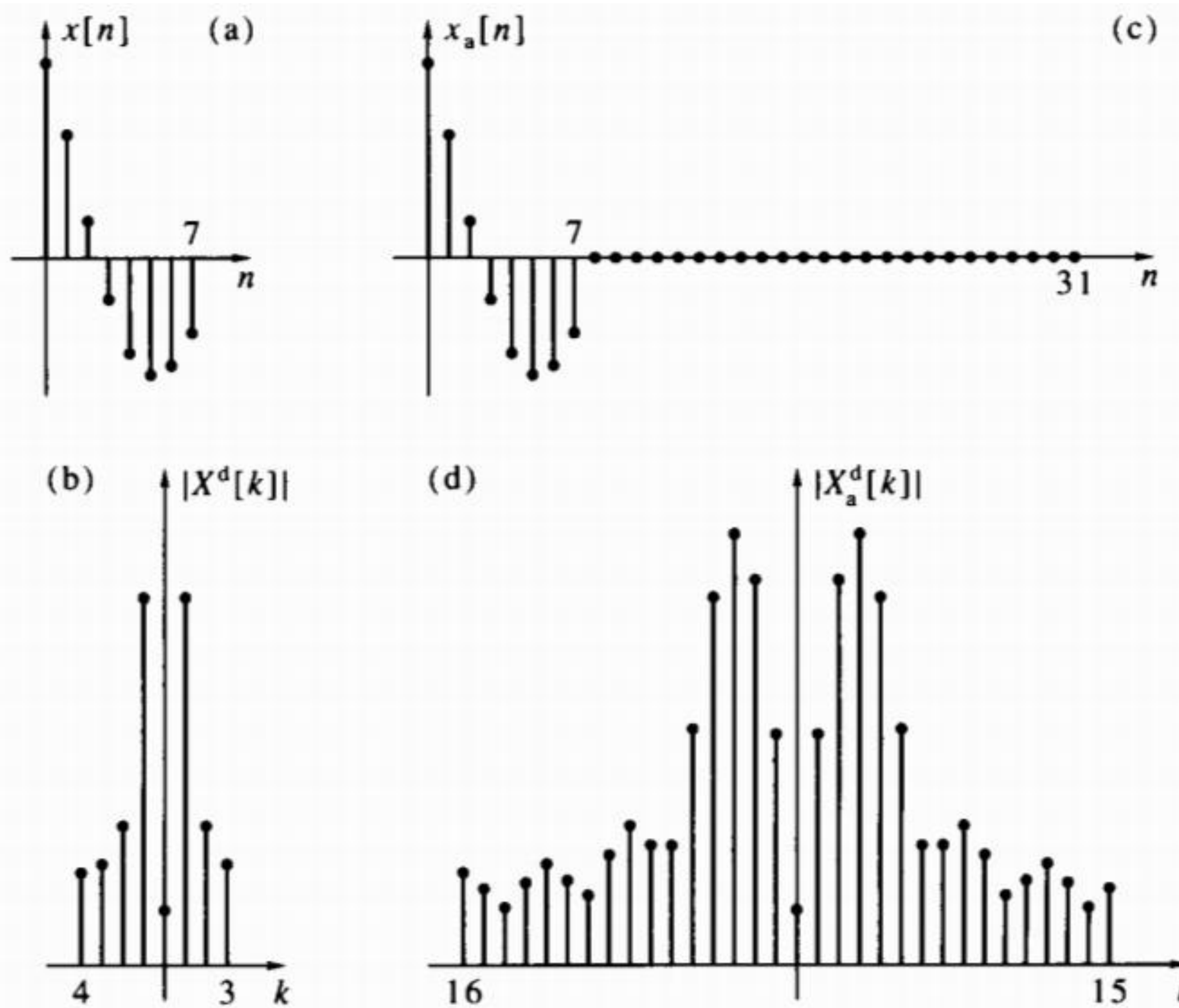


Figure 4.7 Increasing the DFT length by zero padding: (a) a signal of length 8; (b) the 8-point DFT of the signal (magnitude); (c) zero padding the signal to length 32; (d) the 32-point DFT of the zero-padded signal.