List of references in Signal Processing

1) Charles W. Therrien, "Discrete Random Signals and Statistical Signal Processing" Prentice Hall, 1992

2) J.G. Proakis and D. Manolakis, "Digital Signal Processing", MacMillan, second Edition, 1992

3) A. Oppenheim, R. Schafer, "Discrete-time Signal Processing", Prentice Hall 1989, 1999

4) J. G. Proakis et al., "Advanced Digital Signal Processing", McMillan, 1992

5) M. H. Hayes, "Statistical Digital Signal Processing and Modeling", John Wiley & Sons, Inc, 1996.

6). S. Haykin, "Adaptive Filter Theory", Prentice Hall, Third Edition, 1996

7) C.L. Nikias and A. Petropulu, "Higher-Order Spectra Analysis", Prentice Hall, 1993

8) S. Haykin, Editor, "Blind Deconvolution", Prentice Hall, 1994

9) C.L. Nikias and M. Shao, "Signal Processing with α-stable distributions and applications" New York: John Wiley & Sons, Inc. 1995.

University of Toronto
Department of Electrical and Computer Engineering

THE SAMPLING THEOREM

Question: How often must I sample an analog signal in order not to lose information from it.

Answer: If the highest frequency contained in a signal $x_a(t)$ is Ω_0 and the signal is uniformly sampled at a rate $\Omega_s \geq 2\Omega_0$, then $x_a(t)$ can be exactly recovered from its sample values using the interpolation function

$$g(t) = \frac{\sin\left[\Omega_s t/2\right]}{\Omega_s t/2}$$

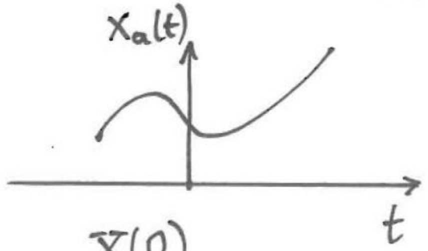
and then, $x_a(t) = \sum_{k=-\infty}^{\infty} x_a(kT)\, g(t-kT)$, where $\{x_a(kT)\}$ are the samples of $x_a(t)$ taken with rate $T = 2\pi/\Omega_s$. see.

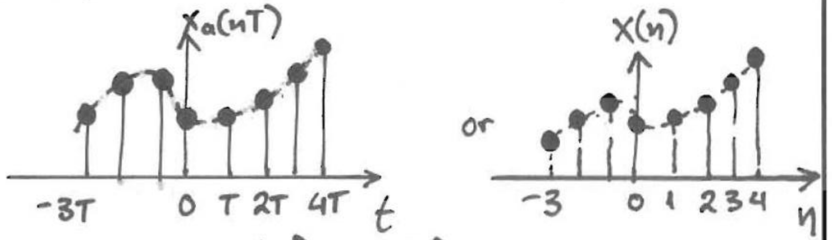
Comments:
- "Aliasing" that is spectral overlapping is caused by inadequate sampling
- Aliasing is not reversable
- Assuming no aliasing has occured $x_a(t)$ can be reconstructed via ideal LP filtering
- In practice: always choose a rate $\Omega > 2\Omega_0$; there is a need for antialiasing filtering; there are no "ideal" LP filters.

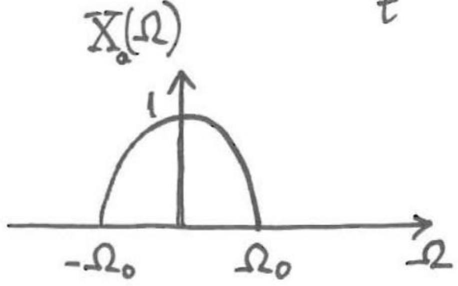
ANALOG TO DIGITAL CONVERSION (A/D)

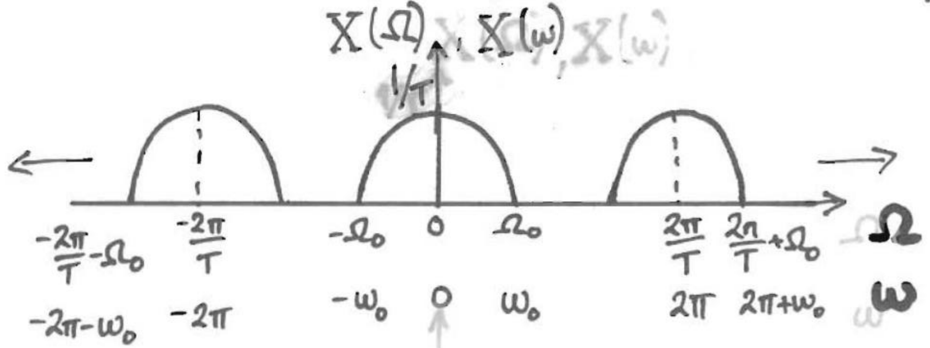
- **Ideal Sampling (uniform)**: take one sample every T seconds (sampling period, $F_s = \frac{1}{T}$ cycles/sec)

Analog signal: $x_a(t) \xrightarrow{F_s}$ Discrete signal: $x(n) = x_a(nT)$, $-\infty < n < \infty$









$$x_a(t) \xleftrightarrow{\text{F.T.}} X_a(\Omega)$$

$$x_a(nT) \xleftrightarrow{\text{DTFT}} X(\Omega) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X_a\left(\Omega + k\frac{2\pi}{T}\right)$$

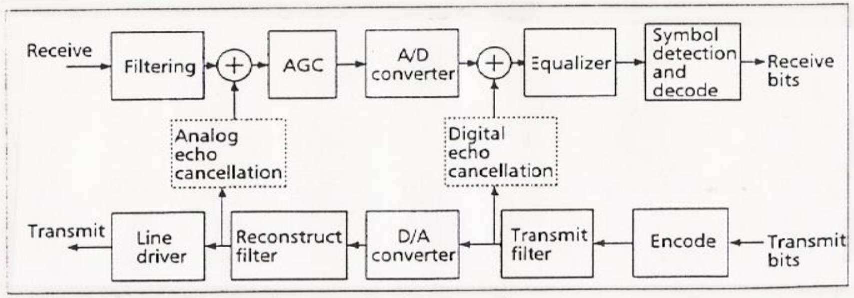
$$x(n) \xleftrightarrow{\text{DTFT}} X(\omega) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X_a(\omega + k2\pi)$$

$$\omega = \Omega T \quad , \quad f = \frac{F}{F_s}$$

ω: Normalized frequency in rads/sample

Ω: Real frequency in rads/sec

ANALOG TO DIGITAL CONVERSION (A/D)



■ **Figure** *A generic digital communication transceiver.*

* Analog and digital signal processors/units are used simultaneously.

* Since 1990 the system design approach places more emphasis on digital processing

* Multiple A/D, D/A conversions may be required

Type	Number of bits	Signal bandwidth	Sampling rate	Latency
Flash	6–8	$<f_s/2$	150–300 ms/s	>1 clock cycle
Two-step	≈ 10–12	$<f_s/2$	10–100 ms/s	>2 clock cycles
Pipeline	10–15	$<f_s/2$	1–100 ms/s	> # of stages
Algorithmic	12–14	$<f_s/2$	10 ks/s to< 1 ms/s	> # of bits
Succ. approx.	10–14	$<f_s/2$	100 ks/s to 5 ms/s	> # of bits
Lowpass Σ–Δ	14–20	20–1000 kHz	2–50 MHz	N/A
Bandpass Σ–Δ	8–15	30–1250 kHz	2–200 MHz	N/A

■ **Table 1.** *Typical state-of-the-art performance of CMOS- and BiCMOS-based A/D converters.*

Nyquist rate converters: quantize input samples every $1/f_s$ sec, ($f_s >$ Nyquist rate); generate one digital sample from a single analog sample (good performance in the 10-12 bit/sample range)

Noise shaping converters: oversample analog signal; generate one digital sample by weighting many input analog samples; (Good performance in the 13-15 bit range)

* Digital converters for signals with bandwidths greater that 100 KHz are commonly used.

"Data Converters for Communication Systems"
From: IEEE Communications Magazine, October 1998, page 113

THE CONCEPT OF FREQUENCY FOR DISCRETE TIME SIGNALS

$x(n) = e^{j\omega n}$, $n = 0, \pm 1, \pm 2, \ldots$

- ω: rads/sample
- periodic only iff $f = \frac{\omega}{2\pi}$ is rational that is $f = \frac{K}{N}$ where K, N co-prime integers then $x(n) = x(n + \ell N)$
- Let $\omega_1 = \omega_2 + 2K\pi$. Then $e^{j\omega_1 n} = e^{j\omega_2 n}$
  Thus distinct discrete exponentials (complex) can be obtained only in intervals of 2π
  i.e., $-\pi \leq \omega \leq \pi$ or $0 \leq \omega \leq 2\pi$ ......
  $-\frac{1}{2} \leq f \leq \frac{1}{2}$ $\quad 0 \leq f \leq 1$
- The highest rate of oscillation for a discrete complex exponential is achieved for $\omega = \pi$ or $f = 1/2$

$x_K(n) = e^{jK\frac{2\pi}{N}n}$, $n = 0, \pm 1, \pm 2, \ldots$ N: integer

N: integer

- periodic in both K and n with period N samples that is
  $$x_{K+N}(n+N) = x_K(n)$$
- Thus, there are only N distinct periodic exponentials (complex) of period N
- Highest oscillation rate for $K = \frac{N}{2}$
- Important relation.
  $$\frac{1}{N}\sum_{n=0}^{N-1} e^{jK\frac{2\pi}{N}n} = \begin{cases} 1 & \text{for } K = 0, \pm N, \pm 2N \ldots \\ 0 & \text{otherwise.} \end{cases}$$

DTFT, DTFS, DFT

Given a discrete signal $x(n)$, $n=0,\pm1,\pm2,\ldots$ form $\tilde{x}(n) = \sum_{\ell=-\infty}^{\infty} x(n+\ell N)$

Ⓘ Discrete time Fourier transform pair (DTFT)

$$\left[\begin{aligned} X(\omega) &= \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n} \\ x(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega) e^{j\omega n} \end{aligned}\right]$$

$x(n)$ discrete non-periodic

$X(\omega)$ continuous periodic (2π)

Ⓘ Discrete time Fourier Series pair. (DTFS or DFS)

$$\left[\begin{aligned} \tilde{X}(k) &= \sum_{n=0}^{N-1} \tilde{x}(n) e^{-j\frac{2\pi kn}{N}} \\ \tilde{x}(n) &= \frac{1}{N}\sum_{k=0}^{N-1} \tilde{X}(k) e^{j\frac{2\pi kn}{N}} \end{aligned}\right]$$

for $n,k = 0,1,\ldots,N-1$ or over $\langle N\rangle$

$\tilde{x}(n)$, $\tilde{X}(k)$ discrete and periodic with period N samples

- N is known a-priori or can be chosen arbitrarily
- Sufficient condition for existence $\sum_{n=-\infty}^{\infty} |x(n)|^2 < \infty$

N-DFT (Discrete Fourier Transform)

- Given a discrete signal $x(n)$, $n=0,\pm1,\pm2,\dots$ with DTFT $X(\omega)$

1) Form the signal $\tilde{x}(n) = \sum_{\ell=-\infty}^{\infty} x(n+\ell N)$, N: chosen arbitrarily

2) Form the signal $\tilde{X}(k) = X(\omega)\Big|_{\omega=\frac{2\pi k}{N}}$, $k=0,\pm1,\pm2,\dots$

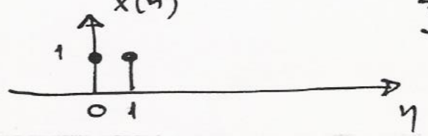
Then,

N-DFT pair

$$\left[\begin{aligned} X(k) &= \sum_{n=\langle N\rangle} \tilde{x}(n)\, e^{-j\frac{2\pi k n}{N}} \\ x(n) &= \frac{1}{N}\sum_{k=\langle N\rangle} \tilde{X}(k)\, e^{j\frac{2\pi k n}{N}} \end{aligned}\right] \xRightarrow[\text{length} \le N]{x(n)\text{ of}} \left[\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n)\, e^{-j\frac{2\pi k n}{N}} \\ x(n) &= \frac{1}{N}\sum_{k=0}^{N-1} X(k)\, e^{j\frac{2\pi k n}{N}} \end{aligned}\right]$$

for $n,k=0,1,\dots,N-1$ or any period $\langle N\rangle$

- The N-DFT maps one period of $\sum_{\ell=-\infty}^{\infty} x(n+\ell N)$ to one period of $X\left(\omega=\frac{2\pi k}{N}\right)$
- If we do not take sufficient number of (samples per period) in the DTFT domain, then "time aliasing" occurs in time domain (dual of sampling theorem)

Ex. Let 

$$X(\omega) = 2\cos\frac{\omega}{2}\, e^{-j\omega/2} = 1 + e^{-j\omega}$$

Evaluate: 1-DFT, 2-DFT, 3-DFT, 8-DFT

A) Working in frequency domain.

$$|X(k)| = X\left(\omega = \frac{2\pi k}{N}\right), \quad k = 0, \dots, N-1$$

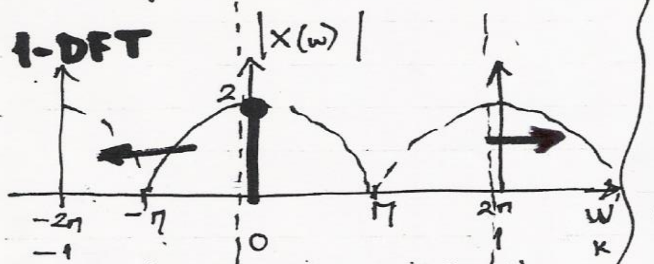
B) Time domain $\xrightarrow{N\text{-DFT}}$ Frequency domain

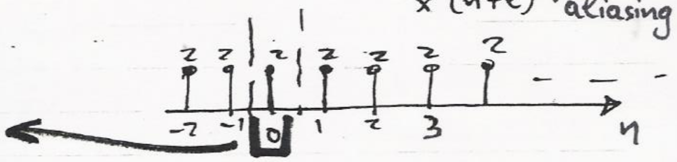
Form $\tilde{x}(n) = \sum_{\ell=-\infty}^{\infty} x(n + \ell N)$

Then

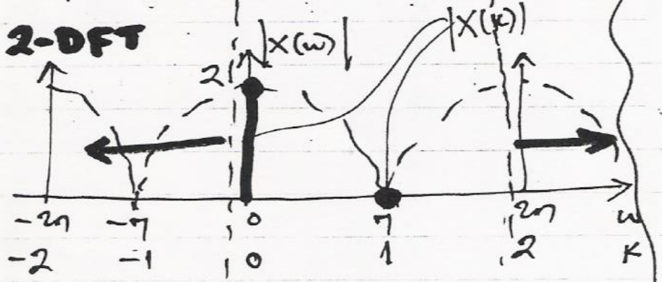
$$X(k) = \sum_{n=0}^{N-1} \tilde{x}(n)\, e^{-j\frac{2\pi k n}{N}}.$$

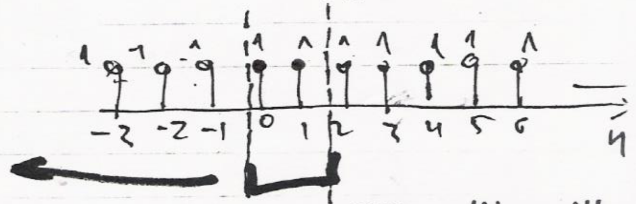
1-DFT



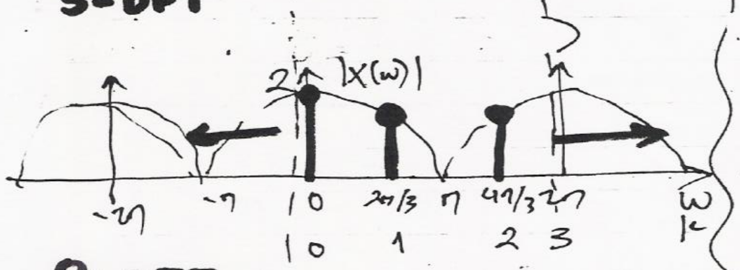


2-DFT

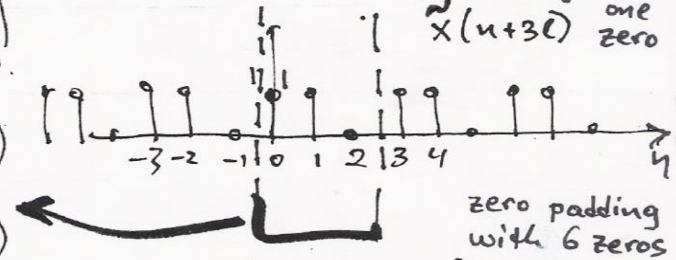




3-DFT

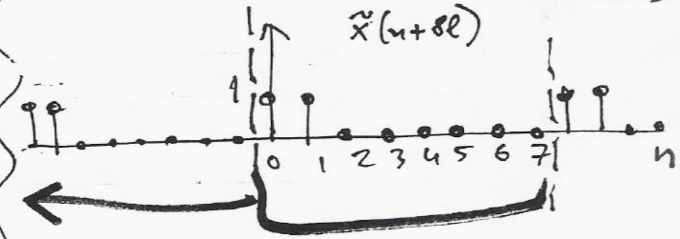


zero padding with one zero



8-DFT

zero padding with 6 zeros



ZERO PADDING – DIGITAL INTERPOLATION

Given $x(n)$ of length less or equal to N and its N-DFT $X(k)$, $k=0,1,\dots,N-1$

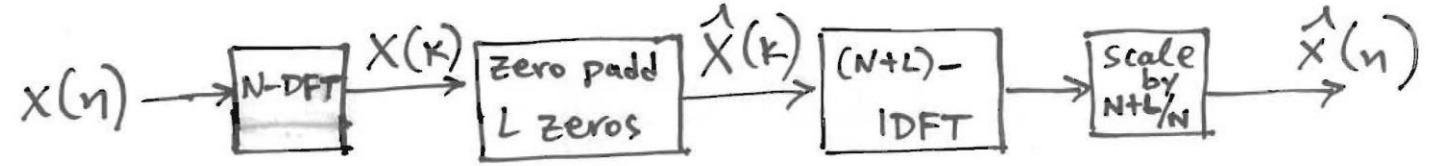
- Zero-padding increases the effective period of the N-DFT by introducing zeros in the "stop-band" of either the signal or the discrete Fourier transform

- By zero padding $x(n)$ with $K \cdot N$ zeros and taking $(K+1)N$-DFT, then K new values are interpolated between (any) two values of the N-DFT

- By zero padding $X(k)$ with $K \cdot N$ zeros (between the samples $k=N/2, N/2+1$) and taking $(K+1)N$-IDFT and multiplying by $(K+1)$, K new values are interpolated between any two values of $x(n)$.

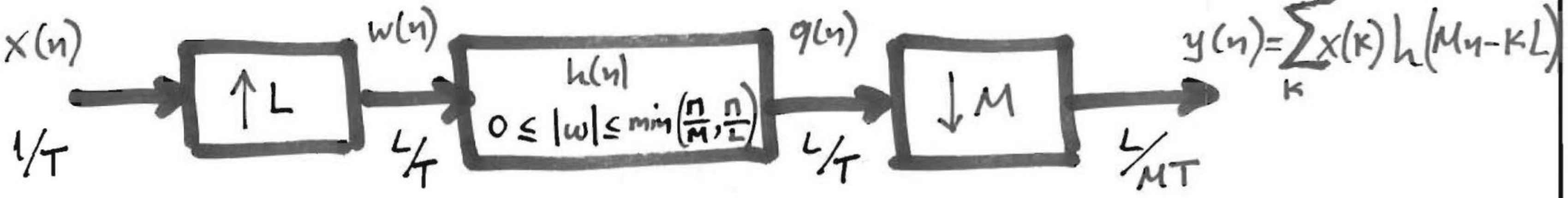
i.e

INTERPOLATION OF L values in $x(n)$:



MULTIRATE SYSTEMS (3)

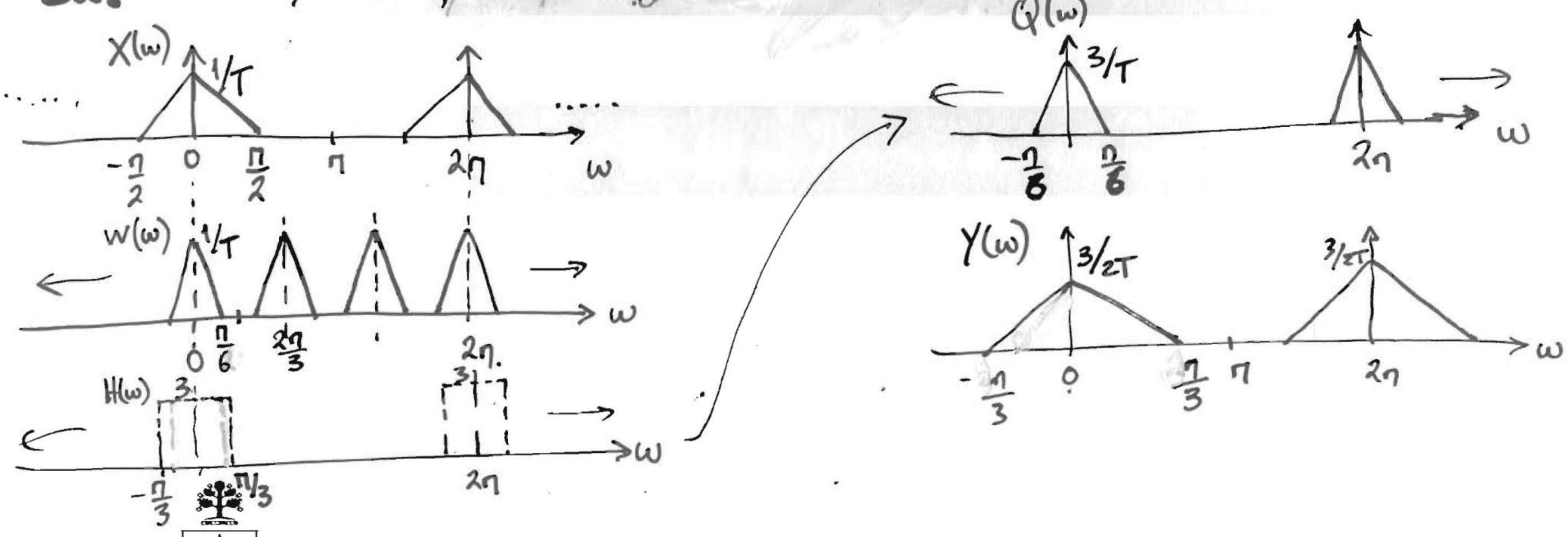
(iii) Change of sampling rate by $\frac{L}{M}$: noninteger



$$Y(\omega) = \frac{1}{M} \sum_{i=0}^{M-1} H\left(\frac{\omega - 2\pi i}{M}\right) X\left(\frac{\omega - 2\pi i}{M} L\right) \xlongequal{\text{no-aliasing}} \frac{1}{M} H\left(\frac{\omega}{M}\right) X\left(\frac{\omega L}{M}\right)$$

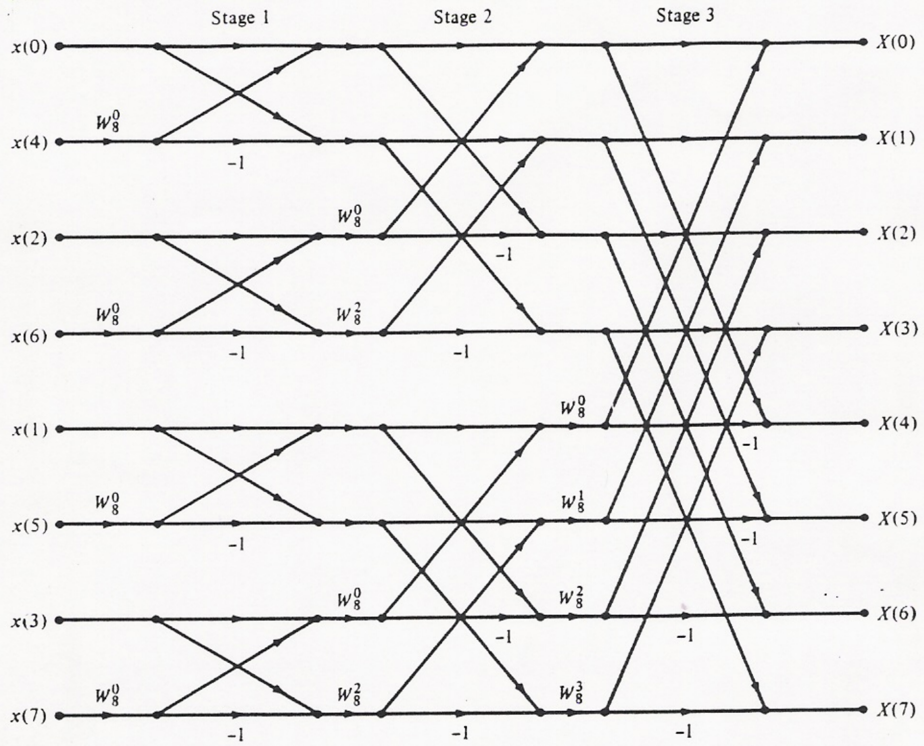
in range $0 \le |\omega| \le \pi$

Ex. $L=3$, $M=2$, no-aliasing



Radix 2 ($N=2^q$) Decimation in time FFT

Input in bit reversed order



Output in normal order

FIGURE Eight-point decimation-in-time FFT algorithm.

Decimation in time FFT is based on the calculation of succesively smaller DFT's by decomposing the signal into successively smaller sequences

<u>In general for $N=2^q$ - FFT.</u> (Decimation in time).

1) $q = \log_2 N$ stages , $N/2$ butterflies / stage

2) m-th stage , 2^{m-1} different butterflies $\left\{ \begin{array}{l} \text{Multiply lower branches of each} \\ \text{butterfly } W_N^{\ell N/2^m} \quad , \ell=0,\dots,2^{m-1}-1 \end{array} \right.$

3) # of $\otimes_c$: $1/\text{butterfly} \times \frac{N}{2}$ butt/stage $\times \log_2 N$ stages $= \frac{N}{2}\log_2 N$

of $\oplus_c$: $2/$ " $\times \frac{N}{2}$ " $\times$ " " $= N\log_2 N$

of storage locations: $2/\text{butterfly} \times \frac{N}{2}$ butt. $= \binom{\text{in place}}{\text{computation}} = N$

4) $2^{\log_2 N-1}$ exponentials stored in look up table

5) $G(N) = \dfrac{\#\text{ of FFT } \otimes}{\#\text{ of DFT } \otimes} = \dfrac{1}{2N}\log_2 N \ll 1$. For ex. $G(8) = \frac{3}{16}$

$G(4096) \simeq 0.0015$

Linear Algebra Relations

Note Title 02/09/2005

→ Vector $\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ $(n \times 1)$ real or complex

→ Transpose: $\underline{x}^T = [x_1, \dots, x_n]$ $(1 \times n)$

→ Hermitian transpose: $\underline{x}^H = (\underline{x}^T)^* = [x_1^*, \dots, x_n^*]$

→ Magnitude of a vector (norm or distance)

Euclidean L_2 norm: $\|\underline{x}\|_2 = \sqrt{\sum_{i=1}^{N} |x_i|^2}$

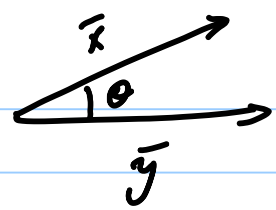
L_1 norm: $\|\underline{x}\|_1 = \sum_{i=1}^{N} |x_i|$

L_∞ norm: $\|\underline{x}\|_\infty = \max_i |x_i|$

→ distance between vectors $d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|$

→ Inner product: $\langle \underline{x}, \underline{y} \rangle = \underline{x}^H \cdot \underline{y} = \sum_{i=1}^{N} x_i^* \cdot y_i$

for Euclidean space: $\langle \underline{x}, \underline{y} \rangle = \|x\| \|y\| \cos\theta$



unit norm vectors $\dfrac{\underline{x}}{\|\underline{x}\|}, \dfrac{\underline{y}}{\|\underline{y}\|}$

→ $\underline{x}, \underline{y}$: orthogonal, $\langle \underline{x}, \underline{y} \rangle = 0$

→ $\underline{x}, \underline{y}$: linearly independed if

$$a_1 \underline{x} + a_2 \underline{y} = 0 \longrightarrow a_1 = a_2 = 0$$

(in this case $\underline{x}, \underline{y}$ can be seen as the generating vectors of a 2-D space where each vector can be obtained by linear combination of $\underline{x}$ and $\underline{y}$).

Ex: Let $x[n] \longrightarrow \boxed{\text{FIR, } h[n]} \longrightarrow y[n]$

$$y[n] = \sum_{i=0}^{N-1} h[i]\, x[n-i] = \underline{h}^T \underline{x}[n]$$

where: $\underline{h}^T = [h_1, \dots, h_N]$, $\underline{x}^T[n] = [x(n), \dots, x[n-N+1]]$

Now, given a matrix $\underline{A}$: $n \times m$

→ square matrix $\underline{A}$: $n \times n$ elements

→ symmetric square matrix: $\underline{A} = \underline{A}^T$

→ Hermitian square matrix: $\underline{A} = \underline{A}^H$

Also: $(A+B)^H = \underline{A}^H + \underline{B}^H$

$(\underline{A}^H)^H = \underline{A}$

$(\underline{A}\underline{B})^H = \underline{B}^H \underline{A}^H$

→ Rank of a matrix: $\rho(\underline{A}) \leq \min(m,n)$ is the number of linearly independent columns!

Note: $\rho(\underline{A}) = \rho(\underline{A}^H\underline{A}) = \rho(\underline{A}\underline{A}^H)$

→ Inverse A^{-1} of a square matrix $\underline{A}$: $\underline{A}^{-1}\underline{A} = \underline{I}$

where $\underline{I} = \begin{bmatrix} 1 & & \underline{0} \\ & 1 & \\ \underline{0} & & \ddots 1 \end{bmatrix}$, $n \times n$ identity matrix

→ Determinant of a matrix $\underline{A}$ $(n \times n)$

$$\det(\underline{A}) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(\underline{A}_{ij})$$

where A_{ij} is the $(n-1)\times(n-1)$ matrix obtained by deleting the i-th row and j-th column of $\underline{A}$

→ An $n \times n$ matrix $\underline{A}$ is invertible iff $\det(\underline{A}) \neq 0$

Also,
$$\det(\underline{A}\underline{B}) = \det(\underline{A})\det(\underline{B})$$
$$\det(\underline{A}^T) = \det(\underline{A})$$
$$\det(a\underline{A}) = a^n \det(\underline{A})$$
$$\det(\underline{A}^{-1}) = 1/\det(\underline{A})$$

→ trace: $\mathrm{tr}(\underline{A}) = \sum_{i=1}^{n} a_{ii}$

→ Linear equations: $\underline{A}\,\underline{x} = \underline{b} \Rightarrow \underline{x} = \underline{A}^{-1}\underline{b}$

(if $\underline{A}$: $n \times m$, $n < m$ $\quad \underline{x} = \underline{A}^H(\underline{A}\underline{A}^H)^{-1}\underline{b}$: minimum norm solution

* The matrix $\underline{A}^H(\underline{A}\,\underline{A}^H)^{-1}$ is called the "pseudoinverse"

* Triangular matrix (upper or lower)

$$\det(\underline{A}) = \prod_{i=1}^{n} a_{ii}$$

* Symmetric matrix $\underline{A} = \underline{A}^T$

* Toeplitz matrix: All elements along each of the diagonals have the same value

&

etc.

* $\underline{A}$: $n \times n$ is orthogonal if $\underline{A}^T\underline{A} = \underline{I} \Rightarrow \underline{A}^{-1} = \underline{A}^T$

Thus, if $\underline{A} = [\underline{a}_1\ \underline{a}_2 \ldots \underline{a}_n]$ then $\underline{a}_i^T \underline{a}_j = \delta(i-j)$

* $\underline{A}$: $n \times n$ is unitary $\underline{A}^{-1} = \underline{A}^H$
(orthogonal & complex)

Quadratic form of a real symmetric or Hermitian matrix A: $n \times n$

$$Q_A(\underline{x}) = \underline{x}^T \underline{A} \underline{x} = \sum_{i,j=1}^{n} \sum^{n} x_i a_{ij} x_j$$

$$Q_A(\underline{x}) = \underline{x}^H \underline{A} \underline{x} = \sum \sum_{i,j=1}^{n} x_i^* a_{ij}' x_j$$

* If $Q_A(\underline{x}) > 0$ then $\underline{A}$ is positive definite ($\underline{A} > 0$) for all $\underline{x} \neq 0$

If $Q_A(\underline{x}) \geq 0$ then $\underline{A}$ is positive semidefinite

* Eigenvalues & Eigenvectors

Solve: $\underline{A} \underline{v} = \lambda \underline{v} \Rightarrow (\underline{A} - \lambda \underline{I}) \underline{v} = 0$

non-trivial solution: $\det(\underline{A} - \lambda \underline{I}) = 0$

Eigenvalues: $\lambda_1, \lambda_2, \ldots, \lambda_n$

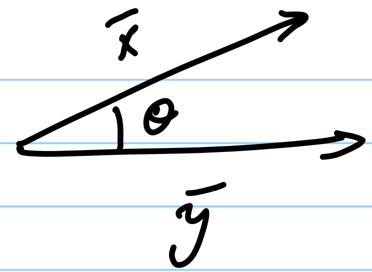
Eigenvectors: $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$

* $\underline{A}$: $n \times n$ singular matrix $\rightarrow$ $\underline{A}\,\underline{v} = 0$ and $\lambda = 0$ is an eigenvalue of $\underline{A}$
* Eigenvectors corr. to distinct eigenvalues are linearly independent
* The eigenvalues of a Hermitian matrix are real
* $\underline{A}$ Hermitian and positive definite $\Leftrightarrow$ all $\lambda_i > 0$
* $\det(\underline{A}) = \prod_{i=1}^{n} \lambda_i$, $tr(\underline{A}) = \sum_{i=1}^{n} \lambda_i$
* $\underline{A}$: Hermitian $\Leftrightarrow$ if $\lambda_i \neq \lambda_j$ then $\langle \underline{v}_i, \underline{v}_j \rangle = 0$
* Eigenvalue decomposition:

Let $\underline{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$

$\underline{V} = [\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n]$

Then, $\underline{A} = \underline{V}\,\underline{\Lambda}\,\underline{V}^H = \sum_{i=1}^{n} \lambda_i \underbrace{\underline{v}_i \underline{v}_i^H}_{\text{Rank 1 Matrix}}$



unit norm vectors

$$\frac{\underline{x}}{\|\underline{x}\|}, \quad \frac{\underline{y}}{\|\underline{y}\|}$$

$\rightarrow$ $\underline{x}, \underline{y}$: orthogonal, $\langle \underline{x}, \underline{y} \rangle = 0$

$\rightarrow$ $\underline{x}, \underline{y}$: linearly independed if

$$a_1 \underline{x} + a_2 \underline{y} = 0 \longrightarrow a_1 = a_2 = 0$$

(in this case $\underline{x}, \underline{y}$ can be seen as the generating vectors of a 2-D space where each vector can be obtained by linear combination of $\underline{x}$ and $\underline{y}$).

: Let $x[n] \longrightarrow \boxed{\text{FIR}, h[n]} \longrightarrow y[n]$

$$y[n] = \sum_{i=0}^{N-1} h[i]\, x[n-i] = \underline{h}^T \underline{x}[n]$$

where: $\underline{h}^T = [h_1, \ldots, h_N]$, $\underline{x}^T[n] = [x[n], \ldots, x[n-N+1]]$

Optimization Theory / Gradient of a scalar

Note Title 02/09/2005

* Let $f(x)$ be a scalar function of a real variable x. Then,

if $f(x)$ is differentiable local & global minima satisfy the conditions

$$\frac{d f(x)}{dx} = 0 \quad \text{and} \quad \frac{d^2 f(x)}{dx^2} > 0$$

if $f(x)$ convex: global minimum.

* Let $f(z)$ be a scalar function of a complex variable z. Then,

if $f(z)$ is differentiable proceed as before

However, in most practical situations $f(z)$ may not be differentiable, being a function of both z and z^*.

Ex: $f(z) = |z|^2 = z z^*$

Solution:

1. Express $f(z) = f(x+jy)$ in terms of real and imaginary parts. Then, minimize w.r.t. x and y

or

2. Treat z and z^* as independent variables. Then, minimize with r.t. both z and z^*.

Ex: to find global & local minima of $|z|^2$

take $\frac{d}{dz}|z|^2 = z^* = 0$, $\frac{d}{dz^*}|z|^2 = z = 0$

* Note: If $f(z)$ is real function of z and z^* it is sufficient to minimize w.r.t. either z or z^* only. (as in above example)

* Note: The justification of the solution 2 is basically based on the observation that the obtained results from such a treatment provide meaningful solutions!!

* Now let $f(\underline{x})$ be a scalar function of a vector $\underline{x}$

Then,

$$\nabla_{\underline{x}} f(\underline{x}) = \frac{d}{d\underline{x}} f(\underline{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(\underline{x}) \\ \vdots \\ \frac{\partial}{\partial x_n} f(\underline{x}) \end{bmatrix}, \quad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$\underline{x}$: minimum or maximum : $\nabla_{\underline{x}} f(\underline{x}) = 0$
(necessary condition)

$\underline{x}$: minimum : Hessian $\underline{H}_x > 0$ (positive definite)

where: $\{\underline{H}_x\}(i,j) = \dfrac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j}$

* Similarly, if $f(\underline{z})$ is a scalar function of a complex vector z, treat z and z^* as independent and proceed as before

Note: If $f(\underline{z}) = f(\underline{z}, \underline{z}^*)$ is real the the stationary points (max or min) are the solutions of the equation $\nabla_{\underline{z}^*} f(\underline{z}, \underline{z}) = 0$

Example:

Let $\underline{z} = [z_1, \ldots, z_n]$ be a complex vector

$\underline{R}$: be a positive definite Hermitian matrix

$\underline{a}$: be a given complex vector

find $\underline{z}$ that minimizes $\underline{z}^H \underline{R}\,\underline{z}$ s.t.c. $\underline{z}^H \underline{a} = 1$

Solution: (possible approach)

Use the Lagrange multiplier λ and minimize the unconstrained function:

real scalar $\rightarrow$ $$Q(\underline{z}, \lambda) = \frac{1}{2}\underline{z}^H \underline{R}\,\underline{z} + \lambda(1 - \underline{z}^H \underline{a})$$

thus:
$$\nabla_{z^*} Q(\underline{z}, \lambda) = \underline{R}\,\underline{z} - \lambda \underline{a} = 0 \qquad (1)$$

$$\Rightarrow \quad \underline{z} = \lambda \underline{R}^{-1} \underline{a}$$

Also,
$$\frac{\partial Q(\underline{z}, \lambda)}{\partial \lambda} = 1 - \underline{z}^H \underline{a} = 0 \qquad (2)$$

(1) $\rightarrow$ (2) :
$$\lambda = \frac{1}{\underline{a}^H \underline{R}^{-1} \underline{a}} \qquad (3)$$

(3) $\rightarrow$ (1) :
$$\underline{z} = \frac{\underline{R}^{-1} \underline{a}}{\underline{a}^H \underline{R}^{-1} \underline{a}} \qquad (4)$$

Q : scalar, a, b : real vectors, $\underline{B}$: real matrix $(N \times N)$

$$\begin{cases} \nabla_{\underline{a}} (\underline{b}^T \underline{a}) = \nabla_{\underline{a}} (\underline{a}^T \underline{b}) = \underline{b} \\ \nabla_{\underline{a}} (\underline{a}^T \underline{B}\, \underline{a}) = (\underline{B} + \underline{B}^T)\, \underline{a} \end{cases}$$

Q : scalar, a, b : complex vectors, $\underline{B}$: complex Matrix

$$\begin{cases} \nabla_{\underline{a}} Q = \frac{1}{2} \left(\nabla_{\underline{a}_r} Q - j \nabla_{\underline{a}_i} Q \right) \\ \nabla_{\underline{a}^*} Q = \frac{1}{2} \left(\nabla_{\underline{a}_r} Q + j \nabla_{\underline{a}_i} Q \right) \\ \nabla_{\underline{a}} (\underline{b}^H \underline{a}) = 0 \\ \nabla_{\underline{a}^*} (\underline{a}^H \underline{b}) = \underline{b} \\ \nabla_{\underline{a}} (\underline{a}^H \underline{B}\, \underline{a}) = \underline{B}^T \underline{a}^* \\ \nabla_{\underline{a}^*} (\underline{a}^H \underline{B}\, \underline{a}) = \underline{B}\, \underline{a} \end{cases}$$