Ψηφιακή Επεξεργασία Ήχου
Μάθημα 3: Μετάβαση στο πεδίο Συχνότητας
(για χρονικά αμετάβλητα φάσματα)

Π.Μ.Σ. «Τεχνολογίες Ήχου και Μουσικής»

Δρ. Χρυσούλα Αλεξανδράκη

Τμήμα Μουσικής Τεχνολογία και Ακουστικής

Ελληνικό Μεσογειακό Πανεπιστήμιο

Fourier Transform (αναλογικό σήμα)

- Ο Μετασχηματισμός Fourier ενός συνεχούς μη-περιοδικού σήματος ορίζεται ως ανάλυση σήματος στο πεδίο της συχνότητας:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\,dt$$

- Ο αντίστροφος του ορίζεται ως σύνθεση σήματος στο πεδίο του χρόνου:

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\,df$$

Discrete Time Fourier Transform (DTFT)

- Ο **Μετασχηματισμός Fourier Διακριτού Χρόνου** ενός σήματος διακριτού χρόνου **x(n)** είναι η αναπαράσταση του σήματος αυτού ως συνδυασμού μιγαδικών εκθετικών ακολουθιών της μορφής $\{e^{-j\omega n}\}$, οπού **ω** μεταβλητή, γνωστή και ως (κυκλική) συχνότητα.

- Ο DTFT της ακολουθίας x(n) ορίζεται ως:
  - $X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}$

- Και ο αντίστροφός του (Inverse - DTFT), ως:
  - $x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} \, d\omega$

Χαρακτηρισμός

- Αποδεικνύεται ότι:
  - O DTFT μιας ακολουθίας x(n), εφόσον υπάρχει, είναι **μοναδικός**.
  - Πρόκειται για μια **μιγαδική**, **συνεχή** και **περιοδική συνάρτηση του ω με περίοδο 2π.**
    - $X(e^{j\omega}) = X(e^{j(\omega+2\kappa\pi)})$, $f(x) = f(x+kT)$, η περίοδος
- Σημειογραφία:
  - Καρτεσιανές Συν/νες: $\boldsymbol{X(e^{j\omega})} = \boldsymbol{X_r(e^{j\omega})}$+j $\boldsymbol{X_{i(e^{j\omega})}}$ **(z = a +j*b), όπου j =** $\sqrt{-1}$
  - Πολικές Συν/νες: $\boldsymbol{X(e^{j\omega})} = |\boldsymbol{X(e^{j\omega})}| e^{j\theta(\omega)}$ (z = |z|e^{φ}) με:
    - Ταυτότητα Euler: $e^{j\varphi}$ = cosφ +j sinφ
    - $|\boldsymbol{X(e^{j\omega})}| = \sqrt{\boldsymbol{X_{r(e^{j\omega})}}^2 + \boldsymbol{X_{i(e^{j\omega})}}^2}$, (|z| = $\sqrt{a^2+b^2}$) λέγεται ***φάσμα μέτρου (magnitude spectrum)***
    - $\theta(\omega) = arctan\frac{X_{i(e^{j\omega})}}{X_{r(e^{j\omega})}}$ = arctan$\frac{b}{a}$, λέγεται ***φάσμα φάσης (phase spectrum)***
- Το φάσμα μέτρου και το φάσμα φάσης είναι πραγματικές συναρτήσεις
  - Φάσμα μέτρου: το συχνοτικό περιεχόμενο ενός σήματος
  - Φάσμα φάσης: η χρονική στιγμή εμφάνισης των συχνοτήτων

Απόδειξη Περιοδικότητας του DTFT

- **DTFT Ορισμός:**
  - $X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}$
- Περιοδικότητα
  - $X(e^{j\omega}) = X(e^{j(\omega+2k\pi)})$
  - $X(e^{j(\omega+2k\pi)}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j(\omega+2k\pi)n} = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}\ e^{-j2k\pi n}$
    - **Όμως:** $e^{-j2k\pi n}$ = cos(-2kπn) + j sin(-2kπn) = 1+j*0 = 1
  - **Άρα:** $X(e^{j(\omega+2k\pi)}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n} = X(e^{j\omega})$
- **Θυμηθείτε ότι στα σήματα διακριτού χρόνου η ω_0 = 2π αντιστοιχεί στη συχνότητα δειγματοληψίας fs**

Σύγκλιση του DTFT

- Καθώς ο DTFT είναι άθροισμα απείρων όρων, ενδέχεται για κάποια σήματα να απειρίζεται
- Αποδεικνύεται ότι:
  - Ο DTFT συγκλίνει (έχει πεπερασμένη τιμή), **αν και μόνο αν**
    - $\sum_{n=-\infty}^{+\infty} |x(n)| < \infty$
  - Δηλαδή εφόσον το κατ' απόλυτη τιμή άθροισμα των δειγμάτων του σήματος είναι πεπερασμένο

Άσκηση

- Να βρεθεί ο DTFT των σημάτων:
  a) $x(n) = \delta(n)$
  b) $x(n) = \delta(n-n_0)$
  - Όπου δ(n) η κρουστική ακολουθία
- $X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} = \sum_{n=-\infty}^{\infty} \delta(n)e^{-j\omega n} = 1\ e^{-j\omega 0} = 1$
- $X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} = \sum_{n=-\infty}^{\infty} \delta(n-n_0)e^{-j\omega n} = 1\ e^{-j\omega n_0} = e^{-j\omega n_0}$

$= \cos(-\omega n_0) + j\sin(-\omega n_0) = \cos(\omega n_0) - j\sin(\omega n_0)$

  - Ποιο είναι το φάσμα μέτρου?
  - Ποιο είναι το φάσμα φάσης?

Ιδιότητες DTFT (1-3/5)

1) Γραμμικότητα:

- Το φάσμα του γραμμικού συνδυασμού δύο σημάτων $x_1(n)$, $x_2(n)$, θα δώσει γραμμικό συνδυασμό των φασμάτων τους.
- Εάν $DTFT\{x_1(n)\} = X_1(e^{j\omega})$ και $DTFT\{x_2(n)\} = X_2(e^{j\omega})$, τότε
  - $DTFT\{ a_1x_1(n) + a_2x_2(n) \} = a_1X_1(e^{j\omega}) + a_2X_2(e^{j\omega})$

2) Χρονική Ολίσθηση

- Εάν ένα σήμα μετατοπιστεί στο χρόνο κατά n_0, τότε το φάσμα μέτρου του παραμένει αναλλοίωτο, ενώ στο φάσμα φάσης προστίθεται η ποσότητα $-\omega n_0$
- Εάν $DTFT\{ x(n) \} = X(e^{j\omega})$, τότε $DTFT\{ x(n- n0) \} = e^{-i\omega n_0}\, X(e^{j\omega})$
- Το συχνοτικό περιεχόμενο ενός σήματος, εξαρτάται μονάχα από τη μορφή του σήματος και όχι από τη θέση του

3) Συχνοτική Ολίσθηση

- Πολ/σμος του σήματος x(n) με $e^{j\omega_0 n}$ ισοδυναμεί με ολίσθηση του φάσματος κατά ω_0
- Σημαντικό για Filter Design
- Εάν $DTFT\{ x(n) \} = X(e^{j\omega})$, τότε $DTFT\{e^{i\omega_0 n}\, x(n) \} = X(e^{j(\omega-\omega_0)})$

Ιδιότητες DTFT (4-5/5)

4) Θεώρημα Συνέλιξης

- Συνέλιξη (*) δύο σημάτων στο πεδίο του χρόνου ισοδυναμεί με πολ/σμο στο πεδίο συχνοτήτων και το αντίστροφο
- Εάν DTFT{x(n)} = $X(e^{j\omega})$ και DTFT{y(n)} = $Y(e^{j\omega})$, τότε
  - DTFT{x(n)* y(n)} = $X(e^{j\omega})\ Y(e^{j\omega})$
- και
  - $X(e^{j\omega})*Y(e^{j\omega})$ = DTFT{x(n)y(n)}

5) Θεώρημα Parseval

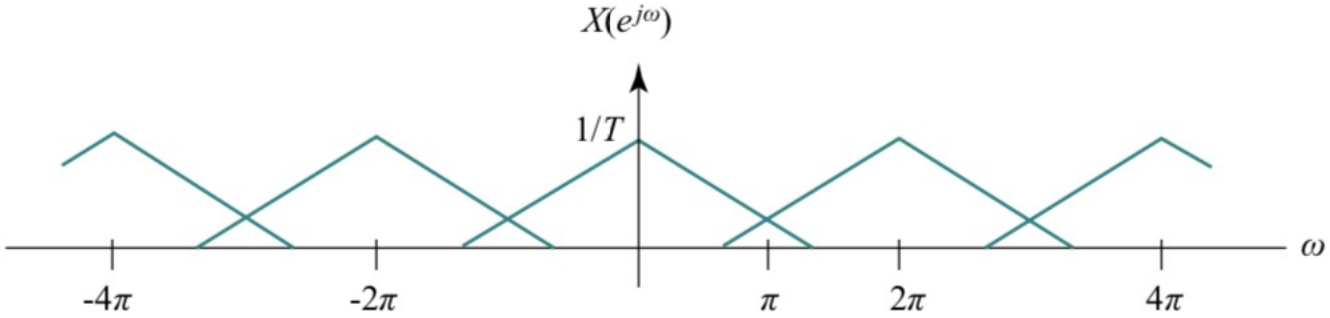
- Η ενέργεια ενός σήματος διατηρείται κατά τη μετάβαση από το πεδίο του χρόνου στο πεδίο συχνότητας.
- Εάν DTFT { x(n) } = $X(e^{j\omega})$, τότε $\sum_{n=-\infty}^{\infty} |x(n)|^2 = \frac{1}{2\pi}\int_0^{2\pi} |X(e^{j\omega})|^2\, d\omega$
- **Ορισμός:**
  - Φάσμα Ενεργειακής Πυκνότητας (Energy Density Spectrum): $|X(e^{j\omega})|^2$

Απόδειξη Θεωρήματος Συνέλιξης

- Υπενθύμιση
  - Διακριτός Μετασχηματισμός Fourier σήματος x(n):
    - $X\left(e^{j\omega}\right) = \sum_{n=-\infty}^{+\infty} x(n)e^{-j\omega n}$
  - Συνέλιξη δύο σημάτων x(n), y(n):
    - $x_1(n) * x_2(n) = \sum_{m=-\infty}^{+\infty} x_1(m)x_2(n-m)$
- Θυμηθείτε ότι:
  - Η συνέλιξη ενός σήματος x(n) με την κρουστική απόκριση h(n) ενός LTI συστήματος, δίνει την έξοδο του συστήματος y(n) όταν ως είσοδος εφαρμοστεί το x(n)
  - Επομένως, το φάσμα της εξόδου του συστήματος Y($e^{j\omega}$) είναι ίσο με τον πολ/σμο X($e^{j\omega}$) H($e^{j\omega}$)

Επιβεβαίωση θεωρήματος δειγματοληψίας

- Είχαμε δει ότι εάν ένα αναλογικό σήμα έχει συχνότητα f, τότε το ψηφιοποιημένο σήμα θα έχει συχνότητα F = f/fs, π.χ.
  - x(t) = Asin(2πft) και t = nT=n/fs => x(n) = Asin(2πfn/fs) = A sin(2πF n) = A sin(Ωn)
    - Με Ω = 2πF => F = 2π/Ω
  - O DTFT είναι περιοδική συνάρτηση του Ω με περίοδο Ω = 2π,
    - Άρα και περιοδική ως προς F με περίοδο F = 1 => f (Hz) = fs
- Επομένως, η περιοδικότητα του $X(e^{j\omega})$ με κυκλική συχνότητα ω=2π αντιστοιχεί σε Hz σε περιοδικότητα με συχνότητα fs, γεγονός που επιβεβαιώνει το θεώρημα δειγματοληψίας.



Discrete Fourier Transform (DFT)

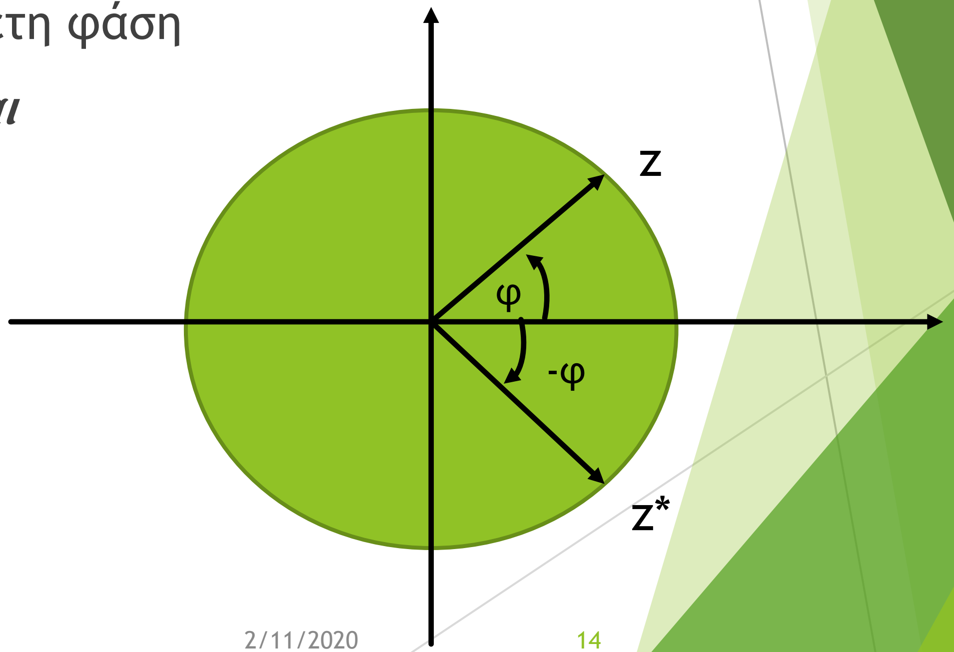
- Ο **Διακριτός Μετασχηματισμός Fourier (DFT)** είναι ο DTFT αφότου υποστεί δειγματοληψία. Δηλαδή είναι μια διακριτή συνάρτηση ως προς τη συχνότητα.
- Καταλληλότερος για DSP (στον υπολογιστή δεν έχω συνεχείς συναρτήσεις)
- Προκύπτει από δειγματοληψία του DTFT (ο οποίος είναι συνεχής συνάρτηση ως προς ω) στο διάστημα [0, 2π]) σε τόσες διακριτές συχνότητες όσες το πλήθος δειγμάτων του x(n), έστω N.
  - **$\omega_k = k(2\pi/N)$, k = 0, 1, …, N-1,** όπου με αντικατάσταση από το DTFT:
    - $X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N}$
    - **Προσοχή:**Γίνεται η παραδοχή ότι η ακολουθία x(n) έχει πεπερασμένο μήκος N, κι επομένως τα όρια του αθροίσματος αλλάζουν, καθώς θεωρείται ότι x(n) = 0 για n >N-1και n<0
  - Ο αντίστροφος του:
    - $x(n) = \frac{1}{N} \sum_{n=0}^{N-1} X(k) e^{j2\pi kn/N}$
  - *Παρατηρήστε ότι x(n) και X(k) είναι και δύο ακολουθίες μήκους N*

Εργαστηριακό Μέρος - Άσκηση Python

- Να υπολογίσετε το μετασχηματισμό DFT ενός τμήματος ψηφιακού ήχου και να απεικονίσετε γραφικά το φάσμα μέτρου και το φάσμα φάσης.
  a) Να αντιστοιχεί σε ένα τμήμα ήχου με σταθερή περιοδικότητα, δηλ. σταθερό τονικό ύψος (νότα)
  b) Τι παρατηρείτε?
  c) Χρησιμοποιήστε τη συνάρτηση numpy.unwrap για την απεικόνιση του φάσματος φάσης
  d) Αλλάξτε το διάστημα κυκλικής συχνότητας ω από [-π, π] σε [0, 2π]. Τι παρατηρείτε?
  e) Αλλάξτε τον οριζόντιο άξονα όπου αντί για ω να αναπαρίσταται ως προς συχνοτητα f (σε Hz)
  f) *Δοκιμάστε τη συνάρτηση numpy.fft.fft*
  g) Να βρείτε τη συχνότητα (Hz) στην οποία εμφανίζεται το φάσμα μέτρου παρουσιάζει τη μέγιστη τιμή του
  h) Να υπολογιστεί το αρχικό σήμα από το φάσμα που βρέθηκε στο προηγούμενο ερώτημα εφαρμόζοντας τον ανάστροφο DFT με χρήση της numpy.fft.ifft
  i) Να συγκριθούν τα δύο σήματα
  j) Να εφαρμοσθεί ο DFT σε φάσμα που έχει υποστεί υποδειγματοληψία.

Συζυγείς Μιγαδικοί (Complex Conjugate)

- Δύο μιγαδικοί λέγονται συζυγείς εάν έχουν ίδιο πραγματικό μέρος και αντίθετο φανταστικό μέρος
  - Ο συζυγής του $z = a + jb$ είναι ο $z^* = a-jb$
  - Σε πολικές συν/νες, εάν $z = re^{\varphi}$, τότε $z^* = re^{-\varphi}$
- Συζυγείς μιγαδικοί έχουν το ίδιο μέτρο αλλά αντίθετη φάση
- Ιδιότητα: ***Το γινόμενο συζυγών μιγαδικών ισούται με το κοινό τους μέτρο στο τετράγωνο***
  - $z\, z^* = a^2 + b^2 = |z|^2 = (a + jb)(a-jb)$



Παρατήρηση: Συμμετρία (Αναδίπλωση)

- Για N δείγματα σήματος x(n) που παίρνει πραγματικές τιμές ο DFT αποτελείται από N/2+1 διαφορετικές τιμές, για τις οποίες
  I. X(0) = πραγματικός
  II. X(1), X(2), …,X(N/2 -1) μιγαδικοί
  III. X(N/2) = πραγματικός
  IV. X(N/2+1), X(N/2+2), …X(N-1) συζυγείς μιγαδικοί των (II) και μάλιστα
     - X(N/2-q) = X*(N/2+q), q = 1, 2,… N/2-1

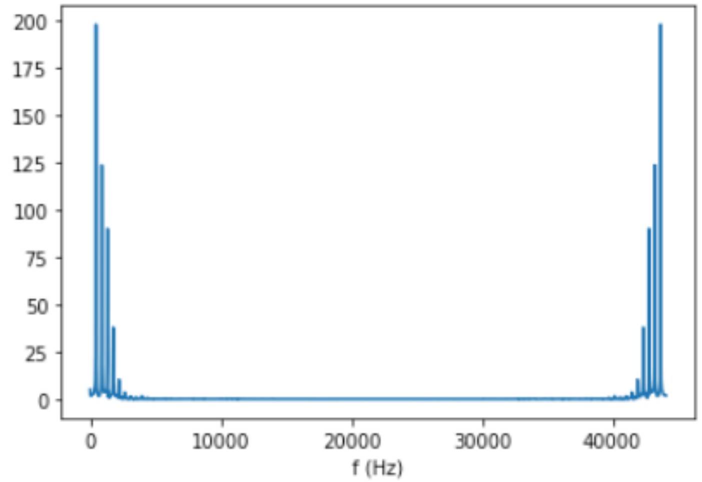
- Να αποδειχθεί μαθηματικά…
  - Υπόδειξη: Αν k1 = 1, 2, …N/2-1 και k2 = N-1, N-2 , …N/2+1, τότε **k2 = N-k1** και εφαρμόζω τη σχέση

- Το γεγονός αυτό είναι άμεση απόρροια του Θεωρήματος Δειγματοληψίας. Παράδειγμα:
  - x(t) = A sin(ωt) ----- (δειγματοληψία) x(n) = A sin(ωnTs) = A sin(ω/Fs n)
  - ω = 2π => αντιστοιχεί σε συχνότητα Fs
  - Στην Fs/2 (ω = π => k=N/2) εμφανίζεται αναδίπλωση του φάσματος

Συμμετρία στην Άσκηση

- Πράγματι, το φάσμα μέτρου και το φάσμα φάσης που είδαμε στην άσκηση είναι:

- Αναδίπλωση γύρω από την fs/2 στο φάσμα μέτρου και στο φάσμα φάσης
- Οι φάσεις για f> fs/2 είναι αντίθετες εκείνων για f<fs/2

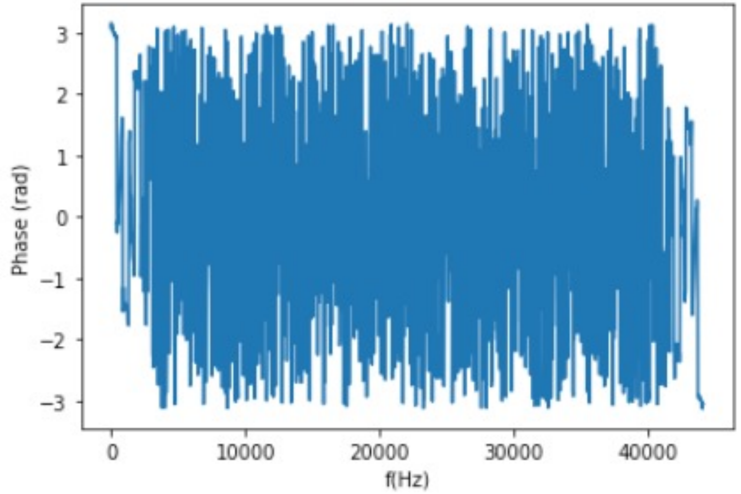



```
phase_spectrum = np.angle(X)

plt.plot( f, phase_spectrum)
plt.ylabel("Phase (rad)")
plt.xlabel('f(Hz)')

print('phase[10] = ',  np.angle(X[10]))
print('phase[N-10] = ', np.angle(X[N-10]))
```

```
phase[10] =  2.9978320422736937
phase[N-10] =  -2.997832042274665
```



Σύνοψη

- Ο DFT είναι:
  - Περιοδική συνάρτηση της συχνότητας f με περίοδο fs (ω με περίοδο 2π)
  - Παρουσιάζει συμμετρία συζυγούς (αναδίπλωση) στη συχνότητα fs/2
- Για κάθε k το X(k) μου δίνει το φάσμα στη συχνότητα
  - $f_k = k\frac{f_s}{N}$

Ο DFT σε μορφή Πινάκων

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ \vdots \\ X(N-2) \\ X(N-1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & e^{-j\frac{2\pi}{N}} & e^{-j\frac{4\pi}{N}} & \cdots & e^{-j\frac{2(N-2)\pi}{N}} & e^{-j\frac{2(N-1)\pi}{N}} \\ 1 & e^{-j\frac{4\pi}{N}} & e^{-j\frac{8\pi}{N}} & \cdots & e^{-j\frac{4(N-2)\pi}{N}} & e^{-j\frac{4(N-2)\pi}{N}} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & e^{-j\frac{2(N-2)\pi}{N}} & e^{-j\frac{4(N-2)\pi}{N}} & \cdots & e^{-j\frac{2(N-2)^2\pi}{N}} & e^{-j\frac{2(N-2)(N-1)\pi}{N}} \\ 1 & e^{-j\frac{2(N-1)\pi}{N}} & e^{-j\frac{4(N-1)\pi}{N}} & \cdots & e^{-j\frac{2(N-1)(N-2)\pi}{N}} & e^{-j\frac{(N-1)^2\pi}{N}} \end{bmatrix}}_{\text{Fourier Transform Matrix}} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-2) \\ x(N-1) \end{bmatrix}$$

Εναλλακτική Αναπαράσταση του DFT

- Συναρτήσεις Βάσης:
  - $W_N^{nk} = e^{-i2\pi kn/N}$
- Ο DFT γίνεται:
  - $X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk} \;\; \mu\varepsilon \; k = 0, \ldots, N-1$

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ \vdots \\ X(N-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w & w^2 & \cdots & w^{(N-1)} \\ 1 & w^2 & w^4 & \cdots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w^{(N-1)} & w^{2(N-1)} & \cdots & w^{(N-1)^2} \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{bmatrix}$$

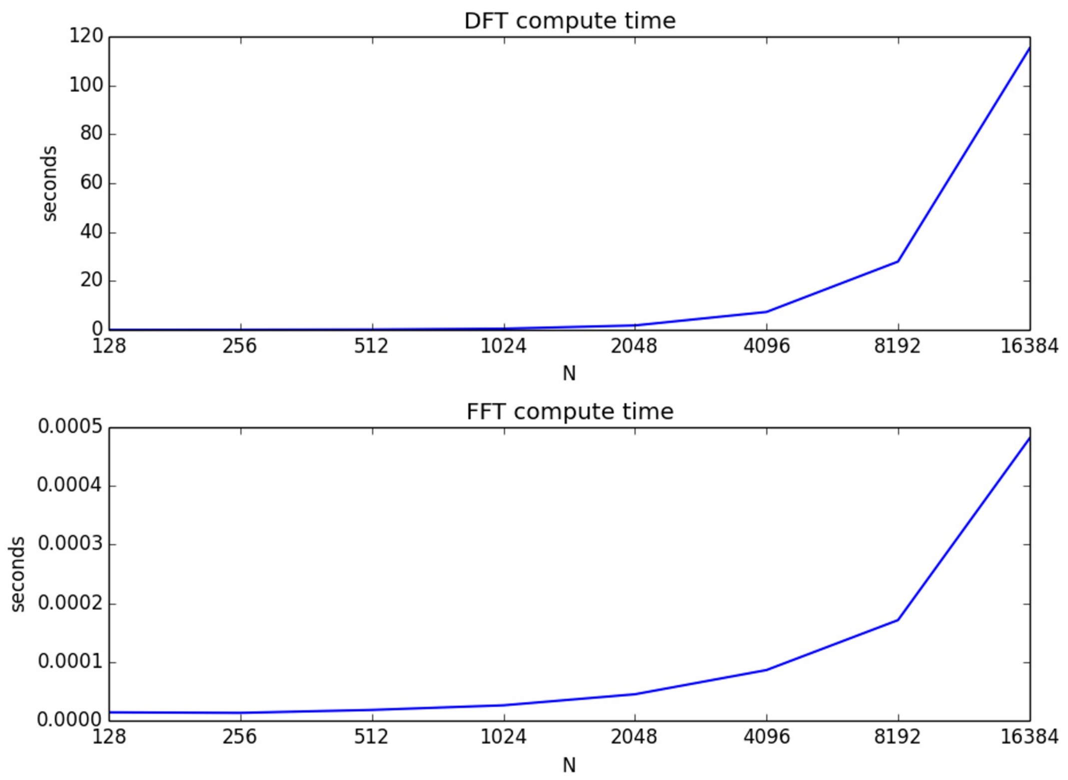
Fast Fourier Transform (FFT)

- Αναφέρεται σε μια σειρά από αποδοτικούς αλγόριθμους υπολογισμού του DFT που βασίζονται σε ιδιότητες των συναρτήσεων βάσης του μετασχηματισμού για να αποφύγουν περιττούς υπολογισμούς:

- Όταν N = 2^x (δύναμη του δύο) αξιοποιούν τα εξής:

1. Periodicity $W_N^{nk} = W_N^{n(k+N)}$
2. Complex Conjugate Symmetry
   - $W_N^{-nk} = \left(W_N^{nk}\right)^* = W_N^{(N-n)k}$

Σύγκριση Πολυπλοκότητας DFT και FFT



- Για N-point DFT η πολυπλοκότητα αλγορίθμου είναι:
  - *$O(N^2)$* του DFT
  - *$O(N \log N)$* του FFT

Η αρχή της απροσδιοριστίας στην ανάλυση σημάτων

- Είσοδος DFT:
  - Σήμα σημείων [x(0), x(1), …x(N-1)], ομοιόμορφα κατανεμημένων διακριτών αριθμών που έχουν ληφθεί σε τακτά χρονικά διαστήματα Ts και έχει συνολική διάρκεια ΔΤ=NTs
- Έξοδος DFT
  - Φάσμα σημείων [X(0), X(1), …X(N-1)], ομοιόμορφα κατανεμημένων στο διάστημα 0-fs.
- Επομένως, η ανάλυση στη συχνότητα είναι
  - Δf=1/ΔΤ = 1/(NTs) = **fs/N**
- Προσέξτε ότι (με σταθερό fs):
  - Όσο αυξάνεται το N (με σταθερή fs) μειώνεται το Δf, δηλαδή αυξάνεται η ανάλυση στη συχνότητα
  - Όμως αύξηση του Ν αντιστοιχεί σε αύξηση της διάρκειας του σήματος το οποίο μετασχηματίζεται
- Επειδή ο DFT θεωρεί ότι το φάσμα είναι σταθερό στο χρόνο (στάσιμο κύμα)
  - Αύξηση της διάρκειας Ν αντιστοιχεί σε μείωση της ανάλυσης στο χρόνο

The Uncertainty Principle

- *An important practical issue in applications of DFT is the issue of **resolution**, that is frequency versus time (or space) resolutions. **Resolution is the ability to resolve the details of a signal in time or frequency or space**. For example a satellite imaging system may have a resolution of 1 metre which means that it cannot capture details or objects smaller than 1 metre. Similarly, the frequency resolution of DFT is the width of each DFT bin between two successive discrete frequencies. **The ability to resolve two closely spaced frequencies is inversely proportional to the DFT length, that is the length of the input signal.***

- Από το βιβλίο του Saeed V. Vaseghi 2007, Multimedia Signal Processing

DFT - Σύνοψη

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N} \quad k = 0, \ldots, N-1, n = 0, \ldots, N-1$$

N: *το μηκος του σήματος που μετασχηματίζεται σε πλήθος δειγμάτων*

n: *διακριτός δείκτης χρόνου (αντιστοιχεί σε χρόνο* $t_n = \frac{n}{f_s}$

k: *διακριτός δείκτης συχνότητας (αντιστοιχεί σε συχν.* $f_k = k\, f_s/N$)

Εργαστηριακό Μέρος

- Να επιβεβαιωθεί η γραμμικότητα και η χρονική ολίσθηση του DFT για ημιτονοειδή σήματα
  - Π.χ. fs = 44100, N = 1000, n0 = 20 (η ολίσθηση σε δείγματα)