

Ψηφιακή Επεξεργασία Ήχου

Μάθημα 1: Εισαγωγή

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

Δρ. Χρυσούλα Αλεξανδράκη

Τμήμα Μουσικής Τεχνολογία και Ακουστικής

Ελληνικό Μεσογειακό Πανεπιστήμιο

Διαλέξεις - Εργαλεία Εκμάθησης

▶ Διαλέξεις

▶ Θεωρία

- ▶ Θα μελετάμε τρόπους επεξεργασίας ψηφιακού ήχου εστιάζοντας στη μαθηματική τους ερμηνεία

▶ Άσκηση

- ▶ Θα υλοποιούμε αυτούς τους τρόπους επεξεργασίας σε κώδικα Python

▶ Εργαλεία Εκμάθησης

- ▶ Πλατφόρμα <http://eclass.hmu.gr>

▶ Λογισμικό - Περιβάλλον υλοποίησης ασκήσεων

- ▶ Google Collaboratory <https://colab.research.google.com/>
 - ▶ Υλοποίηση και εκτέλεση κώδικα Python σε Internet Browser
- ▶ Open Source Software: Audacity (audio editor), Python Libraries and Packages

Αξιολόγηση - Βαθμολογία

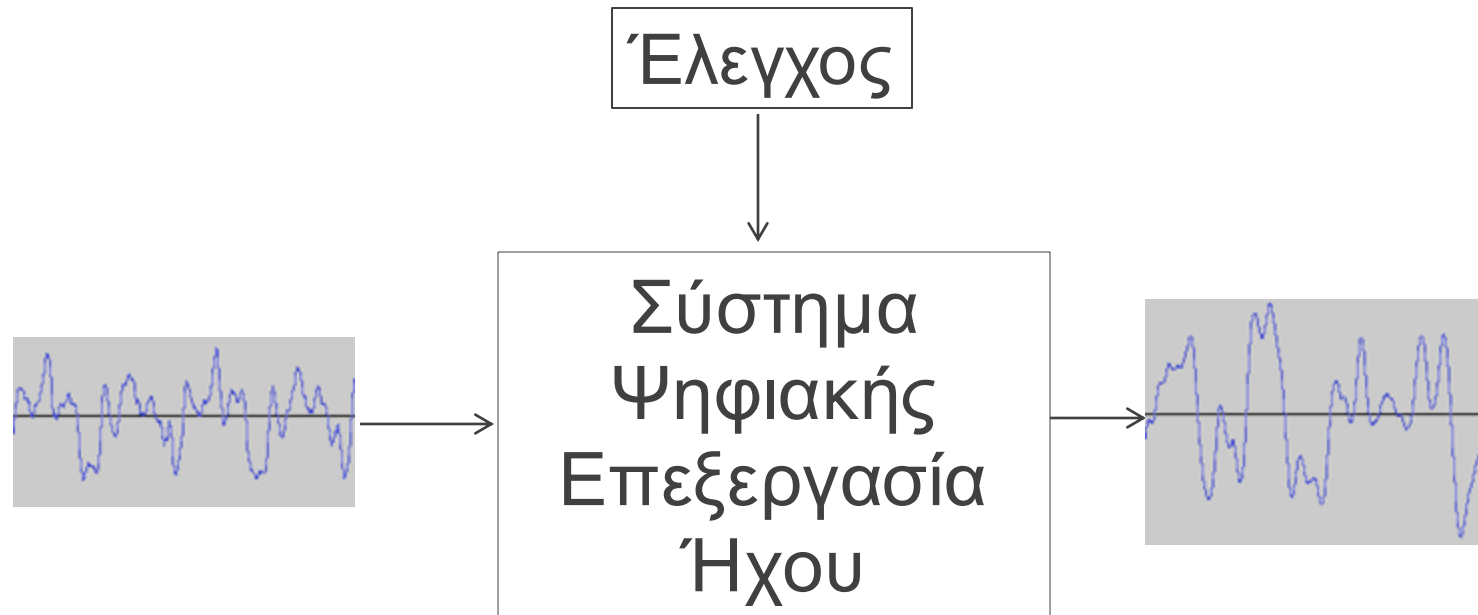
- ▶ Η αξιολόγηση της επίδοσης σας στο μάθημα θα προκύπτει από τις εβδομαδιαίες εργασίες που θα παραδίδετε (E) καθώς και την τελική σας εργασία (T), ως εξής:
 - ▶ **$B = 0.4E + 0.6T$**
- ▶ Όπου:
 - ▶ **B**: ο τελικός βαθμός, ο οποίος απαρτίζεται από:
 - ▶ **40% E**: ο μέσος όρος των βαθμών που θα πάρετε από τις εργασίες που σας έχουν ζητηθεί να εκπονήσετε κατά τη διάρκεια του εξαμήνου
 - ▶ **60% T**: ο βαθμός στην τελική εργασία,
 - ▶ Η τελική σας εργασία επικεντρώνει σε κάποιο **ερευνητικό/αλγοριθμικό πρόβλημα** και απαρτίζεται από:
 - ▶ 40% Γραπτό: ένα κείμενο σε μορφή επιστημονικού άρθρου
 - ▶ 20% Κώδικας: η υλοποίηση αλγοριθμικού προβλήματος
 - ▶ 40% Παρουσίαση: προφορική παρουσίαση του θέματος που διερευνήθηκε

Περίγραμμα Ύλης

- ▶ **Εισαγωγή**
 - ▶ Τύποι Σημάτων, Συχνότητα, Ψηφιοποίηση
- ▶ **Σήματα και Συστήματα Διακριτού Χρόνου**
 - ▶ Διακριτές Ακολουθίες, Κρουστική Απόκριση, Συνέλιξη
- ▶ **Φασματική Ανάλυση Σήματος**
 - ▶ Μετασχηματισμοί Fourier διακριτών σημάτων: DFT, FFT, STFT
- ▶ **Ψηφιακά Φίλτρα**
 - ▶ Φίλτρα πεπερασμένης/άπειρης κρουστικής απόκρισης (FIR, IIR)
- ▶ **Χαρακτηριστικά Ήχου και τρόποι υπολογισμού τους**
 - ▶ Audio Feature Extraction
- ▶ **Αλγόριθμοι ανάλυσης χρονικής δομής ηχητικού σήματος**
 - ▶ Αλγόριθμοι ανίχνευσης έναρξης ηχητικών συμβάντων
 - ▶ Αλγόριθμοι παρακολούθησης ρυθμικού παλμού και εξαγωγής tempo
- ▶ **Αναπαράσταση τονικού ύψους σε μουσικά σήματα**
 - ▶ Χρωματογράφημα
 - ▶ Ανίχνευση τονικού ύψους σε μονοφωνικά σήματα
 - ▶ Εκτίμηση συχορδιών σε πολυφωνικά σήματα

Τι είναι η Ψηφιακή Επεξεργασία Ήχου

- ▶ Οποιαδήποτε εκούσια μεταβολή ψηφιακού ηχητικού σήματος



Τι είναι σήμα

- Ως σήμα ορίζουμε το σύνολο των τιμών που λαμβάνει μια φυσική ποσότητα. Μαθηματικά αυτό εκφράζεται ως συνάρτηση ή ακολουθία μιας ή περισσότερων ανεξάρτητων μεταβλητών (π.χ. του χρόνου, του χώρου, ή και των δύο) Τα σήματα περιέχουν πληροφορία σχετικά με τη συμπεριφορά ή τη φύση ενός φαινομένου, στο χρόνο, στο χώρο ή και στα δύο...



(α)



Χ. Αλεξανδράκη (β)



(γ)

Περιγραφή Σήματος Πολυμεσικού Περιεχομένου:

α) ήχος: $z = f(t)$

β) εικόνα: $z = f(x, y)$

γ) βίντεο: $z = f(x, y, t)$

- Ανεξάρτητες Μεταβλητές: t, x, y
- Εξαρτημένη Μεταβλητή: z

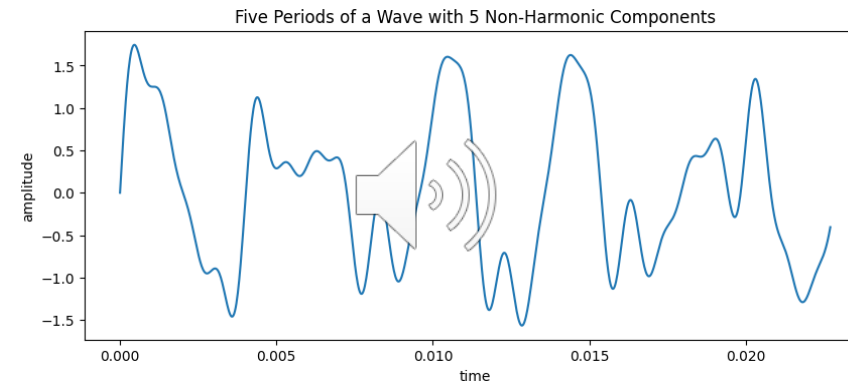
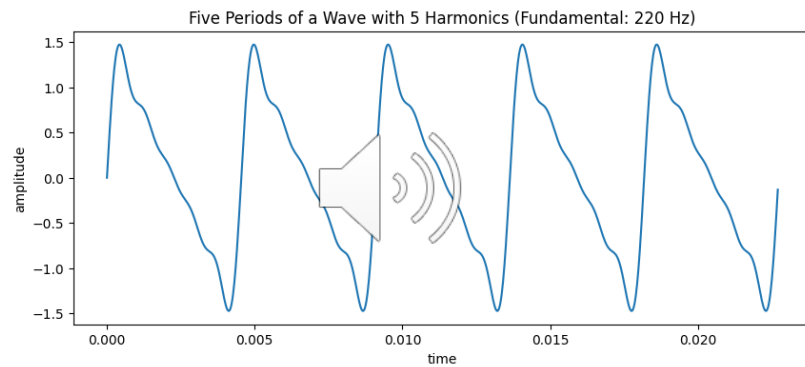
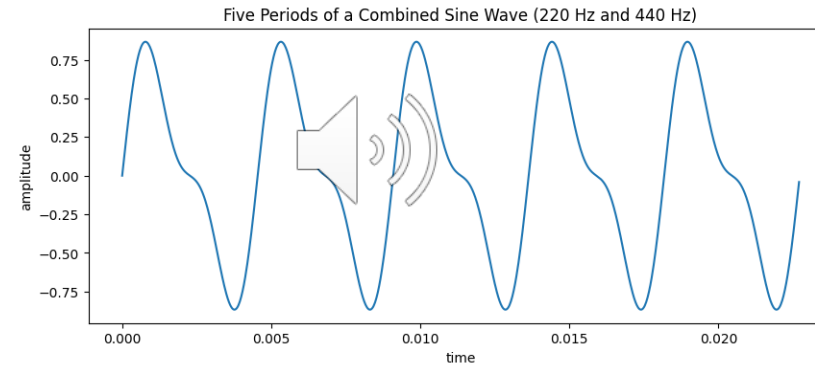
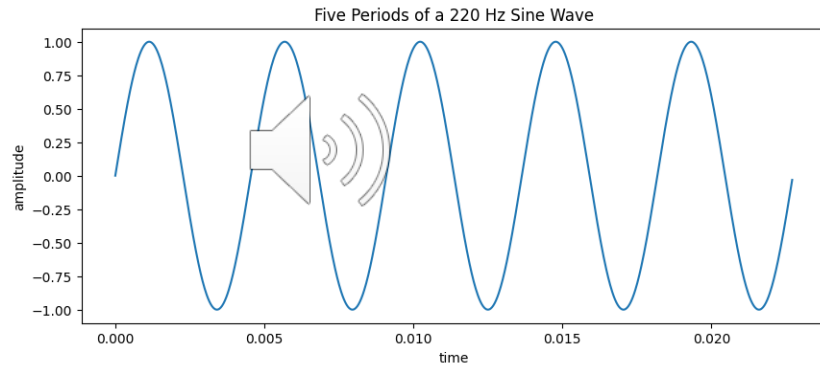
Τι είναι η Ψηφιακή Επεξεργασία Ήχου

- ▶ Η ψηφιακή επεξεργασία σήματος ασχολείται με την ανάπτυξη συστημάτων επεξεργάζονται ηχητικά σήματα με στόχο την:
 - ▶ Αναπαραγωγή - Απόδοση Σήματος
 - ▶ Ανάλυση (π.χ. τι συχνότητες περιλαμβάνει)
 - ▶ Τροποποίηση - Μετασχηματισμός - Συμπίεση (π.χ. εφαρμογή κάποιου φίλτρου)
 - ▶ Κατανόηση και Περιγραφή Περιεχομένου (π.χ. σήμα ομιλίας, σήμα μουσικής, περιβαλλοντικός θόρυβος)
- ▶ Ένα ψηφιακό σήμα είναι ένα σύνολο από αριθμούς.
 - ▶ Το ερώτημα είναι εάν θέλω να τροποποιήσω το ηχητικό αποτέλεσμα, τι πρέπει να κάνω με τους αριθμούς αυτούς?? Παράδειγμα, τι πρέπει να κάνω στους αριθμούς, ώστε:
 - ▶ Να αφαιρέσω το θόρυβο από ένα ήχο?
 - ▶ Να αυξήσω τις χαμηλές συχνότητες (μπάσα)??
 - ▶ Να αυξήσω τη λαμπρότητα ενός ήχου?
 - ▶ Να εντοπίσω το τονικό ύψος ή τα τονικά ύψη που ακούγονται σε ένα ήχο?
 - ▶ Να εντοπίσω τα σημεία αλλαγής νοτών ενός ήχου?
 - ▶ Να μειώσω το μέγεθος ενός αρχείου ήχου

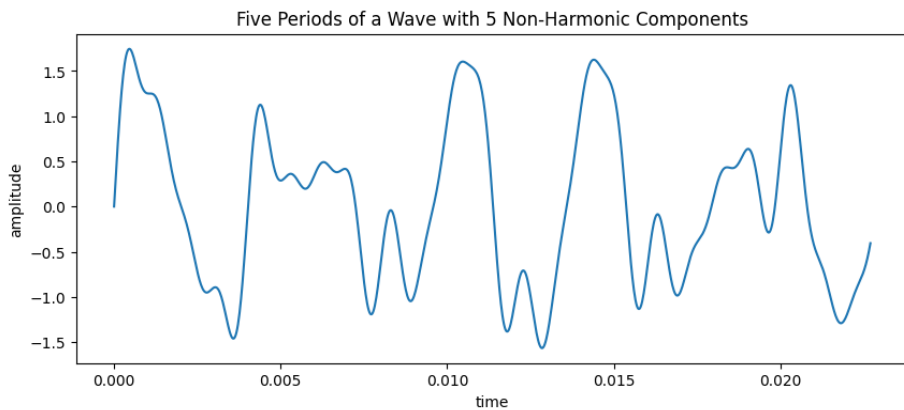
Τί είναι ηχητική συχνότητα

- ▶ Ο ήχος αντιστοιχεί σε μία **περιοδική δόνηση/ταλάντωση** του ατμοσφαιρικού αέρα
 - ▶ **Συχνότητα:** ο ρυθμός αυτής της δόνησης
 - ▶ **Μονάδα μέτρησης:** Hertz (Hz) = 1/sec: ο αριθμός κύκλων/επαναλήψεων της ταλάντωσης ανά δευτερόλεπτο
- ▶ Τιμές ηχητικού φάσματος:
 - ▶ **Ακουστικό φάσμα:** 20Hz έως 20000Hz - συχνότητες που γίνονται αντιληπτές από το ανθρώπινο αυτί, τις ακούμε
 - ▶ **Υπόηχοι:** δεν τις ακούμε αλλά σε ορισμένες περιπτώσεις τις αισθανόμαστε
 - ▶ **Υπέρηχοι (ultrasounds):** χρήσιμοι στην ανίχνευση αντικειμένων και των διαστάσεών τους
- ▶ https://en.wikipedia.org/wiki/Audio_frequency

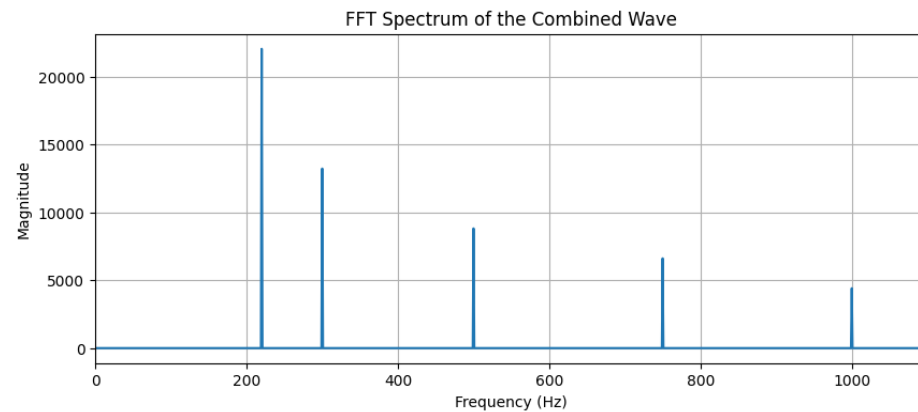
Πολυπλοκότητα Δόνησης και κατά τύπους περιοδικότητα



Αναπαράσταση ηχητικού σήματος στο πεδίο του χρόνου και στο πεδίο της συχνότητας



- ▶ Κυματομορφή
 - ▶ Πλάτος συναρτήσει του χρόνου



- ▶ Φάσμα Μέτρου
 - ▶ Μέτρο συναρτήσει της συχνότητας
 - ▶ Το φάσμα είναι μιγαδική συνάρτηση που

Τροφή για σκέψη

1. Πότε ένας ήχος ονομάζεται αρμονικός;
2. Ποια φυσικοί ήχοι προσεγγίζουν αρμονικούς ήχους;
3. Πώς διαφέρουν αντιληπτικά οι αρμονικοί από τους αρμονικούς ήχους;
4. Γιατί οι μη-αρμονικοί ήχοι δεν παρουσιάζουν περιοδικότητα;

Πλάτος - Ένταση - Στάθμη Έντασης Ήχου

▶ Πλάτος Σήματος (Amplitude)

- ▶ Η μέγιστη μετατόπιση από τη θέση ισορροπίας του μέσου που διαταράσσεται λόγω κυματικής ταλάντωσης. Συνήθως είναι κάποια τιμή μεταξύ 0 και 1

▶ Ηχητική Ένταση (Sound Intensity)

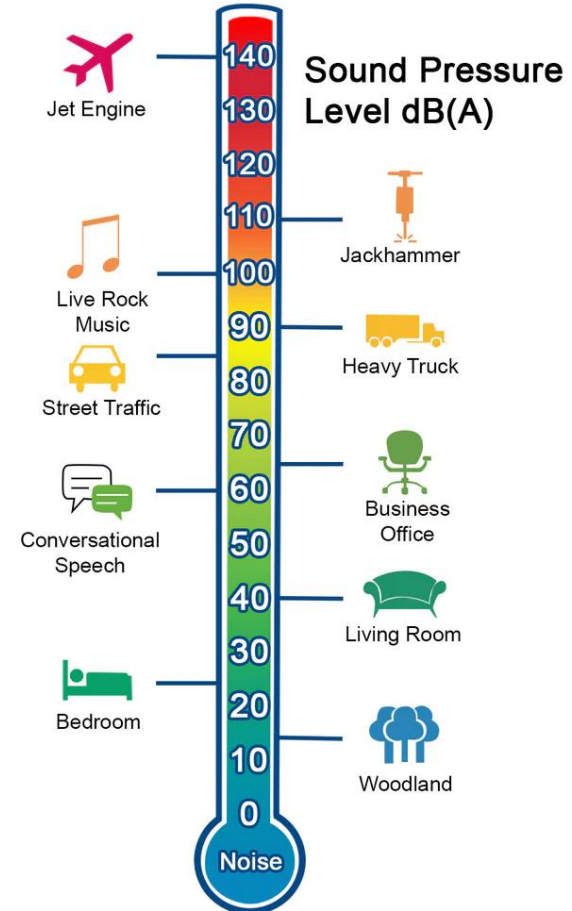
- ▶ Ισχύς ανά μονάδα επιφάνειας $I = P/A$ (Watt/m²)
- ▶ Ένα πολύ αδύναμο σήμα (κατώφλι ακοής) έχει $I_0 = 10^{-12} \text{ W/m}^2$
- ▶ Ένας πολύ δυνατός ήχος (όριο πόνου) έχει $I = 1 \text{ W/m}^2$

▶ Στάθμη Έντασης (Sound Level)

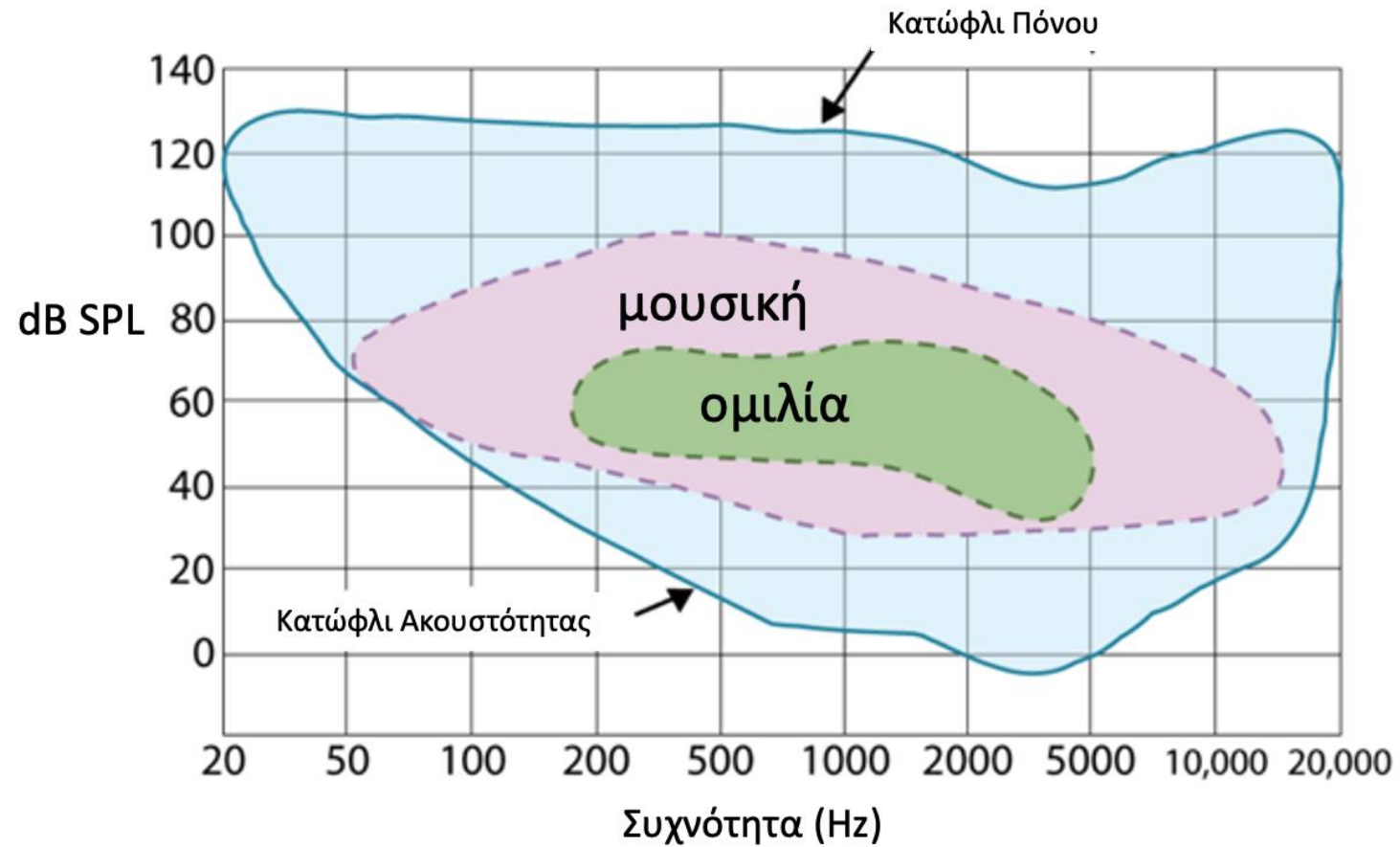
- ▶ Συγκριτικό μέγεθος: πόσο πιο δυνατός είναι ο ήχος συγκριτικά με το κατώφλι ακοής $L = 10 \log_{10} \left(\frac{I}{I_0} \right) \text{ dB}$

Decibel

- ▶ Ενδεικτικά
 - ▶ Ψίθυρος → περίπου 30 dB
 - ▶ Κανονική ομιλία → γύρω στα 60 dB
 - ▶ Κυκλοφορία στον δρόμο → περίπου 80 dB
 - ▶ Συναυλία → 100 dB και πάνω
 - ▶ Απογείωση αεροπλάνου → 120 dB
- ▶ Γιατί χρησιμοποιούμε λογαριθμική κλίμακα;
 - ▶ Μας βοηθούν να εκφράζουμε **πολύ μεγάλες διαφορές έντασης** με απλούς αριθμούς
 - ▶ Ανταποκρίνονται **στον τρόπο που ακούει ο άνθρωπος**
 - ▶ Είναι **πιο πρακτικά** για να συγκρίνουμε και να μετράμε ήχους στην καθημερινή ζωή, στη μουσική και στην τεχνολογία.



Φασματική έκταση ηχητικών σημάτων



Παλμοκωδική διαμόρφωση - Pulse Code Modulation (PCM)

Ο βασικός τρόπος ψηφιοποίησης ηχητικού σήματος

Σήματα Συνεχούς και Διακριτού Χρόνου

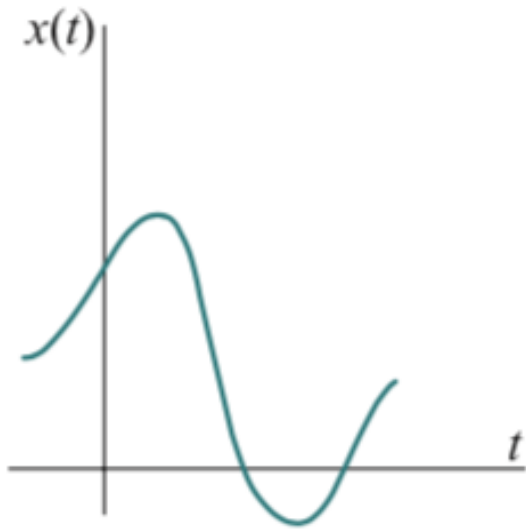
▶ Σήματα Συνεχούς Χρόνου (Continuous Time Signals):

- ▶ Η ανεξάρτητη μεταβλητή είναι συνεχής, δηλαδή τα σήματα αυτά ορίζονται για οποιαδήποτε τιμή της ανεξάρτητης μεταβλητής (π.χ. του χρόνου). Η εξαρτημένη μεταβλητή (δηλ. το πλάτος - amplitude) του σήματος, είναι και αυτή συνεχής. Τα σήματα αυτά αναφέρονται και ως σήματα **συνεχούς χρόνου συνεχούς πλάτους** ή **αναλογικά σήματα**.
- ▶ Συμβολισμός: **$x(t)$** , όπου **t** πραγματικός αριθμός.

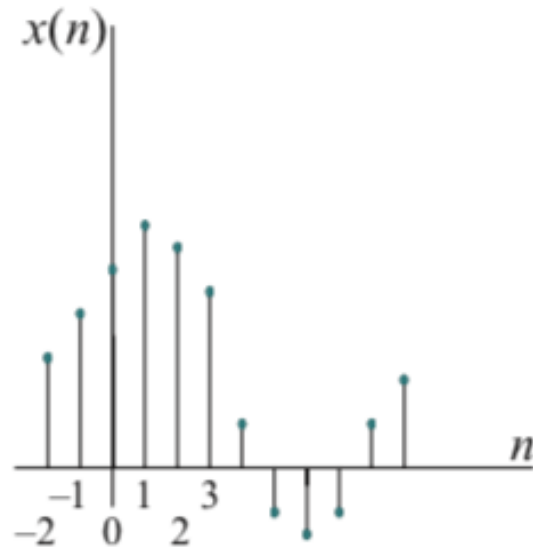
▶ Σήματα Διακριτού Χρόνου (Discrete Time Signals):

- ▶ Η ανεξάρτητη μεταβλητή είναι διακριτή, δηλαδή τα σήματα αυτά ορίζονται μόνο για συγκεκριμένες τιμές της ανεξάρτητης μεταβλητής. Με άλλα λόγια, η ανεξάρτητη μεταβλητή παίρνει τιμές από ένα διακριτό σύνολο αριθμών. Η εξαρτημένη μεταβλητή, δηλαδή το πλάτος του σήματος, είναι συνεχής. Γι' αυτό και τα σήματα αυτά αναφέρονται και ως σήματα **διακριτού χρόνου συνεχούς πλάτους**.
- ▶ Εάν και η εξαρτημένη μεταβλητή παίρνει διακριτές τιμές τότε τα σήματα λέγονται **διακριτού χρόνου διακριτού πλάτους** ή **ψηφιακά σήματα**
- ▶ Συμβολισμός: **$x(n)$** όπου **$n = 0, 1, 2, 3 \dots$** φυσικός αριθμός

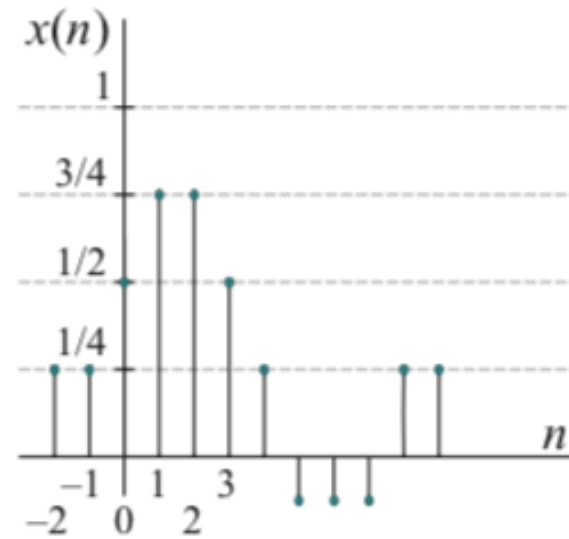
Τύποι Σημάτων και Συμβολισμός



αναλογικό σήμα



σήμα διακριτού χρόνου συνεχούς πλάτους



ψηφιακό σήμα

Ψηφιοποίηση Σήματος - Analogue to Digital Conversion (ADC)

1. Δειγματοληψία (Sampling):

- ▶ Αναλογικό => διακριτού χρόνου συνεχούς πλάτους
- ▶ $x_a(t) \rightarrow x(nT_s)$, δηλαδή $t = nT_s$ και T_s : σταθερή περίοδος δειγματοληψίας
- ▶ Εισάγει **σφάλμα δειγματοληψίας**: χάνονται διακυμάνσεις πλάτους που μπορεί να συμβούν σε χρονικό διάστημα T_s

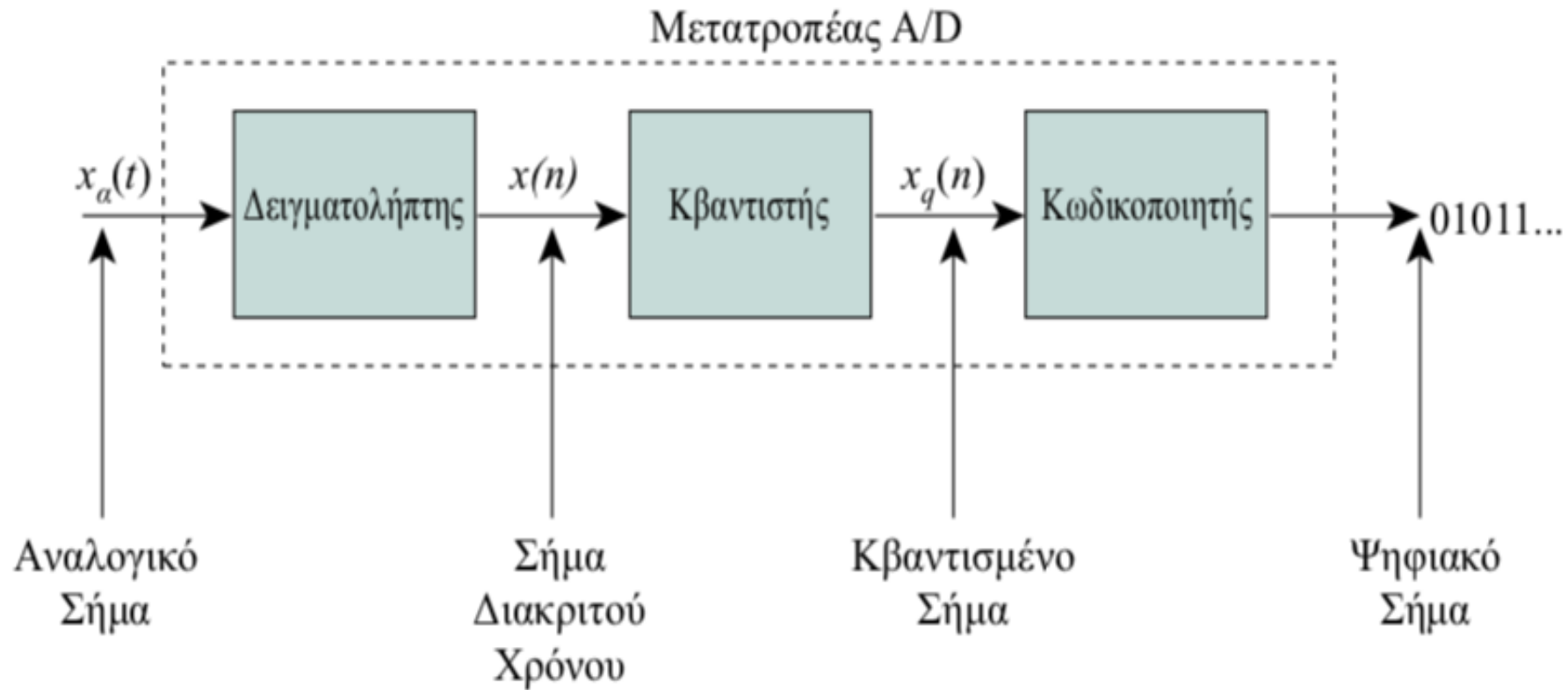
2. Κβαντοποίηση (Quantisation):

- ▶ Διακριτού χρόνου συνεχούς πλάτους => Ψηφιακό
- ▶ $x(n) \rightarrow x_q(n)$, $q = 1, 2, 3$, ορίζει πεπερασμένο πλήθος τιμών πλάτους (στοιβάδες) που ισαπέχουν
- ▶ Εισάγει **σφάλμα κβαντοποίησης**: οι διαφορές ανάμεσα στην ακριβή τιμή πλάτους $x(n)$ και στην πλησιέστερη στοιβάδα $x_q(n)$

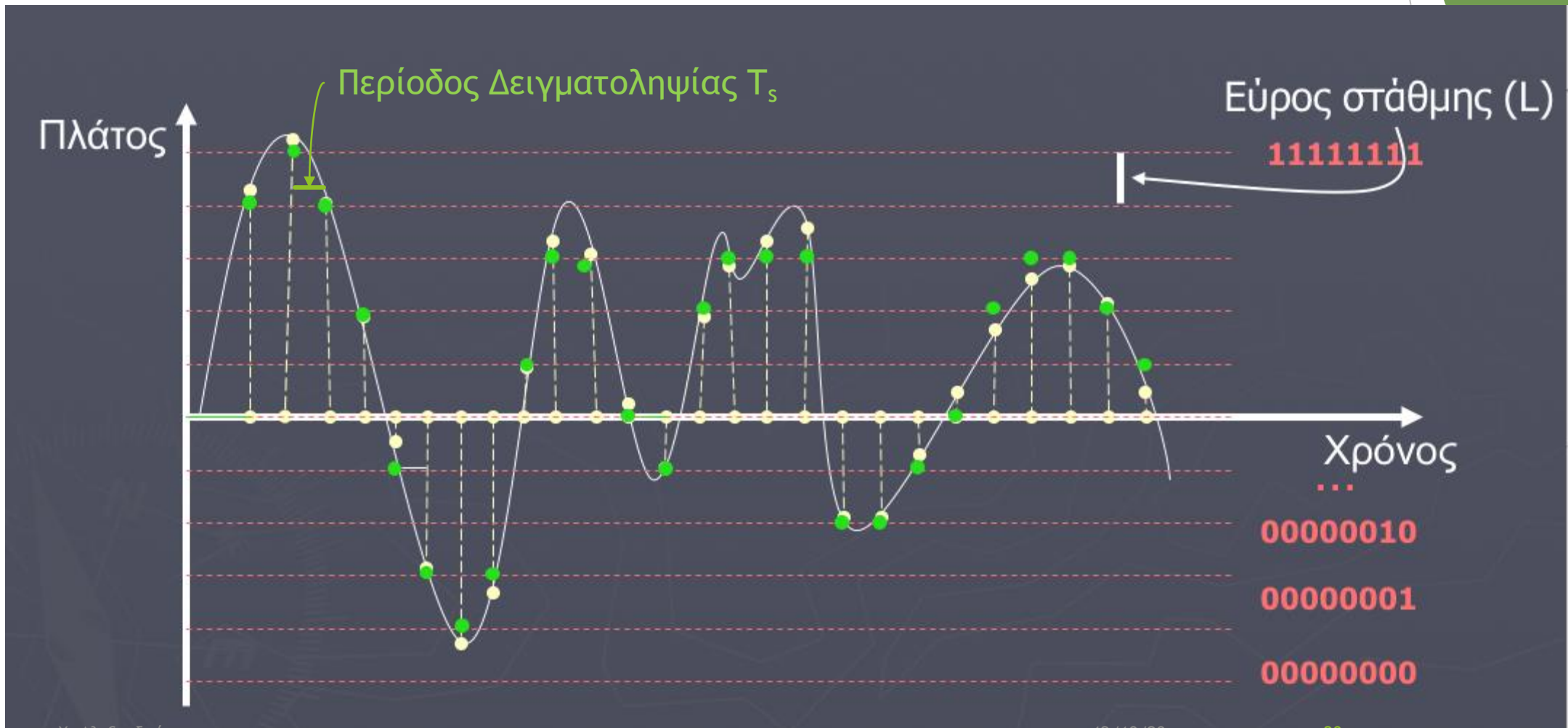
3. Κωδικοποίηση (Coding):

- ▶ Κάθε στοιβάδα $x_q(n)$ αντιστοιχίζεται σε ένα δυαδικό αριθμό αποτελούμενο από d -bits (π.χ. $d = 8, 16, 32$ κλ.π, ανάλογα με το επιθυμητό πλήθος στοιβάδων
 - ▶ Π.χ. $2^{16} = 65536$, δηλαδή οι ακέραιες τιμές πλάτους $-32768, \dots, 0, \dots, +32767$

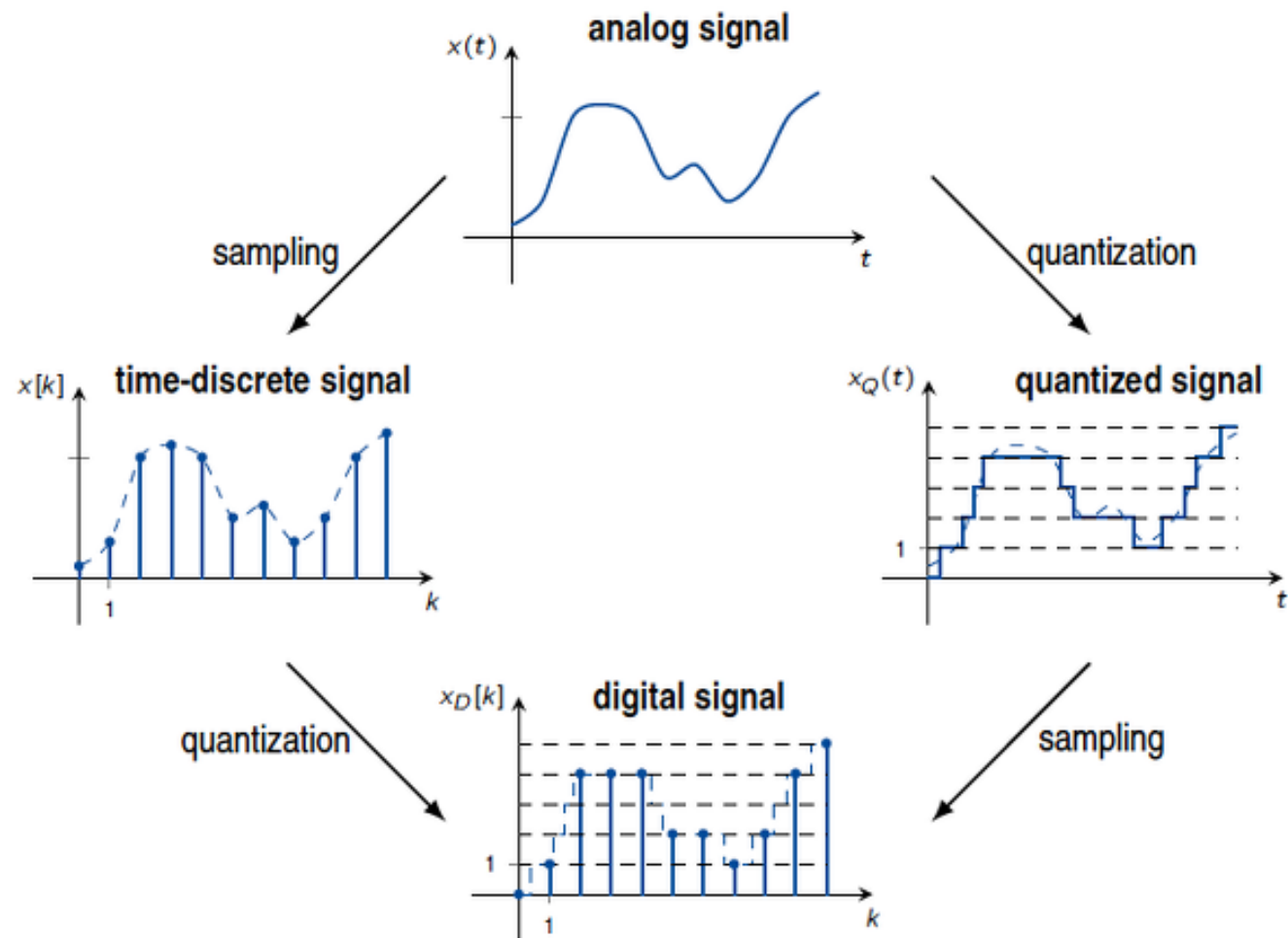
Analogue to Digital Converter (ADC)



Παράδειγμα



Σειρά Διαδικασιών

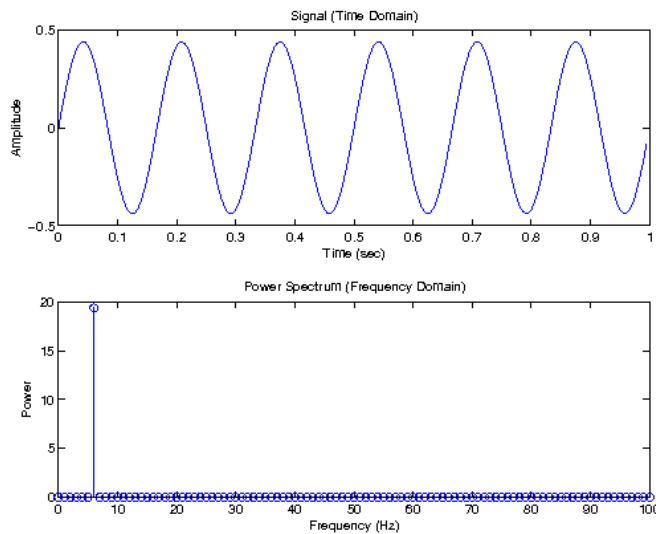


Τι είναι φάσμα;

- ▶ Η αναπαράσταση που προκύπτει όταν σε ένα σήμα, π.χ. $z = f(t)$ κάνω αλλαγή μεταβλητής $z = g(v)$, όπου t (sec χρόνος) και v (Hz συχνότητα)

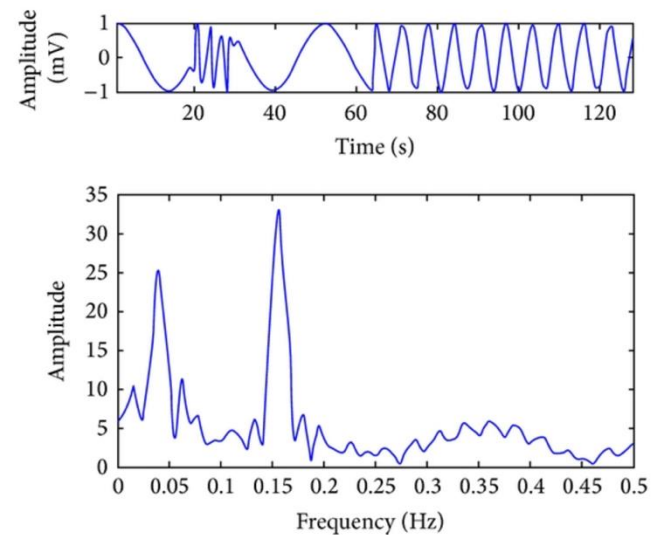
Ημιτονοειδές Σήμα

$$z = f(t) = A \sin(\omega t)$$
$$z = g(v) = \begin{cases} A & \text{για } v = \omega/2\pi \\ 0 & \text{για } v \neq \omega/2\pi \end{cases}$$



Σύνθετο Σήμα

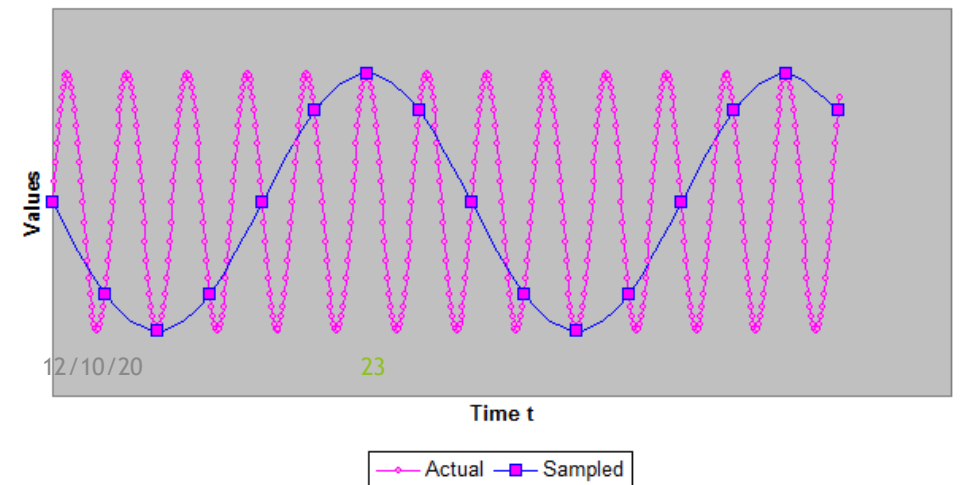
Κάθε σήμα μπορεί να αναπαρασταθεί ως άθροισμα ημ/συν-ημιτονοειδών συναρτήσεων στο μιγαδικό επίπεδο και μέσω αυτής της αναπαράστασης μπορούμε να συμπεράνουμε τις συνχότητες ενός τέτοιου σήματος



Υπολογισμός Φάσματος

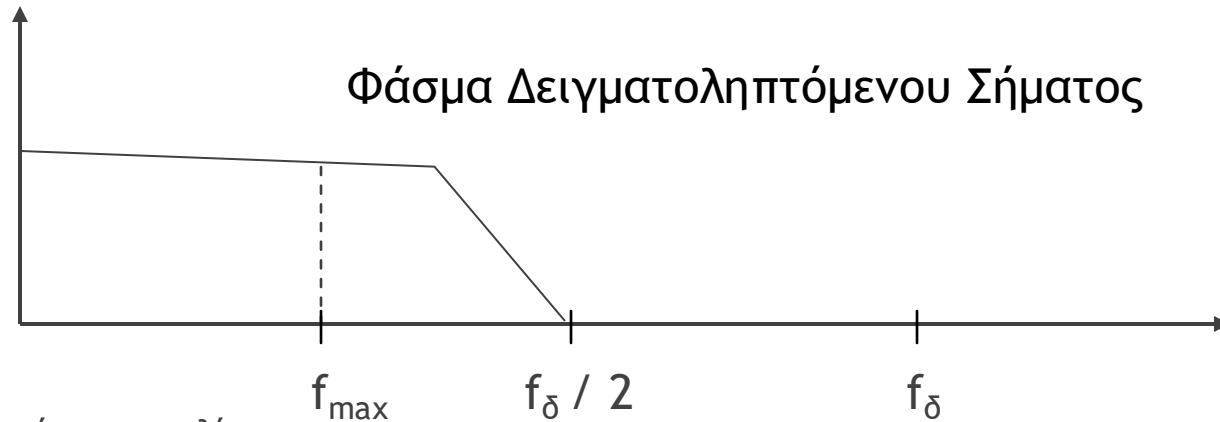
- ▶ Γίνεται με διάφορους μαθηματικούς μετασχηματισμούς:
 - ▶ Ενδεικτικά Fourier Transform, Laplace Transform και οι παραλλαγές τους, Wavelet Transform κ.α.
 - ▶ Θα ασχοληθούμε εκτενώς σε επόμενα μαθήματα
- ▶ Είναι γενικά αντιστρεπτή διαδικασία:
 - ▶ Ο αντίστροφος μετασχηματισμός όταν εφαρμόζεται σε ένα φάσμα δίνει το αρχικό σήμα, π.χ. Inverse Fourier Transform
- ▶ Κατά την ψηφιοποίηση αναλογικού σήματος (και ειδικότερα κατά τη δειγματοληψία) χρειάζεται προσοχή ώστε να μην παραμορφωθεί το φάσμα με την εμφάνιση ψευδών συχνοτήτων (βλέπε σχήμα)

Aliasing example



Θεώρημα Δειγματοληψίας (Nyquist-Shannon)

- ▶ Για την πλήρη περιγραφή ενός αναλογικού σήματος του οποίου το φάσμα συχνοτήτων είναι περιορισμένο σε ένα εύρος ζώνης με μέγιστη συχνότητα f_{max} , αρκεί η περιοδική λήψη δειγμάτων τα οποία να λαμβάνονται με συχνότητα τουλάχιστον ίση προς $f_s \geq 2f_{max}$ (δηλ. με περίοδο $T_s \leq 1/2f_{max}$)
 - ▶ για να μη χαθεί πληροφορία θα πρέπει να παίρνουμε τουλάχιστον δύο δείγματα ανά περίοδο (της υψηλότερης συχνότητας του σήματος)



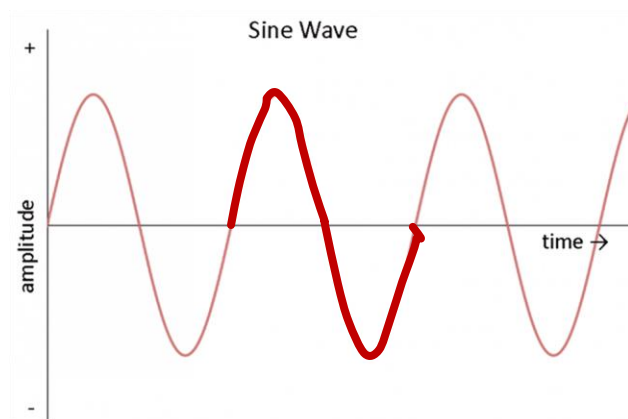
- ▶ Π.χ. Μετάδοση σήματος ομιλίας
 - ▶ 0-4000Hz => $f_\delta \geq 8000\text{Hz}$
- ▶ Παραμόρφωση λόγω Δειγματοληψίας (Aliasing Distortion)
 - ▶ Απουσία συχνοτήτων υψηλότερων του μισού της συχνότητας δειγματοληψίας
- ▶ «μπουκωμένος» ήχος

Ημιτονοειδή Σήματα Διακριτού και Συνεχούς Χρόνου

Περιγραφή Ημιτονοειδούς Σήματος

▶ Αναλογικό

- ▶ $x(t) = A \cos(2\pi ft) = A \cos(\omega t)$
- ▶ f : συχνότητα αναλογικού σήματος
 - ▶ Πόσες φορές το δευτερόλεπτο επαναλαμβάνεται το σήμα
- ▶ $\omega = 2\pi f$: κυκλική συχνότητα αναλογικού



▶ Ψηφιακό:

- ▶ Δειγματοληψία \Leftrightarrow Αντικατάσταση Μεταβλητής
 - ▶ $t = n T_s = n / f_s$
- ▶ $x(n) = A \cos(2\pi f n / f_s) = A \cos(2\pi F n)$
- ▶ $F = f / f_s$: Συχνότητα Ψηφιακού Σήματος
 - ▶ Ανα πόσα δείγματα n επαναλαμβάνεται το σήμα
- ▶ $\Omega = 2\pi F$: Κυκλική Συχνότητα Ψηφιακού

▶ Παράδειγμα:

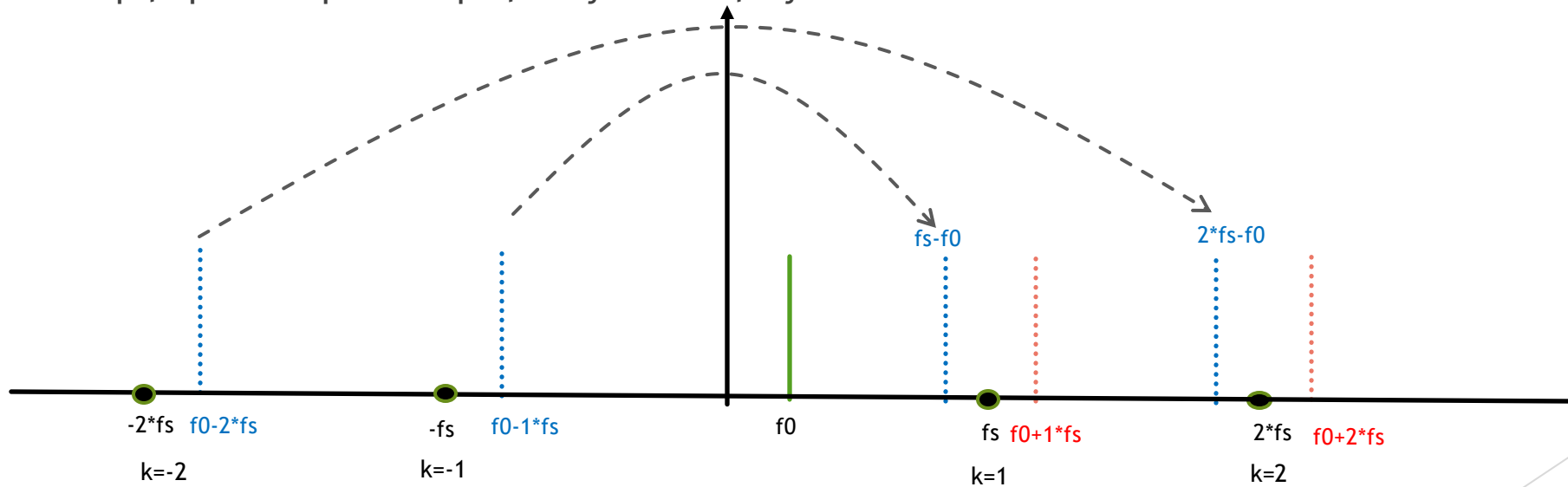
- ▶ $f = 441\text{Hz}$, $f_s = 44100\text{Hz} \Rightarrow F = 1/100$, $T = 100$ δείγματα
- ▶ Ερμηνεία:
 - ▶ **Αναλογικό:** Η περίοδος του ημιτόνου αντιστοιχεί σε $1/441 = 2.27\text{ms}$ δευτερόλεπτα
 - ▶ **Ψηφιακό:** Η περίοδος ημιτόνου αντιστοιχεί σε 100 δείγματα

Παρουσία ‘Ψευδών Συχνοτήτων’ σε ψηφιοποιημένο σήμα

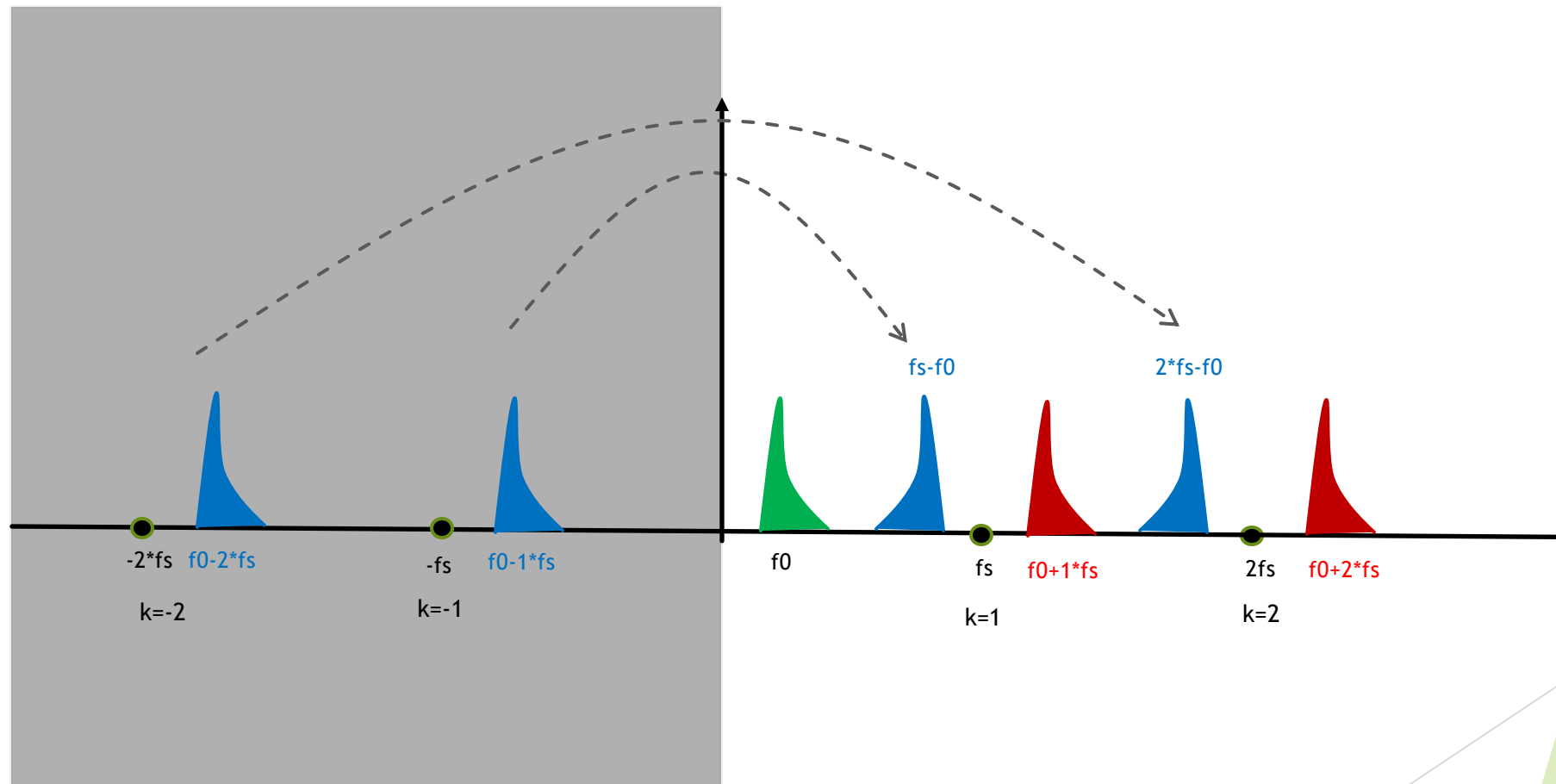
- ▶ Έστω αναλογικό σήμα
 - ▶ $x(t) = A \cos(2\pi f_0 t + \theta)$
- ▶ Έπειτα από ψηφιοποίηση ($t = n/f_s$)
 - ▶ $x(n) = A \cos(2\pi f_0 n/f_s + \theta)$
- ▶ Εύκολα συμπεραίνουμε ότι:
 - ▶ $x(n) = A \cos(2\pi n * f_0 / f_s + \theta + 2\pi k n)$ με $k = \pm 1, \pm 2, \pm 3, \dots$
 - ▶ $x(n) = A \cos(2\pi n * f_0 / f_s + 2\pi k n * f_s / f_s + \theta) = A \cos(2\pi f_k / f_s n + \theta)$
 - ▶ ‘όπου $f_k = f_0 + k f_s$
- ▶ Συνεπώς:
 - ▶ Όταν ένα αναλογικό σήμα (βασικής συχνότητας f_0) ψηφιοποιείται, κατά τη μετατροπή του πίσω σε αναλογικό προκύπτει ένα σήμα που περιλαμβάνει όλες τις συχνότητες της μορφής $f_k = f_0 + k f_s$ με $k = \pm 1, \pm 2, \pm 3, \dots$

Αρνητικές Συχνότητες και Αναδίπλωση

- ▶ Γύρω από τα ακέραια πολλαπλάσια της f_s εμφανίζονται είδωλα της συχνότητα του αρχικού σήματος
- ▶ Η συχνότητα είναι μια φυσική ποσότητα που έχει πάντα θετικές τιμές.
- ▶ Όμως για ευκολία υπολογισμού πολλές φορές στις μαθηματικές πράξεις χρησιμοποιούμε και αρνητικές συχνότητες



Φάσμα σύνθετου σήματος μετά από δειγματοληψία



Ημιτονοειδές Σήμα Διακριτού Χρόνου (ΗΣΔΧ)

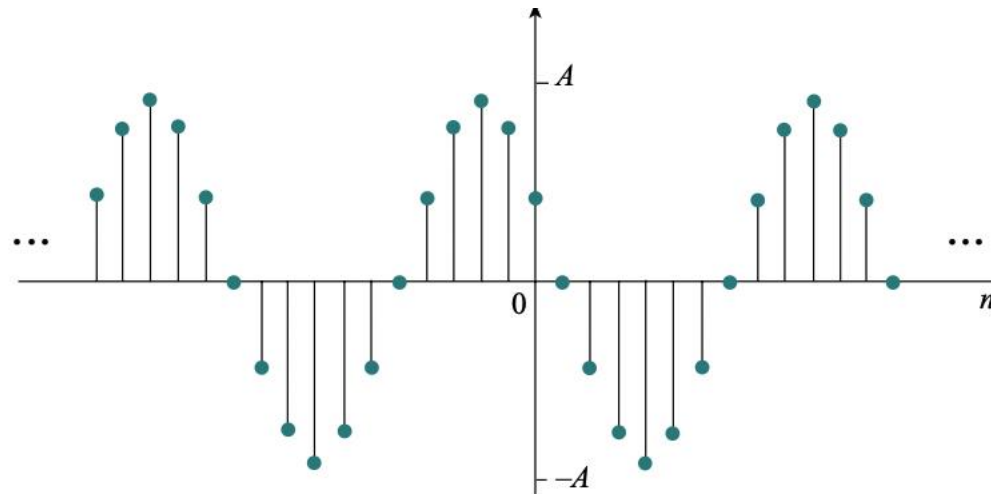
► Ορισμός

► $x(n) = A \cos(\Omega n + \theta), -\infty < n < \infty$

► όπου n αριθμός δείγματος, Ω η συχνότητα σε ακτίνια (rad) ανά δείγμα και θ η φάση σε ακτίνια

► εάν θέσουμε $\Omega = 2\pi F$, τότε $x(n) = A \cos(2\pi F n + \theta), -\infty < n < \infty$

► όπου n αριθμός δείγματος, F η συχνότητα σε Hz (κύκλους ανά δείγμα)



Ιδιότητες ΗΣΔΧ

1. Ένα ΗΣΔΧ είναι περιοδικό **αν και μόνο αν** η συχνότητα του είναι ρητός αριθμός
 - ▶ Δηλαδή όταν υπάρχουν ακέραιοι m, k ώστε $F = m/k$
 - ▶ Απόδειξη: Έστω ότι υπάρχει N για το οποίο $x(n+N) = x(n)$
2. Δύο ΗΣΔΧ των οποίων οι κυκλικές συχνότητες διαφέρουν κατά 2π είναι ταυτόσημα
 - ▶ $\cos(\Omega n + \theta) = \cos[(\Omega + 2\pi)n + \theta]$
 - ▶ Επομένως, όλα τα $x_k(n) = A \cos(\Omega_k n + \theta)$,
 - ▶ με $\Omega_k = \Omega + 2k\pi$, για κάθε $k = \pm 1, \pm 2, \pm 3, \dots$, ταυτίζονται, ή αλλιώς (διαίρεση με 2π) όλα τα
 - ▶ $F_k = F_0 + k \Leftrightarrow f_k = f_0 + kf_s$ για κάθε $k = \pm 1, \pm 2, \pm 3, \dots$
 - ▶ Επομένως διαφορετικά είναι μόνο τα σήματα για τα οποία:
 - ▶ $-\pi \leq \Omega \leq \pi \Leftrightarrow -\frac{1}{2} \leq F \leq \frac{1}{2} \Leftrightarrow -\frac{1}{2}f_s \leq f \leq \frac{1}{2}f_s$
3. Η μέγιστη συχνότητα ενός ΗΣΔΧ είναι $F = \frac{1}{2}$ κύκλος ανά δείγμα, δηλαδή μια περίοδο να έχει δύο δείγματα:
 - ▶ Με άλλα λόγια η μέγιστη συχνότητα του αναλογικού σήματος είναι $f_s/2$

Άσκηση 1

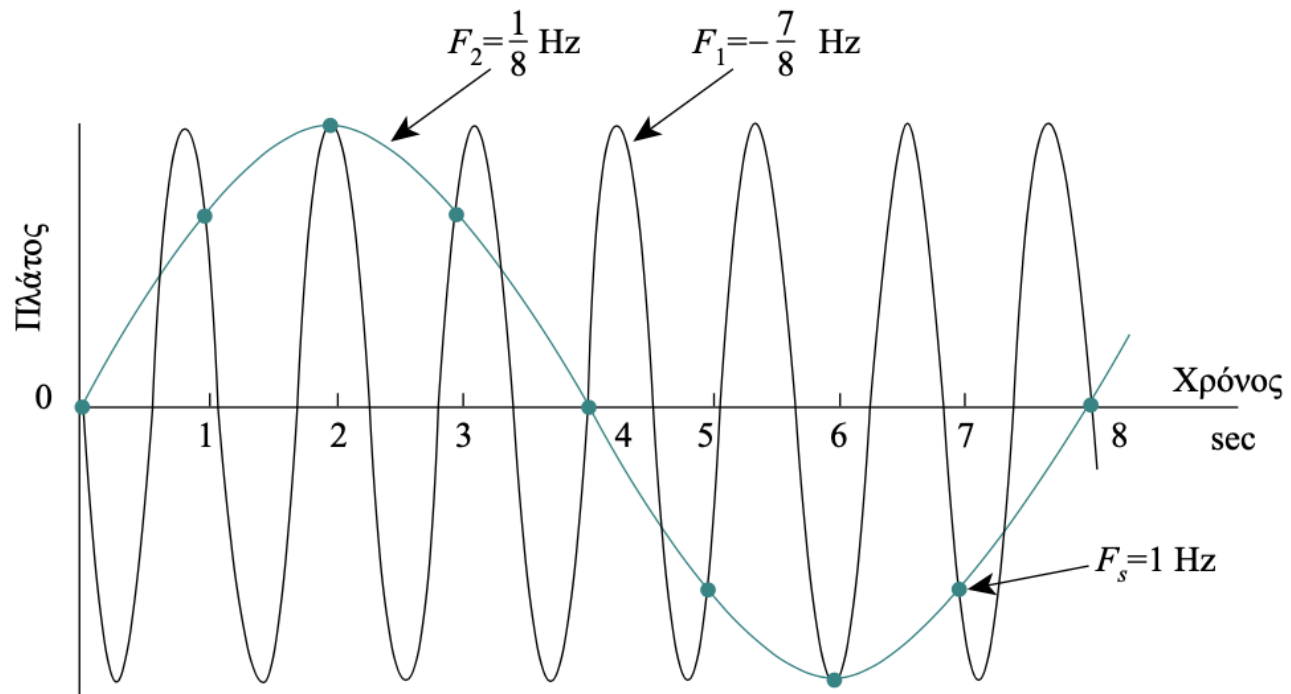
- ▶ Δίνονται τα αναλογικά σήματα
 - ▶ $x_1(t) = \sin(2\pi\frac{1}{8}t)$ και $x_2(t) = \sin(2\pi\frac{-7}{8}t)$
- ▶ Ποια τα σήματα διακριτού χρόνου που θα προκύψουν μετά τη δειγματοληψία αυτών με ρυθμό $f_s = 1$ Hz;

Σχηματικά (Άσκηση 1)

Σχήμα 1.7
Δειγματοληψία
αναλογικών σημά-
των συχνότητας

$$F_1 = \frac{1}{8} \text{ Hz},$$
$$F_2 = -\frac{7}{8} \text{ Hz}$$

με ρυθμό $F_s = 1 \text{ Hz}$



Σύνοψη - Συμπεράσματα (1/2)

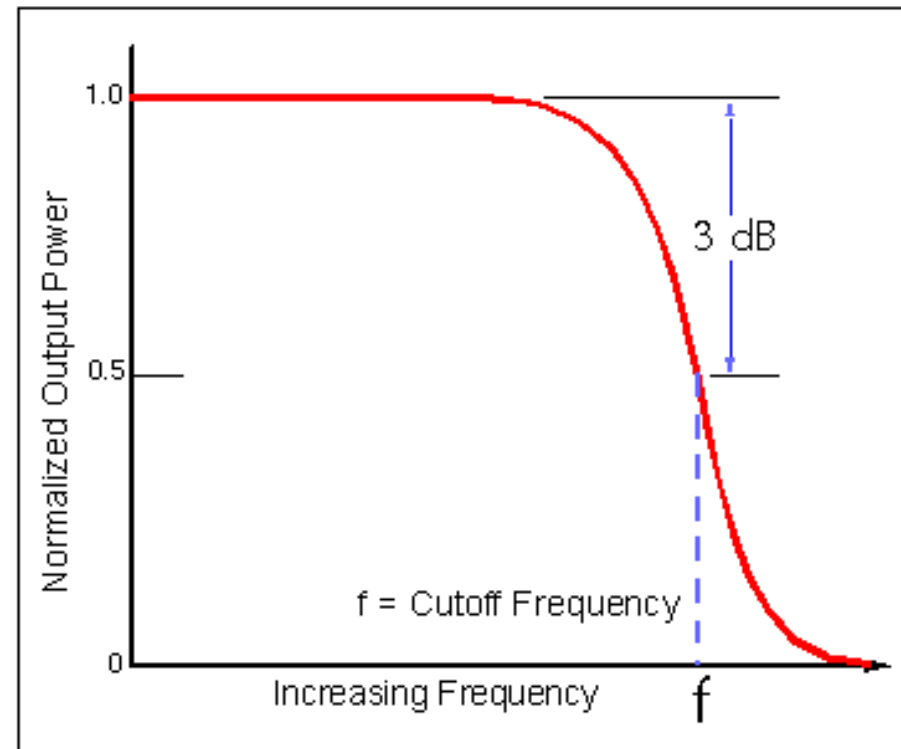
- ▶ Η ψηφιακή αναπαράσταση αναλογικών σημάτων είναι μια **μαθηματική επινόηση** του επιστημονικού χώρου της **Θεωρίας της Πληροφορίας**:
 - ▶ Η θεωρία πληροφορίας είναι τμήμα των εφαρμοσμένων μαθηματικών και ασχολείται με την ποσοτικοποίηση της πληροφορίας ώστε να καθίσταται ικανή για αποθήκευση, μετάδοση και επεξεργασία από υπολογιστικά συστήματα
- ▶ Ένα ψηφιακό σήμα ήχου είναι ένα **σύνολο από αριθμούς $x(n)$** που παρατάσσονται στο χρόνο σε σταθερή μεταξύ τους απόσταση: $T_s = 1/f_s$
- ▶ Οι αριθμοί αυτοί δεν αντιπροσωπεύουν μοναδικό αναλογικό σήμα, αλλά ένα πλήθος τέτοιων σημάτων
 - ▶ Ή μάλλον το άθροισμα πολλαπλών διαφορετικών σημάτων
 - ▶ Γεγονός που οφείλεται στην περιοδικότητα που εμφανίζεται στις κυκλικές συχνότητες του ψηφιακού σήματος:
 - ▶ $\Omega_k = \Omega_0 + 2k\pi \Leftrightarrow F_k = F_0 + k \Leftrightarrow f_k = f_0 + kf_s$, όπου $k = \pm 1, \pm 2, \dots$

Σύνοψη - Συμπεράσματα (2/2)

- ▶ Διαπιστώνουμε έτσι ότι:
 - ▶ Για κάθε συχνότητα του αναλογικού σήματος f_0 παίρνουμε ένα σύνολο συχνοτήτων της μορφής $f_k = f_0 \pm k f_s$
 - ▶ Εάν για κάθε συχνότητα f_0 στο αναλογικό σήμα ισχύει ότι, $f_0 \leq f_s/2$, τότε το φάσμα από 0 έως $f_s/2$ (Hz) αντιπροσωπεύει το φάσμα του αρχικού/αναλογικού σήματος
 - ▶ Έτσι μπορώ, κατά την μετατροπή του ψηφιακού σήματος πίσω σε αναλογικό να εφαρμόσω ένα φίλτρο χαμηλής διέλευσης με συχνότητας αποκοπής $f_s/2$, ώστε η έξοδος του φίλτρου να έχει μόνο τις συχνότητες του αναλογικού
 - ▶ **Προσοχή:** Η μετατροπή αναλογικού σε ψηφιακό σήμα δεν είναι πάντα αντιστρεπτή διαδικασία!

Πραγματικό φάσμα

- ▶ Αν το αρχικό μου σήμα (πριν την ψηφιοποίηση) έχει συχνότητα f_0
- ▶ Για να πάρω το *πραγματικό φάσμα* του (δηλαδή να αγνοήσω τις ψευδείς συχνότητες) πρέπει να βάλω ένα **βαθυπερατό φίλτρο (Low Pass Filter)** με συχνότητα αποκοπής $f_s/2$
 - ▶ Γιατί $f_s/2$ και όχι f_s ?
 - ▶ Λόγω των αρνητικών k



low pass filter

Βιβλιογραφία

- ▶ Σκόδρας Α. και Αναστασόπουλος Β., Ψηφιακή Επεξεργασία Εικόνων και Σημάτων, Πάτρα 2003

Εργαστηριακό Μέρος

Χ. Αλεξανδράκη

12/10/20

39

Περιβάλλον Εξάσκησης

- ▶ Γλώσσα Προγραμματισμού Python:
 - ▶ Είναι ελεύθερη και ανοιχτού κώδικα
 - ▶ Εξαιρετικά δημοφιλής
 - ▶ Εύκολη στην εκμάθηση
 - ▶ Η πλέον διαδεδομένη γλώσσα για προγραμματισμό εφαρμογών επεξεργασίας ηχητικού σήματος
 - ▶ Σήμερα θα ξεκινήσουμε με βασικές εργασίες ανάγνωσης/εγγραφής αρχείου ήχου
- ▶ Google Collaboratory
 - ▶ Ένα web περιβάλλον προγραμματισμού σε γλώσσα Python
 - ▶ Τα περισσότερα από τα πακέτα της Python είναι προ-εγκατεστημένα ώστε να μην απαιτείται εγκατάσταση στον ΗΥ
 - ▶ Καθώς η Python και τα πακέτα της διατίθενται free και open source, σε πολλές περιπτώσεις δημιουργούνται ασυμβατότητες μεταξύ διαφορετικών εκδόσεων των πακέτων και ανάλογα με το λειτουργικό σας σύστημα

Python Packages

- ▶ Τα πακέτα της python που θα χρησιμοποιήσουμε στο μάθημα:
 - ▶ Χειρισμός πινάκων, [NumPy](#):
 - ▶ Προεγκατεστημένο στο Google Colab
 - ▶ NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.
 - ▶ Δημιουργία γραφικών παραστάσεων, [Matplotlib](#):
 - ▶ Προεγκατεστημένο στο Google Colab
 - ▶ Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK+.
 - ▶ Επιστημονικός υπολογισμός, [SciPy](#)
 - ▶ Προεγκατεστημένο στο Google Colab
 - ▶ SciPy is a free and open-source Python library used for scientific computing and technical computing. SciPy contains modules for optimization, linear algebra, integration, interpolation, special functions, FFT, signal and image processing, ODE solvers and other tasks common in science and engineering
 - ▶ Ανάγνωση και εγγραφή αρχείων ήχου, [SoundFile](#)
 - ▶ Προεγκατεστημένο στο Google Colab
 - ▶ SoundFile can read and write sound files. SoundFile represents audio data as NumPy float arrays.

Αρχεία Ήχου

- ▶ Wav file:
 - ▶ Sound file format χωρίς απώλειες για την αποθήκευση ήχων σε σκληρό δίσκο
 - ▶ Κάθε δείγμα ήχου αποθηκεύεται ως ακέραιος αριθμός 16 bit (μερικές φορές επίσης ως 24 bit ακέραιος ή 32 bit float).
 - ▶ Τα αρχεία ήχου που θα χρησιμοποιούμε στις εργασίες είναι:
 - ▶ μονοφωνικά, έχουν ρυθμό δειγματοληψίας 44100Hz και ανάλυση δείγματος 16bit
 - ▶ Στις εργασίες θα διαβάζουμε αρχεία αυτά στην Python σε δεκαδική μορφή:
 - ▶ με εύρος στο διάστημα [-1.0, 1.0]
 - ▶ Μπορείτε να βρείτε τέτοια αρχεία στο eclass -> Εγγραφα
 - ▶ Εάν θελήσετε να χρησιμοποιήσετε άλλο δείγμα θα πρέπει να το μετατρέψετε στο παραπάνω format
 - ▶ Με κάποιο εργαλείο όπως το Audacity

Εργαστηριακές Ασκήσεις

- ▶ Να υλοποιήσετε σε Python τα ακόλουθα προγράμματα
 - ▶ **Άσκηση 1.1** Υπολογισμός και γραφική απόδοση των διακριτών τιμών μιας περιόδου ημιτονοειδούς σήματος. Εκτελέστε τον κώδικα για διαφορετικό πλήθος δειγμάτων, τι διαπιστώνετε?
 - ▶ **Άσκηση 1.2** Υπολογισμός ημιτονοειδούς σήματος δεδομένης διάρκειας (dur), πλάτους (amp) και συχνότητας (f0). Το σήμα να αποδοθεί γραφικά και να αποθηκευθεί στο αρχείο `output_audio/sinewave.wav`
 - ▶ **Άσκηση 1.3** Ανάγνωση και γραφική απόδοση ηχητικού αρχείου (π.χ. `input_audio/piano.wav`). Αλλάξτε τον οριζόντιο άξονα ώστε να απεικονίζει χρόνο
- ▶ Στοιχειώδεις επεμβάσεις σε ηχητικό σήμα που μπορείτε να υλοποιήσετε για περεταίρω εξάσκηση
 - ▶ Audio Reversal
 - ▶ Signal Gain/Attenuation
 - ▶ Normalization
 - ▶ DC Removal
 - ▶ Down-sampling/Up-sampling
- ▶ Ποιες από τις παραπάνω διεργασίες μπορούν να επιτευχθούν σε πραγματικό χρόνο?
 - ▶ Real-time (or online) signal processing refers to digital signal processing techniques that apply on an audio signal at the same time as this is acquired (e.g. from a microphone)
 - ▶ RT DSP are causal processes, i.e. they do not require to know the future of a signal