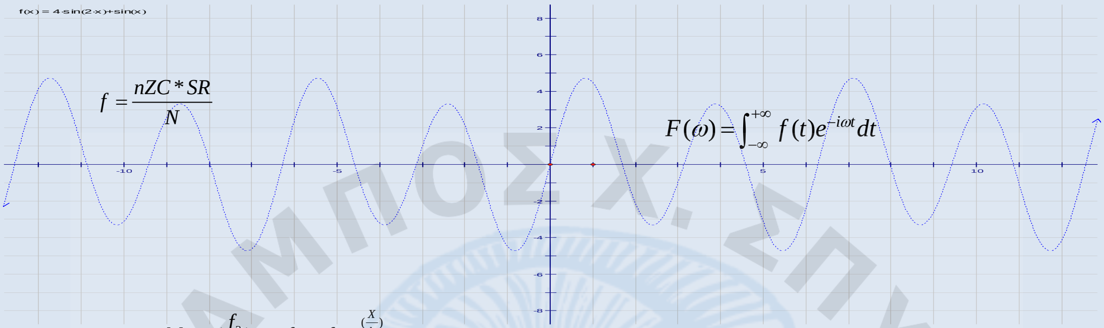
ΕΝΩΣΗ ΕΛΛΗΝΩΝ ΦΥΣΙΚΩΝ
Συνέδριο Μαρτίου – Απριλίου 2006

Χόρδισμα Οργάνων

με την μέθοδο των Zero Crossings



$$X = k \log_2(\frac{f_2}{f_1}) \Rightarrow f_2 = f_1 \cdot 2^{(\frac{X}{k})}$$

Μιχαήλ Γ. Χουρδάκης[1] - Χαράλαμπος Χ. Σπυρίδης[2]

1 Μουσικολόγος, Υποψήφιος Διδάκτωρ Τμήματος Μουσικών Σπουδών Πανεπιστημίου Αθηνών.
2 Καθηγητής Μουσικής Ακουστικής, Πληροφορικής –Διευθυντής Εργαστηρίου Μουσικής Ακουστικής τεχνολογίας, Τμήματος Μουσικών Σπουδών Πανεπιστημίου Αθηνών.

– 1 –

Κατά την εισήγησή μας θα περιγραφούν τα πλεονεκτήματα των time-domain μεθόδων και ειδικότερα των Zero Crossings, για την ανίχνευση του τονικού ύψους ενός μουσικού ήχου και θα αναλυθεί η εφαρμογή της στο χόρδισμα των ευρωπαϊκών και των βυζαντινών μουσικών οργάνων.

Προκειμένου να καταστεί αντιληπτή η μέθοδος των Zero Crossings, θα παρουσιασθούν τα μειονεκτήματα των υπαρχουσών μεθόδων ανίχνευσης του τονικού ύψους ενός μουσικού ήχου, οι οποίες βασίζονται στην ανάλυση σε frequency – domain, δηλαδή στο πεδίο συχνότητας/εντάσεως, με το μετασχηματισμό Fourier ή τις παραλλαγές του.

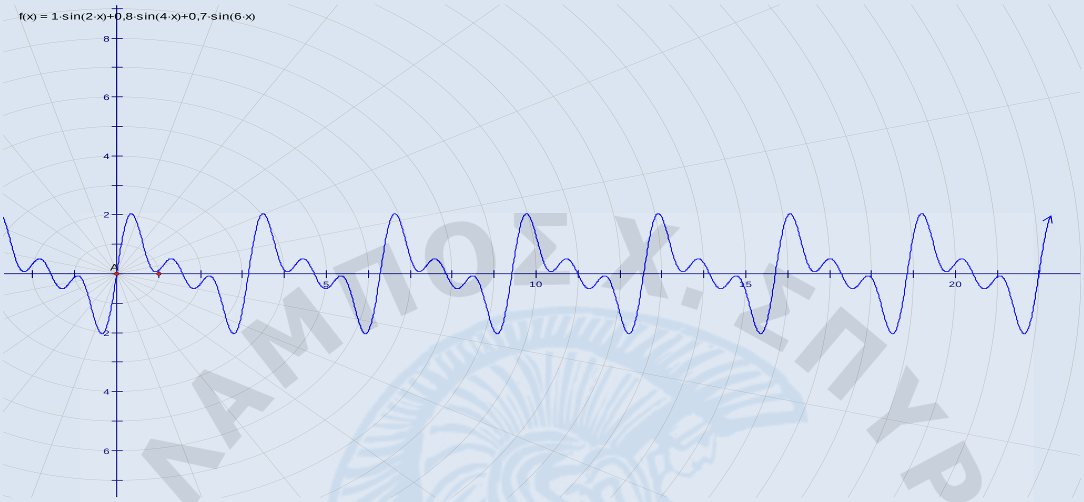
(00:40)

– 2 –

Το σήμα που παράγεται από κάποιο ηχογόνο σώμα (φωνή, όργανο), έχει την μορφή:

Το ηχητικό σήμα

$$S(t) = A_0 \sin(2\pi f_0 t) + A_1 \sin(2\pi f_1 t) + \ldots A_\nu \sin(2\pi f_\nu t)$$



οπου Α1, Α2...Αn είναι το πλάτος των ταλαντώσεων, που προκύπτουν, και f1,f2...fn είναι η συχνότητά τους. Ένα ψηφιακό σήμα που έχει ηχογραφηθεί με συχνότητα δειγματοληψίας S, θα περιέχει, σύμφωνα με το θεώρημα Niquist, ταλαντώσεις των οποίων η συχνότητα δεν θα ξεπερνάει το S/2.

Το σήμα αυτό είναι περιοδικό, εφ' όσον συχνότητες fn n=2, 3, 4,... ήταν ακέραια πολλαπλάσια της αρχικής f_1, αφού η για την συνάρτηση $S(t)$ ισχύει $S(t) = 0, t \in \mathbb{Z}$, αφού $\sin(x) = 0 \Leftrightarrow \sin(\nu x) = 0, \nu \in \mathbb{N}$.

(01:40)

Ο συνεχής μετασχηματισμός Fourier F(ω) (ή για συντομία CFT – Continuous Fourier Transform) μιας συνεχούς συνάρτησης f(t) είναι μία συνάρτηση συχνότητας/εντάσεως , και ορίζεται ώς εξής (Δείκτης).

Στο ψηφιακό σήμα, δεν μπορούμε να εφαρμόσουμε τον CFT, αλλά εφαρμόζουμε τον Διακριτό Μετασχηματισμό (DFT) (Δείκτης):

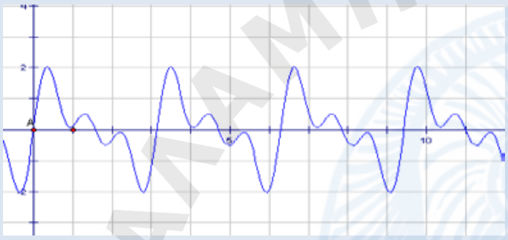
Μετασχηματισμός Fourier.

Ο συνεχής μετασχηματισμός Fourier:

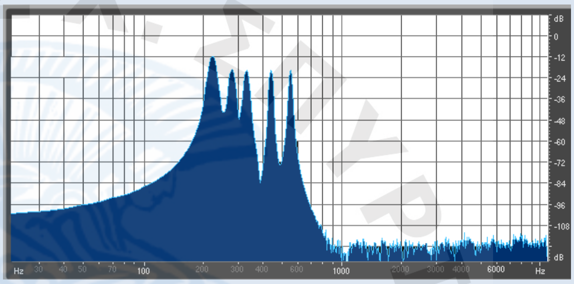
$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt$$

Ο διακριτός μετασχηματισμός Fourier:

$$F(\omega) = \sum_{t=-\infty}^{+oo} f(t) e^{-i\omega t}$$



Σήμα



Μετασχηματισμός (Συνεχής/Διακριτός)

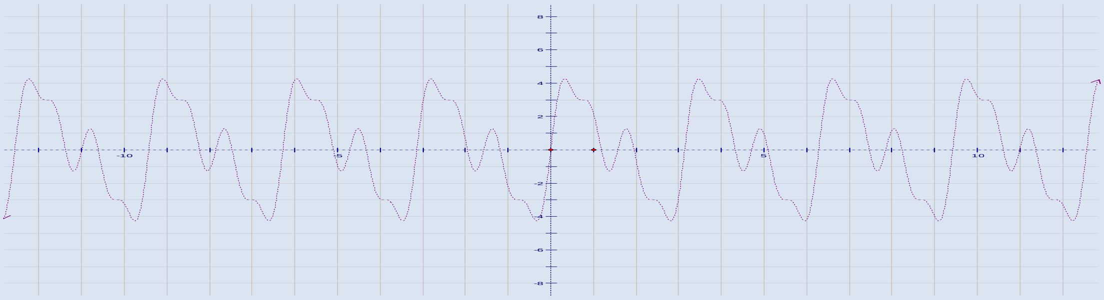
Ο παραπάνω μετασχηματισμός Fourier, δίνει όλες τις συχνότητες του σήματος και τις εντάσεις τους. Για να υπολογιστεί, λοιπον, η συχνότητα του σήματος σε ένα συγκεκριμένο σημείο, θα υπέθετε κανείς ότι αρκεί να βρεθεί ο μετασχηματισμός Fourier στην θέση του σήματος, που μας ενδιαφέρει.

(03:00)

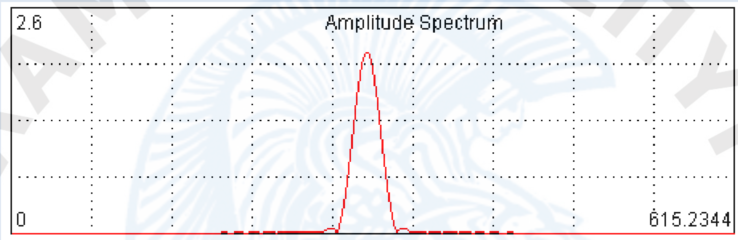
Σύμφωνα, όμως, με το θεώρημα της αβεβαιότητας του Heisenberg , είναι αδύνατο να προσδιορίσουμε την συχνότητα σε ένα σημείο του σήματος, αλλά χρειαζόμαστε τουλάχιστον ένα τμήμα αυτού του σήματος (window).

Επομένως, θα πρέπει να επιλεγεί ένα μικρό τμήμα του σήματος και να εφαρμοσθεί σε αυτό ο μετασχηματισμός Fourier.
Για παράδειγμα, σε μία περιοδική συνάρτηση ο μετασχηματισμός έχει καλά αποτελέσματα για την ανάλυση που απαιτείται.

Περιοδική συνάρτηση $f(t) = 3\sin(2\pi t) + 2\sin(4\pi t) + \sin(8\pi t)$



Διάγραμμα Fourier

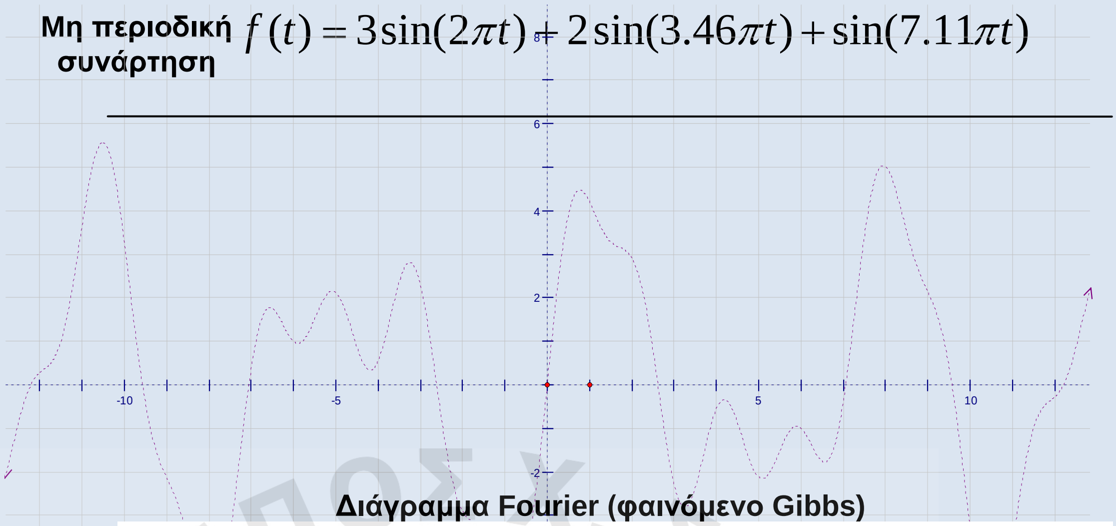


(04:00)

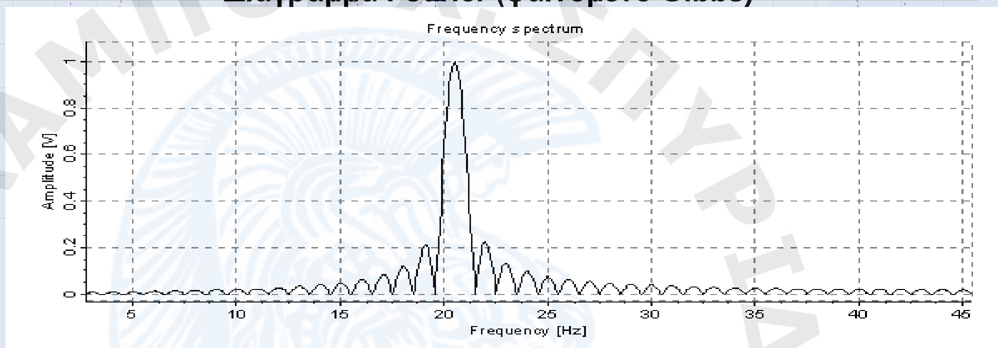
Η χρήση του παραπάνω μετασχηματισμού Fourier επιβάλλει τους παρακάτω περιορισμούς:

- Ο μετασχηματισμός Fourier είναι κατάλληλος για περιοδικά σήματα. Στην πράξη, κανένα σήμα δεν είναι περιοδικό. Ετσι, ο μετασχηματισμός Fourier παράγει ανακρίβειες (γνωστές σαν leaks) , δηλαδή στοιχεία μη ακριβή για κοντινές συχνότητες. Το φαινόμενο αυτό (φαινόμενο Gibbs) παρουσιάζεται επειδή στην επιλογή του τμήματος, που χρειαζόμαστε, αναγκαστικά κόβονται ορισμένες από τις συχνότητες του σήματος, ακριβώς επειδή το σήμα δεν είναι περιοδικό.
  Για παράδειγμα , η συνάρτηση

Μη περιοδική συνάρτηση $f(t) = 3\sin(2\pi t) + 2\sin(3.46\pi t) + \sin(7.11\pi t)$



Διάγραμμα Fourier (φαινόμενο Gibbs)



δεν είναι περιοδική, και έτσι, λόγω του φαινομένου Gibbs ο μετασχηματισμός Fourier έχει leakage γύρω από τις συχνότητες που πρέπει να ανιχνευτούν.

- Ο δεύτερος περιορισμός είναι πιο σημαντικός. Ο διακριτός μετασχηματισμός Fourier, που χρησιμοποιείται στους Η/Υ, δέχεται σαν είσοδο Χ δείγματα του σήματος και επιστρέφει Χ μιγαδικούς αριθμους.

Διακριτός μετασχηματισμός Fourier

Input — Ψηφιακό σήμα $f(t)$

Με N πληθος ψηφίων σε δειγματοληψία SR, διάρκειας N/SR sec

Output — Πίνακας μιγαδικών 1xN $A_z[N]$

Για κάθε $z = \mathrm{A}_z[i], i \in [0, N-1]$

όπου $$f_i = \frac{iSR}{N}$$

ισχύει $$If_i = |z|$$

Το μέτρο του καθενός μιγαδικου αριθμού z, υποδηλώνει την Ένταση X διαφορετικών συχνοτήτων, οι οποίες βρίσκονται από τον αλγόριθμο $f_i = \frac{iSR}{N}$
όπου SR η συχνότητα δειγματοληψίας του σήματος. Από την περιγραφη αυτή γίνεται αμέσως σαφές ότι ο αλγόριθμος μπορεί να μας δώσει πληροφορίες μόνο για τις συχνότητες, οι οποίες προκύπτουν από τον παραπάνω τύπο για ακέραιο i και όχι για οποιαδήποτε συχνότητα από 0 μέχρι SR/2, που περιέχεται στο σήμα. Η ενέργεια των υπολοίπων συχνοτήτων διοχετεύεται στις συγκεκριμένες συχνότητες, που έχει ο διακριτός μετασχηματισμός. Αυτό, σε συνάρτηση και με το leakage που είδαμε προηγουμένως, ουσιαστικά μας απαγορεύει να χρησιμοποιήσουμε τον μετασχηματισμό Fourier για μετρήσεις ακριβείς, πράγμα αναγκαίο στο χόρδισμα των μουσικών οργάνων. Για παράδειγμα: Έστω μία χορδή Λα που πρέπει να χορδιστεί σε συχνότητα 440 Hz. Ας υποθέσουμε ότι είναι χορδισμένη στην συχνότητα 441 Hz. Εξαιτίας των περιορισμών του Fourier, η ενέργεια της συχνότητας 441 Hz θα διοχετευτεί στην συχνότητα 440 Hz, και έτσι ο Fourier θα μας δώσει υπόδειξη ότι η χορδή έχει συχνότητα 440 Hz, αντί για την 441 Hz που έχει κανονικά.

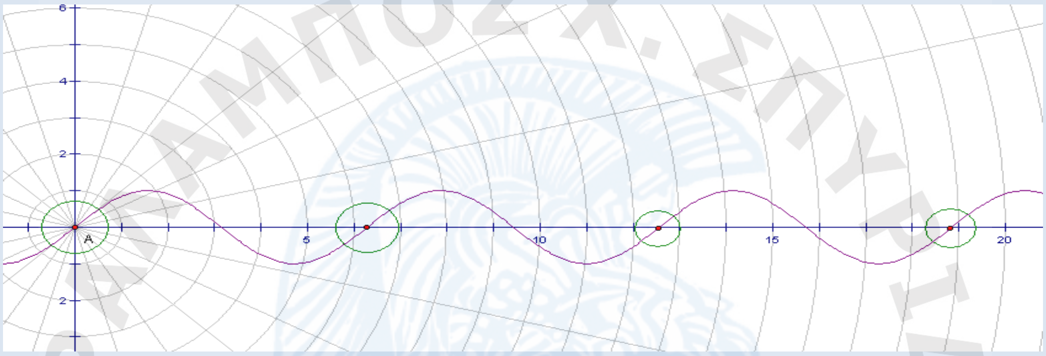
<7:30 ΜΕΧΡΙ ΕΔΩ>

Time domain

(07:30)

ΟΡΙΣΜΟΣ των ZC.

Zero Crossings



$$f(z)=0$$

Για κάθε $z_1 \in [z-\delta, z), \delta > 0$ **ισχύει** $f(z_1)<0$

Για κάθε $z_2 \in (z, z+\delta], \delta > 0$ **ισχύει** $f(z_2)>0$

Έστω μία συνεχής συνάρτηση $f(x) \to \Re, x \in \Re$.
Ορίζουμε σαν **Zero Crossing** κάθε σημείο (z,f(z)) για το οποίο ισχύει

- $f(z)=0$
- **για κάθε** $z_1 \in [z-\delta, z), \delta > 0$ ισχύει $f(z_1)<0$
- **για κάθε** $z_2 \in (z, z+\delta], \delta > 0$ ισχύει $f(z_2)>0$

Έστω μία συνάρτηση g = f με πεδίο ορισμού το Q.

Ορίζουμε σαν **Zero Crossing της συνάρτησης g,** κάθε σημείο (z,f(z)) το οποίο είναι Zero Crossing στην f. Το σημείο z, δεν είναι απαραίτητο να ανήκει στην g, αφού αυτή είναι διακριτή.

Η μελέτη του σήματος στο πεδίο του χρόνου, γίνεται με την εφαρμογή ορισμένων ψηφιακών φίλτρων στο σήμα και με την στατιστική ανάλυσή του σε σημεία έτσι, ώστε να προκύψει ένα απλοποιημένο σήμα, το οποίο θα περιέχει είτε μόνη τη βασική συχνότητα, είτε θα είναι εύκολος ο προσδιορισμός της.

Τα βασικά πλεονεκτήματα των αλγορίθμων, που βασίζονται στο Time Domain, είναι:

1.Δίνουν καλύτερα αποτελέσματα σε μη περιοδικά σήματα, επειδή σαρώνουν το σήμα στον άξονα x-x' (άξονας χρόνου), κάνοντας εκτιμήσεις για το πώς θα ήταν το σήμα, εάν ήταν περιοδικό.

2. Λειτουργούν ικανοποιητικά ακόμα και σε ελάχιστο μήκος σήματος, επειδή δε χρειάζεται να έχουν ολοκληρωθεί οι ταλαντώσεις πλήρως για την ανάλυσή του.

3. Τα αποτελέσματά τους είναι ακριβή και μάλιστα γίνονται καλύτερα, όσο μεγαλύτερη ταχύτητα δειγματοληψίας χρησιμοποιήσουμε.

Το βασικό μειονέκτημα των αλγορίθμων, που βασίζονται στο Time Domain, είναι η ταχύτητα. Για να λειτουργήσουν ικανοποιητικά, πρέπει το σήμα να φιλτραρισθεί κατάλληλα. Μεταξύ άλλων, χρησιμοποιούνται τα λεγόμενα χαμηλοπερατά φίλτρα (**low-pass)**, δηλαδή φίλτρα τα οποία αποκόπτουν τις συχνότητες του σήματος πάνω από κάποια συγκεκριμένη τιμή συχνότητας.

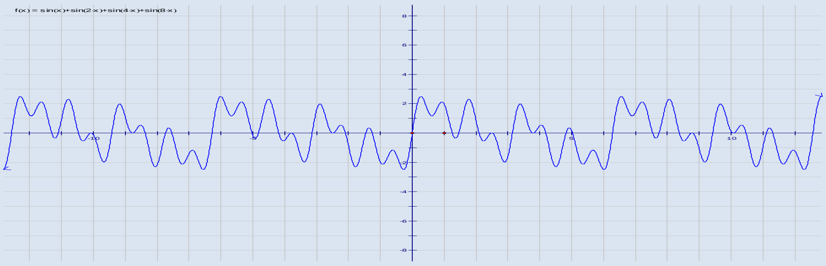
Οι επιλογές των ψηφιακών φίλτρων που έχουμε είναι 2:

- Τα φίλτρα τύπου IIR , δηλαδή τα φίλτρα που βασίζονται στην επανάληψη (recursion).
- Τα φίλτρα τύπου FIR, δηλαδή τα φίλτρα που βασίζονται στην συνέλιξη (convolution)

Τα IIR φίλτρα είναι μεν ταχύτατα, αλλά είναι εξαιρετικά ασταθή και συνήθως η απόδοσή τους δεν επαρκεί για την πλειοψηφία των σημάτων που επεξεργαζόμαστε.

Για την ακρίβεια που χρειαζόμαστε, χρησιμοποιούνται low-pass φίλτρα Kaiser τύπου Windowed – SINC FIR. Αυτά τα φίλτρα, ανάλογα και με τα προβλήματα που υπάρχουν στο σήμα, δίνουν πολύ καλά αποτελέσματα, αλλά είναι χρονοβόρα.

Αρχικό σήμα:

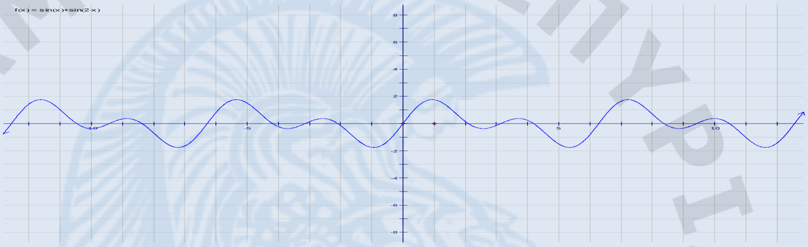


Φίλτρο FIR

$$f(t)*g(t)=\int_{-\infty}^{+\infty} f(\tau)g(t-\tau)dt, \tau \in \square$$

$$s(t)*h(t)=\sum_{k=-\infty}^{+\infty} s_k h_{j-k}, j \in Z+, 0 \leq j < (N+M-1)$$

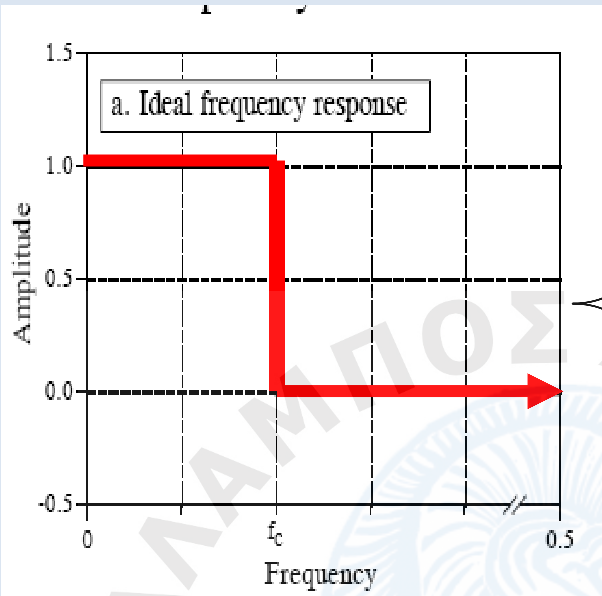
Απλοποιημένο σήμα



Όσο πιο μεγάλο και πολύπλοκο το φίλτρο που θα χρησιμοποιηθεί, τόσο καλύτερη η απόδοση του σήματος και τόσο πιο χρονοβόρα η εφαρμογή του.

Για παράδειγμα, η ιδανική frequency-domain μορφή ενός low-pass φίλτρου έχει την παρακάτω μορφή:

Ιδανικό ψηφιακό φίλτρο lowpass

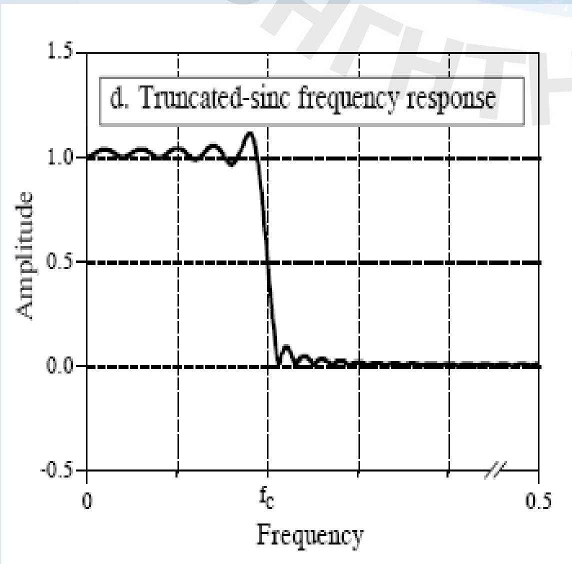


- Passband ripple: 0
- Transition band length: 0
- Stop band ripple: 0.
- Stop band att: -Inf
- Convolution time needed: +Inf
- Coefficients: +Inf

Το παραπάνω φίλτρο πρέπει να κόψει από το σήμα όλες τις συχνότητες πάνω από την fc. Η ιδανική μορφή είναι στην εικόνα όπου όλες οι συχνότητες πριν την fc πολλαπλασιάζονται με το 1.0 (δηλαδή μένουν ανέπαφες) και όλες οι συχνότητες μετά την fc πολλαπλασιάζονται με το 0.0 (δηλαδή εξαφανίζονται).

Στην πράξη όμως, το φίλτρο έχει την παρακάτω μορφή:

Ρεαλιστικό ψηφιακό φίλτρο lowpass

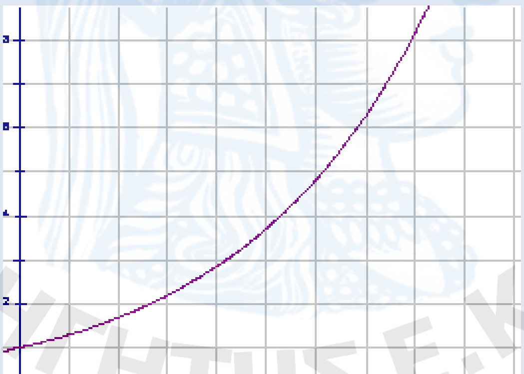


- Passband ripple: (0-δ,0+δ)
- Transition band length: δ > 0
- Stop band ripple: δ > 0.
- Stop band att: δ > -Inf
- Convolution time needed: ~
- Coefficients: N

1. Η pass-band (περιοχή που οι συχνότητες μένουν ανέπαφες) έχει ρυτιδώσεις (ripples), διότι όλες οι συχνότητες δεν πολλαπλασιάζονται με το 1.0 , αλλά με αριθμούς κοντά στο 1.0. Αποτέλεσμα αυτού του γεγονότος είναι να αλλοιώνονται οι συχνότητες που πρέπει να μείνουν αναλλοίωτες.
2. Η ζώνη μετάβασης (transition band) -η περιοχή γύρω από την fc- είναι μεγάλου μήκους. Η ζώνη μετάβασης είναι η περιοχή $[fc-\delta, fc+\delta], \delta > 0$ (το ιδανικό θα ήταν δ = 0, που είναι αδύνατον) και το δ είναι αρκετά μεγάλο έτσι, ώστε πολλές από τις συχνότητες αριστερά από την fc να πολλαπλασιάζονται με αριθμό αρκετά μικρότερο από το 1.0, και πολλές συχνότητες δεξιά της fc να πολλαπλασιάζονται με αριθμό αρκετά μεγαλύτερο από το 0.
3. Η ζώνη αποκοπής (stop-band) έχει χαμηλή εξασθένιση (attenuation), διότι όλες οι συχνότητες δεν πολλαπλασιάζονται με το 0 , αλλά με αριθμούς θετικούς κοντά στο 0. Το stop-band attenuation, λοιπόν, δεν είναι ποτέ άπειρο και ποτέ οι συχνότητες δεν χάνουν εντελώς την ενέργειά τους.

Zero Crossings

Χρόνος φιλτραρίσματος



Ποιότητα φιλτραρίσματος (N->+Inf)

Ένα ψηφιακό φίλτρο είναι καλό, όταν τα παραπανω φαινόμενα ειναι αμελητέα. Για να είναι, όμως, τα φαινόμενα αμελητέα, πρέπει το φίλτρο να είναι μεγάλο, όσον αφορά στο πλήθος των συντελεστών (coefficients), δηλαδή να είναι μεγάλο σε έκταση σήματος και, έτσι, να είναι χρονοβόρα η εφαρμογή του.

Η μέθοδος των Zero Crossings, η οποία εκτός των άλλων χρησιμοποιεί και φίλτρα, είναι πολύ κατάλληλη για την ανάλυση που χρειαζόμαστε, αλλά όχι σε πραγματικό χρόνο. Με την βοήθεια των Η/Υ σήμερα η ανάλυση και το χόρδισμα με τη μέθοδο των Zero Crossings μπορεί να γίνει σε αποδεκτό χρονικό διάστημα. **(15:00) <Πειραματική Επίδειξη>**