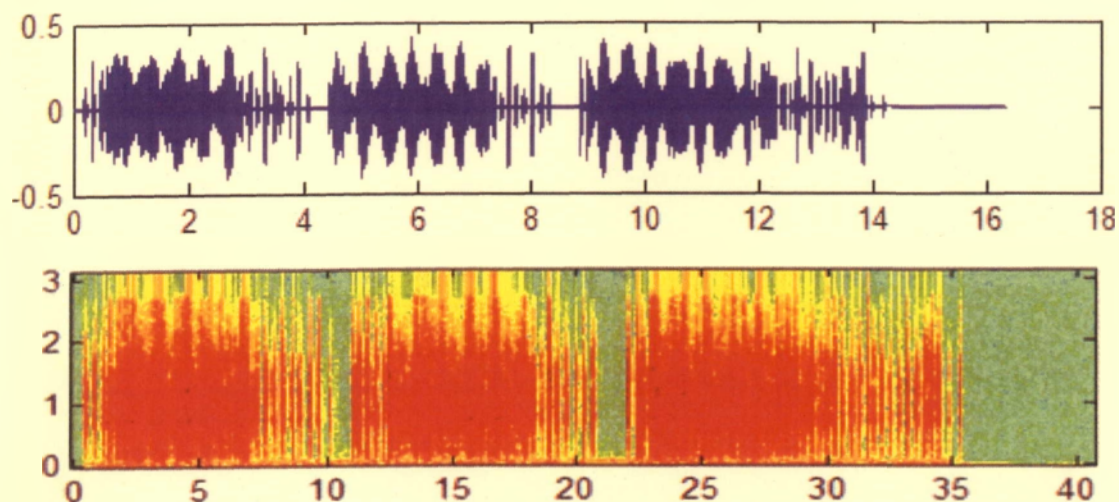


Α.Τ.Ε.Ι ΚΑΛΑΜΑΤΑΣ - ΠΑΡΑΡΤΗΜΑ ΣΠΑΡΤΗΣ



ΤΜΗΜΑ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΤΟΜΕΑΣ  
ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΔΙΚΤΥΩΝ

## Εξαγωγή Χαρακτηριστικών για Αναγνώριση Ήχου



Πτυχιακή Εργασία  
του  
Απόστολου Φαλούκα

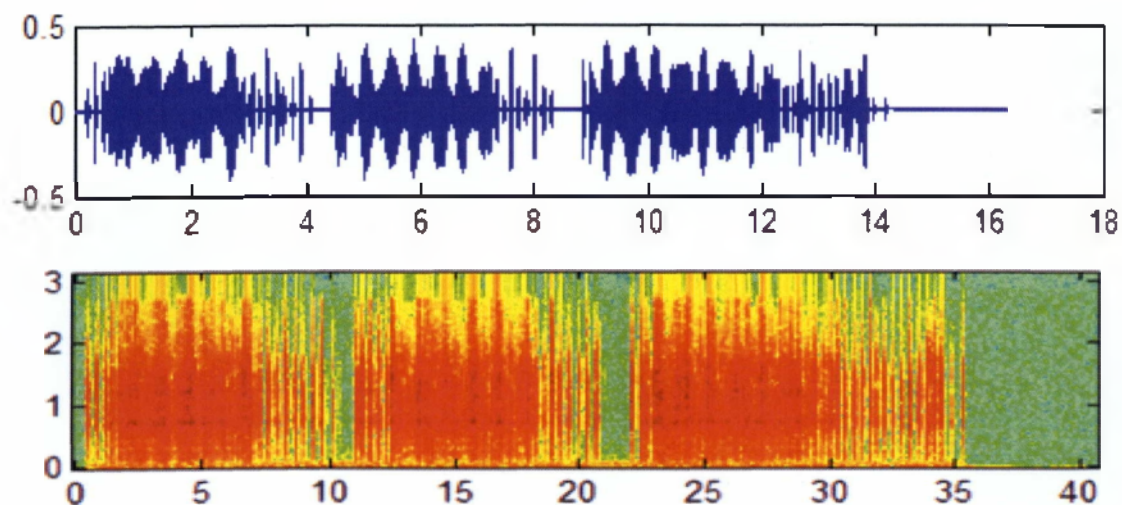
Επιβλέπων Καθηγητής  
Καθ. Ιωάννης Παρασκευάς

Σπάρτη 2010



ΤΜΗΜΑ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΤΟΜΕΑΣ  
ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΔΙΚΤΥΩΝ

## Εξαγωγή Χαρακτηριστικών για Αναγνώριση Ήχου



Πτυχιακή Εργασία  
του  
Απόστολου Φαλούκα

Επιβλέπων Καθηγητής  
Καθ. Ιωάννης Παρασκευάς

Σπάρτη 2010

Θα ήθελα ειλικρινά να ευχαριστήσω τον καθηγητή Ιωάννη Παρασκευά για την καθοδήγηση και την προθυμία του καθ' όλο το διάστημα εκπόνησης της πτυχιακής μου εργασίας. Επίσης θα ήθελα να ευχαριστήσω την οικογένειά μου και όλους εκείνους που με αγάπη στέκονται δίπλα μου.

# Περιεχόμενα

Λίστα Σχημάτων .....	3
Πρόλογος.....	6
Εισαγωγή.....	7
Οργάνωση της Πτυχιακής Εργασίας .....	8
Κεφάλαιο 1.....	10
1.1 Διακριτός Μετασχηματισμός Fourier (DFT).....	10
1.2 Ταχύς μετασχηματισμός Fourier (FFT- Fast Fourier Transform).....	15
1.2.1 Μέτρο και Φάση του σήματος.....	16
1.3 Διαρροή συχνοτήτων ( spectral leakage) .....	18
1.3.1 Παράθυρο τύπου Χάμινγκ (Hamming window).....	20
1.3.2 Παράθυρο τύπου Μπλάκμαν (Blackman window).....	21
1.3.3 Παράθυρο τύπου Χάνινγκ (Hanning window) .....	22
Κεφάλαιο 2.....	23
2.1 Μετασχηματισμοί FT, STFT .....	23
2.2 Παραδείγματα χρονοσυχνοτικής ανάλυσης χρησιμοποιώντας τον αλγόριθμο FFT.....	31
Κεφάλαιο 3.....	36
3.1 Μηχανική Μάθηση .....	37
3.1.1 Μάθηση με Επίβλεψη .....	39
3.1.1.α Μάθηση κατά Περίπτωση.....	41
3.1.2 Άλλες Τεχνικές Μάθησης με Επίβλεψη .....	43
3.1.3 Μάθηση Χωρίς Επίβλεψη .....	43
3.2 Ανακάλυψη Γνώσης σε Βάσεις Δεδομένων .....	43
3.2.1 Τα Στάδια της Ανακάλυψης Γνώσης .....	45
Κεφάλαιο 4.....	49
Συμπεράσματα .....	68
ΒΙΒΛΙΟΓΡΑΦΙΑ .....	69



# Λίστα Σχημάτων

Σχήμα 1.1 Ιδανικός Δειγματολήπτης .....	11
Σχήμα 1.2: Μέτρο του σήματος .....	16
Σχήμα 1.3: Φάση του σήματος .....	17
Σχήμα 1.4: Γραφική παρουσίαση του θεωρήματος Δυϊκότητας .....	19
Σχήμα 1.5: Παράθυρο τύπου Hamming ( $\alpha$ ) στο πεδίο του χρόνου και ( $\beta$ ) στο πεδίο των συχνοτήτων.....	21
Σχήμα 1.6: Παράθυρο τύπου Blackman ( $\alpha$ ) στο πεδίο του χρόνου και ( $\beta$ ) στο πεδίο των συχνοτήτων .....	21
Σχήμα 1.7: Παράθυρο τύπου Hanning ( $\alpha$ ) στο πεδίο του χρόνου και ( $\beta$ ) στο πεδίο των συχνοτήτων.....	22
Σχήμα 2.1: Σήματα στο πεδίο του χρόνου.....	23
Σχήμα 2.2: Σήματα στο πεδίο συχνοτήτων .....	24
Σχήμα 2.3: Λειτουργία Μετασχηματισμού Fourier .....	26
Σχήμα 2.4: Fourier σήματος chirp .....	27
Σχήμα 2.5: STFT σήματος chirp .....	28
Σχήμα 2.6: Διακριτότητα συχνότητας – χρόνου.....	30
Σχήμα 3.1: Προσδιορισμός κατηγορίας με βάση τον 1 και τους 5 κοντινότερους γείτονες.....	43
Σχήμα 3.2: Τα βασικά στάδια της διαδικασίας ανακάλυψης γνώσης.....	45
Σχήμα 4.1: Περιβάλλον Cool edit pro 2.1.....	50
Σχήμα 4.2: Εισαγωγή ήχου Track09_Animal.....	50
Σχήμα 4.3: Track09_Animal.....	51
Σχήμα 4.4: Convert Sample Type ( $\alpha$ ).....	51
Σχήμα 4.5: Convert Sample Type ( $\beta$ ).....	52
Σχήμα 4.6: Το σήμα (ήχος) μετά την μετατροπή.....	52
Σχήμα 4.7: Convert Sample Type 48000Hz, Mono, 16-bit.....	52
Σχήμα 4.8: Track09_Animal1 .....	53
Σχήμα 4.9: Track09_Animal2.....	53
Σχήμα 4.10: Track09_Animal6.....	53
Σχήμα 4.11: Track09_Animal10.....	53

Σχήμα 4.12: Track74_Geophysical.....	54
Σχήμα 4.13: Track74_Geophysical10.....	54
Σχήμα 4.14: Περιβάλλον Matlab.....	55
Σχήμα 4.15: File -> New -> M-file.....	56
Σχήμα 4.16: Editor.....	56
Σχήμα 4.17: Εντολές στο Editor.....	56
Σχήμα 4.18: Κουμπιά Αποθήκευσης και Εκτέλεσης .....	57
Σχήμα 4.19: Γράφημα στο Πεδίο του Χρόνου του σήματος (ήχου) Track09_Animal.wav.....	57
Σχήμα 4.20: Track74_Geophysical στο πεδίο του χρόνου.....	57
Σχήμα 4.21: Track09_Animal & Track74_Geophysical στο πεδίο του χρόνου.....	58
Σχήμα 4.22: Track09_Animal & Track74_Geophysical στο πεδίο του χρόνου.....	58
Σχήμα 4.23: Κώδικας για την εμφάνιση του spectrogram για τον ήχο Track09_Animal.....	59
Σχήμα 4.24: Track09_Animal στο πεδίο των συχνοτήτων.....	60
Σχήμα 4.25: Track74_Geophysical στο πεδίο των συχνοτήτων.....	60
Σχήμα 4.26: Κώδικας εμφάνισης spectrogram της πρόσθεσης των ήχων Track09_Animal & Track74_Geophysical.....	61
Σχήμα 4.27: Track09_Animal & Track74_Geophysical στο πεδίο των συχνοτήτων.....	61
Σχήμα 4.28: Track45_Human (Πυροβολισμός) στο πεδίο του χρόνου.....	62
Σχήμα 4.29: Track90_Geophysical (Θυελλώδεις άνεμοι) στο πεδίο του χρόνου.....	62
Σχήμα 4.30: Track45_Human και Track90_Geophysical στο πεδίο του χρόνου.....	62
Σχήμα 4.31: Track45_Human (Πυροβολισμός) στο πεδίο των συχνοτήτων.....	62
Σχήμα 4.32: Track90_Geophysical (Θυελλώδεις άνεμοι) στο πεδίο των συχνοτήτων.....	63
Σχήμα 4.33: Track45_Human και Track90_Geophysical στο πεδίο των συχνοτήτων.....	63
Σχήμα 4.34: Track15_Animal (Σκύλος) στο πεδίο του χρόνου.....	63
Σχήμα 4.35: Track86_Human (Αεροπλάνο) στο πεδίο του χρόνου.....	63
Σχήμα 4.36: Track15_Animal και Track86_Human στο πεδίο του χρόνου.....	64
Σχήμα 4.37: Track15_Animal (Σκύλος) στο πεδίο των συχνοτήτων.....	64
Σχήμα 4.38: Track86_Human (Αεροπλάνο) στο πεδίο των συχνοτήτων.....	64

Σχήμα 4.39: Track15_Animal και Track86_Human στο πεδίο των συχνοτήτων.....	64
Σχήμα 4.40: Track57_Animal (Κροταλίας) στο πεδίο του χρόνου.....	65
Σχήμα 4.41: Track97_Geophysical (Φωτιά) στο πεδίο του χρόνου.....	65
Σχήμα 4.42: Track57_Animal και Track97_Geophysical στο πεδίο του χρόνου.....	65
Σχήμα 4.43: Track57_Animal (Κροταλίας) στο πεδίο των συχνοτήτων.....	65
Σχήμα 4.44: Track97_Geophysical (Φωτιά) στο πεδίο των συχνοτήτων .....	66
Σχήμα 4.45: Track57_Animal και Track97_Geophysical στο πεδίο των συχνοτήτων.....	66
Σχήμα 4.46: Track09_Animal (Κοράκι) στο πεδίο του χρόνου.....	66
Σχήμα 4.47: Track15_Human (Αυτοκίνητο) στο πεδίο του χρόνου.....	66
Σχήμα 4.48: Track09_Animal και Track15_Human στο πεδίο του χρόνου.....	67
Σχήμα 4.49: Track09_Animal (Κοράκι) στο πεδίο των συχνοτήτων.....	67
Σχήμα 4.50: Track15_Human (Αυτοκίνητο) στο πεδίο των συχνοτήτων .....	67
Σχήμα 4.51: Track09_Animal και Track15_Human στο πεδίο των συχνοτήτων.....	67

## Πρόλογος

«Ηχητικά Τοπία», είναι οι χάρτες που απεικονίζουν το ηχητικό περιεχόμενο μιας περιοχής σε κάποια χρονική στιγμή. Τα ηχητικά χαρακτηριστικά περιλαμβάνουν πληροφορίες που μπορούν να συνδυαστούν με τα οπτικά χαρακτηριστικά ενός τοπίου, με σκοπό την παραγωγή χρήσιμων οικολογικών παρατηρήσεων / δεδομένων, για τομείς περιβαλλοντικού ή οικολογικού ενδιαφέροντος. Αυτά περιλαμβάνουν την παρακολούθηση της άγριας ζωής, τις ανθρώπινες δραστηριότητες των περιοχών αυτών κλπ. όπως εξελίσσονται στο χρόνο. Η παρούσα εργασία, έχει ως σκοπό την εξαγωγή χαρακτηριστικών για την ταξινόμηση περιβαλλοντικών ήχων. Οι αναπαραστάσεις των ηχητικών σημάτων - που λαμβάνονται ως βάση για την εξαγωγή χαρακτηριστικών (feature extraction) με σκοπό την ηχητική ταξινόμηση - είναι στο πεδίο του χρόνου και στο πεδίο των συχνοτήτων. Τα πειραματικά αποτελέσματα επαληθεύουν την προτεινόμενη μέθοδο που ασχολείται με την εξαγωγή χαρακτηριστικών για την ηχητική ταξινόμηση περιβαλλοντικών ήχων.

## Εισαγωγή

«Ακουστική Οικολογία» είναι ένας όρος που καθιερώθηκε για πρώτη φορά από τον R. Murray Schafer, το 1994, [1]. Στον τομέα της έρευνας της «Ακουστικής Οικολογίας» σημαντική είναι η συμβολή της ομάδας Seki [2]. Ο σκοπός της έρευνας τους είναι να εξετάσουν αν οι ηχογραφήσεις περιβαλλοντικών ήχων 'φέρουν' χρήσιμες οικολογικές πληροφορίες και πιο συγκεκριμένα αν τα χαρακτηριστικά που εξάγονται από αυτές τις ηχογραφήσεις μπορούν να χρησιμοποιηθούν ως δείκτες για την την υγεία και για τη δυναμική ισορροπία των βιοτόπων. Πρόσφατα, το ενδιαφέρον για την ακουστική οικολογία έχει αυξηθεί λόγω των δραστηριοτήτων του Παγκόσμιου Φόρουμ για την Ακουστική Οικολογία (WFAE), που ιδρύθηκε στον Καναδά το 1993.

Μέχρι πρόσφατα, η έρευνα σχετικά με το περιβάλλον και την οικολογική αξιολόγηση των τοπίων επικεντρωνόταν μόνο στο οπτικό περιεχόμενο τους, π.χ., τα μορφολογικά χαρακτηριστικά ενός βιοτόπου. Ωστόσο, η ακουστική οικολογία έχει δείξει ότι ο ήχος του τοπίου μπορεί να χρησιμοποιηθεί ως μια πολύτιμη πρόσθετη ροή πληροφορίας με σκοπό τον χαρακτηρισμό ή την παρακολούθηση περιοχών οικολογικού ενδιαφέροντος, π.χ., βιότοποι. Συγκεκριμένα, ο ήχος παρέχει έναν πρόσθετο οικολογικό δείκτη για τις εν λόγω περιοχές, για σκοπούς που περιλαμβάνουν την παρακολούθηση της άγριας πανίδας ή τις διάφορες ανθρώπινες δραστηριότητες όπως αυτές εξελίσσονται στο χρόνο.

Μετά την κατάλληλη επεξεργασία των ηχητικών σημάτων με σκοπό την εξαγωγή χρήσιμων χαρακτηριστικών από αυτά, η ηχητική πληροφορία μπορεί να μορφοποιηθεί ώστε να συμπεριληφθεί στον «ηχητικό χάρτη» / «ηχητικό τοπίο». Συγκεκριμένα, τα «ηχητικά τοπία» [3], [4] είναι χάρτες μιας συγκεκριμένης περιοχής που επικεντρώνονται στα ηχητικά - και όχι στα οπτικά - χαρακτηριστικά αυτής της περιοχής σε μια δεδομένη χρονική στιγμή. Τα «ηχητικά τοπία» αποτελούν χρήσιμα εργαλεία για την παρακολούθηση της ισορροπίας της φύσης [5] μέσω της σύγκρισής τους - π.χ., για μια περιοχή του Δικτύου NATURA 2000 (δίκτυο της Ευρωπαϊκής Ένωσης των ζωνών προστασίας της φύσης), με σκοπό την εξαγωγή χρήσιμων περιβαλλοντικών /

οικολογικών συμπερασμάτων. Σε αντίθεση λοιπόν με τους γεωγραφικούς χάρτες που αλλάζουν σπάνια, τα «ηχητικά τοπία» απαιτούν τακτικές ενημερώσεις / ανανεώσεις διότι αλλάζουν με την πάροδο του χρόνου. Είναι σημαντικό να αναφερθεί ότι τα «ηχητικά τοπία» μπορεί να περιλαμβάνουν και ήχους που δεν είναι ανιχνεύσιμοι από το ανθρώπινο αυτί ως εκ τούτου, ο όρος «ακουστική (audio) πληροφορία» αποφεύγεται ενώ προτιμάται ο όρος «ηχητική (sound) πληροφορία».

Η εφαρμογή της αναγνώρισης ήχου, όπως και κάθε άλλη εφαρμογή αναγνώρισης προτύπων (pattern recognition), διακρίνεται σε δύο στάδια: το πρώτο στάδιο είναι η εξαγωγή χαρακτηριστικών (feature extraction), ενώ το δεύτερο στάδιο είναι αυτό της κατηγοριοποίησης (classification). Στο πρώτο στάδιο, χαρακτηριστικά που διαφοροποιούν τη μια ηχητική τάξη (sound class) από την άλλη εξάγονται από κάθε ηχητική καταγραφή (sound recording). Η πτυχιακή εργασία θα επικεντρωθεί στο πρώτο στάδιο και συγκεκριμένα σε μεθόδους εξαγωγής χαρακτηριστικών που είναι κατάλληλες για την εφαρμογή της αναγνώρισης ήχου.

Έτσι, ανάμεσα στα άλλα η πτυχιακή αυτή ασχολείται με τον μετασχηματισμό Fourier και τις χρονοσυχνοτικές κατανομές. Τον μετασχηματισμό Fourier τον χρησιμοποιούμε, επειδή η επεξεργασία των σημάτων πραγματοποιείται όχι μόνο στο πεδίο του χρόνου αλλά και στο πεδίο των συχνοτήτων, ενώ τις χρονοσυχνοτικές κατανομές τις χρησιμοποιούμε γιατί η φασματική πληροφορία των περιβαλλοντικών ήχων με τους οποίους ασχολούμαστε εξελίσσεται (μεταβάλλεται) στον χρόνο. Επίσης σε αυτή την πτυχιακή αναφέρονται και οι κυριότερες μέθοδοι μηχανικής μάθησης αφού ασχολούμαστε με την εξαγωγή χαρακτηριστικών για αναγνώριση περιβαλλοντικών ήχων.

## Οργάνωση της Πτυχιακής Εργασίας

Η πτυχιακή εργασία με την οποία έχω ασχοληθεί αποτελείται από 4 Κεφάλαια:

Το πρώτο Κεφάλαιο ασχολείται με την θεωρητική υποδομή του DFT και του FFT και με πρακτικά ζητήματα του FFT όπως η Διαρροή συχνοτήτων (spectral leakage) καθώς και με τις πιο γνωστές συναρτήσεις παραθύρου. Σε συνέχεια του πρώτου Κεφαλαίου, το δεύτερο Κεφάλαιο ασχολείται με τις χρονοσυχνοτικές κατανομές.



Στο τρίτο Κεφάλαιο παρουσιάζονται οι σημαντικότερες μεθοδολογίες που έχουν αναπτυχθεί στο πλαίσιο της μηχανικής μάθησης για την πρόσκτηση γνώσης από τα δεδομένα δημιουργώντας μοντέλα ή βρίσκοντας συσχετίσεις μεταξύ τους ενώ γίνεται αναφορά και στην ανακάλυψη γνώσης σε βάσεις δεδομένων που αποτελεί μια εφαρμογή της μηχανικής μάθησης.

Στο τέταρτο Κεφάλαιο παρουσιάζονται τα βήματα επεξεργασίας των σημάτων (ήχων), τα πειράματα, τα αποτελέσματα και τέλος τα συμπεράσματα.

# Κεφάλαιο 1

## 1.1 Διακριτός Μετασχηματισμός Fourier (DFT)

Στο κεφάλαιο αυτό θα προσδιορίσουμε μαθηματικά τις σχέσεις αναλογικών και διακριτών μετασχηματισμών Fourier (DFT), [6], [7]. Κατόπιν αφού πρώτα παρουσιάσουμε τον αλγόριθμο Cooley and Tukey για τον ταχύ υπολογισμό του DFT, παρουσιάζουμε ένα παράδειγμα υπολογισμού του μέτρου και της φάσης ενός σήματος, [8] καθώς και ορισμένες συναρτήσεις παραθύρων, [9].

Η ορολογία που θα ακολουθήσουμε είναι: η αναλογική ή διακριτή συνάρτηση στο μιγαδικό πεδίο Fourier ή  $z$  θα γράφεται με κεφαλαία και με δείκτη  $k$  ενώ η συνάρτηση στο πεδίου χρόνου θα συμβολίζεται με πεζά γράμματα και δείκτη διανύσματος  $n$ . Διατηρείται φυσικά ο συμβολισμός της αναλογικής συχνότητας με  $\Omega$ (rad/sec) και της ψηφιακής με  $\omega$ (rad/sample). Αρχίζουμε με μία σύντομη επανάληψη των μετασχηματισμών Fourier για αναλογικά σήματα. Όταν ένα αναλογικό σήμα  $f(t)$  είναι περιοδικό με περίοδο  $\Lambda$  (συχνότητα  $\Omega = 2\pi/\Lambda$ ), τότε μπορεί να αναπτυχτεί σε άθροισμα ημιτονοειδών συναρτήσεων με τη σειρά Fourier:

$$f(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\Omega t} \quad (1.1)$$

$$C_n = \frac{1}{\Lambda} \int_{-\frac{\Lambda}{2}}^{\frac{\Lambda}{2}} f(t) e^{-jn\Omega t} dt \quad (1.2)$$

όπου:

$C_n$ : το φάσμα συχνοτήτων

$\Lambda, \Omega$ : η περίοδος και κυκλική συχνότητα του  $f(t)$ .

Για ένα μη περιοδικό σήμα  $f(t)$ , με την προϋπόθεση ότι το ολοκλήρωμα του  $|f(t)|$  είναι πεπερασμένο, το ζευγάρι μετασχηματισμού και αντίστροφου μετασχηματισμού Fourier είναι:

$$F(j\Omega) = \int_{-\infty}^{\infty} f(t)e^{-j\Omega t} dt \quad (1.3)$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t)e^{-j\Omega t} d\Omega = \int_{-\infty}^{\infty} f(t)e^{-j\Omega t} df \quad (1.4)$$

Στη συνέχεια θα αναφερθούμε σε διακριτά σήματα που προκύπτουν από τη διαδικασία δειγματοληψίας όπως αυτή απεικονίζεται στο σχήμα 1.1 και θα καταλήξουμε στην έκφραση Διακριτού Μετασχηματισμού Fourier με τις εξισώσεις (1.10,1.11) ή (1.12,1.13).



Σχήμα 1.1 Ιδανικός Δειγματολήπτης

Στην ιδανική περίπτωση που παρουσιάζει το σχήμα 1.1, το σήμα μετριέται στιγμιαία σε τακτά χρονικά διαστήματα  $nT$ . Το σήμα εξόδου μπορεί να θεωρηθεί αναλογικό σήμα αποτελούμενο από το άθροισμα κρουστικών συναρτήσεων  $\delta(t - nT)$  με πλάτος το δείγμα του αρχικού σήματος τη χρονική εκείνη στιγμή, δηλαδή

$$x_s(t) = \sum_{n=-\infty}^{\infty} x(nT)\delta(t - nT) \quad (1.5)$$

Το φάσμα συχνοτήτων του διακριτού σήματος  $x_s(t)$  είναι:

$$X_s(f) = \int_{-\infty}^{\infty} x_s(t) e^{-j2\pi f t} dt \quad (1.6)$$

αντικαθιστώντας την (1.5) στην (1.6) έχουμε:

$$X_s(f) = \sum_{n=-\infty}^{\infty} x(nT) \int_{-\infty}^{\infty} \delta(t - nT) e^{-j2\pi f t} dt = \sum_{n=-\infty}^{\infty} x(nT) e^{-j2\pi f nT} \quad (1.7)$$

Η εξίσωση (1.7) ονομάζεται Διακριτός Μετασχηματισμός Fourier (DFT, Discrete Fourier Transform) για μία μη περιοδική σειρά δειγμάτων  $x(nT)$ , και για τον υπολογισμό του απαιτείται μόνο η γνώση των δειγμάτων  $x(nT)$ .

Το φάσμα συχνότητας  $X_s(f)$  είναι περιοδικό ως προς  $f$  με περίοδο  $f_s$ , συνεπώς η (1.7) μπορεί να θεωρηθεί σειρά Fourier μίας περιοδικής  $X_s(f)$ , όπου  $x(nT)$  είναι οι συντελεστές της σειράς Fourier. Σύμφωνα με τις εξισώσεις (1.1), (1.2) ισχύει η σχέση,

$$x(nT) = \frac{1}{f_s} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} X_s(f) e^{j2\pi f T n} df = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_s(w) e^{jwn} dw \quad (1.8)$$

που εκφράζει ένα διακριτό σήμα  $x(nT)$  σαν άθροισμα διακριτών ημιτονοειδών σημάτων.

Σε πρακτικούς υπολογισμούς φάσματος ο αριθμός δειγμάτων  $x(nT)$  περιορίζεται σε πεπερασμένο αριθμό  $N$ , δηλαδή  $n = 0, 1, \dots, N-1$ . Αυτό ονομάζεται παράθυρο στον χρόνο (time window). Στην περίπτωση αυτή χωρίζοντας τη συχνότητα σε διακριτά σημεία απέχοντα ίσες αποστάσεις  $\Delta f$  μεταξύ τους, όπου  $f = (k\Delta f)$  η (1.7) παίρνει τη μορφή:

$$X_s(k\Delta f) = \sum_{n=0}^{N-1} x(nT) e^{-j2\pi(k\Delta f)nT} \quad (1.9)$$

Επειδή ο μετασχηματισμός Fourier είναι περιοδικός με περίοδο  $f_s$  και επειδή θέλουμε  $N$  σημεία ανά περίοδο,

$$\Delta f = \frac{\text{περίοδος}}{N} = \frac{f_s}{N} = \frac{1}{(TN)}$$

Αντικαθιστώντας την τελευταία εξίσωση στην (1.9) και για ευκολία γράφοντας  $X(k)$  αντί του  $X_s(k\Delta f)$  και  $x(n)$  αντί του  $x(nT)$  έχουμε:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\left(\frac{2\pi}{N}\right)kn} \quad (1.10)$$

Με το ίδιο σκεπτικό η (1.8) γίνεται:

$$x(nT) = \frac{1}{f_s} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} X_s(f) e^{j2\pi f T n} df \approx \frac{1}{f_s} \sum_{k=-\frac{f_s}{2}}^{\frac{f_s}{2}} X_s(k\Delta f) e^{j2\pi n T k \Delta f} \Delta f$$

Χρησιμοποιώντας  $\Delta f = f_s/N$

$$x(n) = \frac{1}{f_s N} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} X(k) e^{j\left(\frac{2\pi}{N}\right)nk}$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j\left(\frac{2\pi}{N}\right)nk} \quad (1.11)$$

Το ζευγάρι των εξισώσεων (1.10), (1.11) αναφέρεται σαν DFT και IDFT (Inverse Discrete Fourier Transform) αντίστοιχα. Για κάπως ευκολότερη παρουσίαση έχει καθιερωθεί η χρήση της εκθετικής συνάρτησης  $W$ , όπου  $W_N = e^{\frac{j2\pi}{N}}$ . Έτσι ο DFT (1.10) παρουσιάζεται συνήθως στη μορφή:

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{kn} \quad 0 \leq k \leq N-1 \quad (1.12)$$

Ενώ ο IDFT (1.11) παίρνει την μορφή:

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)W_N^{-kn} \quad 0 \leq k \leq N-1 \quad (1.13)$$

Όπου

$$W_N = e^{-\frac{j2\pi}{N}}, \quad X(k) = DFT[x(n)] \text{ και } x(n) = IDFT[X(k)]$$



## 1.2 Ταχύς μετασχηματισμός Fourier (FFT- Fast Fourier Transform)

Αρχίζουμε με τον υπολογισμό των αριθμητικών πράξεων που απαιτούνται για την υλοποίηση του DFT μίας ακολουθίας μήκους  $N$  συμφωνά με τον τύπο:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk}, \quad 0 \leq k \leq N-1$$

όπου  $W_N^{nk} = e^{-j(\frac{2\pi}{N})nk}$ , και  $X(k) = DFT[x(n)]$  είναι οι εξής.

το κάθε ένα δείγμα DFT  $X(k)$  απαιτεί  $N$  μιγαδικούς πολλαπλασιασμούς και  $(N-1)$  προσθέσεις. Επομένως ο υπολογισμός  $N$  δειγμάτων του DFT,  $X(k)$ ,  $k = 0, N-1$ , απαιτεί  $N^2$  μιγαδικούς πολλαπλασιασμούς και  $N(N-1)$  μιγαδικές προσθέσεις. Επειδή ο χρόνος εκτέλεσης του πολλαπλασιασμού είναι πολύ μεγαλύτερος της πρόσθεσης, οι αλγόριθμοι για τον ταχύ υπολογισμό του DFT επικεντρώνονται σε διάφορα τεχνάσματα που έχουν σαν αποτέλεσμα τη μείωση των πολλαπλασιασμών.

Έχουν παρουσιαστεί πολλές εργασίες στον τομέα αυτό κάθε μία από τις οποίες προσφέρει μία μικρή εξοικονόμηση χρόνου για τον υπολογισμό του DFT. Εδώ θα ασχοληθούμε με την κυριότερη μέθοδο που δημοσίευσαν οι Cooley και Tukey το 1965.

Για να υπολογιστεί η απόκριση συχνότητας ενός σήματος υπολογίζουμε ξεχωριστά το μέτρο και την φάση (ο μετασχηματισμός Fourier δίνει ως αποτέλεσμα μιγαδικούς αριθμούς).

Παρακάτω λοιπόν θα παρουσιάσουμε ένα παράδειγμα που δείχνει τον υπολογισμό του μέτρου και της φάσης ενός σήματος που αποτελεί το άθροισμα 2 ημιτόνων διαφορετικών συχνοτήτων.

### 1.2.1 Μέτρο και Φάση του σήματος

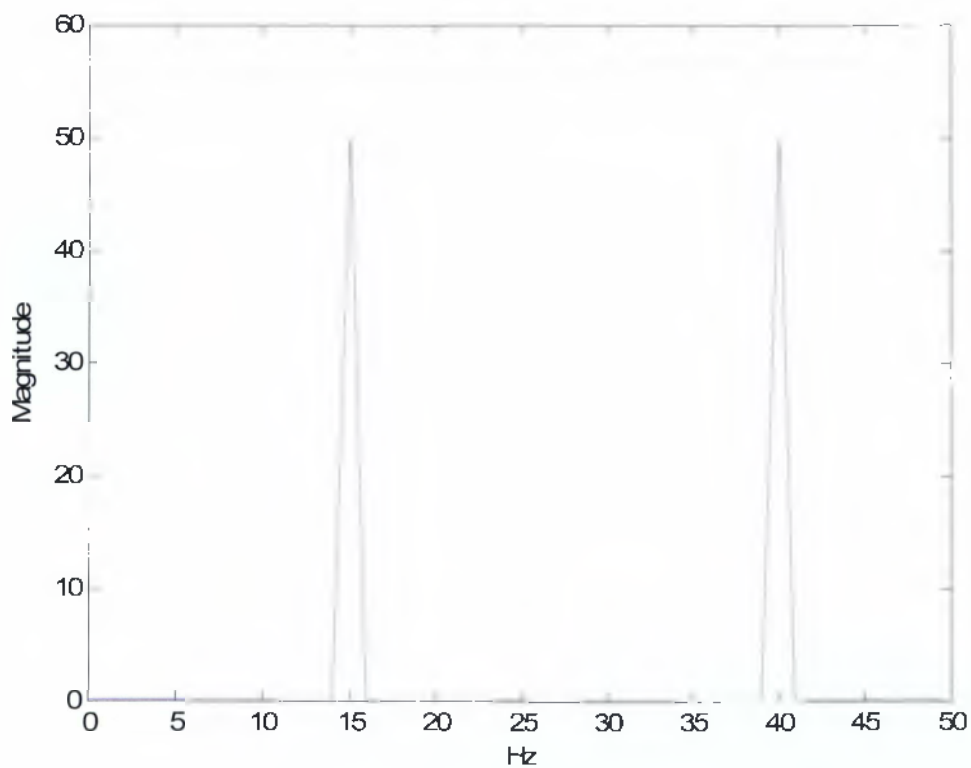
Στο ακόλουθο παράδειγμα θα υπολογίσουμε το μέτρο και τη φάση του σήματος  $x = \sin(2\pi \cdot 15 \cdot t) + \sin(2\pi \cdot 40 \cdot t)$ ; το οποίο είναι το άθροισμα 2 ημιτόνων με διαφορετικές συχνότητες. Για τον υπολογισμό του DFT του σήματος  $x$  θα χρησιμοποιήσουμε την εντολή  $y = \text{fft}(x)$ , και για τον υπολογισμό του μέτρου την εντολή  $m = \text{abs}(y)$ .

```
f=0:50;
```

```
plot(f,m);
```

```
ylabel('Magnitude'); xlabel('Hz');
```

Όπως φαίνεται στο σχήμα 1.2 οι συχνότητες των δύο ημιτόνων είναι 15 και 40 Hz.

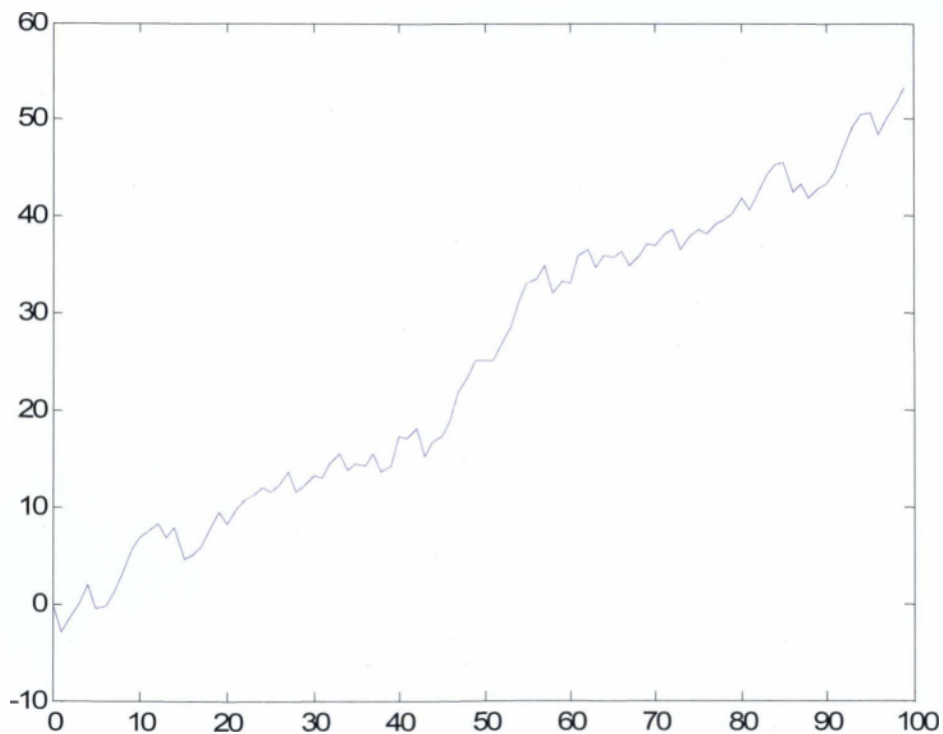


Σχήμα 1.2: Μέτρο του σήματος

Κατόπιν για να υπολογίσουμε τη φάση χρησιμοποιούμε την εντολή `angle` η οποία ορίζεται ως:

$$f(\omega) = \tan^{-1} \left| \frac{\text{Im}(S(\omega))}{\text{Re}(S(\omega))} \right|$$

όπου  $0 < f(\omega) < 2\pi$ ,  $S(\omega)$  είναι το μιγαδικό φάσμα Fourier του σήματος και  $\text{Im}(\cdot)$ ,  $\text{Re}(\cdot)$  είναι το φανταστικό και το πραγματικό μέρος του, αντίστοιχα σχήμα 1.3,[10].



Σχήμα 1.3: Φάση του σήματος

### 1.3 Διαρροή συχνοτήτων ( spectral leakage)

Το φαινόμενο διαρροής συχνοτήτων οφείλεται στο γεγονός ότι για να χρησιμοποιηθεί ο DFT για τη φασματική ανάλυση ενός σήματος, πρέπει να περιοριστεί το μήκος δείγματος  $N$  του σήματος σε κάποια πεπερασμένη τιμή. Δηλαδή η ανάλυση διεξάγεται σε ένα τμήμα του σήματος το οποίο έχει αποκοπεί από το αρχικό σήμα. Η απότομη αυτή διακοπή του σήματος δημιουργεί διαρροή συχνοτήτων η οποία διορθώνεται μερικώς με κατάλληλη επιλογή συνάρτησης παράθυρου (window function), όπως θα δούμε στη συνέχεια. Είναι γνωστό ότι για να χρησιμοποιηθεί ο  $N$ -point FFT για φασματική ανάλυση μιας ακολουθίας  $x(n)$  πρέπει να γίνει διάσπαση της ακολουθίας σε τμήματα  $N$  δειγμάτων.

$$x(n) = \underbrace{\{x(0), x(1), \dots, x(N-1)\}}_{J=0 \text{ N-point FFT}}, \underbrace{\{x(N), x(N+1), \dots, x(2N-1)\}}_{J=1 \text{ N-point FFT}}, \underbrace{\{..\}}_{J=2},$$

Η επιλογή του κάθε  $N$ -μήκους τμήματος της ακολουθίας μπορεί εκφραστεί μαθηματικά σα γινόμενο της ακολουθίας  $x(n)$  επί της συνάρτησης βάρους  $w_J(n)$  όπου:

$$w_J(n) = \begin{cases} 1 & JN \leq n \leq (J+1)N - 1 \\ 0 & \text{διαφορετικά} \end{cases}$$

Η  $w(n)$  λέγεται συνάρτηση ορθογώνιου παράθυρου (rectangular or uniform window) λόγω του σχήματος της. Τώρα θα κάνουμε μία μικρή παρένθεση, για να υπενθυμίσουμε το θεώρημα Δυϊκότητας (Duality theorem). Κατά το θεώρημα αυτό, το γινόμενο δυο συναρτήσεων στο πεδίο συχνότητας αντιστοιχεί με συνέλιξη στο πεδίο χρόνου και αντίστροφα ο πολλαπλασιασμός δύο συναρτήσεων στο πεδίο χρόνου αντιστοιχεί σε συνέλιξη φάσματος των δύο συναρτήσεων στο πεδίο συχνότητας. Τα παραπάνω μπορεί να εκφραστούν μαθηματικά ως εξής:

Πεδίο Συχνότητας

$$Y(\omega) = X(\omega) \cdot H(\omega)$$

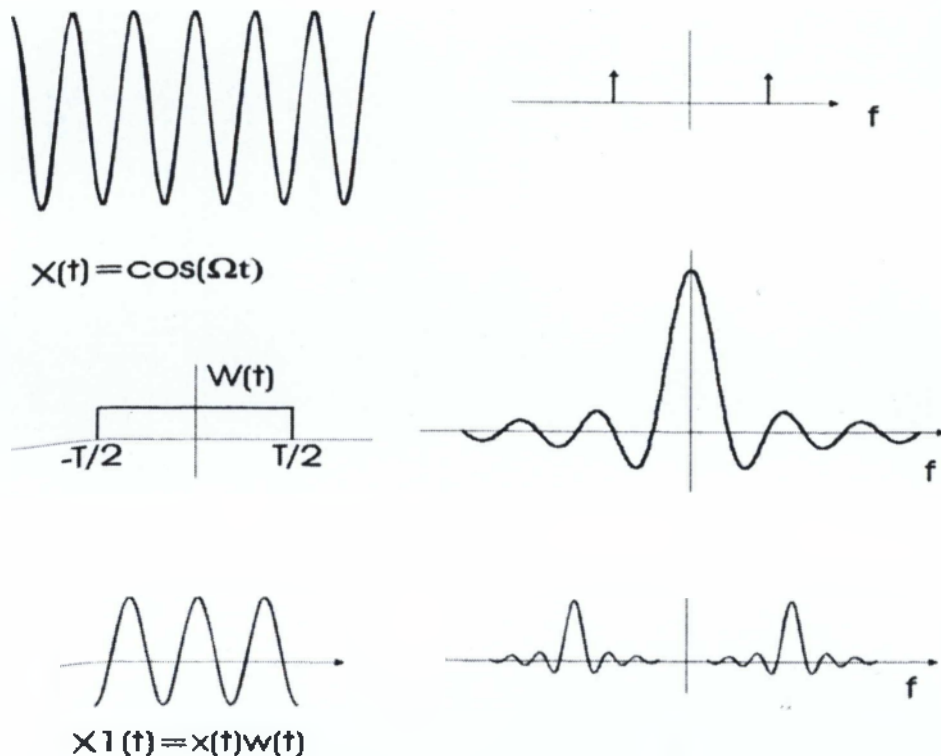
$$Y(\omega) = X(\omega) * H(\omega)$$

Πεδίο Χρόνου

$$y(n) = x(n) * h(n)$$

$$y(n) = x(n) \cdot h(n)$$

Συνεπώς το γινόμενο  $x(n)w(n)$  αντιστοιχεί στη συνέλιξη των  $X(\omega)$  και  $W(\omega)$  στο πεδίο συχνοτήτων. Η συνάρτηση ορθογώνιου παράθυρου έχει φάσμα τύπου  $\frac{\eta\mu x}{x}$  και προφανώς επηρεάζει το παρατηρούμενο φάσμα της  $x(n)$ , εφόσον αυτό που βλέπουμε δεν είναι  $X(\omega)$  αλλά  $X(\omega) * W(\omega)$  (συνέλιξη  $X(\omega)$  με  $W(\omega)$ ) όπως παρουσιάζει το σχήμα 1.4.



Σχήμα 1.4: Γραφική παρουσίαση του θεωρήματος Δυσκότητας

Αν η συνάρτηση παράθυρου  $w(t)$  ήταν 1 για όλα τα  $t$ , τότε φυσικά το παρατηρούμενο φάσμα θα ήταν ίδιο με το πραγματικό αφού δε θα υπήρχε αποκοπή του σήματος. Ο μετασχηματισμός Fourier της  $w(t) = 1$  είναι  $W(\Omega) = \delta(\Omega)$ , όπου  $\delta(\Omega)$  η κρουστική συνάρτηση. Έτσι, αν ήταν δυνατόν να κατασκευαστεί κάποια συνάρτηση παράθυρου πεπερασμένης διάρκειας με φάσμα το δυνατόν "πλησιέστερο" του  $\delta(\Omega)$ , τότε το παρατηρούμενο φάσμα του τμήματος του  $x(n)$  θα "πλησίαζε" το πραγματικό φάσμα του σήματος.

Σε απάντηση του ερωτήματος αυτού και στην προσπάθεια να παραμείνει το φάσμα συχνότητας  $X(\omega)$  όσο γίνεται αναλλοίωτο μετά τη συνέλιξη του με το  $W(\omega)$ , έχουν δημοσιευτεί πολλές συναρτήσεις παράθυρου. Οι πιο γνωστές από αυτές είναι:

- Rectangular or Uniform Window
- Hamming Window
- Blackman Window
- Hanning Window

Παρακάτω θα δούμε μερικά παραδείγματα:

### 1.3.1 Παράθυρο τύπου Χάμινγκ (Hamming window)

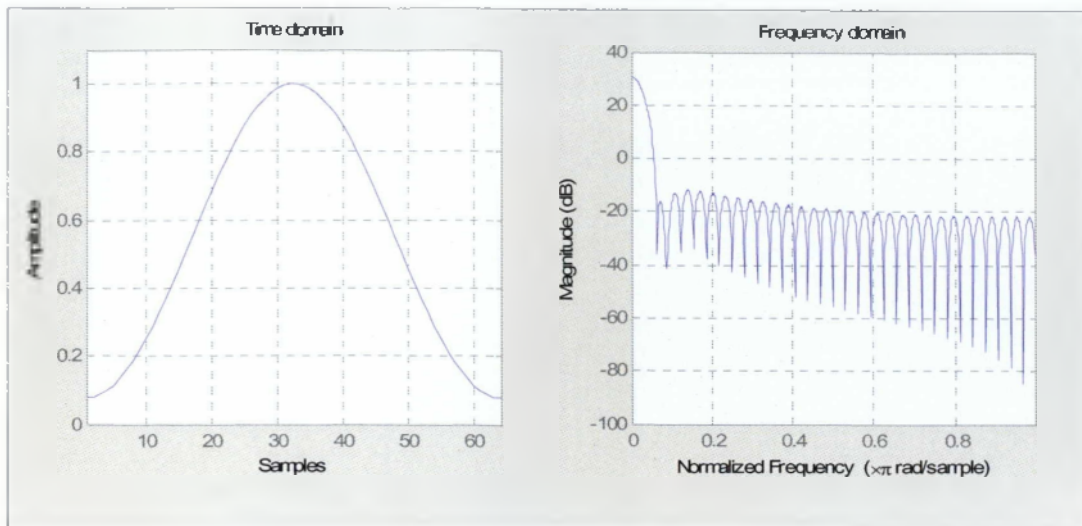
$$w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) & 0 \leq n \leq N-1 \\ 0 & \text{διαφορετικά} \end{cases}$$

Για να δημιουργήσουμε ένα παράθυρο Hamming 64 σημείων και να επιδείξουμε το αποτέλεσμα χρησιμοποιούμε το εργαλείο WVTool του Matlab και τις παρακάτω εντολές:

```
L=64;
```

```
wvtool(hamming(L))
```





Σχήμα 1.5: Παράθυρο τύπου Hamming (α) στο πεδίο του χρόνου και (β) στο πεδίο των συχνοτήτων

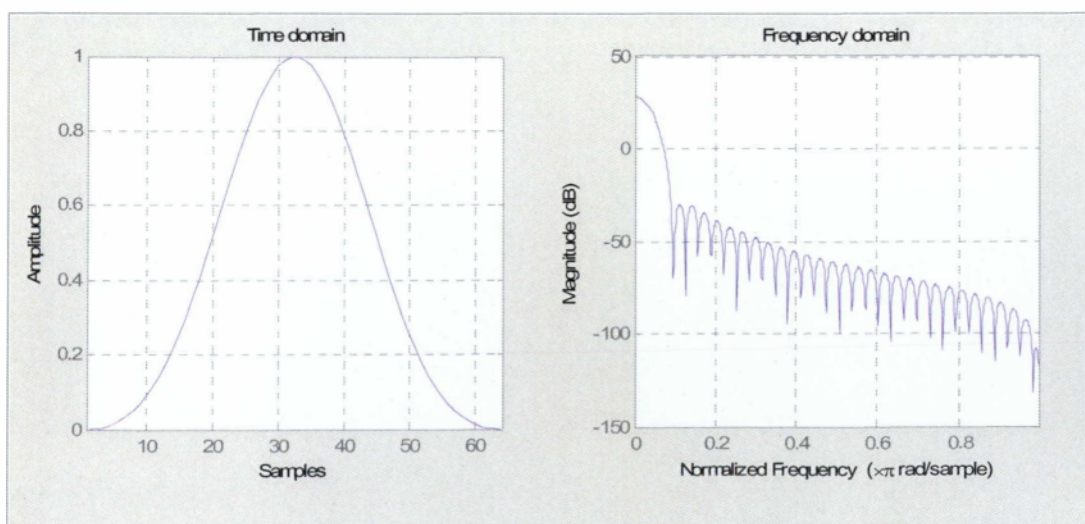
### 1.3.2 Παράθυρο τύπου Μπλάκμαν (Blackman window)

$$w(n) = \begin{cases} 0.42 - 0.5\cos\left(\frac{2\pi n}{N-1}\right) + 0.08\cos\left(\frac{4\pi n}{N-1}\right) & 0 \leq n \leq N-1 \\ 0 & \text{διαφορετικά} \end{cases}$$

Για να δημιουργήσουμε ένα παράθυρο Blackman 64 σημείων χρησιμοποιούμε το εργαλείο WWTTool του Matlab και τις παρακάτω εντολές:

`L=64;`

`wntool(blackman(L))`



Σχήμα 1.6: Παράθυρο τύπου Blackman (α) στο πεδίο του χρόνου και (β) στο πεδίο των συχνοτήτων

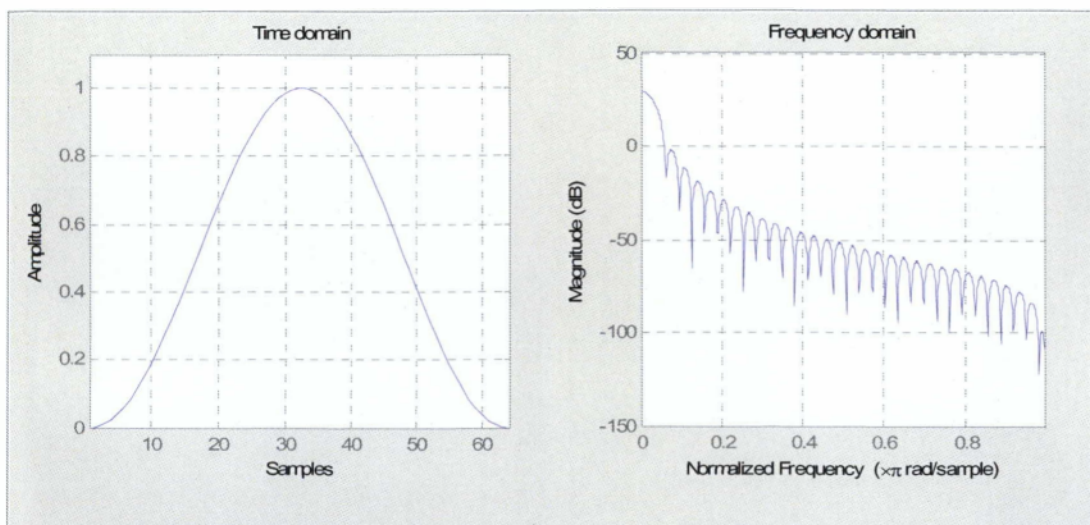
### 1.3.3 Παράθυρο τύπου Χάνινγκ (Hanning window)

$$\begin{cases} \frac{1}{2} \left[ 1 - \cos\left(\frac{2\pi n}{N-1}\right) \right] & 0 \leq n \leq N-1 \\ 0 & \text{διαφορετικά} \end{cases}$$

Για να δημιουργήσουμε ένα παράθυρο Hanning 64 σημείων χρησιμοποιούμε το εργαλείο `WVTool` του Matlab και τις παρακάτω εντολές:

```
L=64;
```

```
wvtool(hann(L))
```



Σχήμα 1.7: Παράθυρο τύπου Hanning (α) στο πεδίο του χρόνου και (β) στο πεδίο των συχνοτήτων

Μετά την επιλογή της συνάρτησης παράθυρου ο υπολογισμός του FFT εφαρμόζεται στην ακολουθία που προκύπτει από το γινόμενο  $w(n)x(n)$ . Να σημειωθεί ότι η συνάρτηση παράθυρου έχει μήκος  $N$ , ώστε να πολλαπλασιάζει μόνο τους μη μηδενικούς όρους της  $x(n)$ .

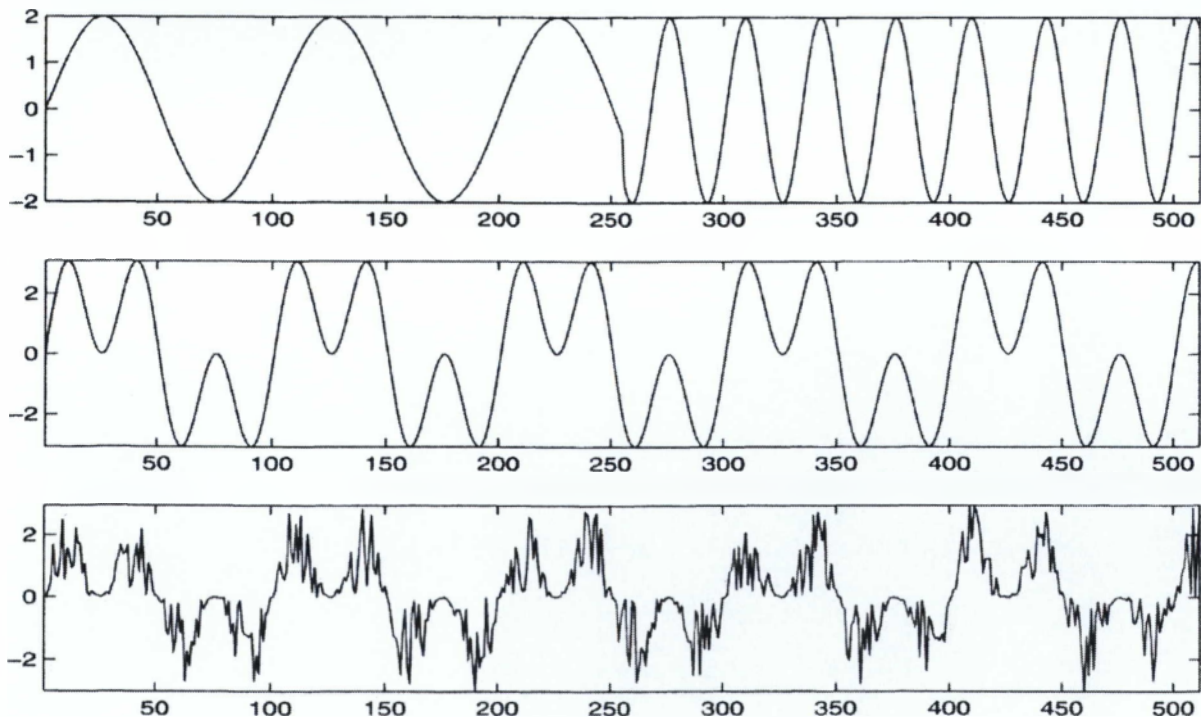
Τέλος, πρέπει να σημειωθεί ότι το φαινόμενο *leakage* είναι αλληλένδετο με το φαινόμενο επικάλυψης. Η επέκταση του φάσματος από τη διαρροή συχνοτήτων μπορεί να δημιουργήσει συχνότητες που υπερβαίνουν τη συχνότητα Nyquist, προκαλώντας έτσι προβλήματα επικάλυψης.

## Κεφάλαιο 2

Σε αυτό το Κεφάλαιο αναπτύσσουμε την έννοια των χρονοσυχνοτικών κατανομών (time-frequency distributions) και παρέχονται ορισμένα παραδείγματα φασματογραμμάτων (spectrogram) για σήματα που έχουμε δημιουργήσει οι ίδιοι, [11].

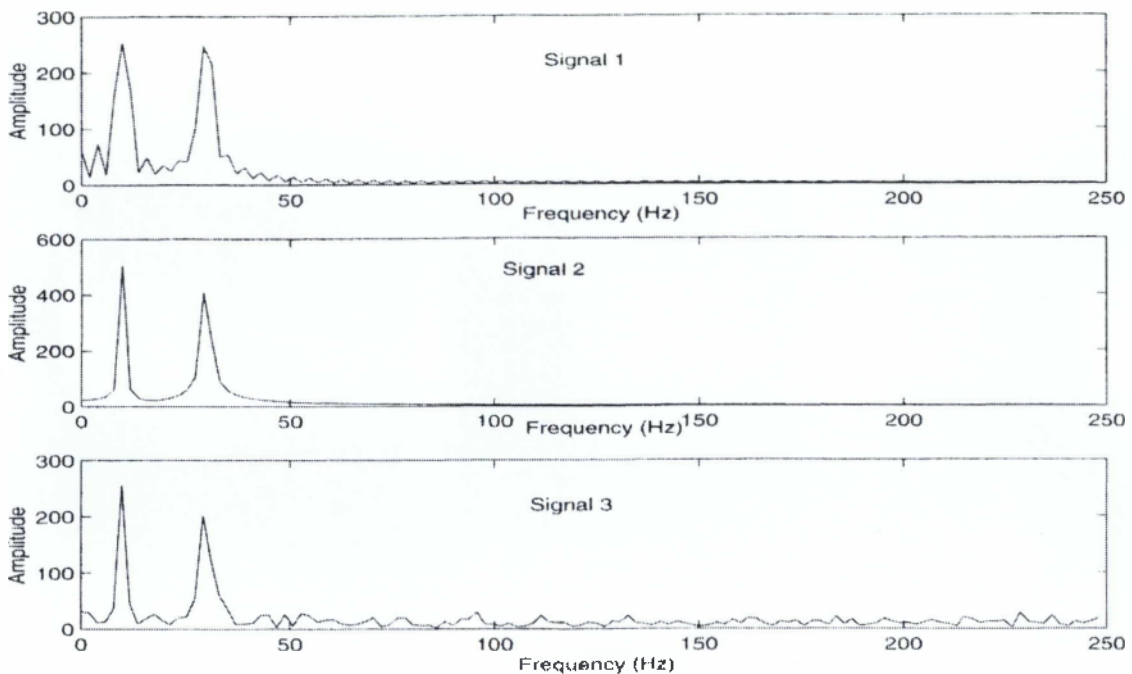
### 2.1 Μετασχηματισμοί FT, STFT

Πολλές φορές, είναι απαραίτητο να επεξεργαστούμε κάποιο σήμα που μεταφέρει μέσα του μια πληροφορία. Η πληροφορία αυτή μπορεί να είναι κωδικοποιημένη με διάφορους τρόπους, είτε σαν συχνότητα, είτε σαν πλάτος. Στην μία περίπτωση, έχουμε μια διαμόρφωση στο πεδίο της συχνότητας, και στην άλλη, στο πεδίο του χρόνου. Δυστυχώς, το πεδίο του χρόνου συνήθως αποδεικνύεται πολύ περίπλοκο όταν θέλουμε να απεικονίσουμε έστω και τα πιο απλά σήματα. Έτσι, πρέπει να χρησιμοποιήσουμε μια άλλη απεικόνιση, που θα μας βοηθήσει να ερμηνεύσουμε καλύτερα και απλούστερα το σήμα μας. Ευτυχώς, υπάρχουν διάφοροι μετασχηματισμοί, που μας βοηθούν να μεταφέρουμε, να μεταφράζουμε, το σήμα που μας ενδιαφέρει σε άλλη, πιο απλή απεικόνιση.



Σχήμα 2.1: Σήματα στο πεδίο του χρόνου

Για παράδειγμα, τα τρία διαφορετικά σήματα που βλέπουμε στο σχήμα 2.1, αποτελούνται από δύο ημιτονικές συνιστώσες. Στην πρώτη εικόνα, οι δύο συνιστώσες (οι δύο συχνότητες, δηλαδή) διαδέχονται η μία την άλλη. Στο δεύτερο παράδειγμα, συνυπάρχουν συνεχώς και οι δύο μαζί. Το τρίτο παράδειγμα κατασκευάστηκε προσθέτοντας θόρυβο στο σήμα του δεύτερου παραδείγματος. Είναι προφανές ότι και η μικρότερη παρεμβολή θορύβου μπορεί να δυσκολέψει πολύ την απεικόνιση του σήματος στο πεδίο του χρόνου. Μπορούμε να δοκιμάσουμε να μετατρέψουμε τις κυματομορφές αυτές στο πεδίο της συχνότητας, χρησιμοποιώντας τον πολύ γνωστό μετασχηματισμό Fourier. Το αποτέλεσμα φαίνεται στο σχήμα 2.2.



Σχήμα 2.2: Σήματα στο πεδίο συχνοτήτων

Παρατηρούμε ότι οι δύο συνιστώσες του σήματος υπάρχουν ξεκάθαρες στο διάγραμμα της συχνότητας. Ακόμα, βλέπουμε ότι η ύπαρξη του θορύβου στην τρίτη περίπτωση δεν επηρέασε σημαντικά την μορφή του φάσματος, δίνοντάς μας την ευχέρεια να διακρίνουμε τις δύο συνιστώσες. Αυτό που είναι πολύ σημαντικό, είναι το ότι ο μετασχηματισμός Fourier δεν μπορεί να διακρίνει μεταξύ των δύο πρώτων παραδειγμάτων. Δηλαδή, ενώ πρόκειται για διαφορετικά σήματα, η απεικόνισή τους είναι ίδια. Αυτό είναι ένα από τα μειονεκτήματα του μετασχηματισμού αυτού, δηλαδή ότι στο μετασχηματισμένο σήμα δεν υπάρχει η πληροφορία του χρόνου, όπως θα δούμε και



παρακάτω. Δηλαδή, δεν ξέρουμε πότε εμφανίστηκε κάθε συχνότητα. Είτε διαδέχτηκε η μία την άλλη, είτε υπήρχαν μαζί συνεχώς, είναι το ίδιο για τον μετασχηματισμό Fourier. Πριν όμως προχωρήσουμε σε άλλες λεπτομέρειες, ας δούμε μερικούς μετασχηματισμούς και την γενική αρχή λειτουργίας τους.

Γενικά, ένας γραμμικός μετασχηματισμός έχει την μορφή της εξής μαθηματικής σχέσης.

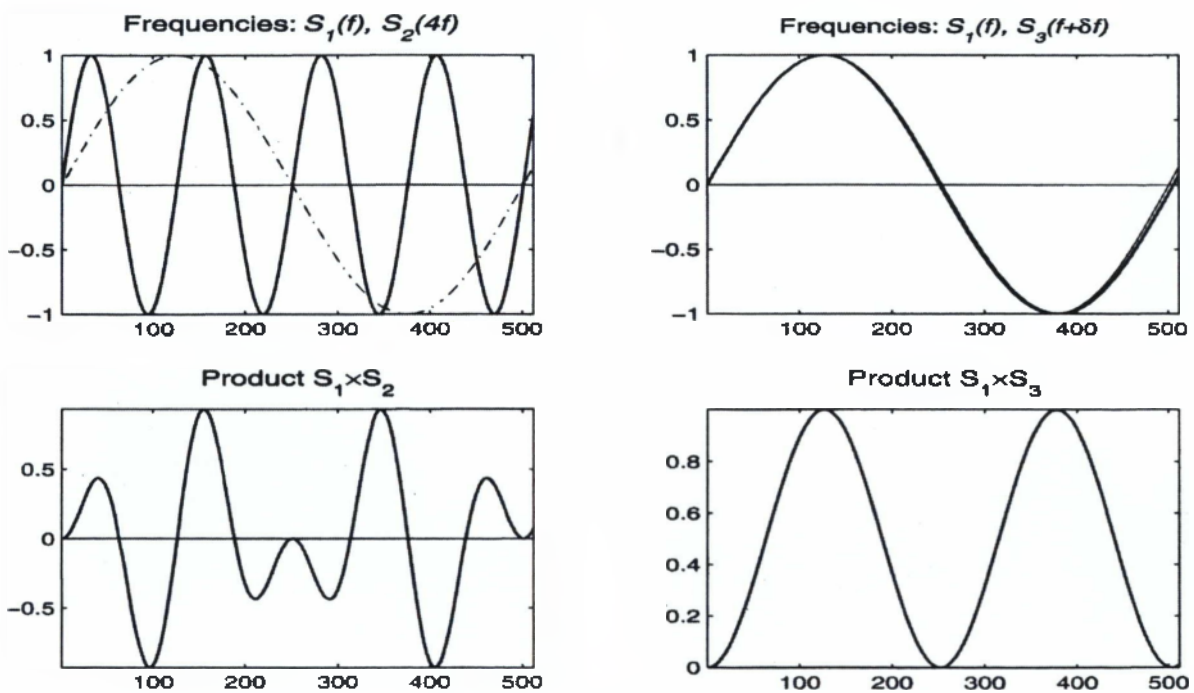
$$F(y_1, \dots, y_m) = \int_x f(x_1, \dots, x_n) g(x_1, \dots, n, y_1 \dots m) \theta_{x_1} \dots \theta_n$$

Στην σχέση αυτή, μπορούμε να διακρίνουμε την μετασχηματισμένη συνάρτηση  $F$ , η οποία ορίζεται με βάση τις μεταβλητές του πεδίου άφιξης. Για παράδειγμα, στον μετασχηματισμό Fourier, αυτή η συνάρτηση εκφράζεται ως προς την συχνότητα. Βασικό συστατικό του μετασχηματισμού είναι η συνάρτηση  $g()$ , η οποία ονομάζεται πυρήνας του μετασχηματισμού. Αυτή ακριβώς είναι η συνάρτηση που ενεργεί τον μετασχηματισμό, την μεταφορά από το ένα πεδίο στο άλλο. Ο πυρήνας είναι ο συνδυαστικός κρίκος ανάμεσα στα δύο πεδία, και γι' αυτό περιέχει μεταβλητές τόσο από το πεδίο ορισμού της αρχικής συνάρτησης, όσο και από το πεδίο ορισμού της μετασχηματισμένης συνάρτησης. Παρατηρώντας την μαθηματική αυτή σχέση, βλέπουμε ότι η συνάρτηση που θέλουμε να μετασχηματίσουμε εμπλέκεται σε ένα ολοκλήρωμα, αφού πρώτα πολλαπλασιαστεί με τον πυρήνα. Πρώτα, η αρχική συνάρτηση πολλαπλασιάζεται με τον πυρήνα, που όπως θα δούμε στην συνέχεια παίζει το ρόλο ενός "προτύπου", κατά κάποιο τρόπο, με το οποίο συγκρίνεται η αρχική συνάρτηση. Το γινόμενο τους περιέχει μεταβλητές και από τα δύο πεδία, δηλαδή και από το πεδίο  $x$  και από το πεδίο  $y$ . Στην συνέχεια, το ολοκλήρωμα πάνω στο πεδίο  $x$  αθροίζει όλες τις μεταβλητές του  $x$  και έτσι το αποτέλεσμα που μένει είναι μια συνάρτηση με μόνο τις μεταβλητές του πεδίου  $y$ .

Εκτός από τους συνεχείς μετασχηματισμούς, έχουμε και τους διακριτούς. Ένας από αυτούς είναι και ο γνωστός μας μετασχηματισμός Z. Στους διακριτούς μετασχηματισμούς, το ολοκλήρωμα αντικαθίσταται από την διακριτή του μορφή, το άθροισμα. Πρέπει εδώ να σημειώσουμε ότι ανάλογα με την εργασία και τον τύπο της πληροφορίας που θέλουμε να εξάγουμε από το σήμα μας, πρέπει να χρησιμοποιήσουμε και τον κατάλληλο κάθε φορά μετασχηματισμό, δηλαδή τον κατάλληλο πυρήνα. Επίσης, πρέπει να ξέρουμε ότι κανένας μετασχηματισμός δεν είναι

ιδανικός, δηλαδή όλοι λίγο-πολύ έχουν ατέλειες που αλλοιώνουν κάποια από τα χαρακτηριστικά του σήματος.

Σαν παράδειγμα για την λειτουργία του μετασχηματισμού, θα δούμε πως ο πυρήνας του μετασχηματισμού Fourier, μια μιγαδική αρμονική συνάρτηση, αναλύει το σήμα σε διάφορες συχνότητες. Οι αρμονικές συναρτήσεις που χρησιμοποιούνται σε αυτό το μετασχηματισμό, δηλαδή ένα ημίτονο με μια καθορισμένη συχνότητα και διαφορά φάσης, έχει την ιδιότητα ότι ορίζεται ακριβώς σε μία και μόνο μία συχνότητα. Όπως φαίνεται στο σχήμα 2.3, όταν πολλαπλασιάσουμε την αρχική συνάρτηση



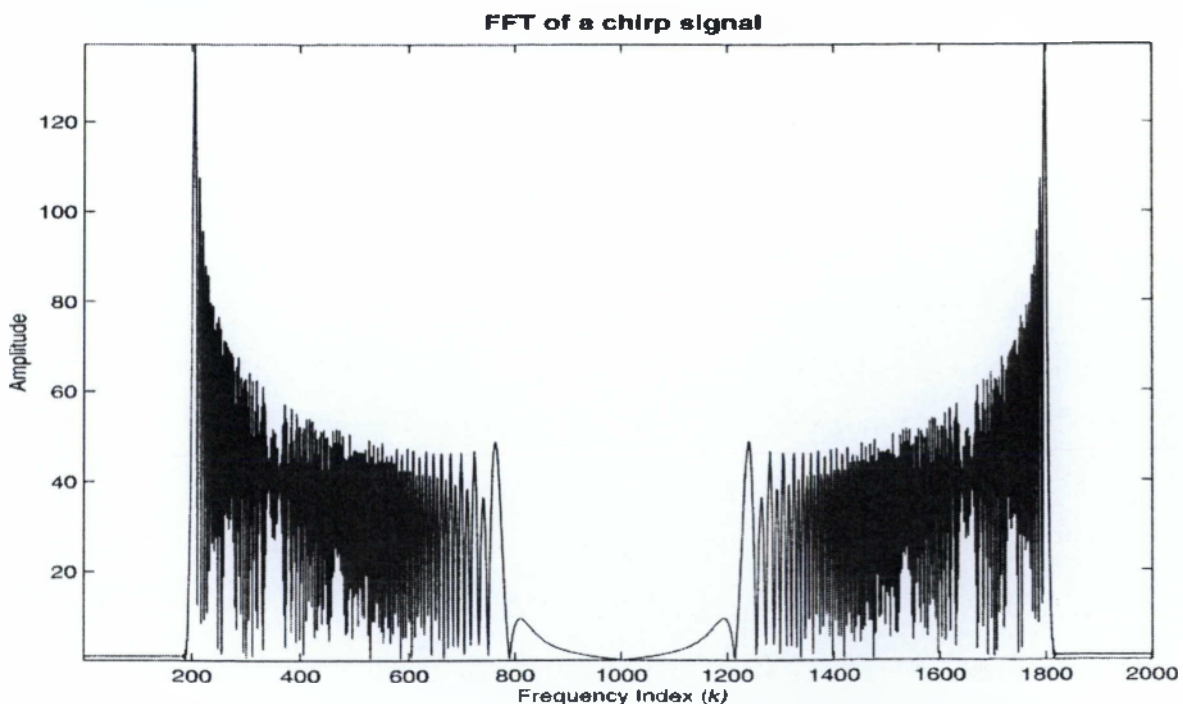
Σχήμα 2.3: Λειτουργία Μετασχηματισμού Fourier

(το ημίτονο χαμηλής συχνότητας στο πάνω αριστερό μέρος) με την συνάρτηση του πυρήνα (που εδώ είναι ένα ημίτονο σε μια καθορισμένη συχνότητα, αρκετά διαφορετική από αυτήν του αρχικού σήματος) παράγεται μια άλλη συνάρτηση. Αν την ολοκληρώσουμε, δηλαδή υπολογίσουμε το εμβαδόν που περικλείεται ανάμεσα στην γραφική παράσταση (κάτω αριστερά) και τον οριζόντιο άξονα, θα δούμε ότι είναι πολύ μικρό. Δηλαδή, το ολοκλήρωμα του μετασχηματισμού έχει μικρή τιμή. Αυτό συμβαίνει επειδή το γινόμενο έχει ίσες θετικές και αρνητικές επιφάνειες. Στην δεύτερη περίπτωση (δεξιά), έχουμε πάλι την αρχική ημιτονοειδή συνάρτηση, και δίπλα της τον πυρήνα,



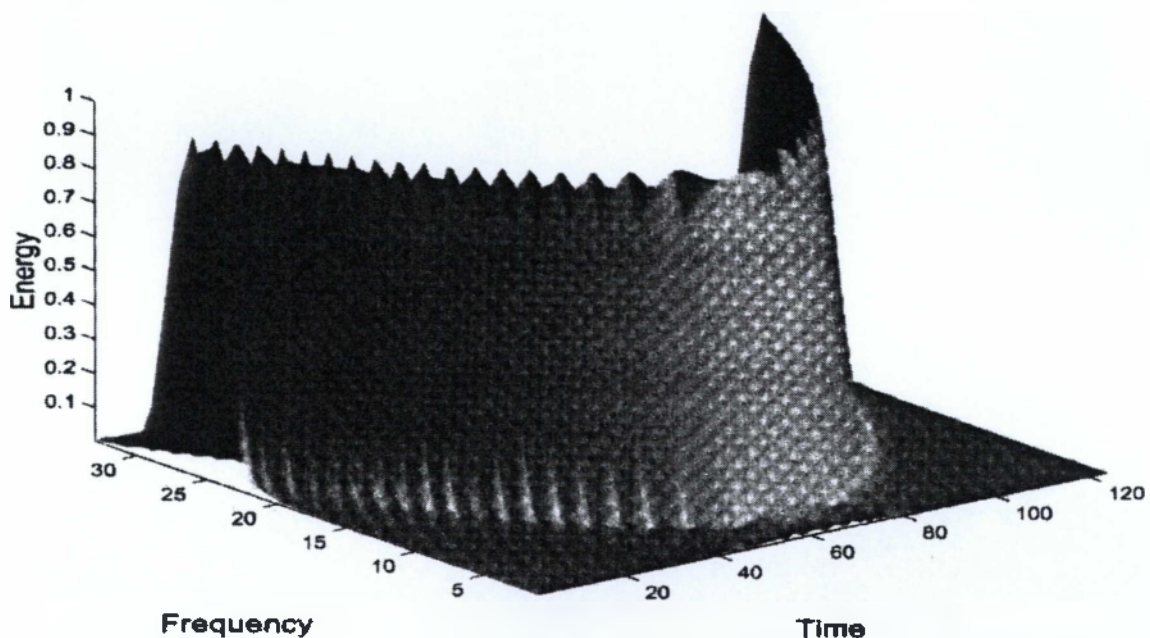
αλλά τώρα υπολογισμένο σε μια συχνότητα παραπλήσια με την συχνότητα της αρχικής συνάρτησης. Τώρα, το γινόμενο των δύο συναρτήσεων σχηματίζει αποκλειστικά θετικές επιφάνειες. Έτσι, το αποτέλεσμα του ολοκληρώματος είναι μεγάλο. Άρα λοιπόν, ο πυρήνας λειτουργεί σαν μέτρο σύγκρισης για τη αρχική συνάρτηση, και το ολοκλήρωμα μας δείχνει τον βαθμό ομοιότητας των δύο συναρτήσεων. Σαρώνοντας όλες τις δυνατές παραλλαγές του πυρήνα, και συγκρίνοντάς τις διαδοχικά με την αρχική συνάρτηση, παράγεται το αποτέλεσμα του μετασχηματισμού. Ανακεφαλιώνοντας λοιπόν, πρώτα, η αρχική συνάρτηση πολλαπλασιάζεται με όλες τις δυνατές παραλλαγές του πυρήνα. Για κάθε παραλλαγή ξεχωριστά (πχ για κάθε συχνότητα στον μετασχηματισμό Fourier υπολογίζεται η ομοιότητα του πυρήνα με την αρχική συνάρτηση, με την βοήθεια του ολοκληρώματος. Στην συνέχεια, τα επιμέρους αποτελέσματα (δηλ. η τιμή του ολοκληρώματος) ενοποιούνται, και απαρτίζουν το σύνολο τιμών της  $F$ , της μετασχηματισμένης δηλαδή συνάρτησης.

Είπαμε πιο πριν ότι κανένας μετασχηματισμός δεν είναι τέλειος. Ένα από τα χαρακτηριστικά τους, είναι το ότι η μετάβαση από το ένα πεδίο στο άλλο μπορεί να ποικίλλει. Δηλαδή, η πληροφορία του αρχικού πεδίου μπορεί να χαθεί εντελώς ή να συνυπάρχει μέσα στο μετασχηματισμένο πεδίο. Αυτό το χαρακτηριστικό ποικίλλει από τον ένα τύπο μετασχηματισμού στον άλλο.



Σχήμα 2.4: Fourier σήματος chirp

Στην εικόνα 2.4, βλέπουμε τον μετασχηματισμό Fourier ενός σήματος που σαρώνει μια μπάνα συχνοτήτων. Δεν είναι προφανές το πώς εξελίσσεται η κυματομορφή, δηλαδή με ποιό τρόπο γίνεται η σάρωση των συχνοτήτων. Έτσι, στο μετασχηματισμένο σήμα έχει χαθεί εντελώς η πληροφορία του χρόνου. Ο χρόνος ήταν το αρχικό πεδίο στο οποίο ζούσε το σήμα μας. Συνεπώς, ο μετασχηματισμός Fourier δεν διατηρεί καθόλου την πληροφορία του χρόνου. Αυτό το είδαμε και πριν, όπου τα σήματα του σχήματος 2.1 είχαν ακριβώς τον ίδιο μετασχηματισμό (σχήμα 2.2), ενώ είχαν εντελώς διαφορετικές κυματομορφές. Παραστατικά, ο μετασχηματισμός Fourier είναι σαν να έχουμε μπροστά μας τις παρτιτούρες ενός κονσέρτου, αλλά με τις νότες γραμμένες ακατάστατα, χωρίς σειρά, επάνω στο χαρτί. Ξέρουμε όλες τις νότες (συχνότητες) που απαρτίζουν το κονσέρτο (σήμα), αλλά δεν ξέρουμε πότε παίχτηκε η κάθε μια. Αυτό είναι ένα πρόβλημα που αντιμετωπιζόταν με τον Short Time Fourier Transform (STFT), αλλά τώρα τα wavelets δίνουν πολύ καλύτερη λύση στο πρόβλημα αυτό.



Σχήμα 2.5: STFT σήματος chirp

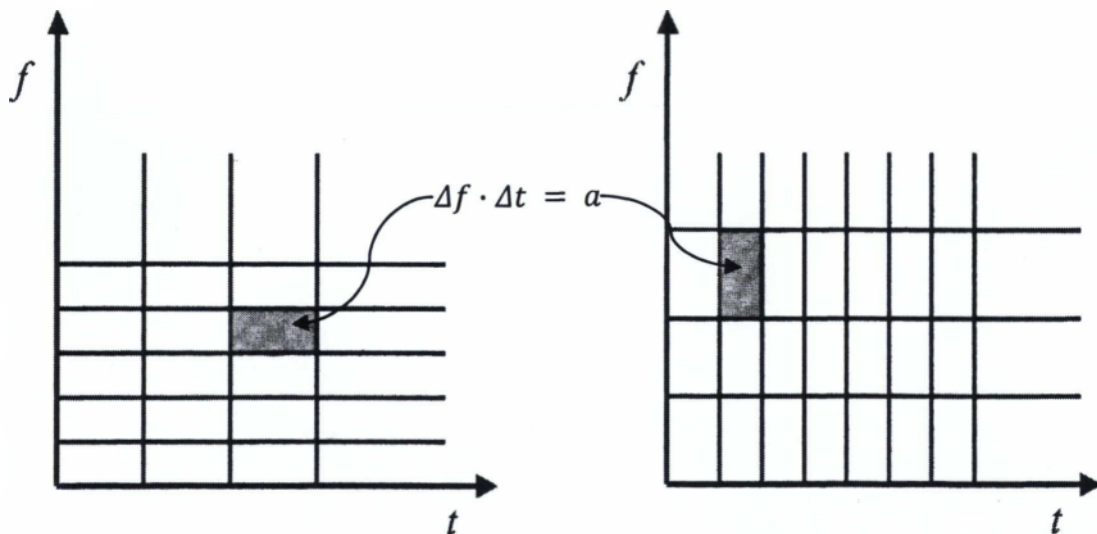
Ο Τμηματικός Μετασχηματισμός (STFT), λειτουργεί 'σπάζοντας' το αρχικό σήμα σε ίσα κομμάτια μικρής χρονικής διάρκειας, και εφαρμόζοντας τον μετασχηματισμό για κάθε ένα κομματάκι χωριστά. Έτσι, έχουμε διαδοχικές εικόνες των συχνοτήτων που περιέχονται στο σήμα μας, όπως περίπου τα καρέ μιας κινηματογραφικής ταινίας. Έτσι, καθώς το σήμα εξελίσσεται, μπορούμε να δούμε πώς αλλάζουν οι συχνότητες που περιέχει.

Στο σχήμα 2.5 φαίνεται το τρισδιάστατο διάγραμμα ενός τέτοιου μετασχηματισμού. Κάθε 'φέτα' που κόβουμε παράλληλα με τον άξονα των συχνοτήτων, είναι ένας 'τοπικός' μετασχηματισμός Fourier. Τοποθετώντας πολλές διαδοχικές τέτοιες φέτες την μία δίπλα στην άλλη, παίρνουμε τον άξονα του χρόνου, που δείχνει την εξέλιξη των συχνοτήτων του σήματος. Τα σημεία που μοιάζουν με βουναλάκια είναι τα σημεία (δηλαδή οι συχνότητες) στα οποία συγκεντρώνεται η ενέργεια του σήματος. Το μετασχηματισμένο σήμα της εικόνας που βλέπουμε εδώ, είναι το ίδιο με το σήμα της προηγούμενης εικόνας, που σαρώνει διάφορες συχνότητες. Εδώ, σε αντίθεση με την προηγούμενη εικόνα, μπορούμε να παρατηρήσουμε την εξέλιξη του σήματος. Βλέπουμε, ότι στην αρχή το σήμα βρίσκεται στις υψηλές συχνότητες, μετά κατεβαίνει στις χαμηλές με γεωμετρικό ρυθμό, και στην συνέχεια ανεβαίνει και πάλι στις υψηλές. Υπάρχουν όμως και προβλήματα στον μετασχηματισμό αυτό.

Όταν το παράθυρο (χρονικό) είναι μεγάλο, μπορούμε να διακρίνουμε με σχετικά καλή ακρίβεια το ποιά ακριβώς συχνότητα είναι παρούσα στο σήμα κάθε στιγμή. Όμως, δεν είναι και πολύ σαφή τα χρονικά όρια στα οποία γίνονται οι εναλλαγές των συχνοτήτων. Υπάρχει μια επικάλυψη στον άξονα του χρόνου, η οποία δημιουργεί την εσφαλμένη εντύπωση πως υπάρχουν ταυτόχρονα δύο συχνότητες στο σήμα. Αυτό συμβαίνει επειδή κάποια από τα τμήματα του σήματος που αποκόβουμε περιέχουν και τις δύο συχνότητες, όχι ταυτόχρονα, αλλά διαδοχικά (όπως στο σχήμα 2.1). Αν μικρύνουμε το μήκος του παραθύρου, θα δούμε ότι ναι μεν μπορούμε να παρατηρήσουμε καλύτερα τις μεταβολές των συχνοτήτων στο χρόνο, αλλά έχουμε χάσει την καλή ανάλυση στον άξονα των συχνοτήτων. Καταλήγουμε λοιπόν στο συμπέρασμα ότι με ένα μικρό παράθυρο έχουμε καλύτερη ανάλυση στο χρόνο, αλλά χειρότερη ανάλυση στην συχνότητα. Αντίθετα, με ένα μεγάλο παράθυρο έχουμε καλή ανάλυση στην συχνότητα και κακή ανάλυση στον χρόνο. Αυτό αναφέρεται ως νόμος της απροσδιοριστίας του *Heisenberg*, που λέει ότι δεν μπορούμε να ξέρουμε ποιά

συγκεκριμένη συχνότητα υπάρχει σε μια χρονική στιγμή, αλλά μπορούμε μόνο να ξέρουμε ποιιά ζώνη συχνοτήτων υπάρχει σε ένα χρονικό διάστημα. Αυτή η αρχή διατυπώνει και την μαθηματική ανισότητα, που λέει ότι το γινόμενο  $\Delta f \Delta t$  είναι πάντα μεγαλύτερο ή ίσο από  $1/4\pi$ . Στο σχήμα 2.6 βλέπουμε παραστατικά την διάσπαση του πεδίου χρόνος/συχνότητα πάνω στο οποίο αναπτύσσεται ο μετασχηματισμός STFT. Εκτός από το γεγονός ότι ο μετασχηματισμός υπόκειται στην αρχή απροσδιοριστίας, παρατηρούμε ότι ολόκληρο το επίπεδο τεμαχίζεται σε διακριτά τμήματα, όλα ίσα μεταξύ τους. Αυτό σημαίνει πως σε όλο το πεδίο έχουμε σταθερή διακριτότητα, σταθερή δηλαδή ανάλυση στους άξονες της συχνότητας και του χρόνου. Αν προσπαθήσουμε να βελτιώσουμε την ανάλυση στον χρόνο, θα χειροτερέψει η ανάλυση στην συχνότητα, και αντίστροφα. Αυτό είναι ένας ακόμα περιορισμός που ξεπερνιέται με την χρήση του μετασχηματισμού Wavelet.

Ένα άλλο πρόβλημα με τον μετασχηματισμό αυτό, είναι ότι η τελική μορφή



Σχήμα 2.6: Διακριτότητα συχνότητας - χρόνου

$$\Delta t \cdot \Delta f \geq \frac{1}{4\pi}$$

του φάσματος που παίρνουμε, αλλάζει όχι μόνο με το μήκος, αλλά και με τον τύπο της συνάρτησης που χρησιμοποιούμε για να απομονώσουμε τα τμήματα του σήματος. Άλλο αποτέλεσμα θα δώσει ένα παράθυρο Hamming, άλλη ένα Kaiser, κλπ.



## 2.2 Παραδείγματα χρονοσυχνοτικής ανάλυσης χρησιμοποιώντας τον αλγόριθμο FFT

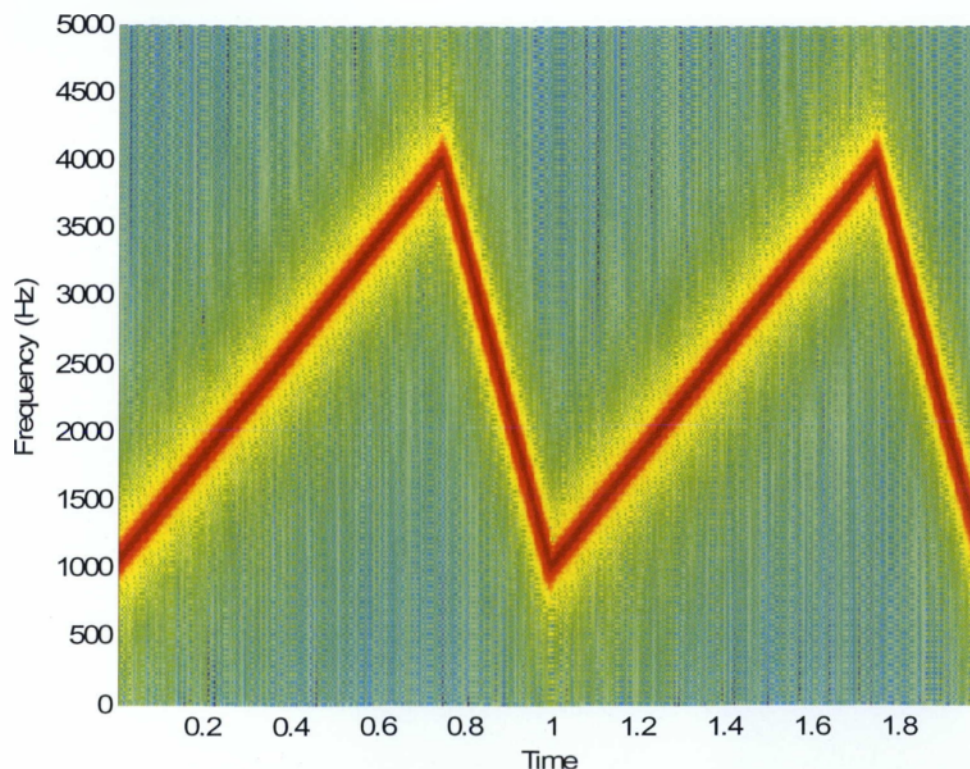
Η εργαλειοθήκη (toolbox) επεξεργασίας σήματος της γλώσσας MATLAB παρέχει μια συνάρτηση που ονομάζεται φασματόγραμμα (spectrogram), και περιγράφει την εξέλιξη του μέτρου (magnitude) του σήματος στον χρόνο. Συγκεκριμένα η συνάρτηση Spectrogram μιας ακολουθίας είναι το μέτρο του χρονικά εξαρτημένου μετασχηματισμού Fourier (Short-Time Fourier Transform, STFT) έναντι του χρόνου.

Ακολουθούν ορισμένα παραδείγματα φασματογραμμάτων για σήματα που έχουμε συνθέσει οι ίδιοι.

Για να παρουσιάσουμε το φασματόγραμμα (spectrogram) ενός σήματος FM χρησιμοποιούμε τον παρακάτω κώδικα:

### Παράδειγμα 1

```
fs = 10000;
t = 0:1/fs:2;
x = vco(sawtooth(2*pi*t,.75),[0.1 0.4]*fs,fs);
spectrogram(x,kaiser(256,5),220,512,fs,'yaxis')
```



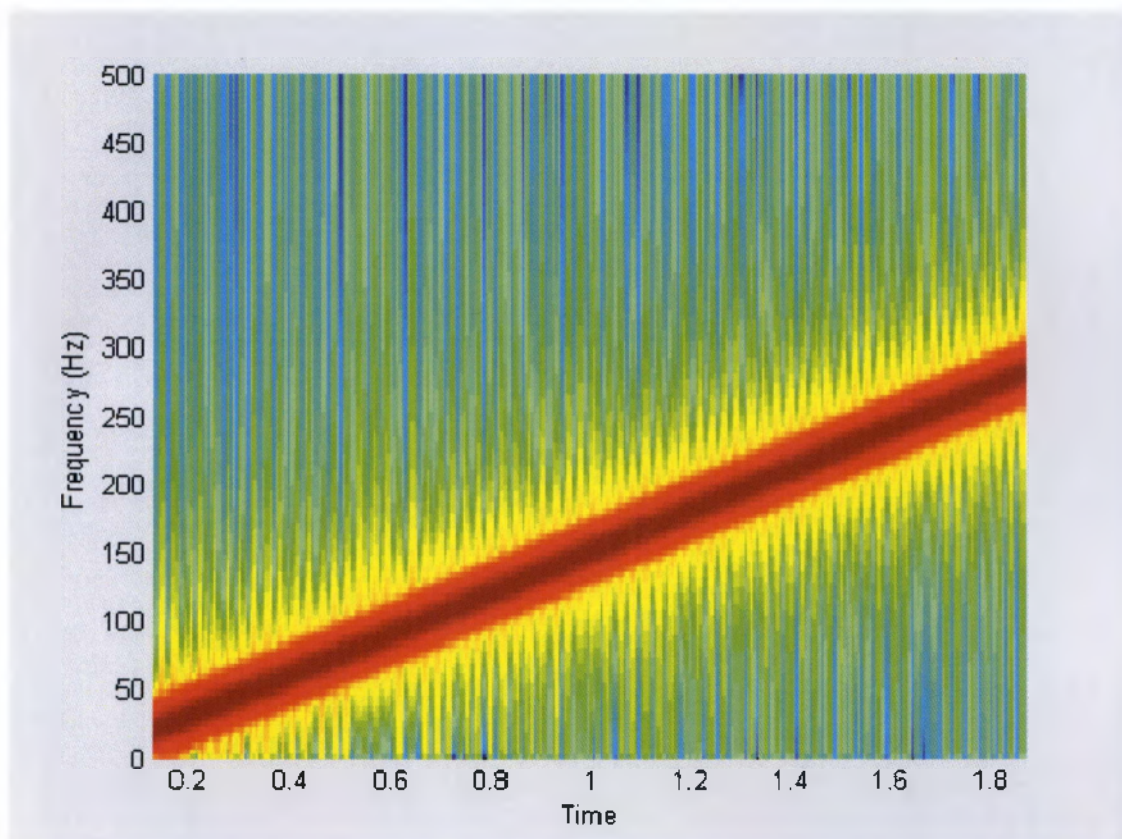
## Παράδειγμα 2

Για να υπολογίσουμε το spectrogram ενός chirp:

```
t = 0:0.001:2;
```

```
y = chirp(t,0,1,150);
```

```
spectrogram(y,256,250,256,1E3,'yaxis')
```



**Παράδειγμα 3**

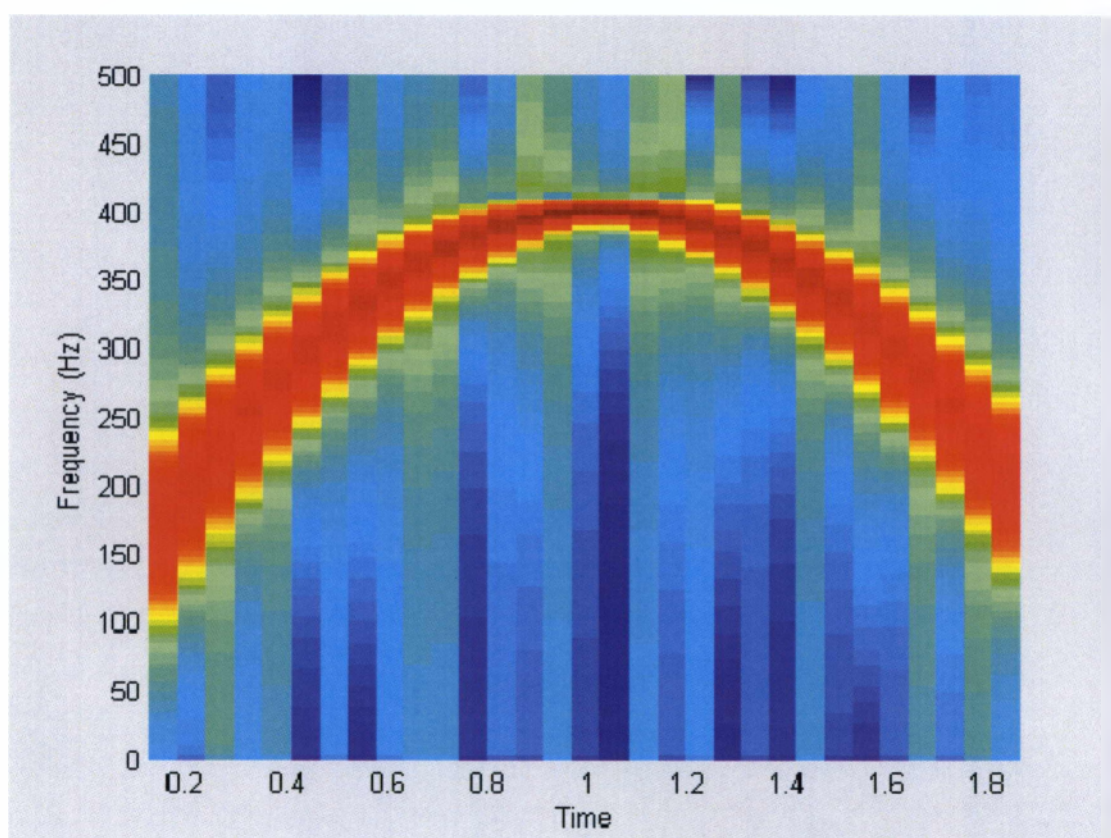
Για να υπολογίσουμε το spectrogram μιας κυρτής εξίσωσης δευτέρου βαθμού:

```
t = -1:0.001:1;
```

```
fo = 100; f1 = 400;
```

```
y = chirp(t,fo,1,f1,'q',[],'convex');
```

```
spectrogram(y,256,200,256,1000,'yaxis')
```





**Παράδειγμα 4**

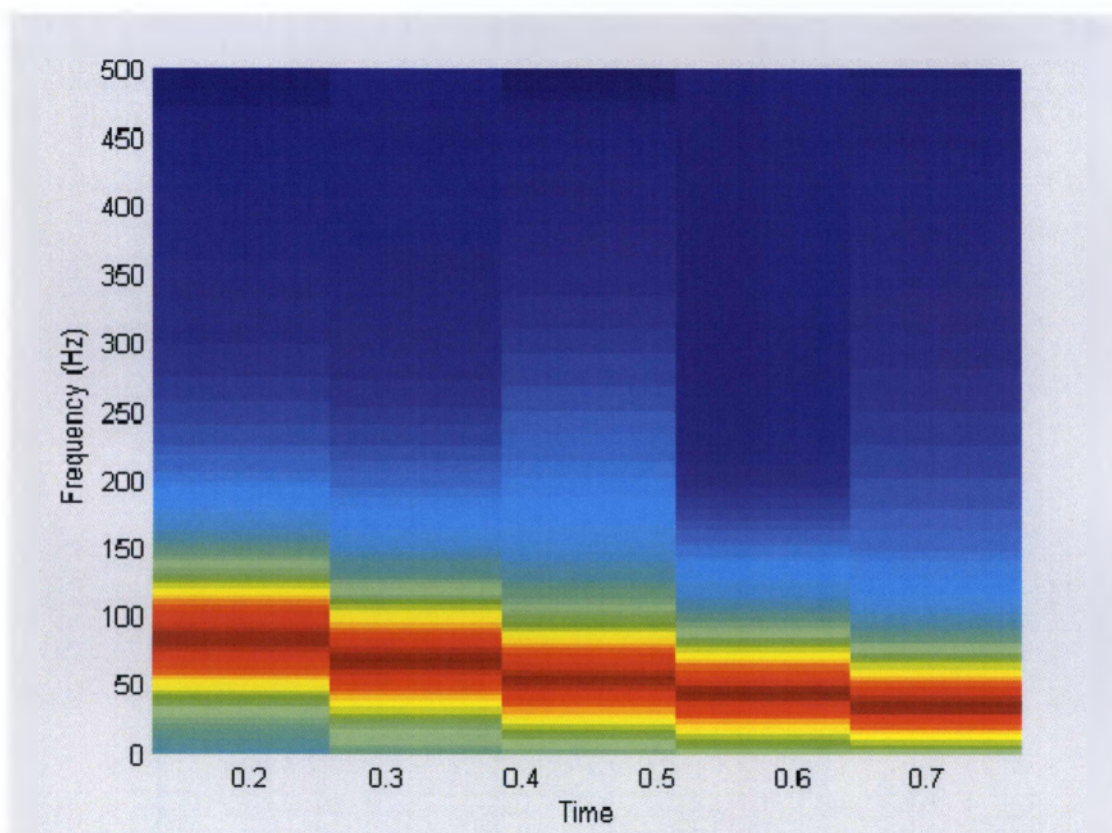
Για να υπολογίσουμε το spectrogram μιας κοίλης εξίσωσης δευτέρου βαθμού:

```
t = 0:0.001:1;
```

```
fo = 100; f1 = 25;
```

```
y = chirp(t,fo,1,f1,'q',[],'concave');
```

```
spectrogram(y,hanning(256),128,256,1000,'yaxis')
```



**Παράδειγμα 5**

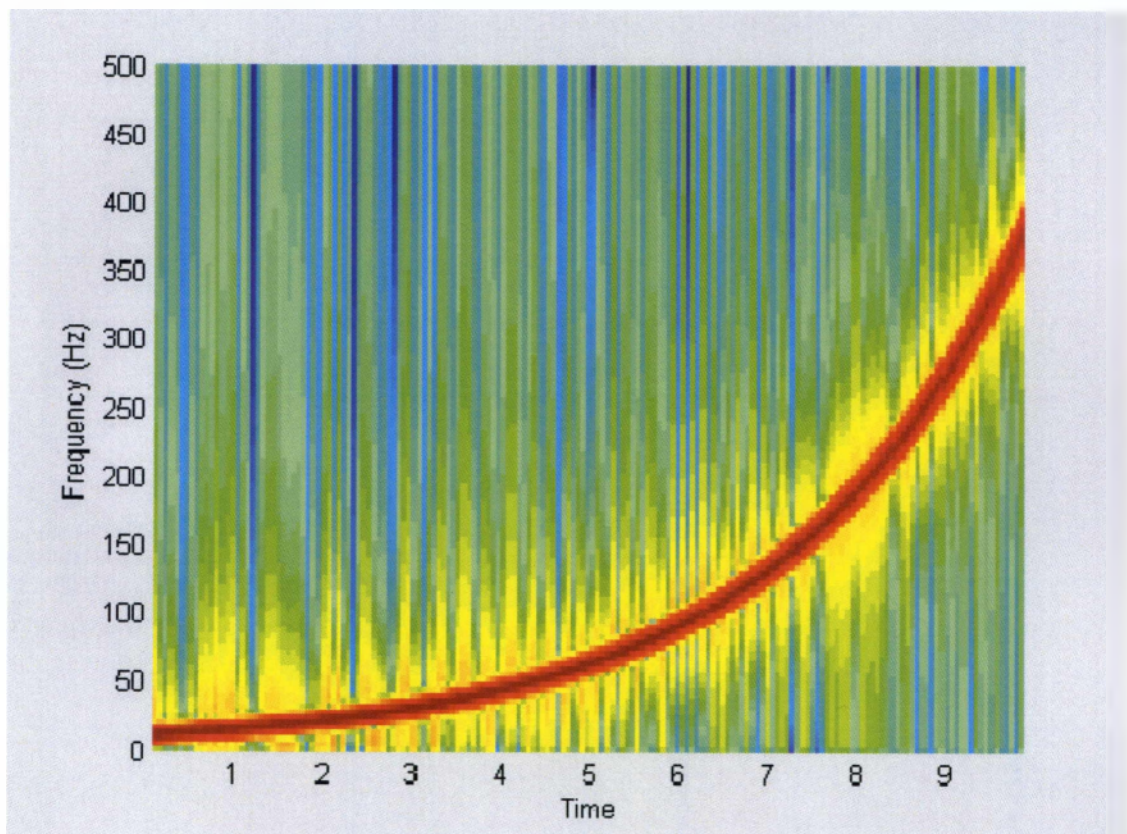
Για να υπολογίσουμε το spectrogram ενός λογάριθμου:

```
t = 0:0.001:10;
```

```
fo = 10; f1 = 400;
```

```
y = chirp(t,fo,10,f1,'logarithmic');
```

```
spectrogram(y,256,200,256,1000,'yaxis')
```



## Κεφάλαιο 3

Η *μηχανική μάθηση* (*machine learning*) είναι ένας από τους παλαιότερους ερευνητικούς τομείς της ΤΝ (Τεχνητής Νοημοσύνης) αν και ο όρος "Μηχανική Μάθηση" εισήχθηκε στις αρχές της δεκαετίας του 1980. Για πολλούς μάλιστα ταυτίζεται με την ΤΝ καθώς η δυνατότητα μάθησης πρέπει να αποτελεί το βασικότερο χαρακτηριστικό μιας "οντότητας" που θέλει να καλείται νοήμων, με την ευρύτερη έννοια του όρου.

Ως ερευνητικός τομέας, η μηχανική μάθηση έχει ασχοληθεί με πλήθος επί μέρους προβλημάτων και έχει προτείνει διάφορες μεθόδους επίλυσης. Η διαφοροποίηση συνίσταται στις στρατηγικές μάθησης, στον τρόπο αναπαράστασης της γνώσης, στην ύπαρξη ή όχι αρχικής γνώσης καθώς και δεδομένων για την εκπαίδευση του συστήματος, κτλ. Η κατασκευή βέβαια μιας μηχανής με γενικευμένη δυνατότητα μάθησης αποτελεί για την ώρα ουτοπία. Αυτό που έχει γίνει όμως ξεκάθαρο είναι ότι η αποτελεσματικότητα ενός τέτοιου συστήματος εξαρτάται άμεσα από την ποσότητα αλλά και την ποιότητα της αρχικής γνώσης με την οποία έχει εφοδιαστεί. Αυτό με τη σειρά του οδηγεί στη διαπίστωση ότι χωρίς καλή κατανόηση του τι είναι αυτό που απαιτείται να "μάθει" το σύστημα, η διαδικασία της μάθησης γίνεται εξαιρετικά δύσκολη.

Τα τελευταία χρόνια έχουν αναπτυχθεί με επιτυχία πολλές εφαρμογές της μηχανικής μάθησης, όπως προγράμματα που μαθαίνουν να εντοπίζουν χρήση κλεμμένων πιστωτικών καρτών, συστήματα διήθησης πληροφοριών που μαθαίνουν τις προτιμήσεις των χρηστών και αυτόνομα οχήματα που μαθαίνουν να κινούνται σε δημόσιους δρόμους. Η λεπτομερής κατανόηση του τρόπου λειτουργίας αλγορίθμων επεξεργασίας πληροφοριών για μηχανική μάθηση μπορεί να οδηγήσουν σε καλύτερη κατανόηση των δυνατοτήτων και μειονεκτημάτων της ανθρώπινης μάθησης.

Στο μέρος αυτό παρατίθενται οι σημαντικότερες μεθοδολογίες που έχουν αναπτυχθεί στο πλαίσιο της μηχανικής μάθησης, [12]. Ιδιαίτερη έμφαση δίνεται στα *νευρωνικά δίκτυα* (*neural nets*) τα οποία παρουσιάζουν μεγάλη πρακτική εφαρμογή. Επίσης, γίνεται αναφορά και στην *ανακάλυψη γνώσης σε βάσεις δεδομένων* (*knowledge discovery in databases, KDD*) που αποτελεί μια εφαρμογή της μηχανικής μάθησης.



### 3.1 Μηχανική Μάθηση

Ένα φυσικό ή τεχνητό σύστημα επεξεργασίας πληροφορίας συμπεριλαμβανομένων εκείνων με δυνατότητες αντίληψης, μάθησης, συλλογισμού, λήψης απόφασης, επικοινωνίας και δράσης ονομάζεται *γνωστικό σύστημα (cognitive system)*. Η έννοια της μάθησης σε ένα γνωστικό σύστημα όπως γίνεται αντιληπτή στην καθημερινή ζωή, μπορεί να συνδεθεί με δύο βασικές ιδιότητες:

- την ικανότητά του στην πρόσκτηση γνώσης κατά την αλληλεπίδρασή του με το περιβάλλον, μέσα στο οποίο δραστηριοποιείται, και
- την ικανότητά του να βελτιώνει με την επανάληψη τον τρόπο με τον οποίο εκτελεί μία ενέργεια (και συνεπώς και την απόδοσή του).

Έχουν προταθεί διάφοροι ορισμοί για τη μάθηση, όπως των:

- Simon (1983), *"η μάθηση σηματοδοτεί προσαρμοστικές αλλαγές σε ένα σύστημα με την έννοια ότι αυτές του επιτρέπουν να κάνει την ίδια εργασία, ή εργασίες της ίδιας κατηγορίας, πιο αποδοτικά και αποτελεσματικά την επόμενη φορά"*.
- Minsky (1985), *"μάθηση είναι να κάνουμε χρήσιμες αλλαγές στο μυαλό μας"*.
- Michalski (1986), *"μάθηση είναι η δημιουργία ή η αλλαγή της αναπαράστασης των εμπειριών"*.

Ένα σύστημα με δυνατότητα μάθησης μεταβάλλεται διαρκώς προς το "καλύτερο", όπως και αν αυτό ορίζεται, αναφορικά με τις λειτουργίες που είναι σε θέση να εκτελέσει. Αυτό που είναι δύσκολο να προσδιοριστεί είναι η ακριβής φύση αυτών των αλλαγών καθώς και ο τρόπος με τον οποίο αυτές μπορούν να αναπαρασταθούν. Στη γενικότερη περίπτωση, η μάθηση προσδιορίζεται ως πρόσκτηση επιπλέον γνώσης σε μία κατάλληλη αναπαράσταση. Ένα μη βιολογικό σύστημα με δυνατότητα μάθησης δομεί ή μετασχηματίζει, σε μια αυστηρά καθορισμένη γλώσσα αναπαράστασης, προτάσεις τις οποίες και αποθηκεύει για μελλοντική χρήση. Δηλαδή, η βασική παραδοχή είναι ότι η λειτουργία του συστήματος επιφέρει μεταβολές στη βάση γνώσης του.

Η μάθηση με βάση το παραπάνω πλαίσιο αφορά τα συστήματα που ανήκουν στην συμβολική προσέγγιση της ΤΝ και δεν αποτελεί τη μόνη λύση. Για παράδειγμα, τα τεχνητά νευρωνικά δίκτυα που ανήκουν στη δεύτερη προσέγγιση της ΤΝ, τη μη συμβολική (ή

συνδετική), έχουν δυνατότητα μάθησης μετασχηματίζοντας την εσωτερική τους δομή, παρά καταχωρώντας κατάλληλα αναπαριστάμενη γνώση.

Η βελτίωση του τρόπου εκτέλεσης μιας ενέργειας με την επανάληψη είναι κάτι εξίσου δύσκολο. Εκτός του ότι πρέπει να προσδιοριστούν με ακρίβεια οι παράμετροι που περιγράφουν αυτή τη βελτίωση, μεγάλη προσοχή απαιτεί η επίδραση της μεταβολής αυτής στις υπόλοιπες ενέργειες που μπορεί να εκτελέσει το σύστημα.

### **Ορισμός μηχανικής μάθησης**

Ο άνθρωπος προσπαθεί να κατανοήσει το περιβάλλον του παρατηρώντας το και δημιουργώντας μια απλοποιημένη (αφαιρετική) εκδοχή του που ονομάζεται μοντέλο (*model*). Η δημιουργία ενός τέτοιου μοντέλου, ονομάζεται επαγωγική μάθηση (*inductive learning*) ενώ η διαδικασία γενικότερα ονομάζεται επαγωγή (*induction*). Επιπλέον ο άνθρωπος έχει τη δυνατότητα να οργανώνει και να συσχετίζει τις εμπειρίες και τις παραστάσεις του δημιουργώντας νέες δομές που ονομάζονται πρότυπα (*patterns*). Η δημιουργία μοντέλων ή προτύπων από ένα σύνολο δεδομένων, από ένα υπολογιστικό σύστημα, ονομάζεται μηχανική μάθηση (*machine learning*). Έχουν προταθεί διάφοροι ορισμοί για τη μηχανική μάθηση, όπως των:

- Carbonell (1987), "... η μελέτη υπολογιστικών μεθόδων για την απόκτηση νέας γνώσης, νέων δεξιοτήτων και νέων τρόπων οργάνωσης της υπάρχουσας γνώσης".
- Mitchell (1997), "Ένα πρόγραμμα υπολογιστή θεωρείται ότι μαθαίνει από την εμπειρία  $E$  σε σχέση με μια κατηγορία εργασιών  $T$  και μια μετρική απόδοσης  $P$ , αν η απόδοση του σε εργασίες της  $T$ , όπως μετριοούνται από την  $P$ , βελτιώνονται με την εμπειρία  $E$ ".
- Witten & Frank (2000), "Κάτι μαθαίνει όταν αλλάζει τη συμπεριφορά του κατά τέτοιο τρόπο ώστε να αποδίδει καλύτερα στο μέλλον".

### **Είδη μηχανικής μάθησης**

Έχουν αναπτυχθεί πολλές τεχνικές μηχανικής μάθησης οι οποίες χρησιμοποιούνται ανάλογα με τη φύση του προβλήματος και εμπίπτουν σε ένα από τα παρακάτω δυο είδη:

- μάθηση με επίβλεψη (*supervised learning*) ή μάθηση με παραδείγματα (*learning from examples*),
- μάθηση χωρίς επίβλεψη (*unsupervised learning*) ή μάθηση από παρατήρηση (*learning from observation*).

Στη μάθηση με επίβλεψη το σύστημα καλείται να "μάθει" μια έννοια ή συνάρτηση από ένα σύνολο δεδομένων, η οποία αποτελεί περιγραφή ενός μοντέλου. Ονομάζεται έτσι επειδή θεωρείται ότι υπάρχει κάποιος "επιβλέπων" ο οποίος παρέχει τη σωστή τιμή εξόδου της συνάρτησης, για τα δεδομένα που εξετάζονται.

Αντίθετα, στη μάθηση χωρίς επίβλεψη το σύστημα πρέπει μόνο του να ανακαλύψει συσχετίσεις ή ομάδες σε ένα σύνολο δεδομένων, δημιουργώντας *πρότυπα*, χωρίς να είναι γνωστό αν υπάρχουν, πόσα και ποια είναι.

Τα μοντέλα περιγράφουν το σύνολο των δεδομένων και χαρακτηρίζονται και ως *μοντέλα πρόβλεψης (predictive models)* επειδή προβλέπουν την τιμή μιας μεταβλητής. Εκτός από τις δυνατότητες πρόβλεψης επιπλέον έχουν και κάποιες δυνατότητες πληροφόρησης επειδή δίνουν και ποιοτικές πληροφορίες για τα δεδομένα. Αντίθετα, τα πρότυπα έχουν τοπικό χαρακτήρα, δηλαδή το καθένα περιγράφει ένα μέρος των δεδομένων και χαρακτηρίζονται ως *πρότυπα πληροφόρησης (informative patterns)* επειδή περιγράφουν συσχετίσεις μεταξύ των δεδομένων.

### 3.1.1 Μάθηση με Επίβλεψη

Στη μάθηση με επίβλεψη το σύστημα πρέπει να "μάθει" επαγωγικά μια συνάρτηση που ονομάζεται *συνάρτηση στόχος (target function)* και αποτελεί έκφραση του μοντέλου που περιγράφει τα δεδομένα. Η συνάρτηση στόχος (συμβολίζεται συνήθως με  $c$ ) χρησιμοποιείται για την πρόβλεψη της τιμής μιας μεταβλητής, που ονομάζεται *εξαρτημένη μεταβλητή* ή *μεταβλητή εξόδου*, βάσει των τιμών ενός συνόλου μεταβλητών, που ονομάζονται *ανεξάρτητες μεταβλητές* ή *μεταβλητές εισόδου* ή *χαρακτηριστικά*.

Το σύνολο των διαφορετικών δυνατών τιμών εισόδου της συνάρτησης, δηλαδή το πεδίο ορισμού της, ονομάζεται *σύνολο των περιπτώσεων* ή *στιγμιότυπων (instances)* και συμβολίζεται με  $X$ . Κάθε περίπτωση (ή στιγμιότυπο) περιγράφεται από ένα σύνολο *χαρακτηριστικών (attributes* ή *features)*. Ένα υποσύνολο του συνόλου των περιπτώσεων για τα οποία γνωρίζουμε την τιμή της μεταβλητής εξόδου, ονομάζεται *σύνολο δεδομένων εκπαίδευσης* ή *παραδείγματα* και συμβολίζεται με  $D$ .

Για να προσεγγίσει το σύστημα όσο το δυνατόν καλύτερα τη συνάρτηση στόχο εξετάζει διάφορες εναλλακτικές συναρτήσεις οι οποίες ονομάζονται *υποθέσεις* και συμβολίζονται με  $h$ . Το σύνολο όλων των δυνατών υποθέσεων που το πρόγραμμα μάθησης πρέπει να

εξετάσει προκειμένου να βρει τη συνάρτηση στόχο ονομάζεται *σύνολο υποθέσεων* και συμβολίζεται με  $H$ . Κάθε υπόθεση  $h \in H$ , αναπαριστά είτε μια λογική συνάρτηση  $h: X \rightarrow \{0, 1\}$  ή μια πραγματική συνάρτηση  $h: X \rightarrow R$ .

Η επαγωγική μάθηση στηρίζεται στην *υπόθεση επαγωγικής μάθησης (inductive learning hypothesis)*, σύμφωνα με την οποία κάθε υπόθεση  $h$  που έχει βρεθεί να προσεγγίζει καλά τη συνάρτηση στόχο για ένα αρκετά μεγάλο σύνολο παραδειγμάτων, θα προσεγγίζει το ίδιο καλά τη συνάρτηση στόχο και για περιπτώσεις που δεν έχει εξετάσει.

Στην μάθηση με επίβλεψη διακρίνονται δυο είδη προβλημάτων (learning tasks), τα προβλήματα *ταξινόμησης* και τα προβλήματα *παρεμβολής*. Η *ταξινόμηση (classification)* αφορά στη δημιουργία μοντέλων πρόβλεψης διακριτών τάξεων (κλάσεων/ κατηγοριών) όπως για παράδειγμα η ομάδα αίματος, ενώ η *παρεμβολή (regression)* αφορά στη δημιουργία μοντέλων πρόβλεψης αριθμητικών τιμών. Οι κυριότερες τεχνικές μηχανικής μάθησης με επίβλεψη είναι:

- *Μάθηση εννοιών (Concept Learning)*
- *Δένδρα ταξινόμησης ή απόφασης (Classification or Decision Trees)*
- *Μάθηση Κανόνων (Rule Learning)*
- *Μάθηση κατά Περίπτωση (Instance Based Learning)*
- *Μάθηση κατά Bayes*
- *Γραμμική παρεμβολή (Linear Regression)*
- *Νευρωνικά Δίκτυα (Neural Networks)*
- *Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines, SVMs)*

### **Η μηχανική μάθηση με επίβλεψη ως πρόβλημα αναζήτησης**

Μια από τις περισσότερο συνηθισμένες προσεγγίσεις του προβλήματος της μηχανικής μάθησης με επίβλεψη είναι αυτή της αντιμετώπισής του ως προβλήματος αναζήτησης (search problem). Σύμφωνα με αυτήν την προσέγγιση, η *μηχανική μάθηση* μπορεί να θεωρηθεί ως η αναζήτηση σε ένα χώρο πιθανών υποθέσεων, ώστε να βρεθεί εκείνη που ταιριάζει καλύτερα στα υπό εξέταση δεδομένα και στην πιθανώς προϋπάρχουσα γνώση.

Ένα σύστημα (πρόγραμμα) μηχανικής μάθησης απαιτεί κάποια *γλώσσα αναπαράστασης* του κόσμου του προβλήματος, δηλαδή των υποθέσεων. Όλες οι τεχνικές ανα-



παράστασης γνώσης που έχουν αναφερθεί μπορούν να χρησιμοποιηθούν, ανάλογα βέβαια με την περίπτωση. Για παράδειγμα, προγράμματα με στόχο την ταξινόμηση αντικειμένων μπορούν να αναπαραστήσουν τις υποθέσεις με κατηγορηματικό λογισμό (predicate calculus), πλαίσια (frames), αντικείμενα (objects), κτλ.

Επιπλέον, απαιτείται ένα σύνολο από τελεστές (*operators*) οι οποίοι θα επιτρέψουν στο σύστημα να οδηγηθεί σε μία γενίκευση, έναν ευριστικό κανόνα ή ένα πλάνο που ικανοποιεί τους στόχους του. Τέτοιοι τελεστές μπορούν, μέσω μετασχηματισμών σε αρχικές αναπαραστάσεις, να επιτελέσουν γενικεύσεις, εξειδικεύσεις και ενδεχομένως δομικούς μετασχηματισμούς που θα επιτρέψουν την επίτευξη της μάθησης. Για παράδειγμα, ένα σύστημα που θα κληθεί να μάθει την έννοια της μπάλας έχοντας σαν είσοδο την περιγραφή από διάφορες μπάλες, πρέπει να είναι σε θέση να αντικαταστήσει τις διάφορες τιμές του χρώματος με μια μεταβλητή, καθώς το χρώμα δεν είναι το ιδιαίτερο χαρακτηριστικό της μπάλας και να επικεντρωθεί στο σφαιρικό της σχήμα.

Η γλώσσα αναπαράστασης και οι διάφοροι τελεστές δημιουργούν το σύνολο ή *χώρο των υποθέσεων* μέσα στον οποίο ένα σύστημα μηχανικής μάθησης αναζητά τη "σωστή" υπόθεση. Η πολυπλοκότητα αυτού του χώρου καθορίζει και το βαθμό δυσκολίας του εκάστοτε προβλήματος. Όπως και στα κλασικά προβλήματα αναζήτησης έτσι και στα προβλήματα μηχανικής μάθησης η χρήση ευριστικών μηχανισμών που καθοδηγούν αποτελεσματικά την αναζήτηση είναι απαραίτητη ειδικά σε μεγάλους χώρους υποθέσεων. Ένας τέτοιος ισχυρός ευριστικός μηχανισμός είναι η παρατήρηση ότι διαδοχικές γενικεύσεις μιας υπόθεσης ορίζουν μια διάταξη (*ordering*) στο χώρο των υποθέσεων. Για παράδειγμα, μια τέτοια διάταξη θα μπορούσε να είναι η:

Τίγρης < Σαρκοφάγο < Θηλαστικό < Ζωντανός Οργανισμός

Η χρήση τέτοιων διατάξεων μπορεί να βελτιώσει σημαντικά την αναζήτηση στο χώρο των υποθέσεων, περιορίζοντας σημαντικά το χώρο αναζήτησης.

### 3.1.1.α Μάθηση κατά Περίπτωση

Σε αντίθεση με τις μεθόδους μηχανικής μάθησης που αναφέρθηκαν ως τώρα και οι οποίες κωδικοποιούν τα παραδείγματα εκπαίδευσης σε μια συμπαγή περιγραφή, στη *μάθηση κατά περίπτωση* (*instance-based learning*) τα δεδομένα εκπαίδευσης διατηρούνται αυτούσια. Όταν ένα τέτοιο σύστημα κληθεί να αποφασίσει για την κατηγορία μιας νέας περίπτωσης, εξετάζει εκείνη τη στιγμή τη σχέση της με τα ήδη αποθηκευμένα

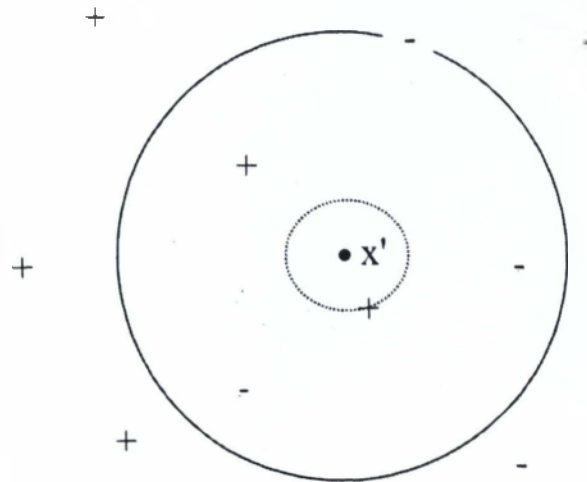
παραδείγματα. Δηλαδή η μέθοδος αυτή αναβάλλει τη μάθηση έως ότου εμφανιστεί μια νέα περίπτωση (στιγμιότυπο) και για το λόγο αυτό ονομάζεται *αναθλητική μάθηση (lazy learning)* σε αντίθεση με τις άλλες οι οποίες μπορεί να χαρακτηριστούν ως *έγκαιρες μέθοδοι μάθησης (eager learners)*, αφού μαθαίνουν το μοντέλο από τα αποθηκευμένα παραδείγματα του συνόλου εκπαίδευσης, χωρίς να περιμένουν την άφιξη της νέας περίπτωσης.

Χαρακτηριστικός αλγόριθμος αυτής της κατηγορίας είναι ο αλγόριθμος των *k-κοντινότερων γειτόνων (k-Nearest Neighbors)*, στον οποίο γίνεται η παραδοχή ότι τα διάφορα παραδείγματα μπορεί να αναπαρασταθούν ως σημεία σε κάποιον  $n$ -διάστατο Ευκλείδειο χώρο  $R^n$  όπου  $n$  ο αριθμός των χαρακτηριστικών (ανεξάρτητων μεταβλητών). Κάθε νέα περίπτωση τοποθετείται στο χώρο αυτό ως νέο σημείο και η τιμή του προσδιορίζεται με βάση το χαρακτηρισμό των  $k$  γειτονικών σημείων. Οι κοντινότεροι γείτονες μιας περίπτωσης υπολογίζονται με βάση τη γνωστή από τη γεωμετρία Ευκλείδεια απόστασή τους.

Για παράδειγμα, η απόσταση μιας νέας περίπτωσης  $x'$  που περιγράφεται από το σύνολο χαρακτηριστικών  $\langle a_1(x'), a_2(x'), a_3(x'), \dots, a_n(x') \rangle$ , όπου  $a_r(x')$  είναι το  $r$  χαρακτηριστικό της και ενός αποθηκευμένου παραδείγματος  $x$ , που περιγράφεται από το σύνολο χαρακτηριστικών  $\langle a_1(x), a_2(x), a_3(x), \dots, a_n(x), y(x) \rangle$ , είναι το άθροισμα των Ευκλείδειων αποστάσεων όλων των χαρακτηριστικών των 2 σημείων. Δηλαδή:

$$d(x, x') = \sqrt{\sum_{r=1}^n (a_r(x) - a_r(x'))^2}$$

Στο Σχήμα 3.1, όπου υπάρχουν παραδείγματα δυο κατηγοριών, η νέα περίπτωση  $x'$  χαρακτηρίζεται ως θετική, αν ληφθεί υπ' όψη μόνο ο πλησιέστερος γείτονας (1- Nearest Neighbor) και ως αρνητική αν ληφθούν υπ' όψη οι πέντε πλησιέστεροι γείτονες (5-Nearest Neighbors) καθώς η πλειοψηφία αυτών έχει αρνητικό χαρακτηρισμό (εξωτερικός κύκλος στη σχήμα).



Σχήμα 3.1: Προσδιορισμός κατηγορίας με βάση τον 1 και τους 5 κοντινότερους γείτονες.

### 3.1.2 Άλλες Τεχνικές Μάθησης με Επίβλεψη

Εκτός από τις τεχνικές μηχανικής μάθησης με επίβλεψη που παρουσιάστηκαν στις προηγούμενες ενότητες, υπάρχουν και άλλες, οι σημαντικότερες από τις οποίες είναι η παρεμβολή, τα νευρωνικά δίκτυα και οι μηχανές διανυσμάτων υποστήριξης.

### 3.1.3 Μάθηση Χωρίς Επίβλεψη

Στη μάθηση χωρίς επίβλεψη το σύστημα έχει στόχο να ανακαλύψει συσχετίσεις και ομάδες από τα δεδομένα, βασιζόμενο μόνο στις ιδιότητές τους. Σαν αποτέλεσμα προκύπτουν πρότυπα (περιγραφές), κάθε ένα από τα οποία περιγράφει ένα μέρος από τα δεδομένα. Παραδείγματα προτύπων πληροφόρησης είναι οι κανόνες συσχέτισης (*association rules*) και οι ομάδες (*clusters*), οι οποίες προκύπτουν από τη διαδικασία της ομαδοποίησης (*clustering*).

## 3.2 Ανακάλυψη Γνώσης σε Βάσεις Δεδομένων

Η ανακάλυψη γνώσης σε βάσεις δεδομένων (*Knowledge Discovery in Databases - KDD*) είναι μία σύνθετη διαδικασία για τον προσδιορισμό έγκυρων, νέων, χρήσιμων και κατανοητών σχέσεων-προτύπων σε δεδομένα. Αν και ως όρος είναι σχετικά πρόσφατος, αποτελεί μια σημαντική εφαρμογή σε πραγματικές συνθήκες και σε μεγάλη κλίμακα των ερευνητικών αποτελεσμάτων της μηχανικής μάθησης και της στατιστικής.

Η διαδικασία ανακάλυψης γνώσης είναι μια ολοκληρωμένη διαδικασία που περιλαμβάνει την επεξεργασία των δεδομένων, την εφαρμογή των αλγορίθμων ανακάλυψης γνώσης και τέλος την ερμηνεία των αποτελεσμάτων. Χρησιμοποιεί τεχνικές από πολλούς τομείς, όπως στατιστική, μηχανική μάθηση, βάσεις δεδομένων, αναγνώριση προτύπων, πράκτορες, επεξεργασία φυσικής γλώσσας, κτλ.

Οι πιο σημαντικές διαφορές μεταξύ ανακάλυψης γνώσης και μηχανικής μάθησης οφείλονται στο γεγονός ότι η πρώτη εφαρμόζεται σε έναν μεγάλο όγκο δεδομένων τα οποία οργανώνονται σε βάσεις δεδομένων που συνήθως έχουν σχεδιαστεί για άλλο σκοπό. Αντίθετα, στη μηχανική μάθηση τα δεδομένα είναι πολύ λιγότερα και προσεκτικά επιλεγμένα ώστε να εξυπηρετούν καλύτερα τον εκάστοτε σκοπό.

Σε μια βάση δεδομένων, ένα σύστημα ανακάλυψης γνώσης προσπαθεί να δημιουργήσει μοντέλα ή να ανακαλύψει πρότυπα τα οποία αξιολογούνται συνήθως μέσω κάποιων κριτηρίων ποιότητας. Επειδή όμως η δημιουργία όλων των δυνατών περιγραφών (μοντέλων ή προτύπων) είναι υπολογιστικά αδύνατη, το πρόβλημα της ανακάλυψης γνώσης σε βάσεις δεδομένων είναι ουσιαστικά ένα πρόβλημα αναζήτησης: η εύρεση της καλύτερης περιγραφής από το σύνολο των δυνατών περιγραφών. Η γνώση που προκύπτει υποστηρίζεται από τη βάση δεδομένων, χωρίς αυτό όμως να σημαίνει ότι ισχύει οπωσδήποτε και στον πραγματικό κόσμο. Για παράδειγμα, από μία βάση δεδομένων με καταγεγραμμένα μόνο κόκκινα αυτοκίνητα, θα προκύψει η γνώση ότι "όλα τα αυτοκίνητα είναι κόκκινα". Αυτό, αν και είναι σωστό για τον κόσμο της βάσης δεδομένων, δεν ισχύει προφανώς στην πραγματικότητα.

Τεχνικές ανακάλυψης γνώσης χρησιμοποιούνται ευρέως σε τομείς όπως οικονομία, ιατρική, μάρκετινγκ, κτλ. Μερικά παραδείγματα εφαρμογής είναι:

- Εταιρία κινητής τηλεφωνίας θέλει να προβλέψει ποιοι από τους συνδρομητές της δε θα ανανεώσουν τη συνδρομή τους, ώστε ενδεχομένως να τους κάνει κάποια περισσότερο ελκυστική προσφορά.
- Εταιρία πωλήσεως ηλεκτρονικών συσκευών θέλει να μελετήσει τις αγοραστικές συνήθειες των πελατών της, ώστε να προγραμματίσει ανάλογα την επόμενη διαφημιστική της καμπάνια.
- Ιατρικά εργαστήρια θέλουν να συσχετίσουν ασθένειες με χαρακτηριστικά των ασθενών όπως τόπος διαμονής, διατροφικές συνήθειες, παλαιότερες ασθένειες, κτλ., ώστε να οδηγηθούν σε ιατρικές ανακαλύψεις.

Η ανακάλυψη γνώσης σε βάσεις δεδομένων απαιτεί συνήθως τη συνδρομή ενός ειδικού του τομέα όπου θα εφαρμοστεί η διαδικασία (για παράδειγμα ενός ιατρού). Πρόκειται για αλληλεπιδραστική και επαναληπτική διαδικασία, στη διάρκεια των επιμέρους σταδίων της οποίας ο ειδικός σε θέματα ανακάλυψης γνώσης καλείται να πάρει συγκεκριμένες αποφάσεις. Τυπικά μεγέθη του όγκου δεδομένων στα οποία εφαρμόζονται διαδικασίες ανακάλυψης γνώσης είναι εκατομμύρια εγγραφές και δεκάδες GBytes.

### 3.2.1 Τα Στάδια της Ανακάλυψης Γνώσης

Η ανακάλυψη γνώσης αρχίζει με την κατανόηση του τομέα στον οποίο θα εφαρμοστεί και τον προσδιορισμό του στόχου της από τη σκοπιά του χρήστη των αποτελεσμάτων. Ο ειδικός επί θεμάτων ανακάλυψης γνώσης πρέπει να συνεργαστεί με τον ειδικό του τομέα ώστε το πρόβλημα να καθοριστεί με αρκετή ακρίβεια και να είναι επιλύσιμο, τα αποτελέσματα να είναι μετρήσιμα και να είναι δυνατή η εφαρμογή τους (για παράδειγμα σε αποδεκτά χρονικά όρια). Τα επιμέρους στάδια απεικονίζονται στο Σχήμα 3.2 και περιγράφονται στη συνέχεια.



Σχήμα 3.2: Τα βασικά στάδια της διαδικασίας ανακάλυψης γνώσης.

Πρέπει να σημειωθεί ότι τις περισσότερες φορές κάποια από τα επί μέρους βήματα είναι αναγκαίο να επαναληφθούν, καθώς στην πορεία ενδέχεται να προκύψουν προβλήματα που σχετίζονται με τις αρχικές επιλογές και τα οποία δεν ήταν δυνατό να εντοπιστούν αρχικά.

#### Επιλογή

Στο στάδιο της επιλογής, δημιουργείται το σύνολο δεδομένων στο οποίο θα εφαρμοστούν οι αλγόριθμοι ανακάλυψης γνώσης. Επειδή τα δεδομένα είναι σχεδόν πάντα οργανωμένα για άλλη χρήση και οι αλγόριθμοι που εκτελούν την ανακάλυψη γνώσης συνήθως δεν μπορούν να εφαρμοστούν σε πολλαπλούς πίνακες δεδομένων, απαιτείται η εξαγωγή των



δεδομένων από αυτούς και η οργάνωσή τους σε απλούστερες δομές. Συνήθως αυτή η απαίτηση καλύπτεται από *συστήματα αποθήκευσης δεδομένων (data warehouse)* τα οποία παρέχουν στους αλγόριθμους ανακάλυψης γνώσης μία ευκολότερα προσβάσιμη *όψη (view)* των δεδομένων.

### **Προεπεξεργασία**

Στο στάδιο της *προεπεξεργασίας (preprocessing)* των δεδομένων αντιμετωπίζονται περιπτώσεις ελλιπών δεδομένων (για παράδειγμα, άδεια πεδία), πεδίων με τιμές που ουσιαστικά τα καθιστούν κενά (για παράδειγμα, Οδός = Άγνωστο), πεδίων με τιμές που υπονοούν (κατά σύμβαση) κάτι άλλο (για παράδειγμα, καταχώριση της ημερομηνίας "1/1/1900" σε πεδίο ημερομηνίας που απαιτούσε τιμή αλλά αυτή δεν ήταν διαθέσιμη), κτλ. Λόγω της φύσης των εργασιών που πραγματοποιούνται, το στάδιο αυτό ονομάζεται και *στάδιο καθαρισμού των δεδομένων (data cleaning)*.

### **Μετασχηματισμός**

Στο στάδιο του *μετασχηματισμού (transformation)* τα δεδομένα μετασχηματίζονται ώστε να διευκολύνουν την ανακάλυψη γνώσης. Τέτοιοι μετασχηματισμοί μπορεί να περιλαμβάνουν:

- την ομοιόμορφη κωδικοποίηση της ποιοτικά ίδιας πληροφορίας (για παράδειγμα, ενοποίηση ενός πεδίου με τίτλο *salary* σε έναν πίνακα με το πεδίο *payment* σε κάποιον άλλο πίνακα),
- τη μείωση του αριθμού των υπό εξέταση χαρακτηριστικών (*dimensionality reduction*) με επιλογή ορισμένων εξ' αυτών (*feature selection* ή *attribute selection*),
- τη *διακριτοποίηση (discretization)*, δηλαδή τη μετατροπή συνεχόμενων αριθμητικών τιμών σε διακριτές, κτλ.

Η *επιλογή χαρακτηριστικών* είναι μια απαραίτητη διαδικασία γιατί συνήθως τα δεδομένα περιλαμβάνουν χαρακτηριστικά τα οποία είναι είτε περιττά στη διαδικασία ανακάλυψης γνώσης (όπως για παράδειγμα τα ονόματα των ασθενών ή των πελατών μιας επιχείρησης) ή ακατάλληλα. Αν και οι περισσότεροι αλγόριθμοι σχεδιάζονται για να μαθαίνουν ποια είναι τα πιο σημαντικά χαρακτηριστικά για τη λήψη μιας απόφασης, όπως για παράδειγμα στα δένδρα απόφασης όπου επιλέγεται σε κάθε κόμβο το πιο κατάλληλο χαρακτηριστικό για να γίνει ο διαχωρισμός, εντούτοις η ύπαρξη μεγάλου αριθμού μη-σχετικών ή ακατάλληλων χαρακτηριστικών μειώνει σημαντικά την απόδοση των αλγορίθμων. Αυτό οφείλεται στο γεγονός ότι σε κάποια φάση της δημιουργίας του



δένδρου, αποφασίζεται να γίνει διακλάδωση βάσει του χαρακτηριστικού αυτού, οδηγώντας τελικά σε λανθασμένη ταξινόμηση.

Επιπλέον, η μείωση του αριθμού των διαστάσεων των δεδομένων, αφαιρώντας ορισμένα ακατάλληλα χαρακτηριστικά, μπορεί να βελτιώσει την απόδοση των αλγορίθμων όσον αφορά την ταχύτητά τους αλλά και την ποιότητα της εξαγόμενης γνώσης, εστιάζοντας στα πιο σχετικά/σημαντικά χαρακτηριστικά. Η επιλογή χαρακτηριστικών μπορεί να γίνει με αυτόματες μεθόδους αλλά ο καλύτερος τρόπος είναι να γίνει από τον ειδικό του τομέα, που κατανοεί καλύτερα τη σημασία κάθε χαρακτηριστικού.

*Διακριτοποίηση* είναι η διαδικασία μετασχηματισμού ενός χαρακτηριστικού που παίρνει (συνεχείς) αριθμητικές τιμές, σε ένα πεπερασμένο σύνολο διακριτών διαστημάτων. Η διαδικασία αυτή απαιτείται είτε γιατί ο αλγόριθμος δεν μπορεί να χειριστεί αριθμητικά δεδομένα ή δεν τα χειρίζεται σε ικανοποιητικό βαθμό, όπως ορισμένες στατιστικές μέθοδοι ομαδοποίησης και ο απλός ταξινομητής Bayes.

### **Επιλογή αλγορίθμου και εφαρμογή του**

Στο στάδιο αυτό καθορίζεται το είδος της γνώσης που θα αναζητηθεί, κάτι που έμμεσα προσδιορίζει και την κατηγορία αλγορίθμου που θα χρησιμοποιηθεί. Η εφαρμογή του αλγορίθμου είναι ένα καθαρά υπολογιστικό στάδιο, στο οποίο γίνεται η ουσιαστική ανακάλυψη της γνώσης από τα δεδομένα. Περιγράφεται και με τον όρο *εξόρυξη σε δεδομένα (data mining)*, ο οποίος συχνά χρησιμοποιείται καταχρηστικά για να περιγράψει ολόκληρη τη διαδικασία ανακάλυψης γνώσης.

Υπάρχουν δύο είδη γνώσης που προκύπτουν ως αποτελέσματα μιας διαδικασίας ανακάλυψης γνώσης: τα *μοντέλα πρόβλεψης (predictive models)* και τα *πρότυπα πληροφόρησης (informative patterns)*. Πολλές φορές προτιμούνται αλγόριθμοι που έχουν ισχυρά χαρακτηριστικά πληροφόρησης, αν και έχουν χειρότερη δυνατότητα πρόβλεψης, διότι τα αποτελέσματά τους γίνονται πιο εύκολα αντιληπτά από τον ειδικό του τομέα. Οι αλγόριθμοι που χρησιμοποιούνται συχνά προέρχονται από τον τομέα της μηχανικής μάθησης, προσαρμοσμένοι κατάλληλα στις απαιτήσεις της διαδικασίας της ανακάλυψης γνώσης σε βάσεις δεδομένων.

### **Ερμηνεία και αξιολόγηση**

Στο στάδιο αυτό γίνεται *ερμηνεία (interpretation)* και *αξιολόγηση (evaluation)* της γνώσης που προκύπτει, πιθανώς με υποβοήθηση γραφικών απεικονίσεων των περιγραφών ή και

των δεδομένων που περιγράφονται (*visualization*). Η γνώση που παράγεται μπορεί να χρησιμοποιηθεί σε ένα σύστημα γνώσης, περίπτωση κατά την οποία ίσως να απαιτείται και η επίλυση πιθανών συγκρούσεων (*conflicts*) με την υπάρχουσα γνώση.

Αξίζει τέλος να αναφερθεί ότι πολλές φορές, τα αποτελέσματα που προκύπτουν μετά από έναν κύκλο εφαρμογής των παραπάνω βημάτων, δίνουν ερέθισμα για νέου είδους αναζητήσεις, δηλαδή οδηγούν σε επανάληψη ολόκληρης της διαδικασίας με σκοπό την ανακάλυψη νέας γνώσης.

## Κεφάλαιο 4

Σε αυτό το Κεφάλαιο θα περιγράψουμε την διαδικασία που απαιτείται για την επεξεργασία των ηχητικών σημάτων. Θα δούμε ποιές εντολές χρειάζονται για να δημιουργήσουμε τα γραφήματα κάθε ήχου στο πεδίο του χρόνου καθώς και τις εντολές για τη δημιουργία των αντίστοιχων φασματογραμμάτων (spectrogram). Ολοκληρώνοντας, παρέχονται παραδείγματα φασματογραμμάτων (spectrogram) καθώς και γραφημάτων στο πεδίο του χρόνου όπου προκύπτουν ύστερα από πρόσθεση (σύνθεση) διαφορετικών σημάτων με σκοπό την εξαγωγή χρήσιμων συμπερασμάτων.

Γενικά, οι ήχοι χωρίζονται στις τρεις ακόλουθες βασικές κατηγορίες:

**Human-Related** → ήχοι από την φύση των ανθρώπων

**Geophysical** → ήχοι από την φύση της γης

**Animal-Related** → ήχοι από την φύση των ζώων

Για κάθε μια από τις παραπάνω κατηγορίες ήχων χρησιμοποίησα 3 ηχητικές τάξεις (classes) τις οποίους επεξεργάστηκα με την βοήθεια του προγράμματος **cool edit pro 2.1** με τον παρακάτω τρόπο:

## Ήχοι

### Animal - related sounds

Track09\_Animal → Κοράκι

Track15\_Animal → Σκύλος

Track57\_Animal → Κροταλιάς

### Geophysical

Track74\_Geophysical → Καταρράκτης

Track90\_Geophysical → Άνεμοι

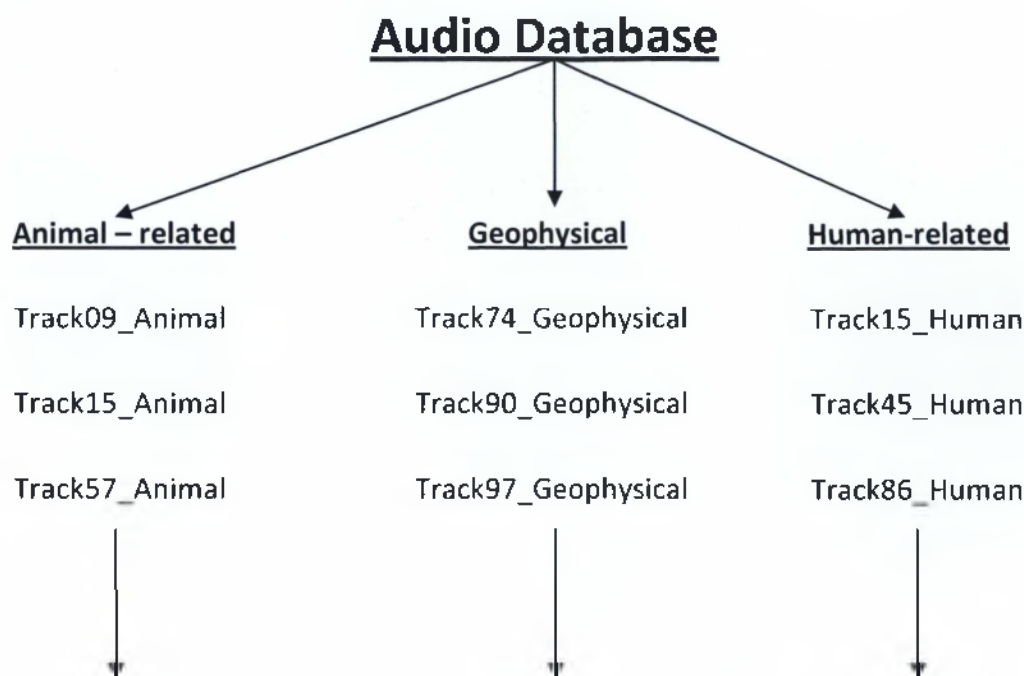
Track97\_Geophysical → Φωτιά

### Human-related sounds

Track15\_Human → Αυτοκίνητο

Track45\_Human → Πυροβολισμός

Track86\_Human → Αεροπλάνο

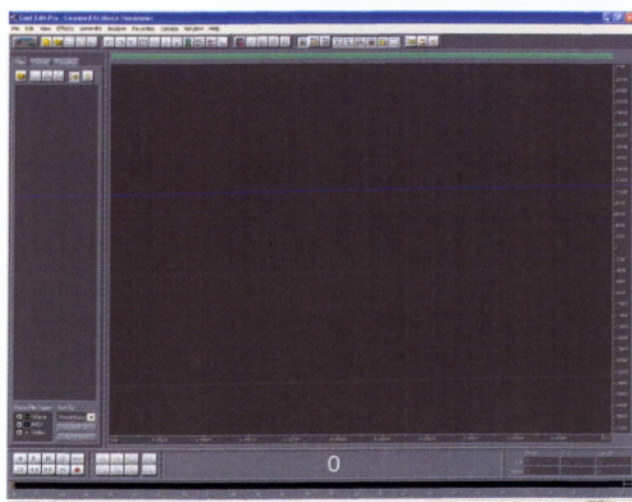


### Οι παραπάνω ήχοι χωρήσικαν περίπου σε 20 τμήματα

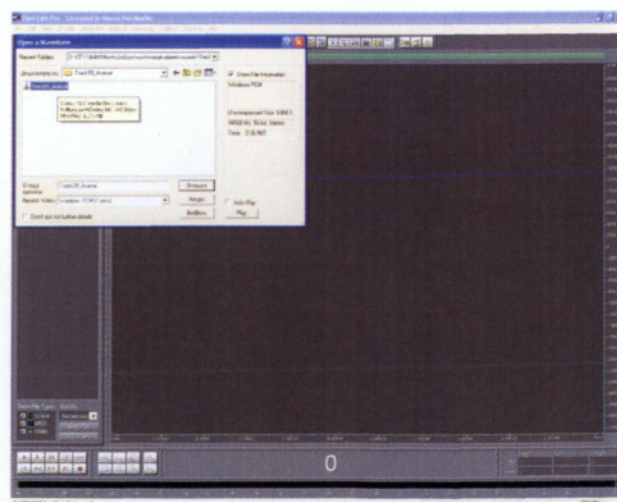
Track09_Animal(part 1-17)	Track74_Geophysical(part 1-15)	Track15_Human(part 1-15)
Track15_Animal(part 1-20)	Track90_Geophysical(part 1-15)	Track45_Human(part 1-15)
Track57_Animal(part 1-15)	Track97_Geophysical(part 1-15)	Track86_Human(part 1-15)

Η διαδικασία αυτή έγινε με την χρήση του Cool Edit Pro 2.1 όπως θα δούμε παρακάτω:

### Cool Edit Pro 2.1 Βήμα 1<sup>ο</sup> Εισαγωγή Ήχων

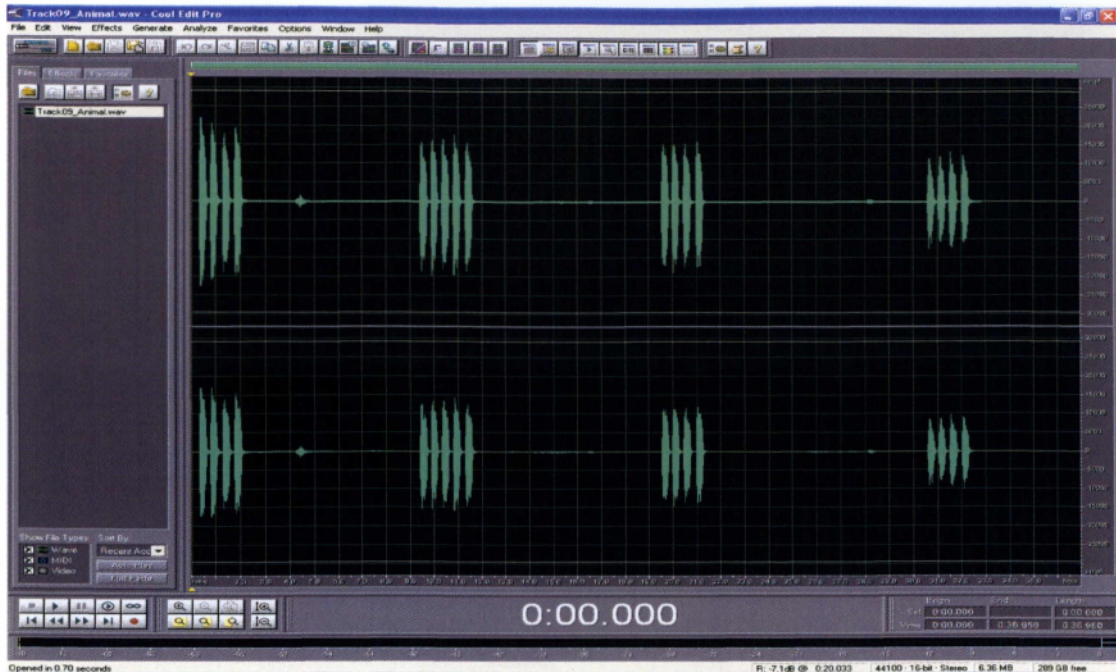


Σχήμα 4.1: Περιβάλλον Cool edit pro 2.1



Σχήμα 4.2: Εισαγωγή ήχου Track09\_Animal



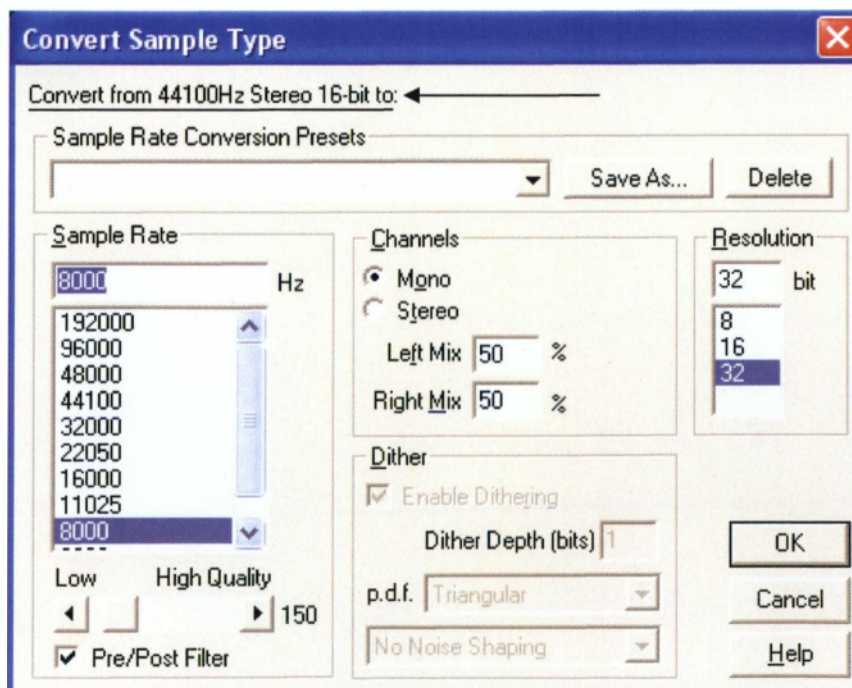


Σχήμα 4.3: Track09\_Animal

Παραπάνω βλέπουμε τον ήχο Track09\_Animal ό οποίος είναι το κρώξιμο από ένα κοράκι.

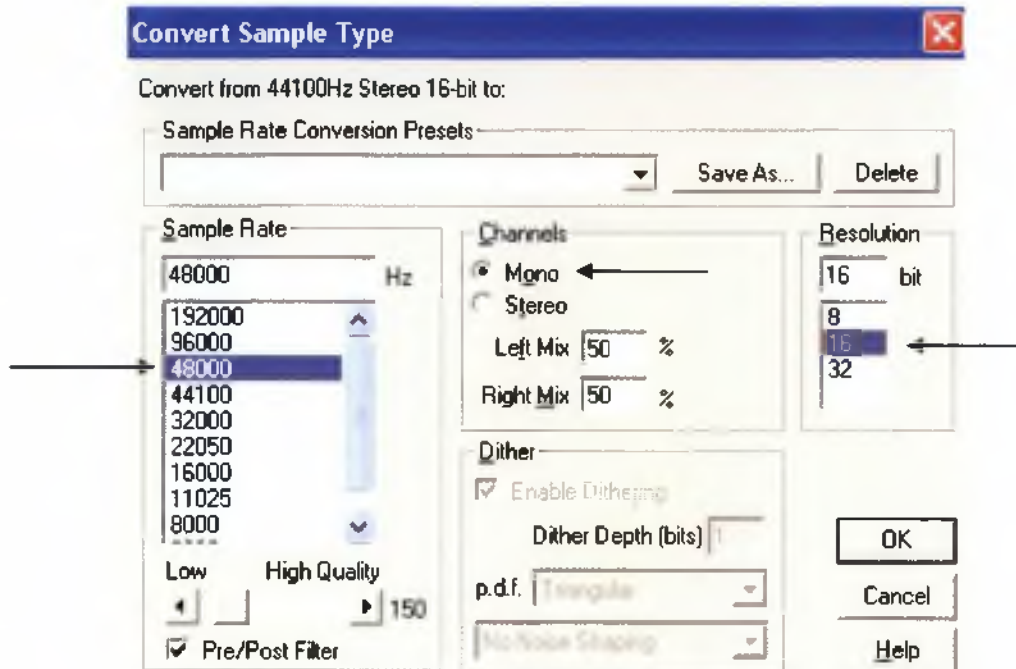
## Cool Edit Pro 2.1 Βήμα 2<sup>ο</sup> Μετατροπή Ήχων

Πατώντας το κουμπί F11 μπορούμε να διακρίνουμε κάποια χαρακτηριστικά για το σήμα (ήχο) μας.



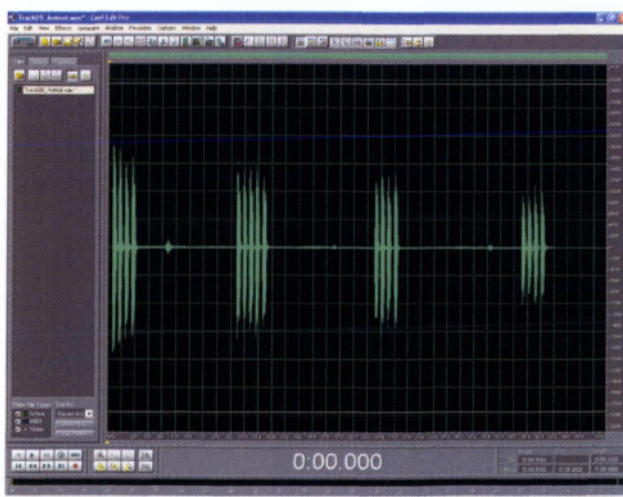
Σχήμα 4.4: Convert Sample Type

Όπως μπορούμε να διακρίνουμε παραπάνω ο ρυθμός δειγματοληψίας (Sample Rate) του σήματός μας είναι 44100Hz, είναι στερεοφωνικό (stereo) ενώ η ανάλυσή (Resolution) του είναι 16-bit.

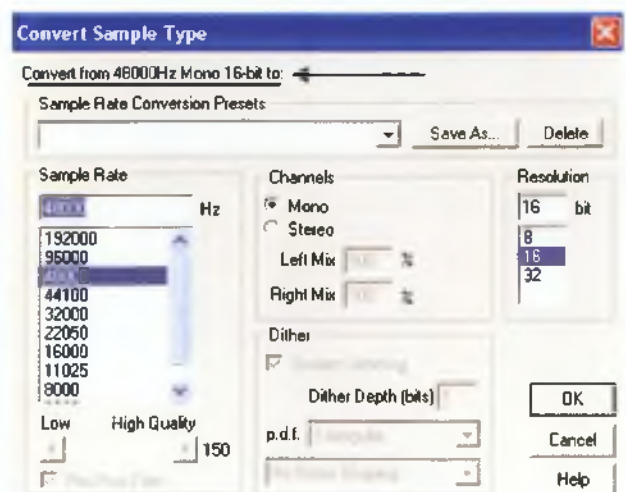


Σχήμα 4.5: Convert Sample Type

Το σήμα (ήχος) θα μετατραπεί από στερεοφωνικό (stereo) σε μονοφωνικό (mono) με ρυθμό δειγματοληψίας (Sample Rate) από 44100Hz σε 48000Hz και με ανάλυση 16-bit. Παρακάτω μπορούμε να δούμε πως γίνεται αυτή η μετατροπή.



Σχήμα 4.6: Το σήμα (ήχος) μετά την μετατροπή



Σχήμα 4.7: Convert Sample Type 48000Hz, Mono, 16-bit



## Cool Edit Pro 2.1 Βήμα 3<sup>ο</sup> Συλλογή Δειγμάτων

Στην συνέχεια προχωράμε στην δημιουργία περίπου 20 δειγμάτων (segmentation) από την κάθε ηχητική τάξη χωριστά. Η συλλογή θα γίνει επιλέγοντας μικρά τμήματα (segments) από τον κάθε ήχο όπως μπορούμε να διακρίνουμε παρακάτω:



Σχήμα 4.8: Track09\_Animal1



Σχήμα 4.9: Track09\_Animal2



Σχήμα 4.10: Track09\_Animal6



Σχήμα 4.11: Track09\_Animal10

Στους ήχους που δεν διακρίνονται εύκολα όπως για παράδειγμα στον καταρράκτη (Track74\_Geophysical)



Σχήμα 4.12: Track74\_Geophysical

θα χρησιμοποιήσουμε σαν κέντρο το σημείο που έχει το υψηλότερο πλάτος στο πεδίο του χρόνου και θα πάρουμε συγκεκριμένο αριθμό δειγμάτων και από τις δύο πλευρές όπως παρακάτω:



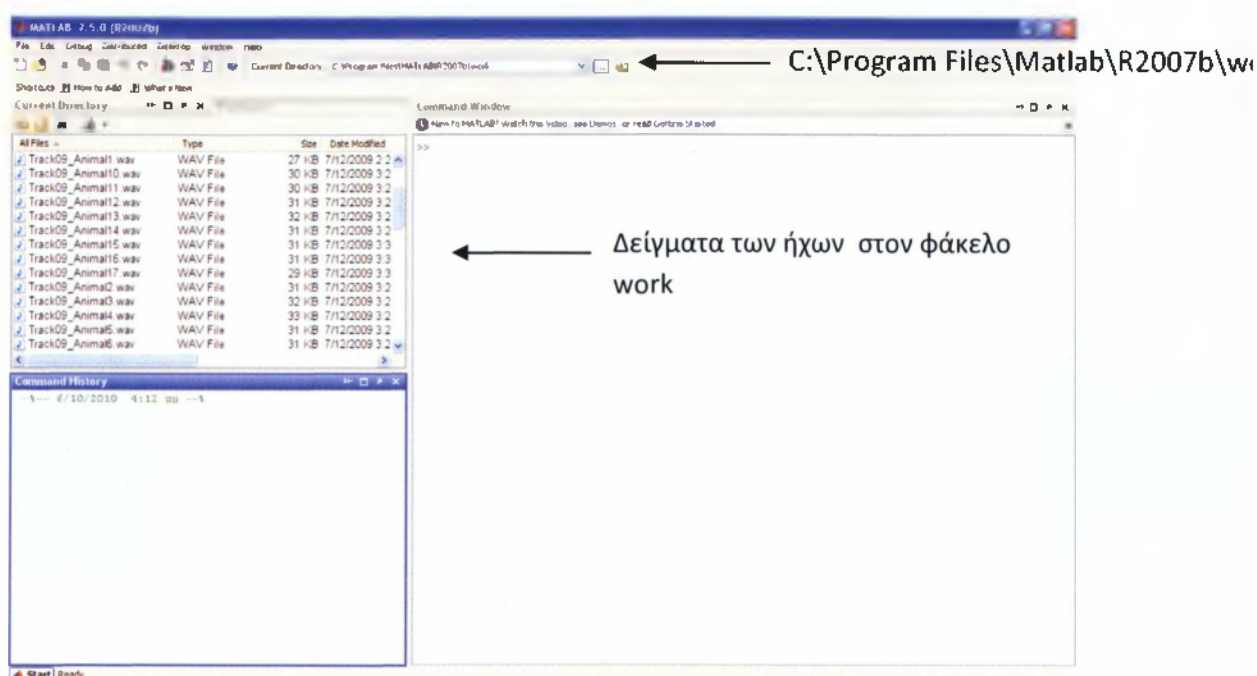
Σχήμα 4.13: Track74\_Geophysical10



Την παραπάνω διαδικασία θα την ακολουθήσουμε για όλους τους ήχους που έχουμε. Εφόσον συλλέξουμε όλα τα δείγματα θα χρησιμοποιήσουμε το Matlab για τα πειράματά μας.

## Matlab Βήμα 1<sup>ο</sup> Εισαγωγή Ήχων

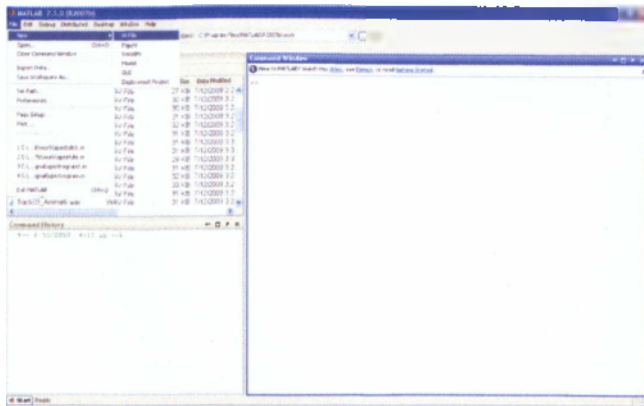
Αρχικά θα πάρουμε όλα τα δείγματα που έχουμε και θα τα τοποθετήσουμε στον φάκελο work ο οποίος βρίσκεται στην τοποθεσία C:\Program Files\Matlab\R2007b\work όπως βλέπουμε παρακάτω:



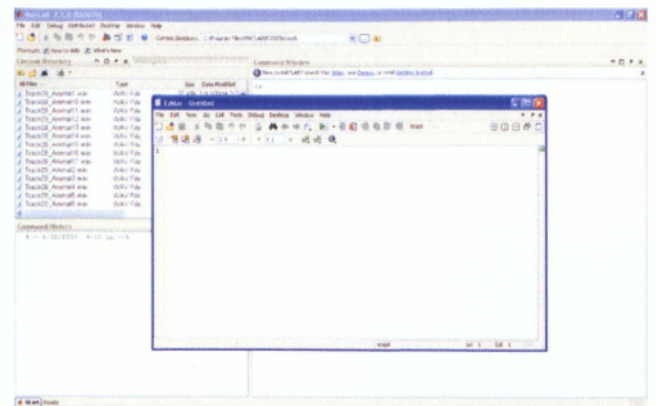
Σχήμα 4.14: Περιβάλλον Matlab

## Matlab Βήμα 2<sup>ο</sup> Δημιουργία Γραφήματος στο πεδίο του Χρόνου

Στην συνέχεια θα δημιουργήσουμε ένα M-File (editor) αρχείο στο οποίο θα χρησιμοποιήσουμε κάποιες εντολές για να προγραμματίσουμε το Matlab έτσι ώστε να μπορέσουμε να εμφανίσουμε κάποια γραφήματα των σημάτων (ήχων) που έχουμε στο πεδίο του χρόνου και να δούμε τι συμπεράσματα μπορούμε να βγάλουμε από αυτά.



Σχήμα 4.15: File -&gt; New -&gt; M-file



Σχήμα 4.16: Editor

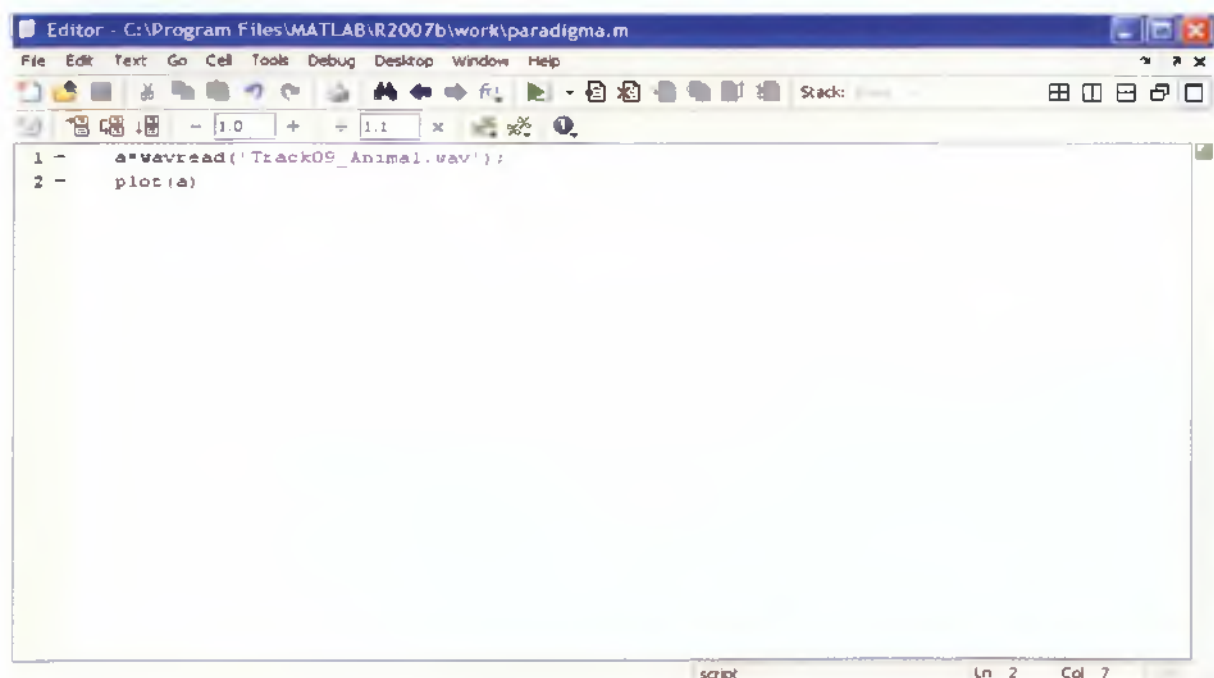
Για να μπορέσουμε να εμφανίσουμε τα γραφήματα θα πρέπει να έχουμε γνώσεις σε κάποιες βασικές εντολές του Matlab όπως θα δούμε παρακάτω:

Αρχικά για να φορτώσουμε το σήμα (ήχο) μας στο περιβάλλον του Matlab θα χρησιμοποιήσουμε την εντολή:

`a = wavread ('όνομα του ήχου')` για παράδειγμα `a = wavread ('Track09_Animal.wav')`

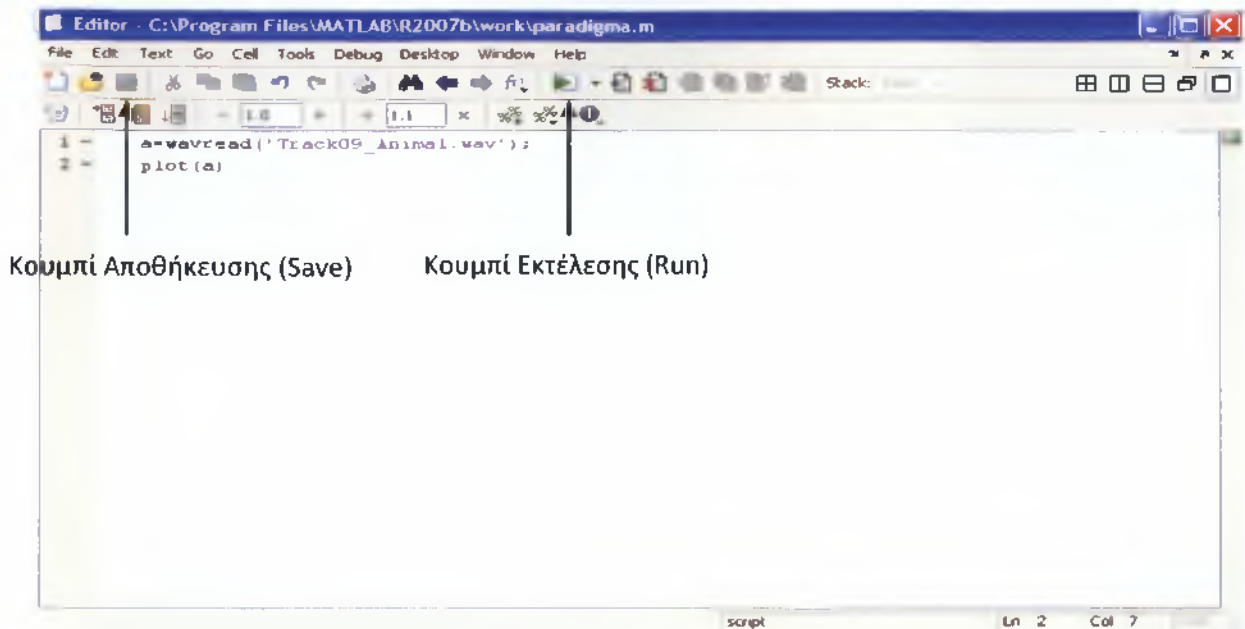
Για να εμφανίσουμε το γράφημα του σήματος μας στο πεδίο του χρόνου θα χρησιμοποιήσουμε την εντολή:

Plot (a)



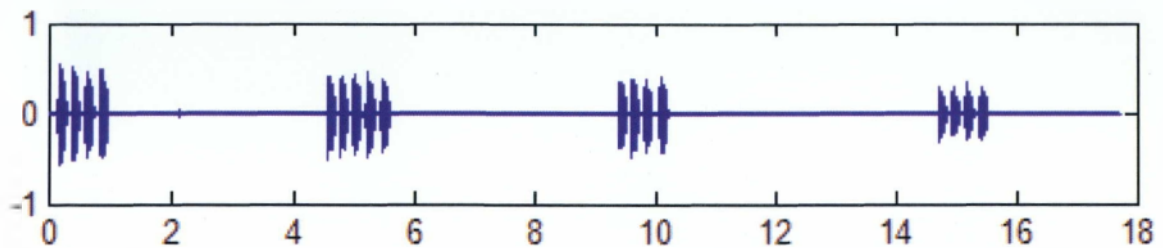
Σχήμα 4.17: Εντολές στο Editor

Για να τρέξουμε το M-File (editor) αρχείο και να δούμε τα αποτελέσματα θα πρέπει να κάνουμε πρώτα αποθήκευση και στην συνέχεια να πατήσουμε το κουμπί Run.

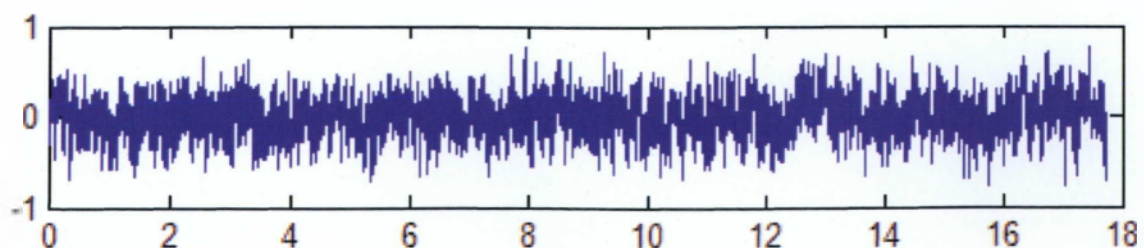


Σχήμα 4.18: Κουμπιά Αποθήκευσης και Εκτέλεσης

Αφού κάνουμε τα παραπάνω ξεχωριστά για τον κάθε ήχο το αποτέλεσμα που θα έχουμε θα είναι το εξής:



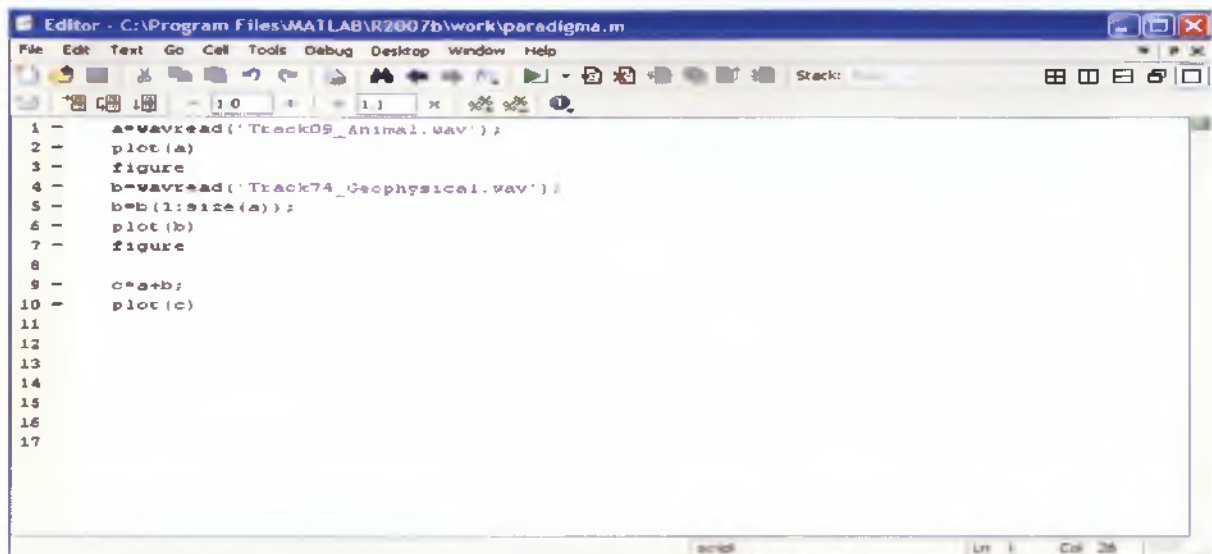
Σχήμα 4.19: Γράφημα στο Πεδίο του Χρόνου του σήματος (ήχου) Track09\_Animal.wav



Σχήμα 4.20: Track74\_Geophysical στο πεδίο του χρόνου



Για να μπορέσουμε να προσθέσουμε τα 2 σήματά (ήχους) μας θα πρέπει αρχικά να τα κάνουμε να έχουν το ίδιο μήκος αλλιώς δεν θα μπορούμε να τα προσθέσουμε. Για να το πετύχουμε αυτό και να εμφανίσουμε το γράφημα των 2 παραπάνω σημάτων (ήχων) στο πεδίο του χρόνου θα χρησιμοποιήσουμε τον παρακάτω κώδικα:



```

1 - a=wavread('Track09_Animal.wav');
2 - plot(a)
3 - figure
4 - b=wavread('Track74_Geophysical.wav');
5 - b=b(1:size(a));
6 - plot(b)
7 - figure
8
9 - c=a+b;
10 - plot(c)
11
12
13
14
15
16
17

```

Σχήμα 4.21: Track09\_Animal & Track74\_Geophysical στο πεδίο του χρόνου

**B = b(1:size(a));** σε αυτή την σειρά του κώδικα μετατρέπουμε το μέτρο του 2<sup>ου</sup> σήματος να είναι ίδιο με το

**C = a+b;** σε αυτή την σειρά του κώδικα προσθέτουμε το 1<sup>ο</sup> με το 2<sup>ο</sup> σήμα (ήχο) μας.

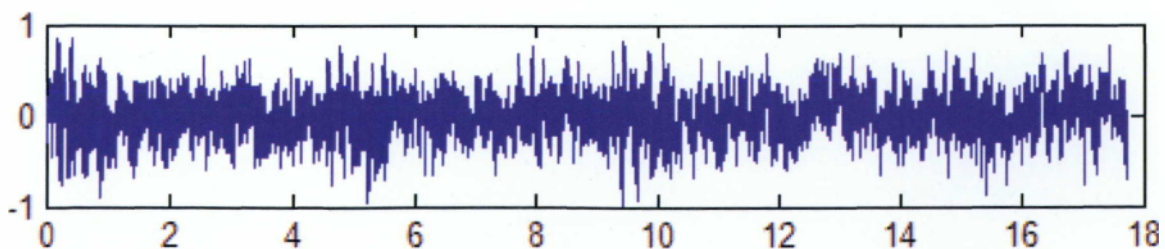
**Plot(b)**

**figure**



Αυτές οι δύο εντολές χρησιμεύουν για την εμφάνιση 1 γραφήματος του αθροίσματος των δύο σημάτων (ήχων) μας.

Μετά από την πρόσθεση των 2 σημάτων στο πεδίο του χρόνου προκύπτει το παρακάτω γράφημα:



Σχήμα 4.22: Track09\_Animal & Track74\_Geophysical στο πεδίο του χρόνου

Εμείς αυτό που επιδιώκουμε όμως είναι το σύστημα μας να μπορεί να αντιληφθεί τον κάθε ήχο ξεχωριστά ακόμα και αν έχουμε πολλούς ήχους μαζί. Αυτό όπως διακρίνουμε

παραπάνω από την πρόσθεση των δύο ήχων δεν μπορούμε να το κάνουμε χρησιμοποιώντας το πεδίο του χρόνου καθώς από αυτό δεν μπορούμε να διακρίνουμε τον κάθε ήχο ξεχωριστά.

Γι αυτό τον λόγο παρακάτω θα μετατρέψουμε τα σήματα (ήχους) μας από το πεδίο του χρόνου στο πεδίο των συχνοτήτων για να δούμε αν μπορούμε να βγάλουμε κάποιο συμπέρασμα.

## Matlab Βήμα 3<sup>ο</sup> Δημιουργία Φασματογράμματος

Προκειμένου να δημιουργήσουμε τα φασματογράμματα των ήχων ακολουθούμε την παρακάτω διαδικασία: κάθε ηχητικό σήμα διαιρείται σε τμήματα (frames) με μήκος 256 δείγματα ανά τμήμα. Κατόπιν, αφού χρησιμοποιήσουμε παράθυρο τύπου Hanning, μετασχηματίζουμε κάθε τμήμα του σήματος στο πεδίο των συχνοτήτων. Κάθε μετασχηματισμένο τμήμα χρησιμοποιείται ως ξεχωριστή γραμμή και τοποθετείται το ένα κάτω από το άλλο μορφοποιώντας με αυτό τον τρόπο ένα πίνακα που ονομάζεται φασματογράμμα μέτρου (magnitude spectrogram).

Για να δημιουργήσουμε τα γραφήματα που απεικονίζουν τα σήματά μας στο πεδίο των συχνοτήτων των σημάτων (ήχων) μας θα πρέπει να προσθέσουμε στον παραπάνω κώδικα:

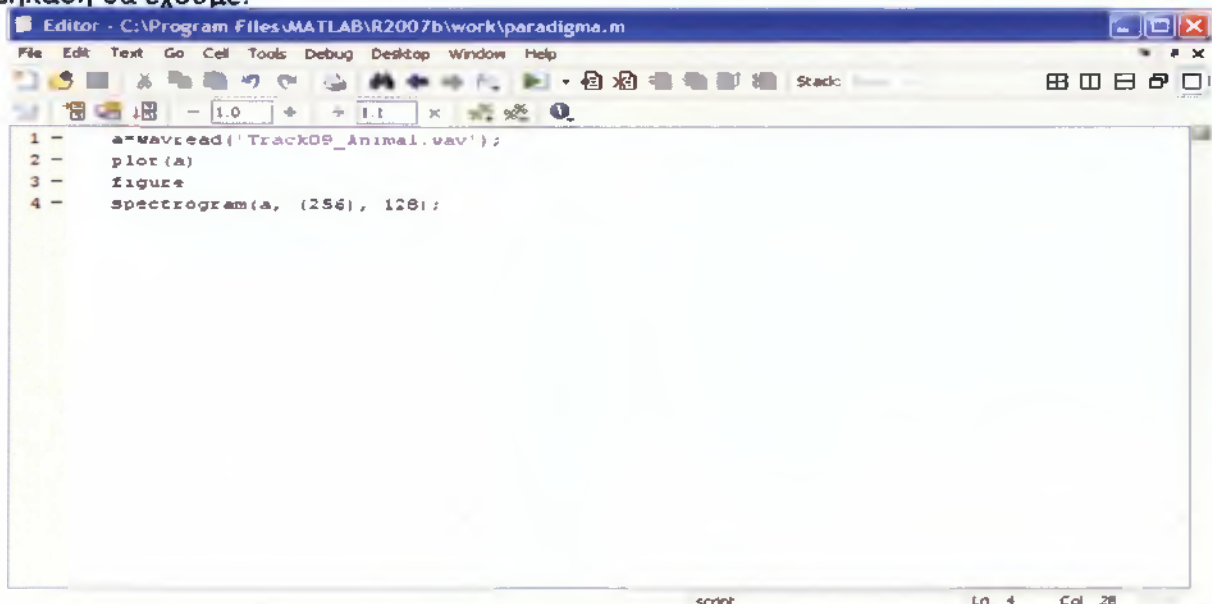
Figure

`Spectrogram(a, (256), 128);` Αντίστοιχα για τον δεύτερο ήχο

Figure

`Spectrogram(b, (256), 128);`

Δηλαδή θα έχουμε:

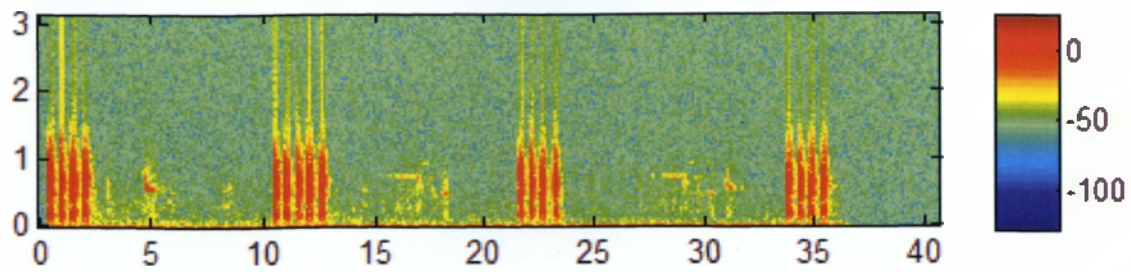


```

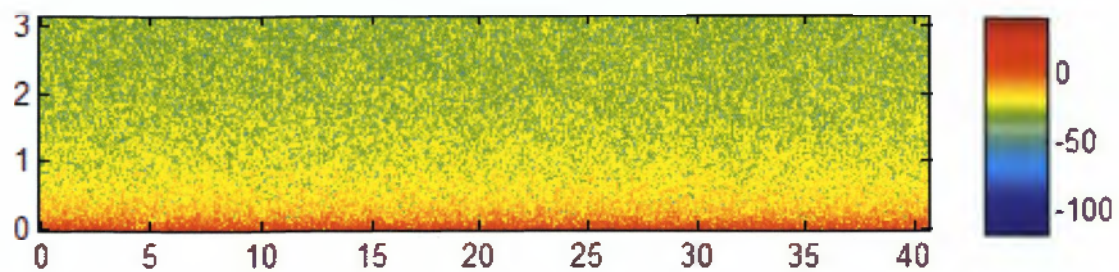
Editor - C:\Program Files\MATLAB\R2007b\work\paradigma.m
File Edit Text Go Cell Tools Debug Desktop Window Help
- 1.0 + - 1.1 x
1 - a=wavread('Track09_Animal.wav');
2 - plot(a)
3 - figure
4 - spectrogram(a, (256), 128);
script Ln 4 Col 28
  
```

Σχήμα 4.23: Κώδικας για την εμφάνιση του spectrogram για τον ήχο Track09\_Animal

Μετά την εκτέλεση του παραπάνω κώδικα για κάθε ήχο ξεχωριστά θα εμφανίσουμε τα παρακάτω **spectrograms** (φασματογράμματα).



Σχήμα 4.24: Track09\_Animal στο πεδίο των συχνοτήτων



Σχήμα 4.25: Track74\_Geophysical στο πεδίο των συχνοτήτων

Για να προσθέσουμε και να εμφανίσουμε το spectrogram των 2 παραπάνω σημάτων (ήχων) θα χρησιμοποιήσουμε τον παρακάτω κώδικα:

**figure**

**Spectrogram(c, (256), 128);**



Αυτές οι 2 εντολές χρησιμεύουν για την εμφάνιση του **spectrogram** (φασματογράμματος) του αθροίσματος των 2 σημάτων (ήχων) μας.

Ο κώδικας που θα έχουμε θα είναι:

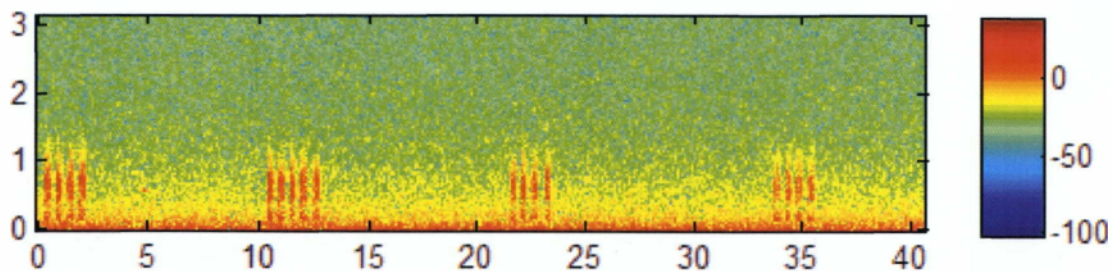
```

Editor - C:\Program Files\MATLAB\R2007b\work\paradigma.m
File Edit Text Go Cell Tools Debug Desktop Window Help
Stack: Base
- 1.0 + ÷ 1.1 × % %
1 - a=wavread('Track09_Animal.wav');
2 - plot(a)
3 - figure
4 - b=wavread('Track74_Geophysical.wav');
5 - b=b(1:size(a));
6 - plot(b)
7 - figure
8
9 - c=a+b;
10 - plot(c)
11 - figure
12
13 - spectrogram(a, (256), 128);
14 - figure
15 - spectrogram(b, (256), 128);
16 - figure
17 - spectrogram(c, (256), 128);
script Ln 4 Col 31 OVR

```

Σχήμα 4.26: Κώδικας εμφάνισης spectrogram της πρόσθεσης των ήχων Track09\_Animal & Track74\_Geophysical

Μετά την πρόσθεση των 2 σημάτων στο πεδίο του χρόνου προκύπτει το παρακάτω φασματόγραμμα:



Σχήμα 4.27: Track09\_Animal & Track74\_Geophysical στο πεδίο των συχνοτήτων

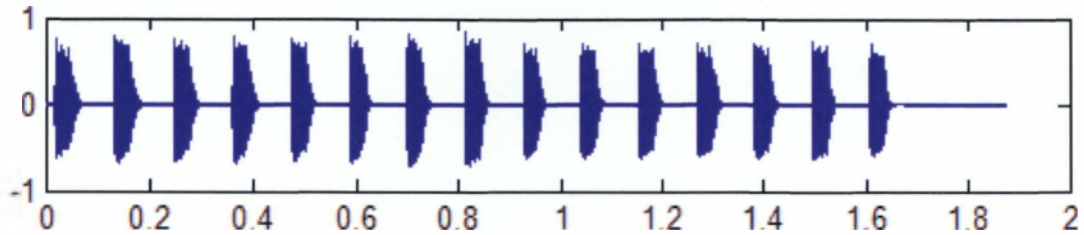
Έτσι όπως μπορούμε να διακρίνουμε παραπάνω μετά την χρήση της χρονοσυχνοτικής κατανομής (φασματόγραμμα) μπορούμε να ξεχωρίσουμε ποιος είναι ο κάθε ένας από τους δυο ήχους πολύ πιο καθαρά συγκριτικά με το πεδίο του χρόνου,[13]. Ομοίως ακολουθούν και άλλα παραδείγματα στα οποία χρησιμοποιώντας το φασματόγραμμα διακρίνουμε πιο καθαρά διαφορετικούς ήχους από το πεδίο των συχνοτήτων σε σύγκριση με το πεδίο του χρόνου.



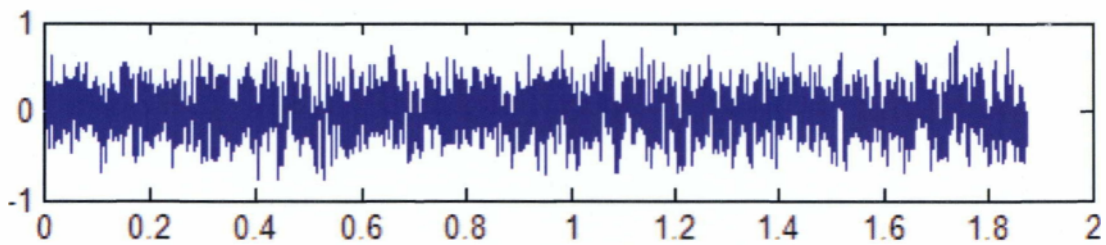
## Παράδειγμα 1

Σε αυτό το παράδειγμα θα χρησιμοποιήσουμε τους ήχους **Track45\_Human** (Πυροβολισμός) και **Track90\_Geophysical** (Θυελλώδεις άνεμοι).

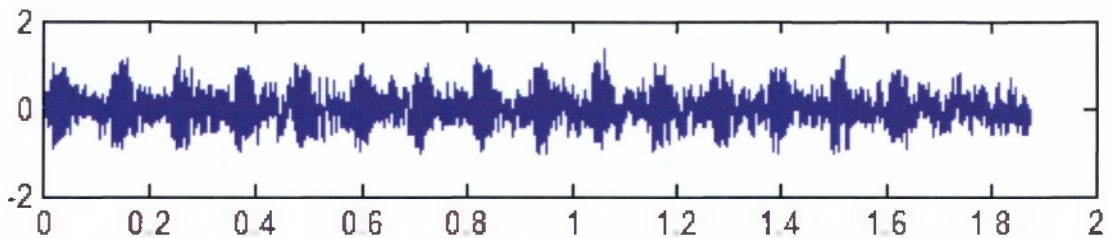
*Γραφήματα των σημάτων (ήχων) στο πεδίο του χρόνου:*



Σχήμα 4.28: Track45\_Human (Πυροβολισμός) στο πεδίο του χρόνου

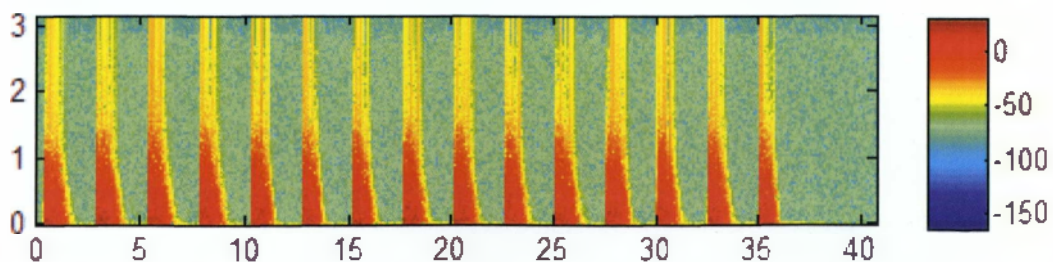


Σχήμα 4.29: Track90\_Geophysical (Θυελλώδεις άνεμοι) στο πεδίο του χρόνου



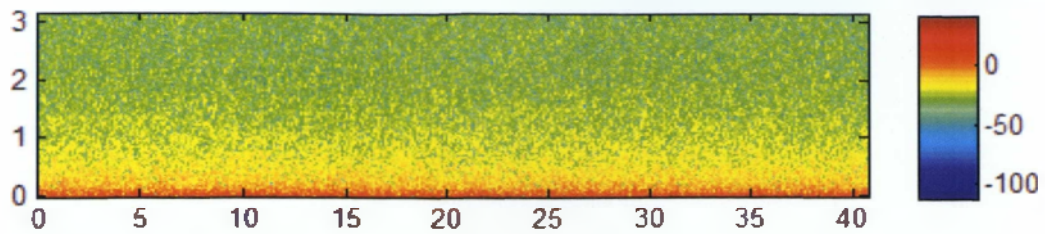
Σχήμα 4.30: Track45\_Human και Track90\_Geophysical στο πεδίο του χρόνου

*Στην συνέχεια θα δούμε πως παρουσιάζονται τα σήματα (ήχοι) στο πεδίο των συχνοτήτων:*

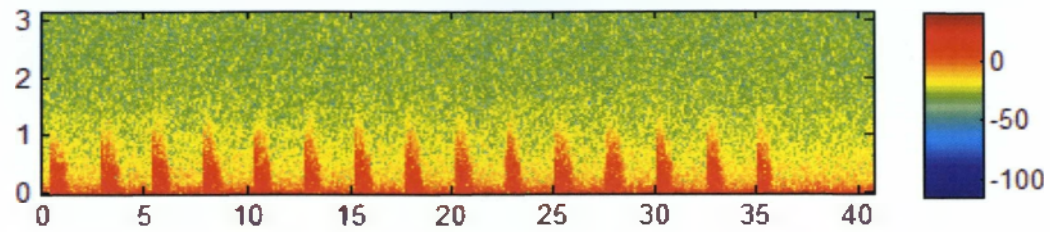


Σχήμα 4.31: Track45\_Human (Πυροβολισμός) στο πεδίο των συχνοτήτων





Σχήμα 4.32: Track90\_Geophysical (Θυελλώδεις άνεμοι) στο πεδίο των συχνοτήτων

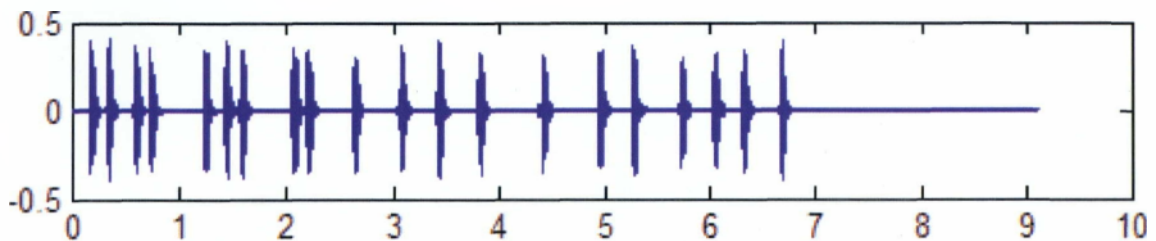


Σχήμα 4.33: Track45\_Human και Track90\_Geophysical στο πεδίο των συχνοτήτων

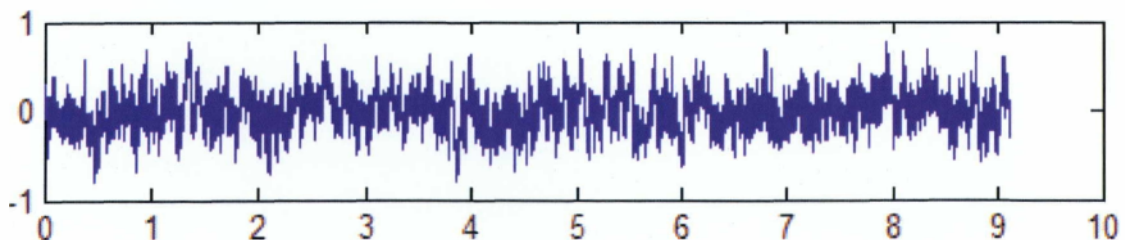
## Παράδειγμα 2

Σε αυτό το παράδειγμα θα χρησιμοποιήσουμε τους ήχους **Track15\_Animal (Σκύλος)** και **Track86\_Human(Αεροπλάνο)**.

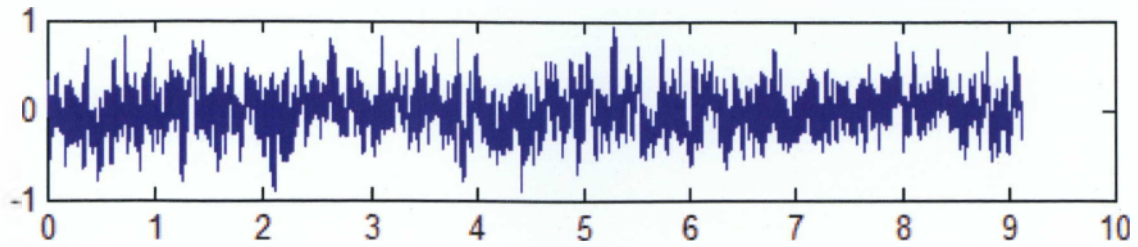
*Γραφήματα των σημάτων (ήχων) στο πεδίο του χρόνου:*



Σχήμα 4.34: Track15\_Animal (Σκύλος) στο πεδίο του χρόνου

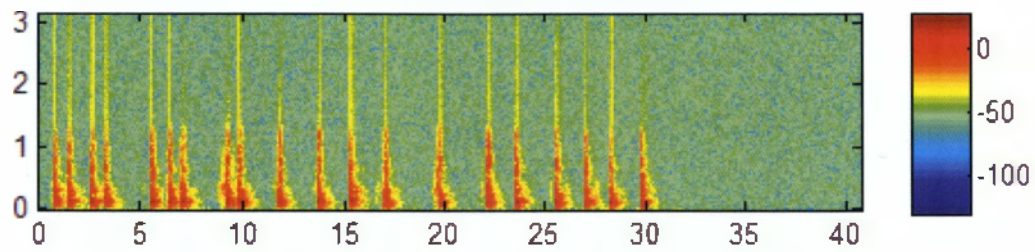


Σχήμα 4.35: Track86\_Human (Αεροπλάνο) στο πεδίο του χρόνου

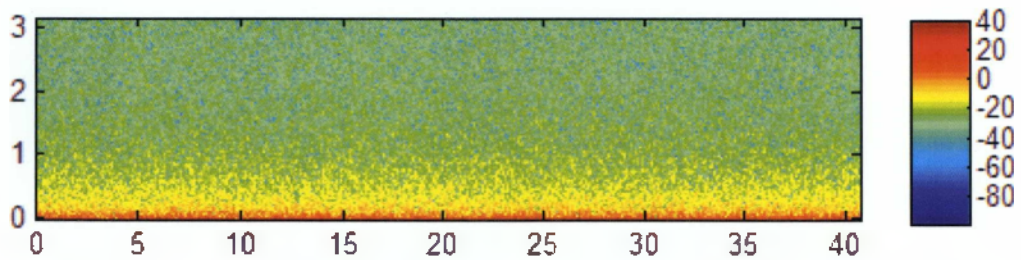


Σχήμα 4.36: Track15\_Animal και Track86\_Human στο πεδίο του χρόνου

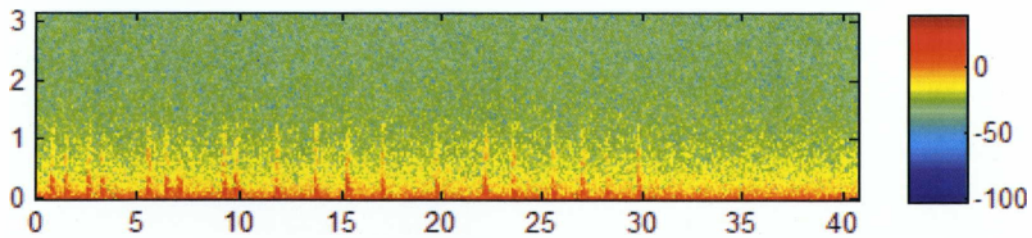
**Τα γραφήματα των σημάτων (ήχων) στο πεδίο των συχνοτήτων:**



Σχήμα 4.37: Track15\_Animal (Σκύλος) στο πεδίο των συχνοτήτων



Σχήμα 4.38: Track86\_Human (Αεροπλάνο) στο πεδίο των συχνοτήτων

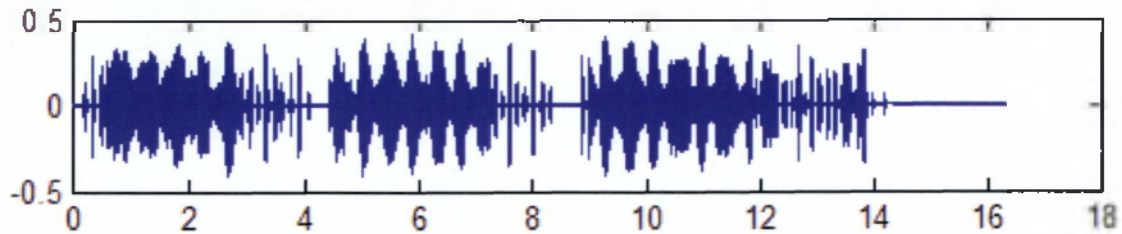


Σχήμα 4.39: Track15\_Animal και Track86\_Human στο πεδίο των συχνοτήτων

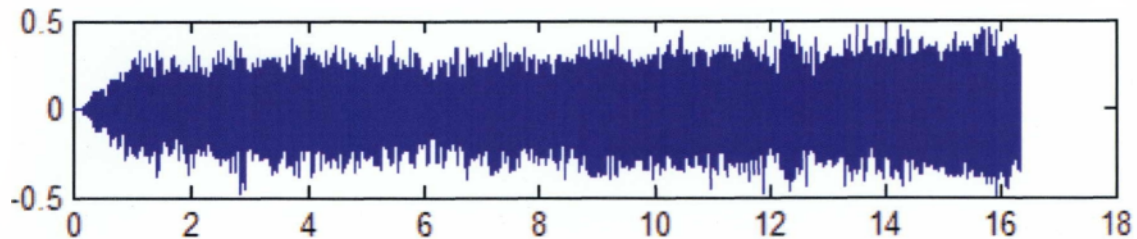
### Παράδειγμα 3

Σε αυτό το παράδειγμα θα χρησιμοποιήσουμε τους ήχους Track57\_Animal (Κροταλίας) και Track97\_Geophysical (Φωτιά).

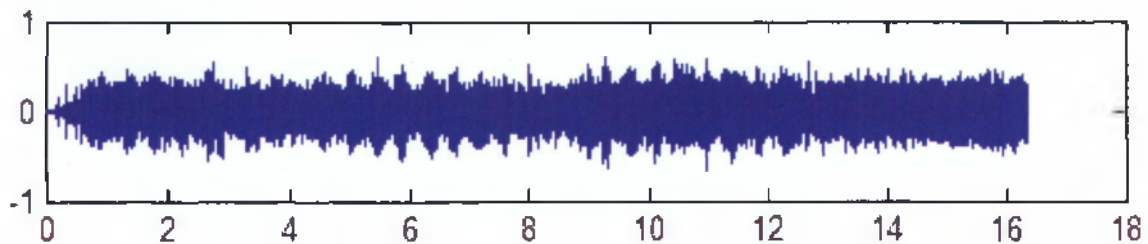
*Γραφήματα των σημάτων (ήχων) στο πεδίο του χρόνου:*



Σχήμα 4.40: Track57 Animal (Κροταλίας) στο πεδίο του χρόνου

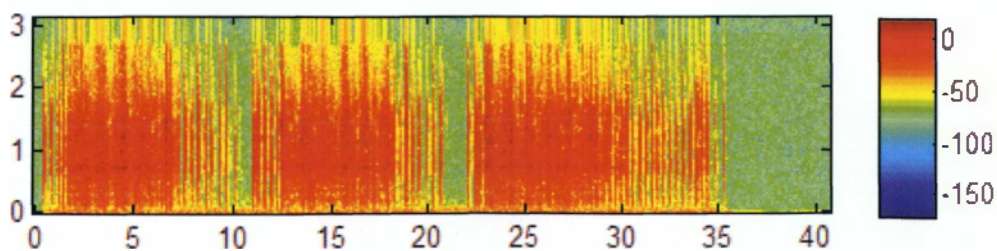


Σχήμα 4.41: Track97 Geophysical (Φωτιά) στο πεδίο του χρόνου



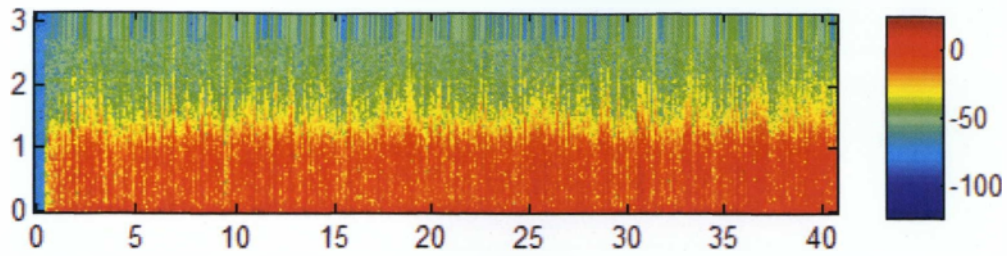
Σχήμα 4.42: Track57\_Animal και Track97\_Geophysical στο πεδίο του χρόνου

*Ενώ στην συνέχεια θα δούμε πως παρουσιάζονται τα σήματα (ήχοι) στο πεδίο των συχνοτήτων:*

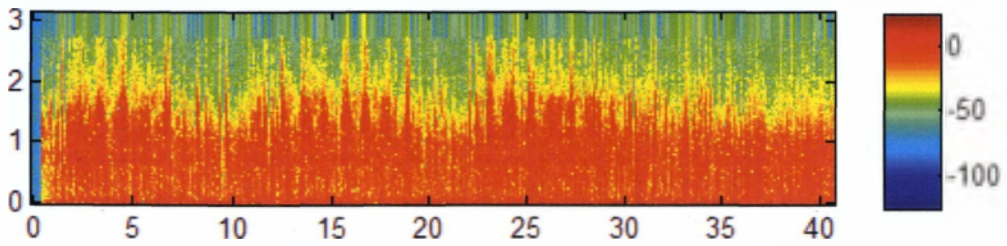


Σχήμα 4.43: Track57\_Animal (Κροταλίας) στο πεδίο των συχνοτήτων





Σχήμα 4.44: Track97\_Geophysical (Φωτιά) στο πεδίο των συχνοτήτων

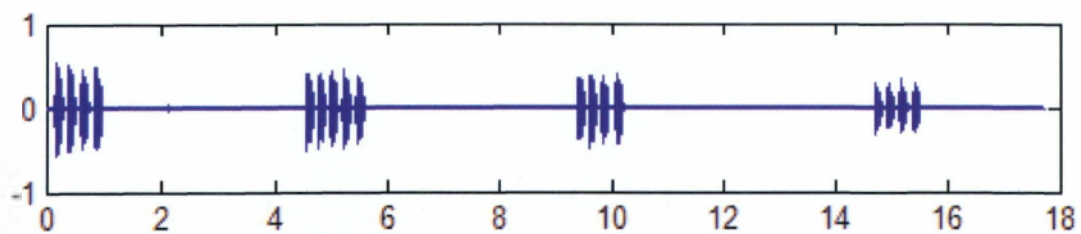


Σχήμα 4.45: Track57\_Animal και Track97\_Geophysical στο πεδίο των συχνοτήτων

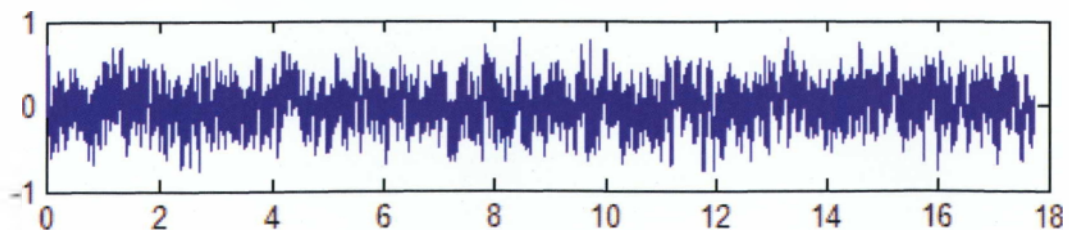
## Παράδειγμα 4

Σε αυτό το παράδειγμα θα χρησιμοποιήσουμε τους ήχους **Track09\_Animal** (Κοράκι) και **Track15\_Human** (Αυτοκίνητο).

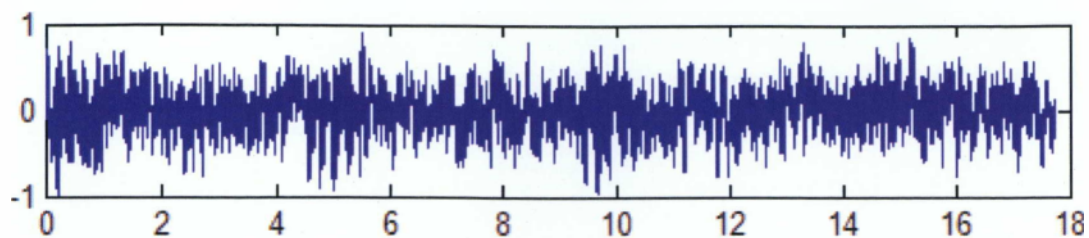
*Γραφήματα των σημάτων (ήχων) στο πεδίο του χρόνου:*



Σχήμα 4.46: Track09\_Animal (Κοράκι) στο πεδίο του χρόνου

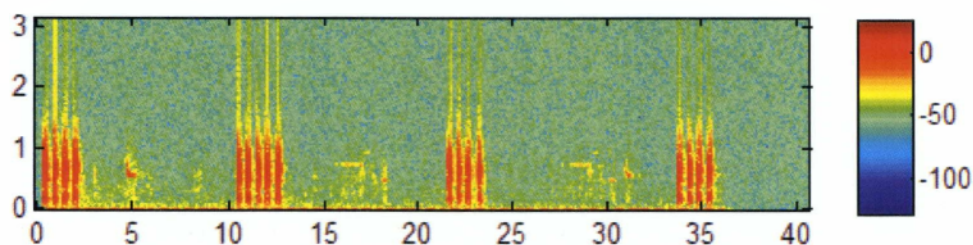


Σχήμα 4.47: Track15\_Human (Αυτοκίνητο) στο πεδίο του χρόνου

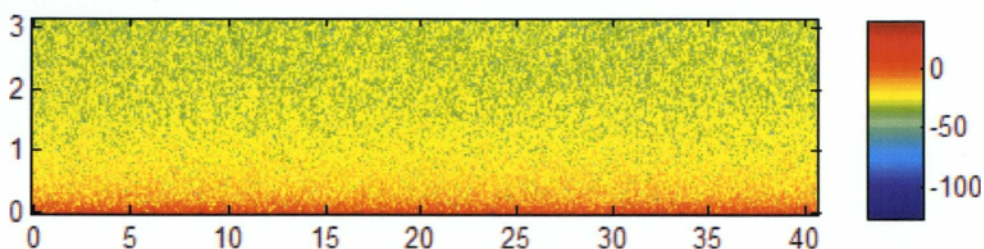


Σχήμα 4.48: Track09\_Animal και Track15\_Human στο πεδίο του χρόνου

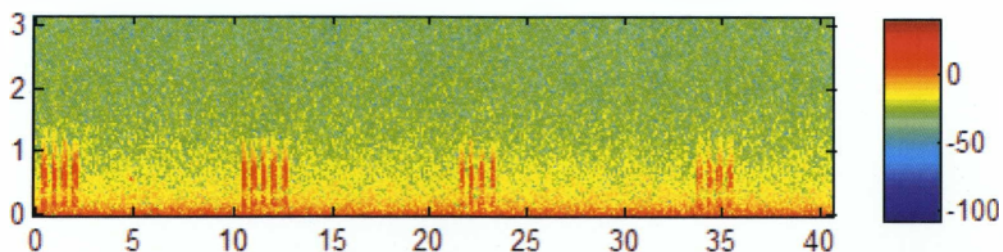
*Ενώ στην συνέχεια θα δούμε πως παρουσιάζονται τα σήματα (ήχοι) στο πεδίο των συχνοτήτων:*



Σχήμα 4.49: Track09\_Animal (Κοράκι) στο πεδίο των συχνοτήτων



Σχήμα 4.50: Track15\_Human (Αυτοκίνητο) στο πεδίο των συχνοτήτων



Σχήμα 4.51: Track09\_Animal και Track15\_Human στο πεδίο των συχνοτήτων



## Συμπεράσματα

Οι χρονοσυχνοτικές κατανομές (time-frequency distributions) είναι ένα σύνολο χρήσιμων τεχνικών που χρησιμοποιούνται για τις περιπτώσεις σημάτων των οποίων οι φασματικές ιδιότητες μεταβάλλονται στο χρόνο (κεφάλαιο 2). Τέτοια σήματα είναι οι ήχοι που χρησιμοποιήθηκαν στα πειράματα αυτού του κεφαλαίου.

Ο μετασχηματισμός Fourier μετατρέπει το σήμα από το πεδίο του χρόνου στο πεδίο της συχνότητας (κεφάλαιο 1). Στην συγκεκριμένη πτυχιακή μετατρέψαμε τα σήματα (ήχοι) από το πεδίο του χρόνου στο πεδίο των συχνοτήτων και χρησιμοποιήσαμε το μέτρο (magnitude) τους προκειμένου να εξάγουμε συμπεράσματα τα οποία δεν μπορούσαν να εξαχθούν από το πεδίο του χρόνου. Συγκεκριμένα, τα πειραματικά αποτελέσματα (κεφάλαιο 4) επιβεβαιώνουν τη χρησιμότητα των χρονοσυχνοτικών κατανομών για τον διαχωρισμό ήχων. Όπως μπορούμε να διακρίνουμε και από τα σχήματα παραπάνω, μετά την μετατροπή των σημάτων (ήχων) μας στο πεδίο των συχνοτήτων από το πεδίο του χρόνου μπορούμε να ξεχωρίσουμε ποίος είναι ο κάθε ήχος πολύ πιο εύκολα συγκριτικά με το πεδίο του χρόνου.

Τα χαρακτηριστικά που εξάγουμε από τις χρονοσυχνοτικές κατανομές κατόπιν μπορούν να χρησιμοποιηθούν για την αυτόματη ταξινόμηση αυτών των ήχων (Κεφάλαιο 3). Μετά την αυτόματη ταξινόμησή τους μπορούμε να αναπτύξουμε τα ηχοτοπία τα οποία είναι ηχητικοί χάρτες περιοχών όπου σε συγκεκριμένα σημεία τους πραγματοποιείται καταγραφή και επεξεργασία με σκοπό την αυτόματη ταξινόμηση των καταγραφόμενων ήχων (εισαγωγή). Οι ηχοχάρτες αυτοί μπορούν να ανανεωθούν περιοδικά και να ελέγχονται σε τακτά χρονικά διαστήματα από ειδικούς σε συγκεκριμένα κέντρα έρευνας (που δεν βρίσκονται απαραίτητα κοντά στην περιοχή καταγραφής των ήχων) με σκοπό την εξαγωγή χρήσιμων συμπερασμάτων σχετικών με την ισορροπία του οικοσυστήματος.

**ΒΙΒΛΙΟΓΡΑΦΙΑ**

- [1] Schafer, R.M., *"The soundscape: Our sonic environment and the tuning of the world"*, Destiny Books, Rochester, Vermont, 1994.
- [2] SEKI Group web site: <http://envirosonic.cevl.msu.edu/seki>, 2008.
- [3] Krause, B, "Bioacoustics, habitat ambience in ecological balance," *Whole Earth Review*, 1987.
- [4] B. Krause B, *Wild soundscapes: Discovering the voice of the natural world*, Wilderness Press, Berkeley, California, 2002.
- [5] Turner, M.G., Gardner, R.H., O'Neill, R.V., *Landscapem ecology in theory and practice: Pattern and process*, Springer-Verlag, New York, Inc., 2001.
- [6] Proakis, J.G and Manolakis D.G, "Digital Signal Processing – Principles, Algorithms and Applications", Prentice Hall, 1996.
- [7] Oppenheim A.V. and Schafer R.W., "Discrete Time Signal Processing, Prentice Hall Signal Processing Series, 1989.
- [8] Orfanidis, S.J, "Introduction to Signal Processing" , Prentice Hall, 1996.
- [9] Kaiser, J.F., "Nonrecursive Digital Filter Design Using the losinh Window Functions" , Proc. 1974 IEEE International Symposium on Circuits and Systems, San Francisco, pp. 20-23, April 1974.
- [10] Tribolet J. M., "A new phase unwrapping algorithm," IEEE Trans. Acoust., Speech, Signal Process. ASSP-25, 170–177 (1977).
- [11] Shie Q. and Dapang C., "Joint Time-Frequency Analysis – Methods and Applications", Prentice Hall, 1996
- [12] Mitchell, T.M, "Machine Learning", McGraw-Hill Higher Education, 1997.
- [13] Paraskevas I., Potirakis S., and Rangoussi M., "Natural Soundscapes and identification of environmental sounds: a pattern recognition approach", Proc. 16<sup>th</sup> Int. Conference on Digital Signal Processing (DSP 2009), Santorini, July 2009.