

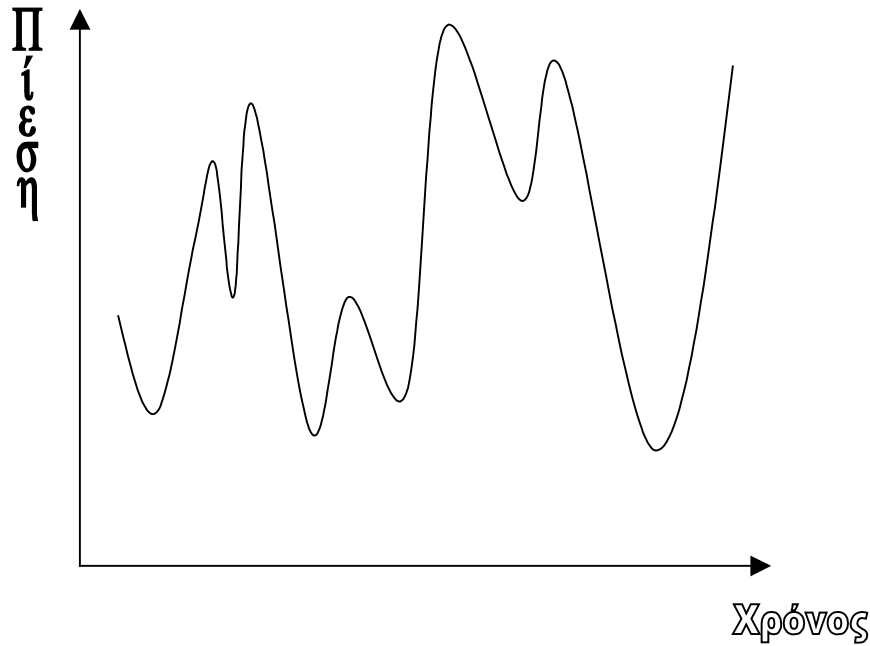
3. ΗΧΗΤΙΚΑ ΣΗΜΑΤΑ

3.1 Τι είναι ήχος;

Από φυσική άποψη ένας ήχος παράγεται από μεταβολές της πίεσης που μεταδίδονται σε ένα μέσο που μπορεί να συμπιεστεί. Παρόλο που ο ακριβής μηχανισμός διαφέρει από περίπτωση σε περίπτωση οι βασικές αρχές της μετάδοσης και παραγωγής ήχων παραμένουν οι ίδιες. Πιο συγκεκριμένα, τα μόρια όλων των φυσικών σωμάτων προτιμούν να κρατούν σταθερές αποστάσεις από όλα τα γειτονικά τους μόρια. Κατά συνέπεια όταν για οποιονδήποτε λόγο τα μόρια σε μία περιοχή συμπιεστούν και επομένως οι αποστάσεις μεταξύ τους μικρύνουν, τα μόρια της περιοχής επιδιώκουν να επανέλθουν στην αρχική τους κατάσταση συμπιέζοντας με τη σειρά τους τα μόρια των γειτονικών τους περιοχών κ.ο.κ. Με αυτό τον τρόπο δημιουργούνται σε ένα μέσο μεταβολές πίεσης που μεταδίδονται με μία ορισμένη ταχύτητα. Οι συγκεκριμένες μεταβολές αποτελούν ένα *ηχητικό κύμα*. Όταν οι μεταβολές αυτές φτάσουν στο αυτί μας θέτουν ένα ολόκληρο μηχανισμό από μεμβράνες και οστά σε ταλάντωση και διεγείρουν κατάλληλα ορισμένα νευρικά κύτταρα με αποτέλεσμα το αίσθημα της ακοής. Συνοψίζοντας, θεωρούμε ότι σώματα σε ταλάντωση και μεταβολές πίεσης που μεταδίδονται με μία ορισμένη ταχύτητα αποτελούν την προέλευση των ήχων. Για παράδειγμα, αν χτυπήσουμε τη χορδή μίας κιθάρας, τότε η χορδή θα αρχίσει να ταλαντώνεται. Κάθε φορά που η χορδή κινείται προς τα πάνω συμπιέζει τα μόρια του αέρα που βρίσκονται πάνω της και τα παραγόμενα ηχητικά κύματα μεταδίδονται μέσω του αέρα μέχρι να συναντήσουν το τύμπανο του αυτιού μας. Οι αρχές λειτουργίας της χορδής είναι κοινές και για άλλες πηγές ήχων. Μια παράδειγμα όταν ένας σαξοφωνίστας φυσάει στο σαξόφωνο του, θέτει σε παλινδρομική κίνηση μία στήλη αέρα μέσα στο όργανο του. Ο σαξοφωνίστας μπορεί να μεταβάλλει την παλινδρόμηση του αέρα ανοίγοντας και κλείνοντας τρύπες που βρίσκονται κατά μήκος του σαξοφώνου.

Η ταλάντωση που εκτελούν τα μόρια ενός σώματος κατά την παραγωγή ενός ήχου μπορεί να είναι αρκετά πολύπλοκη. Για παράδειγμα, η ανθρώπινη φωνή παράγεται από ένα συνδυασμό ταλαντώσεων των φωνητικών χορδών που προκαλούν την ταλάντωση του αέρα που βρίσκεται στους πνεύμονες, στο λαιμό, στο στόμα και στα ιγμόρια. Ο ήχος μίας κιθάρας ή ενός βιολιού προέρχεται από την ταλάντωση μίας χορδής του οργάνου που θέτει σε παλινδρομική κίνηση τον αέρα που βρίσκεται στο αντηχείο του οργάνου. Η χροιά του ήχου που παράγουν τα περισσότερα μουσικά

όργανα εξαρτάται από το σχήμα και τις υπόλοιπες φυσικές ιδιότητες του αντηχείου τους.

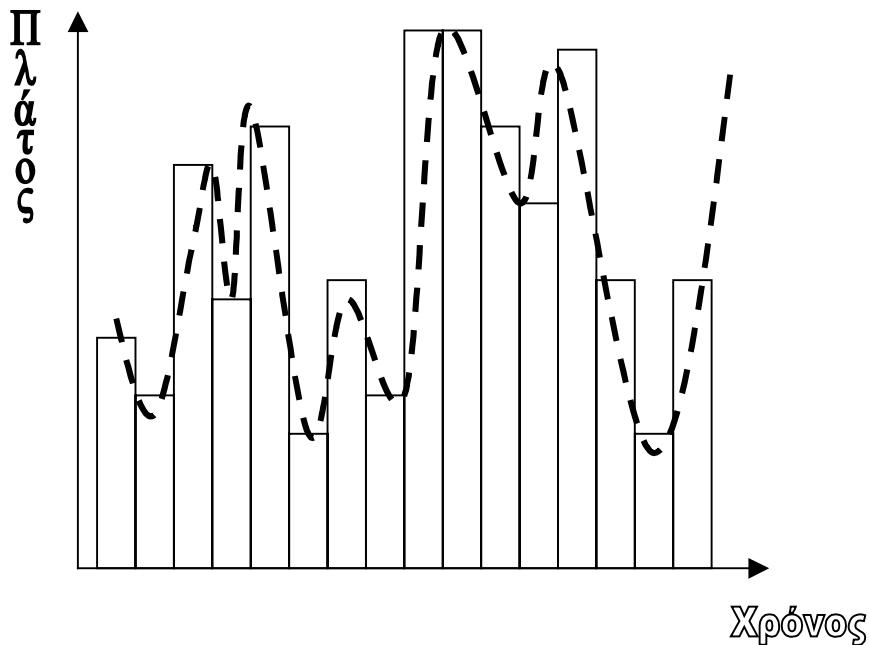


Εικόνα 3.1: Μεταβολές πίεσης στο τύμπανο.

Κάθε ήχος λοιπόν κάποια στιγμή αποτελείται από ένα σύνολο μεταβολών πίεσης. Αν είχαμε κάποιο τρόπο σχεδίασης των συγκεκριμένων μεταβολών της πίεσης που διεγείρουν το τύμπανο του ακροατή θα καταλήγαμε σε ένα διάγραμμα παρόμοιο με αυτό της Εικόνας 3.1. Ενας αρκετά απλός τρόπος για να δημιουργήσουμε ένα τέτοιο διάγραμμα θα ήταν να προσαρμόσουμε ένα κομμάτι μεταλλικό σύρμα σε ένα λεπτό φύλλο χαρτί και να τοποθετήσουμε ένα μαγνήτη σχετικά κοντά. Όταν αρχίσουμε να μιλάμε μπροστά στο χαρτί τότε οι μεταβολές της πίεσης που οφείλονται στη φωνή μας θα αναγκάζουν το χαρτί να κινείται μπρος ή πίσω, μεταβάλλοντας παράλληλα και τη θέση του μεταλλικού σύρματος ως προς το μαγνήτη. Η κίνηση αυτή του σύρματος θα δημιουργήσει ένα μεταβαλλόμενο ηλεκτρικό ρεύμα το οποίο θα είναι ανάλογο με τις μεταβολές της πίεσης του αέρα που προκαλεί η φωνή μας. Ο απλός αυτός μικροφωνικός μηχανισμός δουλεύει και αντίστροφα. Ενα μεταβαλλόμενο ηλεκτρικό ρεύμα μπορεί να μεταβάλλει το μαγνητικό πεδίο του μαγνήτη μεταβάλλοντας και τη θέση του σύρματος ως προς το μαγνήτη και προκαλώντας κατά συνέπεια την κίνηση του χαρτιού. Οι αρχές λειτουργίας της συσκευής που περιγράψαμε αντιστοιχούν στον τρόπο με τον οποίο ηχητικά σήματα μπορούν να μετατραπούν σε ηλεκτρικά σήματα μέχρι τις μέρες μας.

3.2 Ψηφιακός Ηχος

Όπως είναι δυνατό να μετατρέψουμε τις μεταβολές πίεσης που προέρχονται από έναν ήχο σε αναλογικό ηλεκτρικό σήμα με ανάλογο τρόπο είμαστε σε θέση να μετατρέψουμε ένα μεταβαλλόμενο ηλεκτρικό σήμα σε μία σειρά από διακριτές τιμές και αντίστροφα. Επειδή όμως η αναλογική και διακριτή μορφή ενός ήχου είναι διαφορετικές, πάντα χάνουμε πληροφορία κατά την εφαρμογής μίας τέτοιας μετατροπής. Όλη η προσπάθεια σε μία τέτοια περίπτωση είναι να καταλάβουμε τι είδους πληροφορία χάνουμε ώστε να επιλέξουμε τι πληροφορία θα πρέπει να κρατήσουμε κατά την εφαρμογή μίας τέτοιας μετατροπής.



Εικόνα 3.2: Ψηφιακή και αναλογική αναπαράσταση ηχητικού σήματος.

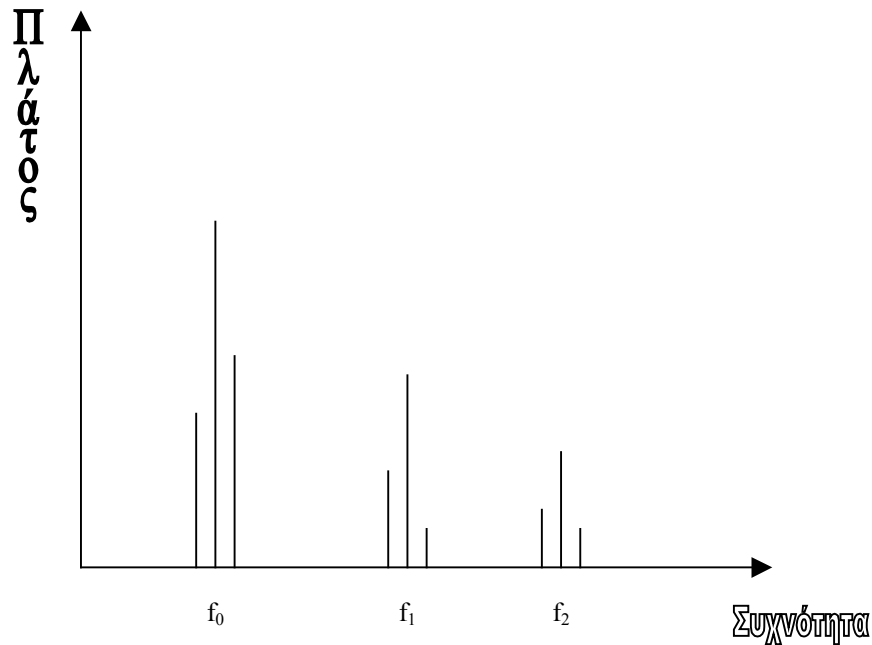
Σε ψηφιακή μορφή ένα ηχητικό σήμα παριστάνεται από μια σειρά από νούμερα (τα οποία ονομάζουμε δείγματα) τα οποία αντιστοιχούν στην πίεση του αέρα ή στην ηλεκτρική τάση σε διαδοχικές χρονικές στιγμές. Η δειγματοληψία ενός αναλογικού σήματος αντιμετωπίζει δύο κύρια προβλήματα που περιγράφονται στην Εικόνα 3.2. Η εικόνα περιγράφει δυο αναπαραστάσεις ενός ηχητικού σήματος. Η διακεκομμένη μαύρη γραμμή περιγράφει την αναλογική μορφή του σήματος. Η ψηφιακή μορφή του σήματος παριστάνεται σαν μία ακολουθία από ορθογώνια. Το πρώτο πρόβλημα προέρχεται από το γεγονός ότι κάθε τιμή του σήματος στη ψηφιακή του μορφή έχει μια ορισμένη χρονική διάρκεια (κάθε ορθογώνιο έχει ένα ορισμένο πλάτος). Το

δεύτερο πρόβλημα είναι ότι οι τιμές του σήματος στη διακριτή του μορφή είναι ορισμένες, με άλλα λόγια υπάρχει ένα πεπερασμένο σύνολο από ύψη που μπορεί να έχει ένα ορθογώνιο. Κατά συνέπεια το ύψος του κάθε ορθογωνίου δεν ταυτίζεται απόλυτα με το ύψος του αναλογικού σήματος κάθε χρονική στιγμή. Τα δύο αυτά προβλήματα είναι οι βασικές πηγές απωλειών κατά τη μετατροπή ενός ήχου από αναλογική σε ψηφιακή μορφή. Τα προβλήματα αυτά μπορούν να αντιμετωπιστούν λιγότερο ή περισσότερο αποτελεσματικά με τη μεταβολή της ψηφιακής αναπαράστασης ενός ήχου αλλά δεν μπορούν ποτέ να εξαλειφθούν. Το πρόβλημα δεν είναι λοιπόν πως μπορούμε να εξαλείψουμε τις απώλειες που αναφέραμε, αλλά δεδομένης της τεχνολογίας που διαθέτουμε πως μπορούμε να κάνουμε ανεκτές τις συγκεκριμένες απώλειες.

Το κύριο πλεονέκτημα της ψηφιακής επεξεργασίας ήχων προέρχεται από το γεγονός ότι η αντιγραφή ψηφιακών σημάτων δεν προσθέτει θόρυβο στο σήμα μας. Αντίθετα η αναλογική επεξεργασία ήχων προσθέτει θόρυβο κατά τη διάρκεια της διαδικασίας αντιγραφής ενός σήματος από το ένα μέσο στο άλλο.

3.3 Συχνότητα & Ύψος

Αν θέλουμε να περιγράψουμε τον ήχο που παράγουν δύο πνευστά μουσικά όργανα όπως η τρομπέτα και η τούμπα θα παρατηρήσουμε ότι παρόλο που και τα δύο είναι παρόμοια όργανα, η τρομπέτα παράγει πιο ψηλούς ήχους από την τούμπα. Το ύψος του ήχου είναι ένα υποκειμενικό χαρακτηριστικό που σχετίζεται με ένα αντικειμενικό χαρακτηριστικό, τη συχνότητα. Όπως έχουμε ήδη αναφέρει η συχνότητα ενός συνημιτονικού ηχητικού σήματος αντιστοιχεί στον αριθμό των κύκλων που εκτελεί η συνάρτηση ανά δευτερόλεπτο και μετρείται σε κύκλους ανά δευτερόλεπτο ή Hertz (Hz). Ο συγκεκριμένος ορισμός της συχνότητας στηρίζεται στο γεγονός ότι το συνημίτονο είναι μία περιοδική συνάρτηση. Στην πράξη πολλοί λίγοι ήχοι στη φύση είναι περιοδικοί επομένως ο ορισμός της συχνότητας που δώσαμε δε θα μπορούσε να εφαρμοστεί για την πλειονότητα των ηχητικών σημάτων αν δεν ίσχυε ότι κάθε ήχος μπορεί να αναλυθεί σε ένα άθροισμα κατάλληλων συνημιτονικών σημάτων. Το γεγονός ότι υπάρχει μία τέτοια ανάλυση μας επιτρέπει να αναφερόμαστε στο *φάσμα συχνοτήτων (frequency spectrum)* ενός ηχητικού σήματος που αποτελεί ένα διάγραμμα του πλάτους που έχει κάθε συχνότητα που περιέχεται στο σήμα μας. Για παράδειγμα, η Εικόνα 3.3 περιγράφει τη μορφή που έχει το φάσμα συχνοτήτων για ένα ηχητικό σήμα. Για να γίνει αντιληπτός ένα ήχος από το ανθρώπινο αυτί θα πρέπει να έχει συχνότητες μεταξύ 20 και 20.000 Hz.



Εικόνα 3.3: Παράδειγμα φάσματος συχνοτήτων για ένα ηχητικό σήμα.

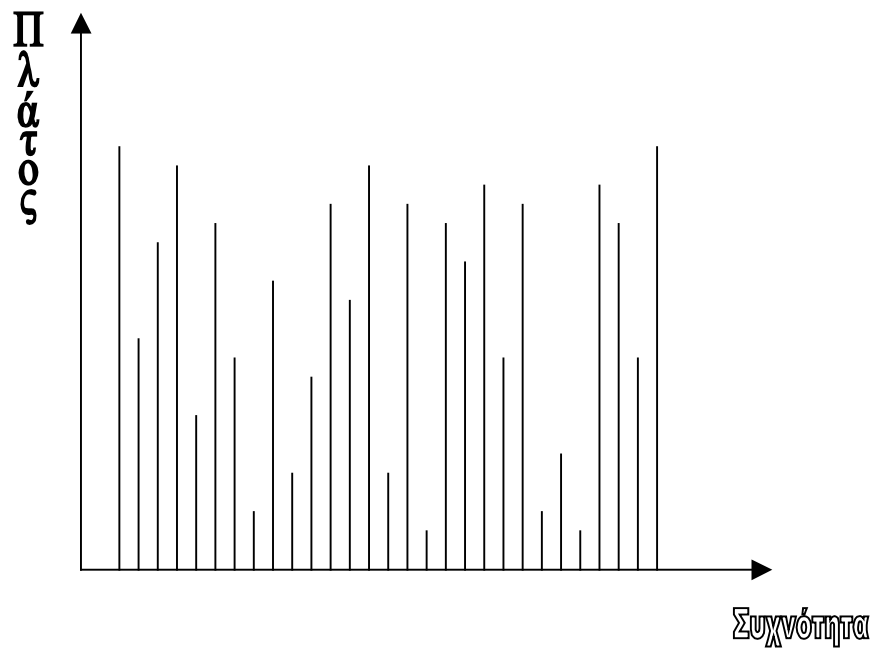
Κάθε περιοδικός ήχος μπορεί να αναλυθεί σε ένα άθροισμα συνημιτονικών συναρτήσεων με διάφορα πλάτη και φάσεις και με συχνότητες οι οποίες είναι ακέραια πολλαπλάσια μίας *θεμελιώδους συχνότητας (fundamental frequency)*. Οι συχνότητες αυτές ονομάζονται *αρμονικές (harmonic frequencies)* και το πλήθος και το σχετικό τους πλάτος είναι σε μεγάλο βαθμό υπεύθυνες για το υποκειμενικό αίσθημα της *χροιάς* ενός ήχου. Σε αρκετές περιπτώσεις η αρμονική συχνότητα με το μεγαλύτερο πλάτος καθορίζει και το ύψος του ήχου και συνήθως πάλι η συχνότητα αυτή είναι η θεμελιώδης συχνότητα.

Η αντιστοιχία μεταξύ ύψους και συχνότητας δεν ισχύει πάντα. Για παράδειγμα το ύψος αρκετά δυνατών ήχων συνήθως είναι χαμηλότερο από ότι το ύψος ασθενών ήχων με την ίδια συχνότητα.

3.4 Θόρυβος

Αν ένα ηχητικό σήμα έχει ένα φάσμα συχνοτήτων παρόμοιο με αυτό της Εικόνας 3.3 τότε μπορούμε να διακρίνουμε τις συχνότητες με το μεγαλύτερο πλάτος και να τις χρησιμοποιήσουμε για να χαρακτηρίσουμε το ηχητικό σήμα. Τι συμβαίνει όμως στην περίπτωση που το φάσμα συχνοτήτων έχει τη μορφή της Εικόνας 3.4; Στην

περίπτωση αυτή δεν υπάρχει κάποια συχνότητα ή σύνολο συχνοτήτων που μπορούν να χρησιμοποιηθούν για το χαρακτηρισμό του ηχητικού σήματος. Ένα ηχητικό σήμα που διαθέτει ένα πλήρες σύνολο συχνοτήτων ονομάζεται *θόρυβος (noise)*. Η τεχνική χρήση του όρου θόρυβος είναι αρκετά διαφορετική από την καθημερινή χρήση του όρου. Ένας δυνατός οξύς ήχος που στην καθημερινή ζωή θα τον χαρακτηρίζαμε ως θόρυβο έχει ένα αρκετά καθορισμένο φάσμα συχνοτήτων και επομένως δεν αποτελεί θόρυβο με την τεχνική σημασία του όρου.



Εικόνα 3.4: Παράδειγμα φάσματος συχνοτήτων για σήμα θορύβου.

Όσο και αν φαίνεται περίεργο ο θόρυβος βρίσκει αρκετές εφαρμογές στη μουσική. Για παράδειγμα, τα περισσότερα κρουστά όργανα παράγουν ήχους με φάσματα συχνοτήτων παρόμοια με αυτό της Εικόνας 3.4.

3.5 Ένταση

Η υποκειμενική αίσθηση της έντασης ενός ήχου είναι συνδεδεμένη με την ισχύ του ηχητικού σήματος που διεγείρει το αυτί μας. Η κύρια αιτία για την οποία ήχοι έχουν διαφορετικές εντάσεις είναι ότι πιέζουν με διαφορετική δύναμη το τύμπανο του αυτιού μας. Σε μία τέτοια περίπτωση, ένας φυσικός θα έλεγε ότι τα ηχητικά κύματα έχουν διαφορετική ισχύ. Όσο πιο μεγάλη ισχύ έχουν τα ηχητικά κύματα τόσο μεγαλύτερη πίεση εξασκούν στο μηχανισμό του αυτιού μας.

Τα ηλεκτρικά σήματα έχουν και αυτά ισχύ, η στιγμιαία τιμή της οποίας είναι ανάλογη του τετραγώνου των διαφορών τάσης που προκαλούν στο μέσο στο οποίο διαδίδονται. Η ολική ισχύ ενός ηλεκτρικού σήματος προέρχεται από την άρθρωση όλων των στιγμιαίων τιμών της ισχύος στη μονάδα του χρόνου. Η ολική ισχύς ενός ηχητικού σήματος σε ψηφιακή μορφή είναι ανάλογη του αρθροίσματος των τετραγώνων των διακριτών τιμών από τις οποίες συντίθεται. Η μέση τιμή της ισχύος στην περίπτωση αυτή είναι ανάλογη της μέσης τιμής του αρθροίσματος των τετραγώνων των διακριτών τιμών από τις οποίες συντίθεται το σήμα μας.

Στις εφαρμογές πολυμέσων δε μας ενδιαφέρει τόσο η απόλυτη τιμή της ισχύος ενός ηχητικού σήματος όσο η σχετική τιμή της ισχύος μεταξύ δύο ηχητικών σημάτων. Η σχετική ισχύς μετριέται σε bels ή πιο συχνά σε decibels (dB) ($1 \text{ dB} = 0.1 \text{ bel}$). Για να συγκρίνουμε την ισχύ δύο ηχητικών σημάτων υπολογίζουμε το λόγο των ισχύων τους. Ο λογάριθμος ως προς 10 του λόγου αυτού είναι η διαφορά των δύο ισχύων σε bels. Αν πολλαπλασιάσουμε το συγκεκριμένο λογάριθμο επί 10 παίρνουμε την διαφορά σε dB. Μαθηματικά έχουμε ότι:

$$(\text{Διαφορά Έντασης})_{\text{dB}} = 10 \times \log_{10} (\text{Ισχύς}_1 / \text{Ισχύς}_2)$$

Κατά σύμβαση για τη μέτρηση της έντασης ενός ήχου χρησιμοποιείται η διαφορά της έντασης του σε dB από ένα ηχητικό σήμα αναφοράς. Το σήμα αναφοράς είναι προσεγγιστικά ο πιο ασθενής ήχος συχνότητας 1000 Hz που μπορεί να γίνει αντιληπτός από το ανθρώπινο αυτί. Με βάση αυτή την κλίμακα ο πιο δυνατός ήχος που μπορούμε να ακούσουμε είναι περίπου 120 dB δυνατότερος (δηλ. ένα εκατομμύριο επί ένα εκατομμύριο φορές πιο δυνατός σε απόλυτη ένταση από τον ήχο αναφοράς) και αντιστοιχεί στον ήχο του κινητήρα ενός αεροπλάνου. Η λογαριθμική βάση υπολογισμού των dB έχει σαν αποτέλεσμα δύο ήχοι καθένας από τους οποίους έχει π.χ. ένταση 60 dB να έχουν ένταση μόλις 63 dB όταν ακούγονται συγχρόνως. Μια αύξηση της έντασης ενός ήχου κατά 1.000.000 φορές αντιστοιχεί σε μία αύξηση της έντασης του κατά 60 dB.

Πέρα από την μέτρηση της έντασης ενός ήχου η κλίμακα σε dB χρησιμοποιείται στη μέτρηση της απώλειας έντασης που έχουν ορισμένα ηλεκτρονικά κυκλώματα ή αλγόριθμοι επεξεργασίας ψηφιακού ήχου. Για παράδειγμα διαφορετικοί ήχοι με την ίδια ισχύ μπορεί να προκύψουν με σχετική διαφορά 10 dB στην ισχύ τους μετά από

μια τέτοια επεξεργασία. Επιπλέον η κλίμακα σε dB χρησιμοποιείται για τη μέτρηση του θορύβου ή της παραμόρφωσης που εισάγεται σε ένα σήμα.

Για παράδειγμα ένας μουσικός οπτικός δίσκος αποθηκεύει τα δείγματα του ήχου σε ακεραίους μήκους 16-bit. Το εύρος των αριθμών που μπορούν να αποθηκευτούν σε ακέραιους αυτής της μορφής είναι από -32.768 ως $+32.767$. Κατά τη διάρκεια της αποθήκευσης οι τιμές του πρωτότυπου σήματος έχουν στρογγυλευτεί στον πλησιέστερο ακέραιο άρα το μέγιστο λάθος που έχει γίνει κατά την αποθήκευση είναι 0.5, το οποίο είναι 2^{-16} φορές μεγαλύτερο από τη μέγιστη τιμή που μπορεί να λάβει το σήμα μας. Επειδή η ισχύς του σήματος είναι ανάλογη του τετραγώνου του πλάτους του, το λάθος στην περίπτωση μας θα έχει 2^{-32} φορές μεγαλύτερη ισχύ από την ισχύ του μεγαλύτερου σήματος μας. Ο λόγος της ισχύος μεταξύ του δείγματος με τη μέγιστη τιμή και του λάθους (θορύβου) στην περίπτωση μας είναι 2^{32} προς 1, ή $10 \times \log_{10}(2^{32}) \approx 96.3$ dB.

Υπάρχουν δύο κύριοι λόγοι για τους οποίους η μέτρηση σε dB αντιστοιχεί στην ανθρώπινη αντίληψη της έντασης ενός ήχου (δηλ. την υποκειμενική ένταση (*loudness*)). Ο πρώτος λόγος είναι ότι η ανθρώπινη ακοή είναι κατά προσέγγιση λογαριθμική δηλ. η διαφορά μεταξύ των υποκειμενικών εντάσεων δύο ήχων δεν εξαρτάται από τη απόλυτη τιμή της διαφοράς αυτής αλλά από το λόγο των ισχύων τους. Παρόλο που δεν είναι απόλυτα σωστό μπορούμε να θεωρήσουμε ότι το 1 dB αντιστοιχεί στην ελάχιστη διαφορά έντασης που μπορεί να γίνει αντιληπτή μεταξύ δύο ήχων.

Ο δεύτερος λόγος προέρχεται από το γεγονός ότι η υποκειμενική ένταση ενός σύνθετου ήχου εξαρτάται από τη σχετική τιμή των συνιστωσών του. Πιο συγκεκριμένα, η ανθρώπινη ακοή εμφανίζει μια συγκεκριμένη μορφή ψευδαίσθησης που είναι γνωστή ως *συγκάλυψη* (*masking*). Αν ένας ήχος αποτελείται από δύο συνιστώσες και ισχύει ότι η ένταση της μίας από αυτές είναι αρκετά μεγαλύτερη από την άλλη, τότε σε αρκετές περιπτώσεις η ασθενέστερη συνιστώσα δε γίνεται αντιληπτή. Ουσιαστικά, η ανθρώπινη ακοή προσαρμόζεται στο δυνατότερο ήχο ενώ ο ασθενέστερος ήχος γίνεται αντιληπτός σαν περισσότερο ασθενής από ότι πραγματικά είναι. Το φαινόμενο γίνεται αρκετά πιο έντονο όταν οι δύο συνιστώσες έχουν κοντινά ύψη. Η συγκάλυψη βρίσκει αρκετές εφαρμογές στη συμπίεση ηχητικών σημάτων. Πιο συγκεκριμένα, αρκετοί μηχανισμοί συμπίεσης ανιχνεύουν και εξαλείφουν στο συμπιεσμένο σήμα τις συνιστώσες εκείνες που θα συγκαλυφθούν

κατά την ακρόαση του ήχου. Το σήμα που προκύπτει είναι απλούστερο και περισσότερο εύκολο να συμπιεστεί.

Η υποκειμενική αντίληψη της έντασης ενός ήχου εξαρτάται και από άλλους παράγοντες όπως λ.χ. το ύψος του ήχου. Πιο συγκεκριμένα, το αυτί μας είναι πιο ευαίσθητο σε ορισμένα εύρη συχνοτήτων από ότι σε άλλα. Η πιο ευαίσθητη περιοχή για το αυτί μας βρίσκεται μεταξύ 2700-3200 Hz με την ευαισθησία να μειώνεται βαθμιαία σε κάθε πλευρά αυτής της περιοχής. Κατά συνέπεια ένας συνημιτονικός ηχητικό σήμα με συχνότητα 3000 Hz και με μία ορισμένη ένταση θα ακούγεται πιο ισχυρός από ένα σήμα της ίδιας έντασης αλλά με συχνότητα 200 ή 8000 Hz.

Επιπλέον το αυτί μας είναι λιγότερο ευαίσθητο σε σύνθετους παρά σε απλούς τόνους. Πιο συγκεκριμένα, δεν αντιλαμβανόμαστε αρκετά εύκολα θορύβους που περιέχουν αρκετές υψηλές συχνότητες. Ορισμένες ψηφιακές τεχνικές, όπως το *dithering*, μετασχηματίζει ορισμένους τύπους λαθών κατά την επεξεργασία ενός ηχητικού σήματος σε θόρυβο υψηλής συχνότητας που γίνεται δύσκολα αντιληπτός.

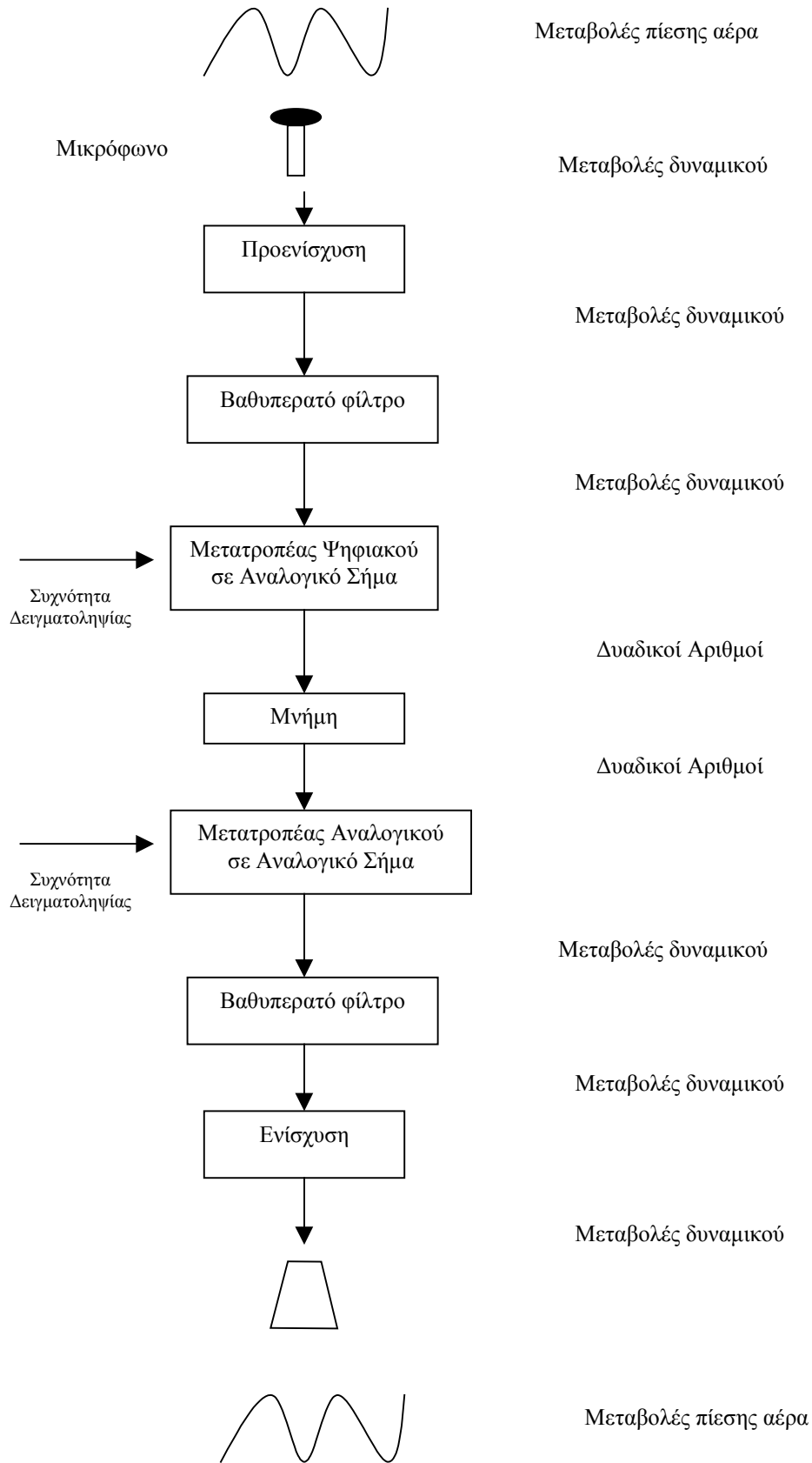
3.6 Στάδια Επεξεργασίας Ψηφιακού Ήχου

Ενα σύστημα ψηφιακού ήχου για να ενσωματωθεί σε ένα σύστημα πολυμέσων θα πρέπει να προσφέρει τις δυνατότητες της καταγραφής, της αποθήκευσης και της απόδοσης (playback) ηχητικών σημάτων. Ενα πλήρες σύστημα που προσφέρει τις δυνατότητες που αναφέραμε αποτελείται από τα στάδια που περιγράφονται στην Εικόνα 3.5 και τα οποία θα αναλύσουμε στη συνέχεια.

Η καταγραφή ενός ηχητικού σήματος γίνεται με τη βοήθεια ενός μικροφώνου που μετατρέπει τα ηχητικά σήματα σε διαφορές δυναμικού μέσω του μηχανισμού που περιγράφηκε στην § 3.1. Επειδή οι διαφορές αυτές έχουν μικρό εύρος για τη διευκόλυνση της επεξεργασίας συνήθως παρεμβάλλεται ένα στάδιο προενίσχυσης του σήματος έτσι ώστε να προκύψει μία περισσότερο ενισχυμένη μορφή του ηλεκτρικού σήματος. Το ηλεκτρικό σήμα μετά το στάδιο της προενίσχυσης τροφοδοτείται σε ένα βαθυπερατό φίλτρο¹. Σκοπός του φίλτρου αυτού είναι η αποκοπή των συχνοτήτων που υπάρχουν στο σήμα μας και που είναι μεγαλύτερες του μισού της συχνότητας δειγματοληψίας έτσι ώστε να αποφύγουμε φαινόμενα αναδίπλωσης κατά τη δειγματοληψία του ηλεκτρικού σήματος. Στη συνέχεια το

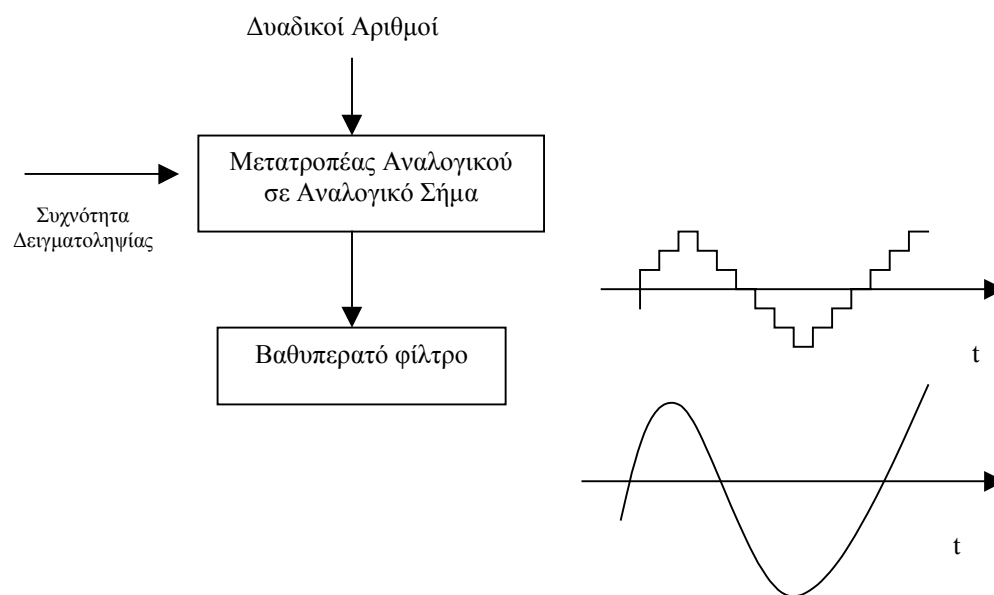
¹ Για τα χαρακτηριστικά και τη λειτουργία των φίλτρων θα αναφερθούμε σε επόμενη ενότητα.

στάδιο της δειγματοληψίας μετατρέπει το αναλογικό ηλεκτρικό σήμα στην είσοδο του σε μια ακολουθία δυαδικών αριθμών η οποία και αποθηκεύεται στη μνήμη του υπολογιστικού συστήματος. Στο σημείο αυτό το ηχητικό σήμα έχει μετασχηματιστεί από αναλογική σε ψηφιακή μορφή.



Εικόνα 3.5: Στάδια Επεξεργασίας Συστήματος Ψηφιακού Ήχου.

Η απόδοση ενός ψηφιακού ηχητικού σήματος ακολουθεί μία αντίστροφη ακολουθία βημάτων από αυτή της καταγραφής. Πιο συγκεκριμένα, το σήμα τροφοδοτείται από τη μνήμη του υπολογιστικού συστήματος σε ένα μετατροπέα ψηφιακού σε αναλογικό σήματος. Ανάλογα με τον τρόπο λειτουργίας του μετατροπέα και τη συχνότητα δειγματοληψίας που χρησιμοποιείται ο μετατροπέας παράγει ένα αναλογικό ηλεκτρικό σήμα που αποτελεί μια προσεγγιστική μορφή του σήματος εισόδου. Η εικόνα 3.6 μας δίνει μια ιδέα της μορφής που έχει το σήμα στην έξοδο του μετατροπέα. Το βαθυπερατό φίλτρο που ακολουθεί σκοπό έχει να εξομαλύνει τη μορφή του ηλεκτρικού σήματος το οποίο ενισχύεται και στη συνέχεια τροφοδοτείται στην είσοδο ενός ηχείου το οποίο και αποδίδει το σήμα.



Εικόνα 3.6: Εξομάλυνση αναλογικού σήματος κατά το στάδιο της απόδοσης.

Τα προβλήματα που αντιμετωπίζει η ψηφιακή επεξεργασία ήχων εντοπίζονται στην επιλογή μίας κατάλληλης συχνότητας δειγματοληψίας τόσο κατά την καταγραφή όσο και κατά την απόδοση ενός ήχου και στην εισαγωγή θορύβου λόγω της πεπερασμένης διακριτικότητα της ψηφιακής αναπαράστασης.

3.6.1 Προβλήματα Δειγματοληψίας

Η χρησιμοποίηση ενός βαθυπερατού φίλτρου πριν από τη δειγματοληψία του αναλογικού σήματος σκοπό έχει να αποκόψει πλήρως συχνότητες που μπορούν να

προκαλέσουν αναδίπλωση κατά τη δειγματοληψία. Από την §2.5 ξέρουμε ότι τέτοιες συχνότητες είναι όσες έχουν τιμή μεγαλύτερη του $\frac{1}{2}$ της συχνότητας Nyquist. Η χρήση τέτοιων φίλτρων εισάγει δύο νέα προβλήματα. Το πρώτο πρόβλημα προέρχεται από το γεγονός ότι δεν υπάρχει ένα τέλειο βαθυτερατό φίλτρο. Με άλλα λόγια δεν υπάρχει ένα φίλτρο που να μπορεί να αποκόψει τέλεια συχνότητες πάνω από ένα επιθυμητό όριο.

Το δεύτερο πρόβλημα προέρχεται από το γεγονός ότι η χρήση βαθυτερατών φίλτρων εισάγει παραμορφώσεις φάσης στο σήμα στο οποίο εφαρμόζεται. Γενικά η επίδραση που έχει η φάση στον τρόπο με τον οποίο αντιλαμβανόμαστε ένα ήχο είναι αρκετά σύνθετη. Το αυτί μας δε μπορεί να αντιληφθεί διαφορές μεταξύ ηχητικών σημάτων που έχουν μια απόλυτη διαφορά φάσης. Παρόλα αυτά οι μεταβολές στη φάση ενός σήματος αποκτούν ιδιαίτερη σημασία σε περίπτωση φιλτραρίσματος ενός ήχου. Πιο συγκεκριμένα, κάθε φίλτρο συνδυάζει σήματα που βρίσκονται εκτός φάσης για να μεταβάλλει τα χαρακτηριστικά τους. Ένα φίλτρο δημιουργεί τέτοιες διαφορές καθυστερώντας το σήμα στην είσοδο του για ένα ορισμένο χρόνο και συνδυάζοντας το σήμα που προκύπτει με την παρούσα είσοδο του για να δημιουργήσει αλληλοαναιρέσεις σε ορισμένες συχνότητες οι οποίες μεταβάλλουν το φάσμα του σήματος εισόδου. Όταν οι συχνότητες οι οποίες φιλτράρονται μεταβάλλονται με το χρόνο τότε το αποτέλεσμα περιγράφεται από το ηχητικό εφέ που είναι γνωστό σα *flanging* ή *phasing*. Επιπλέον έχει παρατηρηθεί ότι συστήματα αναπαραγωγής ήχων τα οποία εισάγουν παραμορφώσεις φάσης στο σήμα εισόδου τους έχουν μια θολή *ηχητική εικόνα (imaging)* δυσκολεύοντας την (νοητική) τοποθέτηση των ηχητικών πηγών στο χώρο.

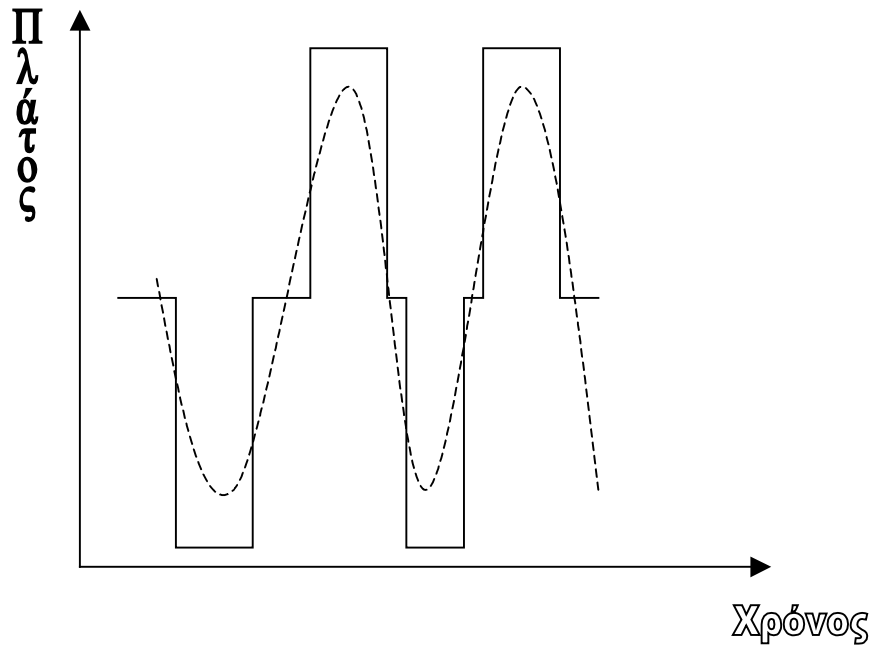
3.6.2 Προβλήματα Διακριτικότητας

Ένα ακόμα σημαντικό πρόβλημα κατά τη διαδικασία της ηχογράφησης και αναπαραγωγής ψηφιακών ήχων είναι ο θόρυβος που προέρχεται από την πεπερασμένη *διακριτικότητα* της ψηφιακής αναπαράστασης (δηλ. από τον αριθμό των bits που χρησιμοποιούνται για την αναπαράσταση ενός σήματος). Ο θόρυβος αυτός είναι γνωστός και σα *θόρυβος κβαντοποίησης (quantization error* ή *quantization noise)*. Πιο συγκεκριμένα, κατά τη σχεδίαση ενός συστήματος ψηφιακού ήχου θα πρέπει να αποφασιστεί ο τρόπος με τον οποίο θα αποθηκεύονται τα δείγματα του ήχου καθώς κάθε υπολογιστικό σύστημα μπορεί να χειριστεί ένα πεπερασμένο σύνολο από αριθμητικές τιμές. Το μήκος της λέξης κλαθε υπολογιστή καθορίζει και το εύρος των τιμών που μπορεί να χειριστεί. Για παράδειγμα, ένα σύστημα που

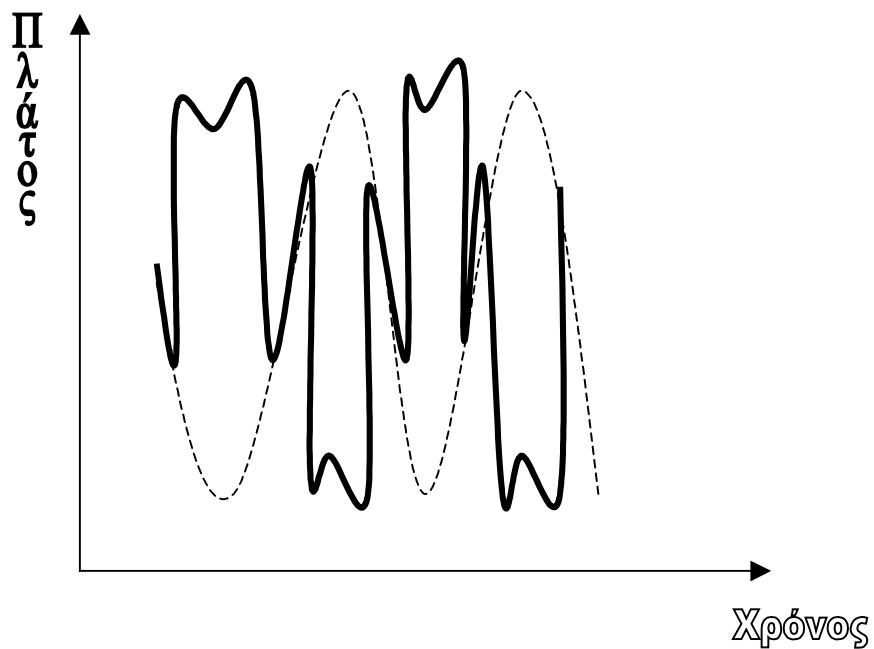
αποθηκεύει ακέραιους σε λέξεις των 8 ή 16 bits μπορεί να χειριστεί είτε 256 είτε 65.536 διαφορετικούς ακεραίους. Ο θόρυβος κβαντοποίησης προέρχεται από το γεγονός ότι τα δείγματα του αναλογικού σήματος που θα χρησιμοποιηθούν δεν είναι κατ' ανάγκη ακέραιοι και συνεπώς η μετατροπή τους σε ψηφιακά δείγματα συνεπάγεται τη στρογγύλευση τους στην πλησιέστερη ακέραια μορφή. Στην περίπτωση αυτή μπορούμε να θεωρήσουμε ότι το ψηφιοποιημένο σήμα αποτελείται από το άρθροισμα του αναλογικού σήματος μας και ενός σήματος θορύβου.

Για να μπορέσουμε να αντιμετωπίσουμε το θόρυβο κβαντοποίησης θα πρέπει να καταλάβουμε πόσο δυνατός είναι σε σύγκριση με το υπόλοιπο σήμα. Ο λόγος σήματος προς θόρυβο (*signal-to-noise ratio* ή *SNR*) μετρά τη σχετική ισχύ του ηχητικού σήματος ως προς το θόρυβο που περιέχεται σε αυτό. Προφανώς ένα σύστημα ψηφιακής επεξεργασίας ήχου θέλουμε να έχει όσο το δυνατόν μεγαλύτερο λόγο σήματος προς θόρυβο. Για παράδειγμα στην §3.5 υπολογίσαμε το μέγιστο λόγο σήματος προς θόρυβο για ένα 16-bit σύστημα επεξεργασίας ίσο με περίπου 96 dB. Σε συστήματα με μικρότερο μήκος λέξης η μέγιστη τιμή του λόγου είναι μικρότερη.

Αν το σήμα μας δεν έχει τη μέγιστη ισχύ τότε ο λόγος σήματος προς θόρυβο θα έχει μικρότερη τιμή από τη μέγιστη τιμή του. Η χειρότερη περίπτωση συμβαίνει όταν το σήμα μας έχει τόσο μικρή ισχύ ώστε να αναγκάζει τις τιμές που προκύπτουν κατά τη δειγματοληψία να αλλάζουν τιμή κατά μία η δύο στάθμες. Για παράδειγμα η Εικόνα 3.7 απεικονίζει ένα συνημιτονικό παλμό το πλάτος του οποίου είναι μικρότερο από 1. Στην περίπτωση αυτή οι τιμές του ψηφιακού σήματος θα εναλλάσσονται μεταξύ των τιμών -1 , 0 και 1 . Η Εικόνα 3.8 απεικονίζει το αναλογικό σήμα και το λάθος που προκύπτει κατά τη στρογγύλευση των τιμών του δείγματος κάθε χρονική στιγμή.



Εικόνα 3.7: Δειγματοληψία ενός αρκετά ασθενούς σήματος. Με διακεκομμένη και συνεχή γραμμή περιγράφονται το αναλογικό και το ψηφιακό σήμα αντίστοιχα.



Εικόνα 3.8: Δειγματοληψία ενός αρκετά ασθενούς σήματος. Με διακεκομμένη και συνεχή γραμμή περιγράφονται το αναλογικό και το σήμα θορύβου που παράγεται κατά τη δειγματοληψία της Εικόνας 3.7 αντίστοιχα.

Επειδή το αυτί μας είναι λιγότερο ευαίσθητο στην παρουσία τυχαίου θορύβου παρά θορύβου παραμορφώσεων, μια διαδικασία με την οποία αντιμετωπίζεται η παρουσία ορισμένου τύπου παραμορφώσεων είναι με τη μετατροπή των παραμορφώσεων αυτών σε θόρυβο υψηλών συχνοτήτων. Η μέθοδος αυτή που ονομάζεται *dithering* συνίσταται στην εισαγωγή μιάς μικρής ποσότητας αναλογικού θορύβου πριν από τη δειγματοληψία. Ο θόρυβος αυτός λέγεται *dither* και προκαλεί τυχαίες μεταβολές σε συχνότητες χαμηλής έντασης οι οποίες περιορίζουν την εμφάνιση των τετραγωνικών αυτών παλμών (και των αρμονικών τους).

Ενα άλλο πρόβλημα που οφείλεται στην περιορισμένη διακριτικότητα της ψηφιακής αναπαράστασης αφορά τις μέγιστες εντάσεις ήχων που μπορούν να αποδοθούν σωστά από το σύστημα. Πιο συγκεκριμένα, το εύρος των εντάσεων που μπορεί να αναπαραχθεί από ένα ψηφιακό σύστημα ορίζεται σαν το *δυναμικό εύρος* (*dynamic range*) του συστήματος. Το δυναμικό εύρος ενός τέτοιου συστήματος μας δίνεται προσεγγιστικά από τον ακόλουθο τύπο:

$$(\text{δυναμικό εύρος})_{\text{dB}} = \text{αριθμός bits λέξης} \times 6.11$$

Όταν το ψηφιακό σύστημα είναι αναγκασμένο να αποδώσει ένα ήχο εντάσεως μεγαλύτερης από το δυναμικό του εύρος, τότε, λόγω υπερχειλίσης των δεδομένων, η θετική τιμή του σήματος μας μετατρέπεται σε μια μεγάλη αρνητική τιμή με ολέθρια αποτελέσματα στην πιστότητα της απόδοσης του ήχου.

3.7 Μέθοδοι Επεξεργασίας Ηχητικών Σημάτων

Στην τρέχουσα ενότητα θα παρουσιάσουμε τις κυριότερες μεθόδους επεξεργασίας σήματος που χρησιμοποιούνται σε ηχητικά σήματα. Οι μέθοδοι αυτές περιλαμβάνουν τη μεταβολή του δυναμικού εύρους, τα διάφορα είδη φίλτρων, τη συνέλιξη και τους διάφορους τρόπους χρονική υστέρησης ενός ηχητικού σήματος.

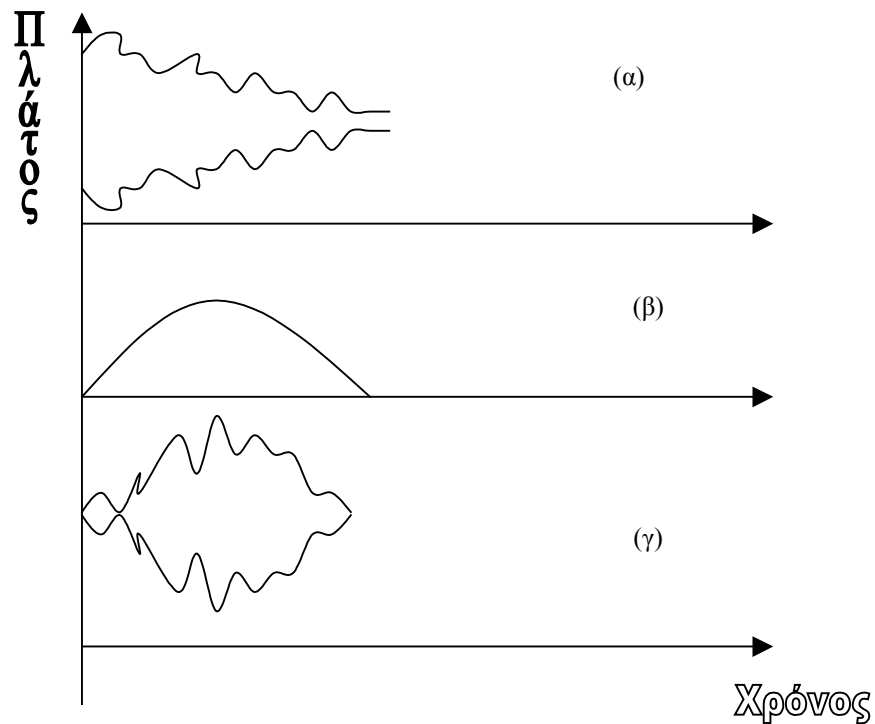
3.7.1 Δυναμική Μεταβολή Εύρους

Οι τεχνικές δυναμικής μεταβολής του εύρους των ηχητικών σημάτων (*dynamic range processing*) μετασχηματίζουν το πλάτος (ένταση) των ηχητικών σημάτων. Οι τεχνικές αυτές χρησιμοποιούνται σε μια σειρά από συσκευές όπως μορφοποιητές περιγράμματος (*envelope shapers*), πύλες θορύβου (*noise gates*), συμπιεστές

(compressors), μεγεθυντές (expanders), κλπ. Οι συγκεκριμένες μέθοδοι εφαρμόζονται σε μια ποικιλία εργασιών που εκτείνονται από αρκετά πρακτικές εφαρμογές όπως λ.χ. η απάλειψη του θορύβου από παλιές ηχογραφήσεις μέχρι τη μορφοποίηση μουσικών σημάτων ή ακόμα της ανθρώπινης φωνής κατά τη μουσική σύνθεση.

3.7.1.1 Μορφοποίηση Περιγράμματος

Τα περισσότερα από τα συστήματα επεξεργασίας ήχων παρέχουν τη δυνατότητα της μορφοποίησης του περιγράμματος των σημάτων που διαχειρίζονται. Η μορφοποίηση μπορεί να περιλαμβάνει μια μικρή μεταβολή της έντασης του σήματος είτε μια επανασχεδίαση της μορφής του σήματος. Για παράδειγμα, η Εικόνα 3.9 περιγράφει τη μορφοποίηση με γραφικό τρόπο ενός ηχητικού σήματος.



Εικόνα 3.9: Μορφοποίηση περιγράμματος: (α) Αρχικό σήμα (β) Επιθυμητό περίγραμμα (γ) Μετασχηματισμός του αρχικού σήματος με βάση το επιθυμητό περίγραμμα.

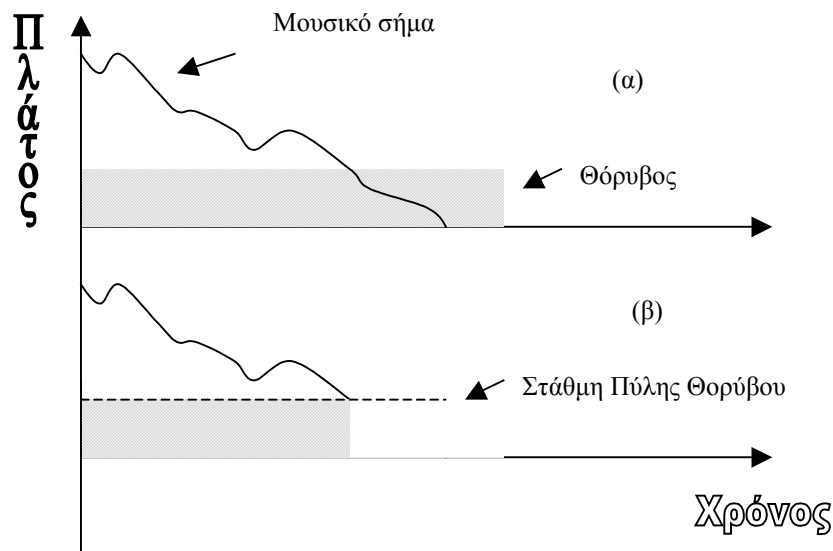
3.7.1.2 Πύλες Θορύβου

Οι πύλες θορύβου προσφέρουν ένα φτηνό τρόπο καθαρισμού μουσικών σημάτων που περιέχουν θόρυβο σταθερής έντασης (π.χ. hiss ή hum). Η πύλη θορύβου συμπεριφέρεται σαν ένας διακόπτης ο οποίος είναι ανοιχτός όταν δέχεται ως είσοδο

ένα μουσικό σήμα υψηλής έντασης και κλείνει όταν η ένταση του σήματος πέφτει κάτω από ένα καθορισμένο όριο αποκόπτοντας έτσι οποιονδήποτε εναπομείναντα θόρυβο (βλ. Εικόνα 3.10).

3.7.1.3 Συμπίεστές

Οι συμπίεστές είναι ενισχυτές στους οποίους το μέγεθος ενίσχυσης (*gain*) του σήματος μεταβάλλεται ανάλογα με την ένταση του σήματος εισόδου. Οι συμπίεστές χρησιμοποιούνται αρκετά συχνά για να διατηρείται η ένταση του σήματος εισόδου σχετικά σταθερή. Στην περίπτωση αυτή ο συμπίεστής μειώνει την ένταση του σήματος εισόδου όταν αυτή ξεπερνά κάποια στάθμη. Ένας αρκετά διαδεδομένος τρόπος περιγραφής της λειτουργίας ενός συμπίεστή αποτελεί η *συνάρτηση μεταφοράς (transfer function)* η οποία περιγράφει τον τρόπο με τον οποίο μια τιμή της έντασης του σήματος εισόδου αντιστοιχίζεται σε μια τιμή της έντασης του σήματος εισόδου (βλ. Εικόνα 3.11).



Εικόνα 3.10: Πύλη Θορύβου: (α) Αρχικό σήμα + Θόρυβος (β) Μετασχηματισμός του αρχικού σήματος με χρήση πύλης θορύβου.

Ο λόγος συμπίεσης (*compression ratio*) ορίζεται ως ο λόγος της μεταβολής της έντασης του σήματος εισόδου ως προς τη μεταβολή του σήματος εξόδου. Για παράδειγμα ένας λόγος συμπίεσης 4:1 καθορίζει ότι μια μεταβολή της έντασης εισόδου κατά 4 dB προκαλεί μια μεταβολή μόνο 1 dB στο σήμα εξόδου. Ακραίοι λόγοι συμπίεσης (π.χ. 10:1) περιγράφονται με τον όρο *limiting* και χρησιμοποιούνται κυρίως σε περιπτώσεις κατά τις οποίες πρέπει να διασφαλιστεί ότι καμία από τις

πηγές παραγωγής ήχων δεν φτάνει στα όρια του δυναμικού της εύρους (π.χ. ζωντανές ηχογραφήσεις).

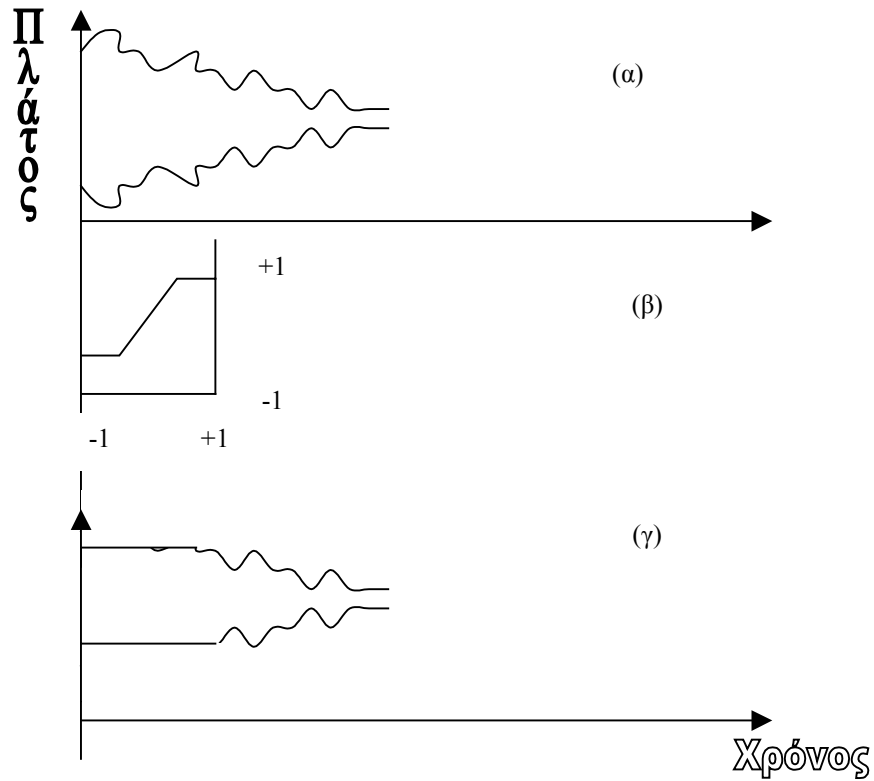
3.7.1.4 Μεγεθυντές

Οι μεγεθυντές (expanders) λειτουργούν αντίστροφα από τους συμπιεστές και μεγεθύνουν μικρές μεταβολές στο σήμα εισόδου. Ο λόγος μεγέθυνσης καθορίζει το βαθμό της μεταβολής του σήματος εισόδου. Για παράδειγμα, ένας λόγος μεγέθυνσης 1:5 δηλώνει ότι μία μεταβολή του σήματος εισόδου κατά 1 dB θα προκαλέσει μια μεταβολή κατά 5 dB του σήματος εξόδου.

Αρκετές από τις μεθόδους απάλειψης θορύβων σε ηχητικά συστήματα χρησιμοποιούν ένα μίγμα συμπιεστών και μεγεθυντών. Για παράδειγμα, το σύστημα Dolby εφαρμόζει ξεχωριστά ζεύγη συμπιεστών και μεγεθυντών για διαφορετικά εύρη συχνοτήτων στο σήμα εισόδου. Ο διαχωρισμός ανά συχνότητες των μεθόδων συμπίεσης διευκολύνει την αντιμετώπιση των προβλημάτων που εισάγει η μεταβολή του δυναμικού εύρους ενός ηχητικού σήματος (βλ. § 3.7.1.5).

3.7.1.5 Προβλήματα κατά τη Μεταβολή του Δυναμικού Εύρους

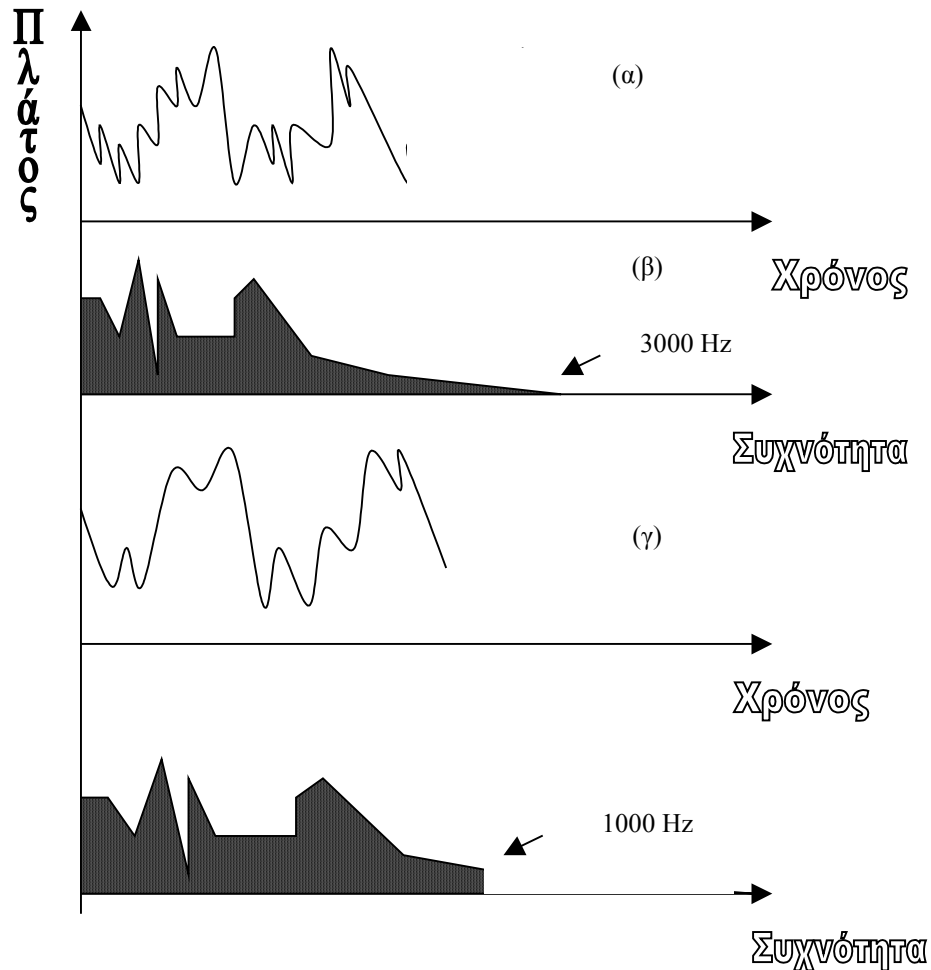
Το μεγαλύτερο πρόβλημα που προκύπτει από τη χρήση των μεθόδων μεταβολής του δυναμικού εύρους που εξετάσαμε προέρχεται από την παραμόρφωση που εισάγουν στις μεταβατικές (transient) περιοχές μίας κυματομορφής όπως π.χ. στις αρχικές απότομες μεταβολές της έντασης (*attack*) και στη βαθμιαία εξασθένιση (*decay*) του σήματος που παράγουν αρκετά μουσικά όργανα και που καθορίζουν σε σημαντικό βαθμό τη χροιά του ήχου τους.



Εικόνα 3.11: Συμπίεση: (α) Αρχικό σήμα (β) Συνάρτηση Μεταφοράς (γ) Μετασχηματισμός του αρχικού σήματος με βάση τη συνάρτηση μεταφοράς.

3.7.2 Ψηφιακά Φίλτρα

Ο όρος φίλτρα αναφέρεται σε συσκευές που ενισχύουν ή αποκόπτουν τμήματα του φάσματος συχνοτήτων που απαρτίζουν ένα ήχητικό σήμα. Υπάρχουν δύο τρόποι με τους οποίους μπορούμε να περιγράψουμε το αποτέλεσμα της εφαρμογής ενός φίλτρου σε ένα σήμα εισόδου: στο *χώρο του χρόνου* (*time domain*) και στο *χώρο της συχνότητας* (*frequency domain*). Για παράδειγμα η Εικόνα 3.12 περιγράφει το αποτέλεσμα της εφαρμογής ενός φίλτρου που αποκόπτει συχνότητες μεγαλύτερες των 1000 Hz σε ένα ηχητικό σήμα.

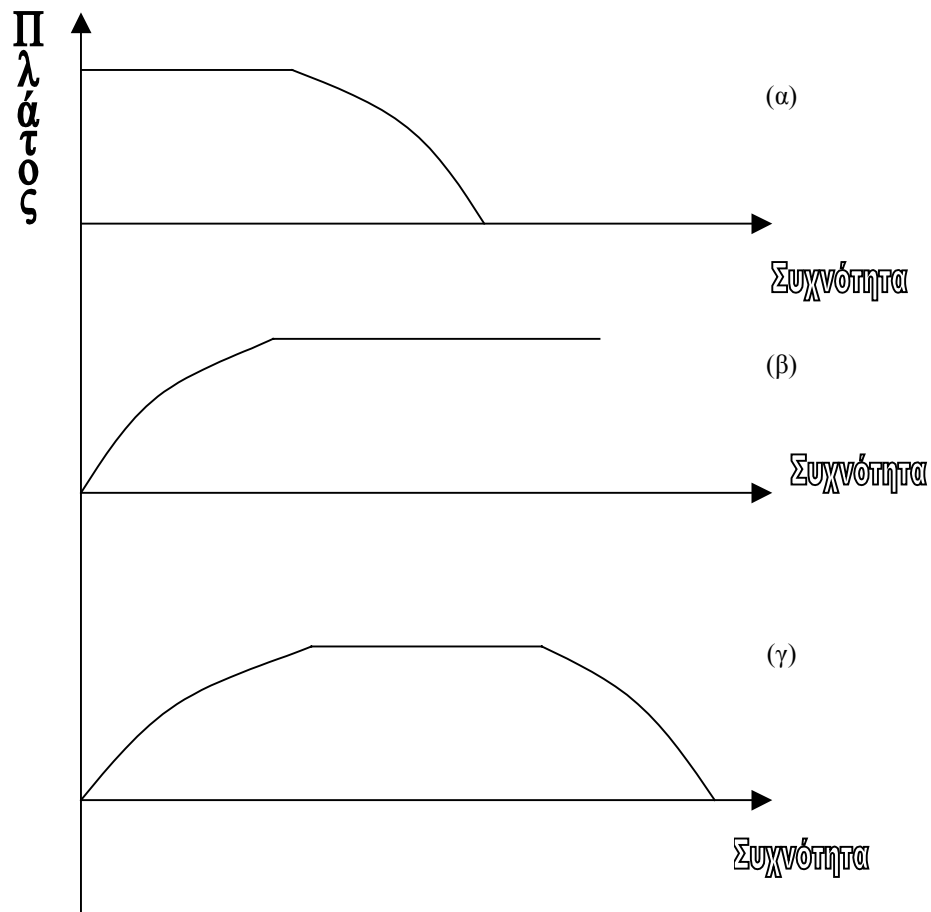


Εικόνα 3.12: Ψηφιακό φίλτρο: (α) Αρχικό σήμα στο πεδίο του χρόνου (β) Φάσμα συχνοτήτων αρχικού σήματος (γ) Φιλτραρισμένο σήμα στο πεδίο του χρόνου (δ) Φάσμα συχνοτήτων φιλτραρισμένου σήματος.

Για να έχουμε μια όσο το δυνατόν πληρέστερη περιγραφή των αποτελεσμάτων που έχει η εφαρμογή ενός σήματος σε ένα ηχητικό σήμα θα πρέπει να χρησιμοποιήσουμε ένα σήμα εισόδου που να περιέχει όλες τις συχνότητες. Τι μορφή θα έχει ένα τέτοιο σήμα:

Από τον προηγούμενο αιώνα είναι γνωστό μέσω της ανάλυσης κατά Fourier ότι υπάρχει μία αντίστροφη σχέση μεταξύ της διάρκειας ενός σήματος και των συχνοτήτων που περιέχει. Διαισθητικά μπορούμε να δούμε ότι ένα συνημιτονικό σήμα άπειρης διάρκειας θα αποτελείται από μία και μόνο συχνότητα. Όσο μικραίνει η διάρκεια ενός τέτοιου σήματος, η πολυπλοκότητα του φάσματος συχνοτήτων που αντιστοιχεί σε αυτό θα αυξάνει. Με άλλα όσο μικραίνει η διάρκεια ενός σήματος τόσο θα αυξάνει ο αριθμός των συνημιτονικών συνιστωσών άπειρης διάρκειας που θα

πρέπει να χρησιμοποιήσουμε για να το περιγράψουμε. Σε ένα ψηφιακό σύστημα, το πιο σύντομο σήμα που υπάρχει αποτελείται από ένα και μόνο δείγμα. Το σήμα αυτό περιέχει το σύνολο των συχνοτήτων και αποτελεί τη ψηφιακή προσέγγιση ενός αναλογικού σήματος απειροελάχιστης διάρκειας που είναι γνωστό ως το *δέλτα του Kronecker* ή η *κρουστική συνάρτηση (unit impulse)*. Αν τροφοδοτήσουμε το σήμα αυτό στη είσοδο ενός γραμμικού, αμετάβλητου στο χρόνο φίλτρου τότε το σήμα εξόδου θα περιγράψει την επίδραση του φίλτρου σε όλο το φάσμα συχνοτήτων με άλλα λόγια θα μας δίνει μια ακριβέστατη περιγραφή της συμπεριφοράς του φίλτρου στο χώρο της συχνότητας,. Η περιγραφή του συγκεκριμένου σήματος εξόδου στο χώρο του χρόνου ονομάζεται *κρουστική απόκριση (impulse response)* του φίλτρου. Η περιγραφή του συγκεκριμένου σήματος εξόδου στο χώρο της συχνότητας ονομάζεται *απόκριση συχνότητας (frequency response)* του φίλτρου.



Εικόνα 3.13: Παραδείγματα απόκρισης συχνότητας φίλτρων: (α) Βαθυπερατό φίλτρο (β) Υψιπερατό φίλτρο (γ) Ζωνοπερατό φίλτρο.

Ανάλογα με τη μορφή που έχει η απόκριση συχνότητας τα φίλτρα διαχωρίζονται (κυρίως) σε *βαθυπερατά (lowpass)*, *υψιπερατά (highpass)* και *ζωνοπερατά (bandpass)* (βλ. Εικ. 3.13). Ένα βαθυπερατό φίλτρο εξασθενεί τις συχνότητες που

είναι μεγαλύτερες από μια ορισμένη συχνότητα. Ένα υπεραπλοποιημένο φίλτρο παρουσιάζει την αντίθετη συμπεριφορά. Τέλος ένα ζωνοπερατό (*bandpass*) φίλτρο εξασθενεί τις συχνότητες που είναι μικρότερες (μεγαλύτερες) από ένα κάτω (άνω) όριο γύρω από μία κεντρική συχνότητα.

Ένα σημαντικό χαρακτηριστικό των κατηγοριών φίλτρων που αναφέραμε είναι η *συχνότητα αποκοπής* (*cutoff frequency*). Η συχνότητα αυτή ορίζεται σαν η συχνότητα εκείνη στην οποία το πλάτος του σήματος πέφτει στο 0.707 της μέγιστης τιμής της. Επειδή η ισχύς του σήματος είναι ανάλογη του τετραγώνου του πλάτους του, στη συχνότητα αποκοπής η ισχύς του σήματος θα είναι η μισή της μέγιστης τιμής της ($0.707^2 = 0.5$). Όσο πιο μικρή είναι η διαφορά μεταξύ της συχνότητας στην οποία αρχίζει η εξασθένηση του σήματος και της συχνότητας αποκοπής τόσο πιο απότομο είναι το φίλτρο.

Όσο πιο απότομο είναι το φίλτρο τόσο πιο μεγάλη σε διάρκεια είναι η κρουστική του απόκριση και αντίστροφα (βλ. Εικόνα 3.14).

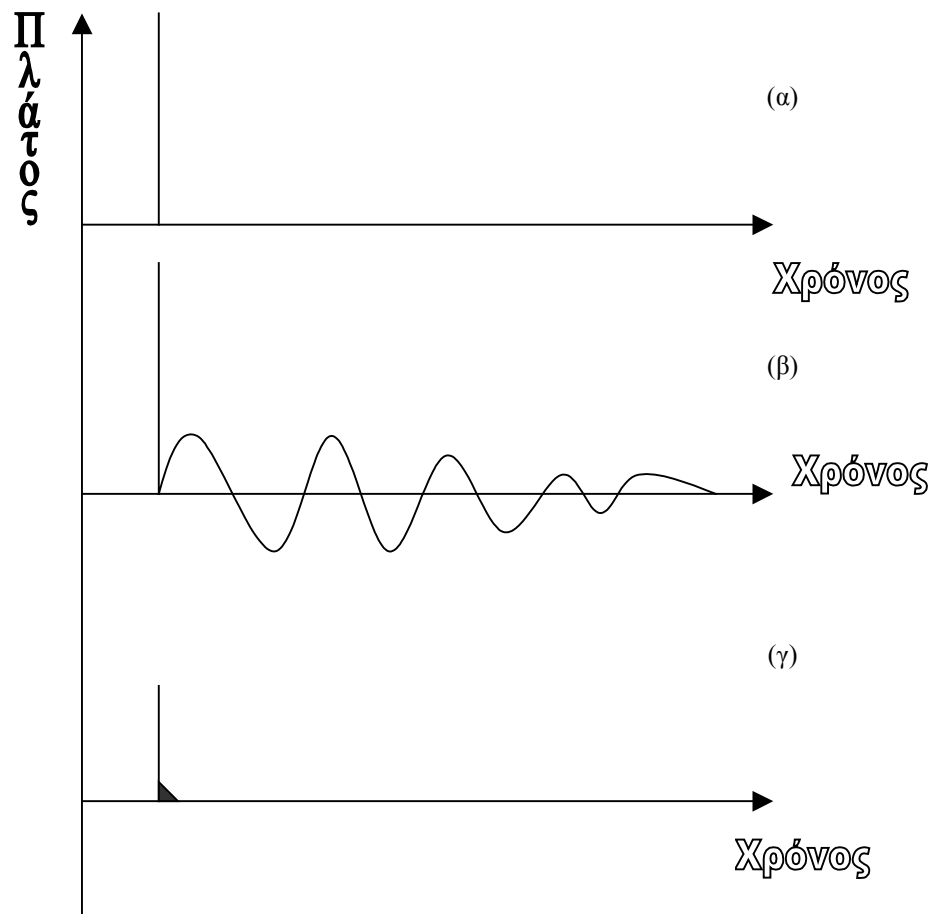
3.7.2.1 Λογισμικό Υλοποίησης Ψηφιακών Φίλτρων

Υπάρχουν δύο τρόποι με τους οποίους υλοποιείται ένα ψηφιακό φίλτρο:

1. Με την χρονική υστέρηση ενός αντιγράφου του σήματος εισόδου και τον συνδυασμό του καθυστερημένου σήματος με το καινούργιο σήμα εισόδου. Στην περίπτωση αυτή μιλάμε για φίλτρα *FIR* (*Finite Impulse Response*).
2. Με την καθυστέρηση ενός αντιγράφου του σήματος εξόδου και τον συνδυασμό του με το σήμα εισόδου. Στην περίπτωση αυτή μιλάμε για φίλτρα *IIR* (*Infinite Impulse Response*).

Η υλοποίηση των ψηφιακών φίλτρων βασίζεται στη χρήση εξισώσεων που περιγράφουν την επίδραση του φίλτρου στα δείγματα του σήματος εισόδου. Οι εξισώσεις αυτές μπορούν να παρασταθούν σε διαγράμματα στα οποία τα βέλη περιγράφουν τη φορά του σήματος, οι απλές αριθμημένες γραμμές περιγράφουν τους συντελεστές που χρησιμοποιούνται στην υλοποίηση του φίλτρου, οι τελείες περιγράφουν σημεία διακλάδωσης του σήματος, τα σύμβολα «+» και «x» περιγράφουν τις πράξεις πρόσθεσης και πολλαπλασιασμού ενώ το γράμμα Δ περιγράφει ένα κύκλωμα καθυστέρησης του σήματος εισόδου κατά μια περίοδο. Για παράδειγμα, η Εικόνα 3.15 περιγράφει την υλοποίηση του βαθυπερατού φίλτρου που χαρακτηρίζεται από την εξίσωση:

$$y[n] = 0.5 \times x[n] + 0.5 \times x[n-1]$$

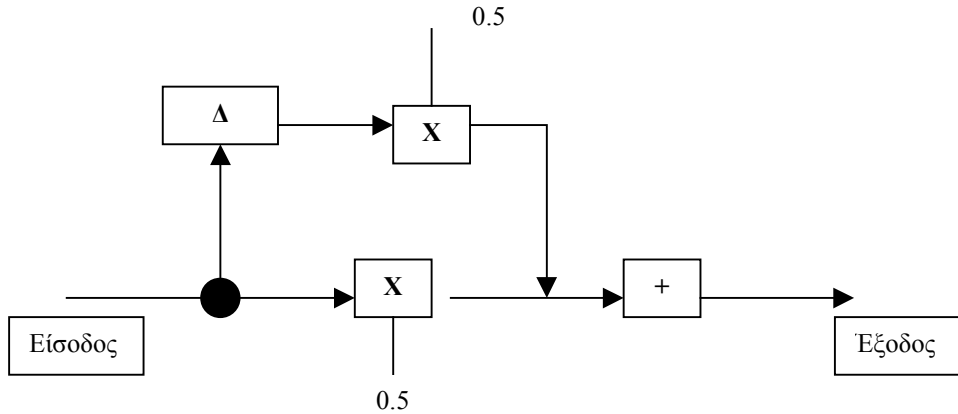


Εικόνα 3.14: Κρουστική απόκριση φίλτρων: (α) Κρουστική συνάρτηση εισόδου (β) Κρουστική απόκριση ζωνοπερατού φίλτρου (20 dB στα 300 Hz) με μικρό εύρος ζώνης (20 Hz) (γ) Κρουστική απόκριση βαθυπερατού φίλτρου (-15 dB στο 1 KHz).

3.7.2.2 Σύγκριση FIR και IIR φίλτρων

Η εφαρμογή ενός φίλτρου σε ένα σήμα μεταβάλλει πέρα από το πλάτος του και τη φάση του. Η εφαρμογή των FIR φίλτρων είναι γενικά προτιμότερη σε ηχητικά σήματα αφού είναι δυνατό να σχεδιαστούν τέτοια φίλτρα με γραμμική απόκριση φάσης. Μια τέτοια απόκριση προλαμβάνει πιθανές παραμορφώσεις φάσης που εισάγει το φιλτράρισμα ενός ήχου.

Επιπλέον, τα FIR φίλτρα είναι περισσότερο σταθερά από τα IIR μιας και δεν περιέχουν ανάδραση. Από την άλλη μεριά, η υλοποίηση των FIR φίλτρων απαιτεί περισσότερες αριθμητικές πράξεις και μνήμη από ότι η υλοποίηση FIR φίλτρων με αντίστοιχα χαρακτηριστικά.



Εικόνα 3.15: Παράδειγμα υλοποίησης βαθυπερατού φίλτρου.

3.7.3 Συνέλιξη

Η συνέλιξη αποτελεί μια από τις θεμελιώδεις λειτουργίες με τις οποίες υλοποιείται η επεξεργασία ενός ηχητικού σήματος. Βασικές μορφές επεξεργασίας ψηφιακού ήχου όπως είναι το φιλτράρισμα, η διαμόρφωση ή η αντήχηση στηρίζονται στη διαδικασία της συνέλιξης για να πετύχουν τα επιθυμητά αποτελέσματα. Για παράδειγμα, ένα γραμμικό, αμετάβλητο στο χρόνο φίλτρο συνελίσει την κρουστική του απόκριση με το σήμα εισόδου για να πετύχει τη δημιουργία ενός φιλτραρισμένου σήματος στην έξοδο.

Η πράξη της συνέλιξης παριστάνεται με το σύμβολο «*». Ο μαθηματικός ορισμός της συνέλιξης δύο πεπερασμένων ψηφιακών σημάτων a , b περιγράφεται από τον τύπο:

$$h[k] = \sum_{n=0}^{N-1} a[n]b[k-n]$$

όπου N είναι ο αριθμός των δειγμάτων του a και η μεταβλητή k λαμβάνει τιμές από 0 έως και L . Η τιμή του L υπολογίζεται από τον τύπο:

$$L = N + M - 2$$

όπου M ο αριθμός των δειγμάτων του σήματος b .

Όπως αναφέραμε και προηγουμένως η συνέλιξη χρησιμοποιείται για την υλοποίηση αρκετών μορφών επεξεργασίας ήχων. Για παράδειγμα, ένας απλός τρόπος με τον οποίο μπορούμε να προσομοιώσουμε την ακουστική εντύπωση και πιο συγκεκριμένα τις *αντηχήσεις* (*reverberations*) που δημιουργεί ένας χώρος (π.χ. ένα άδειο μεγάλο δωμάτιο ή ένα τούνελ) είναι να μετρήσουμε με κατάλληλα όργανα την κρουστική απόκριση του συγκεκριμένου χώρου. Στη συνέχεια μπορούμε να σχεδιάσουμε ένα ψηφιακό φίλτρο το οποίο θα έχει την ίδια κρουστική απόκριση με το χώρο μας. Αν συνελίξουμε ένα οποιοδήποτε ηχητικό σήμα με την κρουστική απόκριση του φίλτρου που κατασκευάσαμε το σήμα εξόδου θα αποκτήσει τα ακουστικά χαρακτηριστικά του χώρου που μας ενδιαφέρει!!

3.7.4 Μορφές Χρονικής Καθυστέρησης

Η χρονική καθυστέρηση (*delay*) ενός ηχητικού σήματος συνίσταται στην κατακράτηση στη μνήμη της ροής του σήματος και στη συνέχεια στη ανάμιξη της με το σήμα που λαμβάνεται από το σύστημα την παρούσα χρονική στιγμή. Μια τέτοια επεξεργασία δημιουργεί μια ποικιλία από ηχητικά εφέ.

Πιο συγκεκριμένα, η χρονική καθυστέρηση που εφαρμόζεται στο σήμα μπορεί να είναι *σταθερή* ή *μεταβλητή*. Στην πρώτη περίπτωση, όταν η διάρκεια της καθυστέρησης είναι μικρή (≤ 10 msec) δημιουργούνται αλλοιώσεις στο φάσμα συχνοτήτων του σήματος εισόδου ανάλογες με αυτές που προκαλούνται από την εφαρμογή ενός βαθυπερατού φίλτρου. Μια μέση καθυστέρηση (≥ 10 msec, ≤ 50 msec) προσθέτει «σώμα» στον ήχο δημιουργώντας την ψευδαίσθηση μίας υψηλότερης υποκειμενικής έντασης από αυτή που πραγματικά διαθέτει ένα ηχητικό σήμα. Τέλος μεγάλες καθυστερήσεις (≥ 50 msec) δημιουργούν το φαινόμενο της ηχούς στο σήμα εισόδου.

Στη φύση η ηχώ δημιουργείται όταν τα ηχητικά σήματα μεταδίδονται από την πηγή τους, ανακλώνται σε μια επιφάνεια και γίνονται αντιληπτά από τον ακροατή αρκετά μετά από τον ήχο που ακολούθησε την ευθεία οδό διάδοσης προς τον ακροατή. Επειδή η ταχύτητα διάδοσης του ήχου στον αέρα είναι περίπου 344 μέτρα το δευτερόλεπτο στους 20°C, μια καθυστέρηση άφιξης του σήματος στον ακροατή ίση με 1 msec αντιστοιχεί σε απόσταση διάδοσης 30 cm. Για να δημιουργηθεί το φαινόμενο της ηχούς θα πρέπει να υπάρξει καθυστέρηση άφιξης του σήματος κατά τουλάχιστον 50 msec. Η συγκεκριμένη καθυστέρηση αντιστοιχεί σε μια απόσταση

περίπου 8 μέτρων από την επιφάνεια ανάκλασης του ηχητικού σήματος ή σε διαδρομή περίπου 16 μέτρων του ηχητικού σήματος από την πηγή του μέχρι τον ακροατή μέσω της επιφάνειας ανάκλασης. Όταν οι αποστάσεις που περιγράψαμε είναι μικρότερες τότε εμφανίζονται φαινόμενα *αντήχησης (reverberation)* κατά την ακρόαση του ήχου.

Η μεταβλητή καθυστέρηση ενός ηχητικού σήματος χρησιμοποιείται για την υλοποίηση τριών εφφέ: *flanging*, *phasing* και *chorus*.

3.8 Ευαισθησία

Το αυτί μας είναι τρομερά ικανό να διακρίνει μεταξύ ενός συνόλου ηχητικών σημάτων που δέχεται συγχρόνως όπως και να αντιλαμβάνεται πολύ μικρές διαφορές στη λήψη του ίδιου σήματος. Για παράδειγμα το αυτί μας είναι ικανό να παρακολουθεί δύο ή περισσότερες συζητήσεις συγχρόνως. Επίσης είναι ικανό να αντιλαμβάνεται διακοπές στην παραγωγή ενός ηχητικού σήματος που αρχίζουν από τα 40 msec. Τέλος το αυτί μας αντιλαμβάνεται παροδικές παραμορφώσεις συχνότητας που αρχίζουν από το 10% της εκπεμπόμενης συχνότητας. Η εκπληκτική διακριτική ικανότητα του αυτιού μας μπορεί να γίνει κατανοητή και από την ευκολία με την οποία αντιλαμβανόμαστε ανεπαίσθητους και ασυνήθιστους ήχους σε οικείους χώρους.

3.9 Τριδιάστατος ήχος

Ο προσδιορισμός της κατεύθυνσης από την οποία εκπέμπεται ένας ήχος στηρίζεται στην εκτίμηση τριών μεγεθών που καθορίζουν τη θέση της ηχητικής πηγής ως προς τη θέση του ακροατή: του *αζιμουθίου (azimuth)* δηλ. της οριζόντιας γωνίας που σχηματίζει η ηχητική πηγή με τη θέση του ακροατή, της *απόστασης* ή της *ταχύτητας* για σταθερές ή κινούμενες πηγές αντίστοιχα και του *ζενίθ (zenith)* δηλ. του ύψους ή της κάθετης γωνίας που σχηματίζει η πηγή με τον ακροατή.

Η εκτίμηση του αζιμουθίου μίας ηχητικής πηγής βασίζεται στα ακόλουθα χαρακτηριστικά της ακρόασης:

- Στο διαφορετικό χρόνο άφιξης ενός ήχου σε καθένα από τα αυτιά μας όταν ο ήχος έρχεται από μια πλευρά.

- Στη διαφορά στο πλάτος των υψηλών συχνοτήτων που γίνονται αντιληπτές σε καθένα από τα αυτιά μας. Η διαφορά οφείλεται στην απορρόφηση ορισμένων υψηλών συχνοτήτων από το κρανίο του ακροατή.
- Στις επιπτώσεις που έχει στο φάσμα συχνοτήτων του ήχου οι ασύμμετρες ανακλάσεις του ηχητικού σήματος στην εξωτερική πλευρά των αυτιών, των ώμων και του κορμού του ακροατή.

Η εκτίμηση της απόστασης στηρίζεται στα παρακάτω χαρακτηριστικά της ακρόασης:

- Στο λόγο της έντασης του άμεσου σήματος προς την ένταση του σήματος που προέρχεται από αντηχήσεις, όταν η ένταση του άμεσου σήματος μειώνεται ανάλογα με το τετράγωνο της απόστασης της πηγής από τον ακροατή.
- Στην απώλεια των συνιστωσών που περιέχουν υψηλές συχνότητες σε ένα ηχητικό σήμα λόγω της απόστασης της ηχητικής πηγής από τον ακροατή.
- Στην απώλεια των λεπτομερειών του ηχητικού σήματος (δηλ. των συχνοτήτων χαμηλής έντασης) λόγω της απόστασης της ηχητικής πηγής από τον ακροατή.

Όταν η απόσταση μεταξύ της ηχητικής πηγής και του παρατηρητή μεταβάλλεται, η εκτίμηση της ταχύτητας και της κατεύθυνσης με την οποία γίνεται η μεταβολή στηρίζεται στη μεταβολή του ύψους του ηχητικού σήματος λόγω του φαινομένου *Doppler* (βλ. § 3.9.2).

Τέλος, η εκτίμηση του ζενίθ προκύπτει από τη μεταβολή του φάσματος συχνοτήτων του σήματος λόγω της ανάκλασης του στους ώμους και στο εξωτερικό τμήμα των αυτιών του ακροατή.

Συστήματα *οικιακού κινηματογράφου (home cinema)* ή εφαρμογές εικονικής πραγματικότητας εκμεταλλεύονται τα χαρακτηριστικά της ανθρώπινης ακοής που περιγράψαμε για να δημιουργήσουν τη ψευδαίσθηση του τριδιάστατου ήχου. Η αληθοφάνεια μιας τέτοιας διεργασίας βελτιώνεται και με την εξομοίωση των ακουστικών χαρακτηριστικών του εικονικού χώρου από τον οποίο εκπέμπεται ο ήχος όπως λ.χ. με τον υπολογισμό των διαφόρων ανακλάσεων που θα έχει ένα ηχητικό σήμα στα διάφορα αντικείμενα σε έναν εικονικό κόσμο.

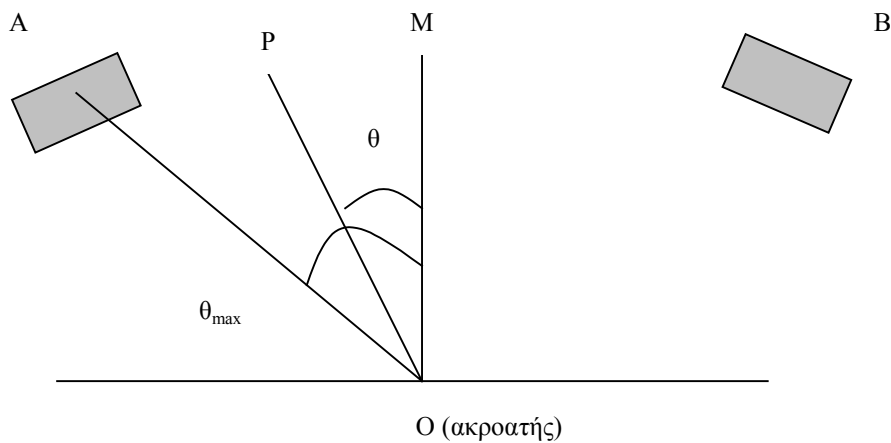
3.9.1 Προσομοίωση Αζιμουθίου

Η προσομοίωση του αζιμουθίου μίας ηχητικής πηγής μπορεί να γίνει με τη χρήση τουλάχιστον δύο ηχείων A, B και τη μεταβολή της έντασης του ήχου που παράγει καθένα από αυτά. Για παράδειγμα ο ακροατής έχει την αίσθηση ότι ένας ήχος προέρχεται από τη θέση στην οποία βρίσκεται το A όταν η ένταση του A έχει μη μηδενική τιμή και η ένταση του B είναι ίση με το μηδέν. Η μεταβολή της οριζόντιας θέσης του ήχου γίνεται δυνατή με την εφαρμογή της μεθόδου της *οριζόντιας μετατόπισης (panning)*.

Πιο συγκεκριμένα, αν υποθέσουμε ότι ο ακροατής και τα δύο ηχεία A, B έχουν τις θέσεις που περιγράφονται στην Εικόνα 3.16 τότε για να τοποθετηθεί νοητικά μια ηχητική πηγή σε ένα σημείο P μεταξύ των ηχείων A, B θα πρέπει να υπολογιστεί η γωνία που σχηματίζει η ευθεία OP που συνδέει τον ακροατή με το P με τη μεσοκάθετο OM στο ευθύγραμμο τμήμα AB. Στην περίπτωση αυτή η ένταση του ήχου σε καθένα από τα ηχεία A, B θα πρέπει να πολλαπλασιαστεί με τους συντελεστές E_A, E_B που προκύπτουν από τους τύπους:

$$E_A = \frac{\sqrt{2}}{2} (\cos(\vartheta) + \sin(\vartheta))$$

$$E_B = \frac{\sqrt{2}}{2} (\cos(\vartheta) - \sin(\vartheta))$$



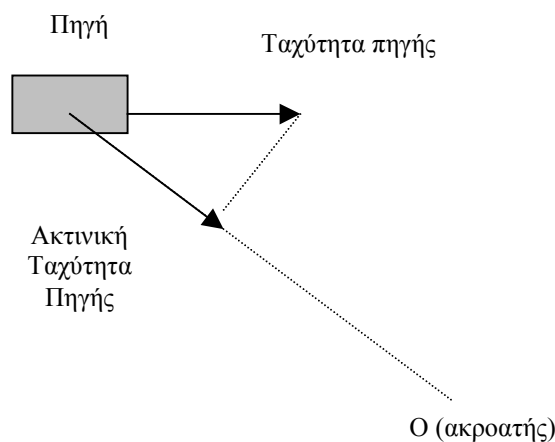
Εικόνα 3.16: Προσομοίωση αζιμουθίου με δύο ηχεία.

3.9.2 Φαινόμενο Doppler

Το φαινόμενο Doppler περιγράφει τη μεταβολή του ύψους του ήχου που παράγεται από μια ηχητική πηγή που κινείται σε σχέση με τη θέση του ακροατή. Για παράδειγμα, όταν στεκόμαστε στην άκρη του δρόμου, ο ήχος της μηχανής ενός αυτοκινήτου που κινείται προς έναν ακροατή με μεγάλη ταχύτητα γίνεται όλο και πιο οξύς όσο το αυτοκίνητο πλησιάζει τον ακροατή, ενώ το ύψος του συνεχώς ελαττώνεται όσο το αυτοκίνητο απομακρύνεται από τον ακροατή. Το ύψος Y του ήχου που γίνεται αντιληπτό στην περίπτωση αυτή είναι ανάλογο της ακτινικής ταχύτητας της πηγής προς τον παρατηρητή και δίνεται από τον τύπο:

$$Y = f \frac{vs}{vs - vr}$$

όπου f η πραγματική συχνότητα του ηχητικού σήματος, vs η ταχύτητα του ήχου και vr η ακτινική ταχύτητα της πηγής σε σχέση με τον ακροατή (βλ. Εικόνα 3.17). Η φυσική εξήγηση του φαινομένου στηρίζεται στην παρατήρηση ότι η απόσταση μεταξύ των ηχητικών κυμάτων που προέρχονται από την πηγή μειώνεται όσο η πηγή πλησιάζει στον παρατηρητή, επομένως η φαινόμενη συχνότητα του ήχου αυξάνεται. Το αντίθετο συμβαίνει όταν η πηγή απομακρύνεται από τον παρατηρητή.



Εικόνα 3.17: Φαινόμενο Doppler.

3.9.3 Προσομοίωση Ζενίθ

Η προσομοίωση του ζενίθ που αντιστοιχεί σε μια ηχητική πηγή επιτυγχάνεται με την εφαρμογή φίλτρων στο ηχητικό σήμα τα οποία μεταβάλλουν το φάσμα συχνοτήτων του σήματος κατά τρόπο ανάλογο των μεταβολών που επιφέρουν οι ανακλάσεις του ήχου στο κεφάλι και στους ώμους του ακροατή. Οι μεταβολές αυτές διαφέρουν από

ακροατή σε ακροατή καθώς εξαρτώνται από τις διαστάσεις των κεφαλιών και των ώμων τους. Η απόκριση συχνότητας των συγκεκριμένων φίλτρων ονομάζεται *συνάρτηση μεταφοράς κεφαλής* (*head-related transfer function* ή *HRTF*) και για την επίτευξη ενός ικανοποιητικού βαθμού πιστότητας θα πρέπει να αντιστοιχεί στα χαρακτηριστικά του συγκεκριμένου ακροατή.

3.10 Μέθοδοι Συμπίεσης Ηχητικών Σημάτων

Οι περισσότερες μέθοδοι συμπίεσης ηχητικών σημάτων ανήκουν στις τεχνικές με απώλειες. Ο κύριος λόγος για τον οποίο οι συγκεκριμένες τεχνικές είναι τόσο δημοφιλείς οφείλεται στο γεγονός ότι σε ένα ηχητικό σήμα συνήθως δεν υπάρχουν *ομοιότυπα* (*patterns*) με ιδιαίτερα μεγάλη συχνότητα εμφάνισης. Πιο συγκεκριμένα, η ψηφιοποίηση οποιουδήποτε αναλογικού σήματος εισάγει παραμορφώσεις που εντοπίζονται κυρίως στα λιγότερο σημαντικά δυαδικά ψηφία του κάθε δείγματος (βλ. § 3.6.2). Ο τυχαίος χαρακτήρας των συγκεκριμένων παραμορφώσεων εμποδίζει την εμφάνιση ομοιοτύπων στη συγκεκριμένη κατηγορία σημάτων.

Ένας επιπλέον λόγος για την περιορισμένη εφαρμογή των τεχνικών χωρίς απώλειες οφείλεται στην ανομοιόμορφη συμπίεση που επιτυγχάνουν σε ένα ηχητικό σήμα. Ο περιορισμός αυτός εμποδίζει τη χρησιμοποίηση των συγκεκριμένων τεχνικών από εφαρμογές που στηρίζονται στην επίτευξη ενός σταθερού ρυθμού μετάδοσης των δειγμάτων ενός ηχητικού σήματος (π.χ. ραδιόφωνο στο Διαδίκτυο). Για παράδειγμα, έστω ότι μια ραδιοφωνική εκπομπή λαμβάνεται από έναν ακροατή στο Διαδίκτυο μέσω ενός modem που μπορεί να μεταφέρει περίπου 3000 bits/sec. Στην περίπτωση αυτή η εφαρμογή θα πρέπει να διασφαλίσει ότι ένα δευτερόλεπτο ήχου θα πρέπει πάντα να χωρέσει σε 3000 δυαδικά ψηφία. Δυστυχώς οι τεχνικές χωρίς απώλειες δεν μπορούν να πετύχουν έναν τέτοιο σταθερό ρυθμό συμπίεσης καθόλη τη διάρκεια της μετάδοσης. Αντίθετα, η πλειονότητα των τεχνικών με απώλειες πετυχαίνει σταθερούς ρυθμούς συμπίεσης. Για παράδειγμα η μέθοδος *IMA ADPCM* που θα περιγράψουμε παρακάτω πετυχαίνει πάντα λόγο συμπίεσης 4:1 σε ψηφιοποιημένα ηχητικά σήματα με μήκος δείγματος 16-bit.

Οι κυριότερες μέθοδοι συμπίεσης σε ηχητικά σήματα εφαρμόζονται σε μια συγκεκριμένη αναπαράσταση του σήματος που είναι γνωστή ως *παλμοκωδική διαμόρφωση* (*Pulse Code Modulation* ή *PCM*). Στη συγκεκριμένη μέθοδο κάθε δείγμα αναπαριστάται με ένα σύνολο παλμών που αντιστοιχούν στο δυαδικό κώδικα που αντιστοιχεί στην τιμή του δείγματος. Η πιστότητα του σήματος που προκύπτει είναι

συνάρτηση του δυναμικού εύρους του δυαδικού κώδικα. Για παράδειγμα, ένα δυναμικό εύρος 128 τιμών θα έχει ως αποτέλεσμα την απάλειψη ήχων στο ψηφιοποιημένο σήμα που έχουν ένταση ίση ή μικρότερη από το $1/128$ της έντασης του δυνατότερου ήχου που μπορεί να αναπαρασταθεί από το σύστημα (βλ. § 3.6.2). Προφανώς μια τέτοια αναπαράσταση εισάγει προβλήματα πιστότητας σε ασθενή σε ένταση ηχητικά σήματα. Για την αντιμετώπιση του συγκεκριμένου προβλήματος είναι δυνατόν να χρησιμοποιηθούν μη γραμμικές μέθοδοι παλμοκωδικής διαμόρφωσης (*non-linear PCM*). Για παράδειγμα, αν το εύρος τιμών του κώδικα σε μια τέτοια μέθοδο είναι 128 τότε μια τιμή δείγματος ίση με 1 μπορεί να μην αντιστοιχεί σε ήχο έντασης ίση με το $1/128$ της έντασης του δυνατότερου ήχου αλλά σε ήχο αρκετά μικρότερης έντασης. Μια τέτοια μέθοδος είναι μη γραμμική καθώς χρησιμοποιεί περισσότερα δυαδικά ψηφία για την αναπαράσταση ασθενών ήχων στους οποίους όπως είδαμε υπάρχουν μεγαλύτερα προβλήματα διακριτικότητας και λιγότερα για τους ήχους μεγάλης έντασης στους οποίους η ευαισθησία της ανθρώπινης ακοής σε σχετικές διαφορές έντασης είναι σημαντικά μειωμένα.

Οι πλέον διαδεδομένες μέθοδοι συμπίεσης ηχητικών σημάτων είναι τεχνικές κωδικοποίησης πηγής και ειδικότερα τεχνικές διαφορικής κβαντοποίησης όπως οι οικογένειες μεθόδων *DPCM* (Differential Pulse Code Modulation), *ADPCM* (Adaptive Differential Pulse Code Modulation) και οι τεχνικές *κωδικοποίησης υποζώνης* (*subband coding*). Πιο συγκεκριμένα, η οικογένεια των *DPCM* τεχνικών αποθηκεύουν τις διαφορές μεταξύ διαδοχικών τιμών του σήματος και όχι τις απόλυτες τιμές των δειγμάτων. Το ίδιο συμβαίνει και στις τεχνικές *ADPCM* με τη διαφορά ότι στις τελευταίες η κλίμακα αναπαράστασης των διαφορών είναι μεταβλητή και προσαρμόζεται σε μία πρόβλεψη του απόλυτου μεγέθους των τιμών των δειγμάτων που συγκρίνονται με βάση τις προηγούμενες τιμές των δειγμάτων αυτών. Τα πρότυπα συμπίεσης *WAVE* της Microsoft και *AIFF-C* της Apple αποτελούν παραδείγματα *ADPCM*² μεθόδων.

Ενα σημαντικό μειονέκτημα των τεχνικών *DPCM* και *ADPCM* οφείλεται στην εξάρτηση του μεγέθους των διαφορών μεταξύ διαδοχικών δειγμάτων ενός ηχητικού σήματος από τη συχνότητα του σήματος. Όπως είναι αναμενόμενο, οι διαφορές μεταξύ διαδοχικών δειγμάτων σε ήχους χαμηλών συχνοτήτων είναι σημαντικά μικρότερες από τις αντίστοιχες διαφορές σε ήχους υψηλών συχνοτήτων. Επομένως η διακριτική ικανότητα της κωδικοποίησης των διαφορών θα πρέπει να προσαρμόζεται

² Ακριβέστερα *IMA ADPCM* δηλ. μια ειδική κατηγορία *ADPCM* μεθόδων.

στο συχνοτικό περιεχόμενο του σήματος κάθε χρονική στιγμή κάτι που δε γίνεται στις δύο τεχνικές που περιγράψαμε. Αντίθετα, οι τεχνικές κωδικοποίησης υποζώνης υποδιαιρούν το ηχητικό σήμα σε δύο ή περισσότερες ζώνες συχνοτήτων και συμπιέζουν καθεμία από αυτές ξεχωριστά. Η χρησιμοποίηση μίας τέτοιας υποδιαίρεσης εκμεταλλεύεται τα φυσιολογικά χαρακτηριστικά της ανθρώπινης ακοής η οποία παρουσιάζει τη μέγιστη ευαισθησία στο εύρος συχνοτήτων μεταξύ 2700-3200 Hz με την ευαισθησία να μειώνεται όσο απομακρυνόμαστε από τη ζώνη αυτή. Κατά συνέπεια μια μέθοδος κωδικοποίησης υποζώνης μπορεί να συμπιέζει δραστικά ήχους με συχνοτικό περιεχόμενο που απέχει αρκετά από τη ζώνη που αναφέραμε και να εφαρμόζει την ελάχιστη συμπίεση σε συχνότητες που ανήκουν στη ζώνη αυτή. Τα αποτελέσματα μίας τέτοιας διαδικασίας δε θα γίνονται εύκολα αντιληπτά από την ανθρώπινη ακοή. Στην κατηγορία των μεθόδων κωδικοποίησης υποζώνης ανήκουν τα πρότυπα συμπίεσης *MPEG audio*, *Dolby AC-2* και *AC-3*, το σύστημα *Sony MiniDisc* και το *RealAudio*.

