

Ιόνιο Πανεπιστήμιο.

Τμήμα Μουσικών Σπουδών.

Μεταπτυχιακό πρόγραμμα σπουδών στις Τέχνες και Τεχνολογίες του Ήχου.

**Ανάπτυξη αλγορίθμων επεξεργασίας σήματος
πραγματικού χρόνου για αναπαραγωγή και ηχογράφηση
τρειςδιάστατων ηχητικών πεδίων.**

Φοιτητής:

Γεώργιος Μαραγκουδάκης

Επιβλέπων καθηγητής:

Ανδρέας Φλώρος

Βαθμολογητές:

Παναγιώτης Βλάμος

Ανδρέας Μιέστρης

Ανδρέας Φλώρος

2007



Ιόνιο Πανεπιστήμιο.
Τμήμα Μουσικών Σπουδών.
Μεταπτυχιακό πρόγραμμα σπουδών στις Τέχνες και Τεχνολογίες του Ήχου.

Τίτλος Διπλωματικής:
Ανάπτυξη αλγορίθμων επεξεργασίας σήματος πραγματικού χρόνου για αναπαραγωγή και ηχογράφηση τρισδιάστατων ηχητικών πεδίων.

Φοιτητής: Γεώργιος Μαραγκουδάκης
Επιβλέπων καθηγητής: Ανδρέας Φλώρος

Βαθμολογητές: Παναγιώτης Βλάμος
Ανδρέας Μιέστρης
Ανδρέας Φλώρος

2007

Κατάλογος περιεχομένων

1 Εισαγωγή.....	11
2 Θεωρία.....	15
2.1. Τι είναι τρισδιάστατος ήχος και με ποιο τρόπο τον αντιλαμβάνεται ο άνθρωπος.....	16
2.1.1 Εντοπισμός της γωνίας προέλευσης του ήχου.....	16
Κριτήρια που βασίζονται στην διαφορά χρόνου πρόσπτωσης και έντασης του ήχου μεταξύ των δύο αυτιών.....	16
Πλευρική θέση των αυτιών (lateralization).....	16
Το φαινόμενο της προπορείας (Precedence ή Hass effect).....	19
Κριτήρια που βασίζονται στην κίνηση του κεφαλιού ή στην κίνηση της πηγής του ήχου.....	20
Κίνηση του κεφαλιού.....	20
Κίνηση της πηγής του ήχου.....	20
Κριτήρια που βασίζονται στις φασματικές τροποποιήσεις που προκαλούνται από την κόγχη.....	21
Ασαφή κριτήρια ITD και IID.....	21
Οι Head Related Transfer Function.....	21
Σπουδαία χαρακτηριστικά των HRTF.....	22
Φασικά χαρακτηριστικά των HRTF.....	22
Εντοπισμός βάση των HRTF.....	22
2.1.2 Υπολογισμός της απόστασης.....	23
Στοιχεία έντασης και ηχηρότητας.....	23
Επίδραση της οικειότητας και της πρόβλεψης.....	24
Φασματικά και Binaural στοιχεία για την απόσταση.....	24
Φασματικές αλλαγές και απόστασης της ηχητικής πηγής (μέτωπο κύματος).....	25
Εντοπισμός του ηχητικού πεδίου μέσα στο κεφάλι του ακροατή (Inside the head localization).....	26
Αντηχήσεις.....	26
2.2. Συστήματα αναπαραγωγής Surround ήχου.....	27
2.2.1 Τρία κανάλια στερεοφωνικού ήχου (3-0).....	28
2.2.2 Surround τεσσάρων καναλιών (3-1 stereo).....	29
Διαμόρφωση των ηχείων.....	29
Περιορισμοί της τετρακάναλης αναπαραγωγής.....	30
2.2.3 5.1 κανάλια surround (3-2 stereo).....	31
Παγκόσμια πρότυπα και σχετική διάταξη ηχείων.....	31
Το LFE κανάλι και η χρήση των Μονάδα αναπαραγωγής χαμηλών συχνοτήτων (subwoofer).....	33
Περιορισμοί της αναπαραγωγής των 5.1 καναλιών.....	33
Ένταση σημάτων στο πρωτόκολλο 5.1 surround.....	33
2.2.4 Πρωτόκολλο 7.1 καναλιών surround.....	34
2.2.5 10.2 κανάλια surround.....	34
2.2.6 Ψηφιακά πρωτόκολλα μεταφοράς πολυκάναλου ήχου.....	35
Dolby Digital.....	35

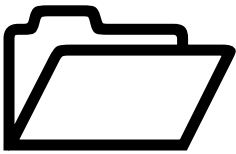
2.2.7 DTS.....	37
SDDS 37	
MPEG surround modes.....	37
Meridian Lossless Packing (MLP).....	38
2.2.8 Αμφιωτική ακουστική (Binaural).....	39
Περίληψη.....	41
2.2.9 Αμφιωτική ακουστική με χρήση μεγαφώνων (Transaural).	42
Περίληψη.....	44
2.2.10 Wave Field Synthesis (WFS), Wave Field Analysis (WFA)	46
Περίληψη.....	48
2.2.11 Ambiophonics.....	49
2.3. Ambisonic.....	50
2.3.1 Φυσική ερμηνεία της μεθόδου Ambisonic.	50
2.3.2 Ambisonic Format	56
A-Format ,ηχογράφιση ηχητικού πεδίου.	56
B-Format.....	58
D-Format.....	58
UHJ-Format ή C-Format.....	58
G-Format.....	59
2.3.3 Πλεονεκτήματα.	59
2.3.4 Μειονεκτήματα.	60
2.3.5 Κωδικοποιητής (Encoder) Ambisonic B Format	60
Encoder B-Format από A-Format (υπάρχον ηχητικό πεδίο).	60
Encoder B-Format από εικονικό ηχητικό πεδίο (σύνθεση).....	63
2.3.6 Ambisonic αποκωδικοποιητής (Decoder).	64
2.3.7 Κατηγορίες ψηφιακών φίλτρων.	70
2.3.8 Κβαντισμός.	71
3Υλοποίηση.....	77
3.1. Encoder B-Format.....	78
3.1.1 Encoder B-Format από A-Format (υπάρχον ηχητικό πεδίο).....	78
Τεκμηρίωση κώδικα Matlab για κωδικοποιητή A-Format ambisonic σε B-Format.....	79
Τεκμηρίωση κώδικα Matlab για την συνάρτηση κβαντοποίησης.....	79
Τεκμηρίωση κώδικα Matlab για την συνάρτηση δημιουργίας θορύβου dither.....	81
3.1.2 Encoder B-Format από εικονικό ηχητικό πεδίο (σύνθεση).....	82
Τεκμηρίωση κώδικα Matlab για την συνάρτηση κωδικοποίησης εικονικού ηχητικού πεδίου σε B-Format Ambisonics.	83
3.2. Ambisonic αποκωδικοποιητής (Decoder).....	86
3.2.1 Δημιουργία συνάρτηση υπολογισμού της κρουστικής απόκρουσης για τα φίλτρα Highpass και Lowpass και τρόπος φιλτραρίσματος σήματος.....	87
3.2.2 Τεκμηρίωση κώδικα Matlab για την δημιουργία του αποκωδικοποιητή B-Format..	89
3.2.3 Τεκμηρίωση κώδικα Matlab για τον σχηματισμό του πίνακα συντελεστών για τη δημιουργία των σημάτων που θα τροφοδοτήσουν τα ηχεία, που θα χρησιμοποιηθούν στη συνέχεια στην αποκωδικοποίηση.....	91
4Αποτιμηση και παρατηρησεις.....	95
5Συμπερασματα.....	103
Βιβλιογραφία.....	105
Ευρετήριο λέξεων.....	107

Παράρτημα 1 Πίνακας 8. Σφαιρικές αρμονικές μέχρι την 3η τάξη.....	113
Παράρτημα 2 Ορισμός και σύμβολα για σφαιρικές συντεταγμένες.....	117
Παράρτημα 3 Η εξίσωση που δίνει το πολικό διάγραμμα οποιαδήποτε μικροφώνου στον χώρο.	119
Παράρτημα 4 Συνάρτηση υπολογισμού του σήματος που λαμβάνει ένα μικρόφωνο.....	121
Παράρτημα 5 Κώδικας Matlab για σχεδιασμό σφαιρικών αρμονικών μέχρι 3ης τάξης.....	123
Παράρτημα 6 Απόδειξη για γωνίες προσανατολισμού των καψών του Soundfield μικροφώνου.	129
Παράρτημα 7 Κώδικας Matlab για δημιουργία κβαντοποιητή με δυνατότητα επιλογής του τύπου κβαντοποίησης, dither , τεχνικών noise shape. Υποστηρίζει πολυκάναλα σήματα εισόδου.....	137
Παράρτημα 8 Κώδικας Matlab για δημιουργία dither με δυνατότητα επιλογής ανάμεσα σε RPDF, TPDF και Highpass TPDF.....	143
Παράρτημα 9 Κώδικας Matlab για κωδικοποίηση του σήματος A-Format Ambisonic σε B-Format.	147
Παράρτημα 10 Κώδικας Matlab για σύνθεση του σήματος B-Format Ambisonic.....	149
Παράρτημα 11 Κώδικας Matlab για αποκωδικοποίηση ηχητικού πεδίου σε διάταξη Ambisonics με δυνατότητα κωδικοποίησης και για Binaural.....	155
Παράρτημα 12 Κώδικας Matlab για τον σχηματισμό του πίνακα των συντελεστών, για την δημιουργία των σημάτων που πρέπει να τροφοδοτηθούν σε κάθε ηχείο, κατά την διαδικασία της αποκωδικοποίησης.....	161

ΠΡΟΛΟΓΟΣ

Αυτή η εργασία δημιουργήθηκε στα πλαίσια του μεταπτυχιακού προγράμματος 'Τέχνες και Τεχνολογίες του Ήχου' του τμήματος Μουσικών σπουδών του Ιονίου Πανεπιστημίου. Η εκπόνηση αυτής πραγματοποιήθηκε την χρονική περίοδο 2006-2007. Σε αυτή την εργασία έχει γίνει μια αναφορά στα πιο γνωστά συστήματα αναπαραγωγής Surround και τρισδιάστατου ήχου, και στην λειτουργία του συστήματος ακοής του ανθρώπου. Τελικά επιλέχθηκε ως καταλληλότερη μέθοδος για την κωδικοποίηση και αποκωδικοποίηση ηχητικών πεδίων, στις τρεις διαστάσεις, το σύστημα Ambisonic, του οποίου γίνεται εκτενής αναφορά των προτύπων (Format) αυτού, της φυσικής του ερμηνείας και των μεθόδων αποσύνθεσης και ανασύνθεσης των ηχητικών πεδίων. Η υλοποίηση του συστήματος έγινε σε περιβάλλον Matlab, και η αποτίμηση αυτού έγινε με προσομοίωση πολυκάναλων συστημάτων αναπαραγωγής με χρήση της μεθόδου Binaural. Στη εργασία συμπεριλαμβάνονται οι κώδικες του Matlab που δημιουργήθηκαν καθώς και εξήγηση αυτών. Η πρωτοτυπία αυτού του συστήματος είναι ότι η αναπαραγωγή του συντεθειμένου ηχητικού πεδίου μπορεί να γίνει μέσω Ambisonic συστήματος με ηχεία, αλλά και από ακουστικά με την μέθοδο Binaural. Επίσης συμπεριλαμβάνονται αποδείξεις για αρκετά θεωρήματα ώστε να γίνει κατανοητή η θεωρία.

Ευχαριστώ τον επιβλέπων καθηγητή της εργασίας κ. Ανδρέα Φλώρο για τις πολύτιμες συμβουλές, βιβλιογραφία, διορθώσεις και υποδείξεις που μου έκανε για την πραγματοποίηση αυτής της εργασίας. Ευχαριστώ επίσης τον κ. Ανδρέα Μνιέστηρη για την υπόδειξη βιβλιογραφίας καθώς και τον κ. Ιωάννη Ζάννο για την υπόδειξη μεθοδολογίας υλοποίησης της εργασίας.



ΕΙΣΑΓΩΓΗ

Πριν γίνει ανάλυση των συστημάτων αναπαραγωγής surround ήχου και των πλεονεκτημάτων και μειονεκτημάτων που έχει το καθένα από αυτά, καλό θα ήταν να γίνει μια ιστορική αναδρομή, ώστε να γίνει εμφανής η εξελικτική πορεία τους (Βλέπε Rumsey [28]).

Η αναπαραγωγή και ηχογράφηση του ήχου ξεκίνησε μονοφωνικά. Στη συνέχεια έγινε αντιληπτό πως αυτό δεν ήταν αρκετό για την ακριβή απεικόνιση των ηχητικών πεδίων. Έτσι αναπτύχθηκαν σταδιακά διάφορα πρότυπα αναπαραγωγής (όπως Stereo και πολυκάναλος ήχος) ώστε να μπορούν να ηχογραφηθούν και να αναπαραχθούν με μεγαλύτερη ακρίβεια τα ηχητικά πεδία. Το 1881 ο

Clement Adler πέτυχε την πρώτη γνωστή μετάδοση στερεοφωνικού ήχου. Τοποθετώντας τηλεφωνικούς δέκτες στα φώτα του προσκηνίου της Όπερας του Παρισιού, και στη συνέχεια αναμεταδίδοντας της λήψεις αυτών μέσω τηλεφωνικών δεκτών στην έκθεση, οι επισκέπτες μπορούσαν να ακούνε την όπερα σε 'πραγματικό χρόνο'. Παρά το γεγονός ότι έκανε την πρώτη αναμετάδοση στερεοφωνικού ήχου, δεν υπήρχαν ακόμα στερεοφωνικές αρχές για την ηχογράφηση και την αναπαραγωγή. Το 1930 δημιουργήθηκαν οι στερεοφωνικές αρχές, βάση των οποίων δημιουργούνταν εικονικές ηχητικές πηγές με την χρήση δύο ηχείων. Τέτοιες τεχνικές δημιουργήθηκαν την ίδια εποχή ανεξάρτητα από τους Alan Blumlein και τα Bell Labs. Οι προσεγγίσεις του ήταν τελείως διαφορετικές. Πρωτοποριακά στα Bell Labs, οι Snow και Steinberg δημιούργησαν ένα σύστημα που αποτελούνταν από τρία κανάλια για αναπαραγωγή σε ευρύ ακροατήριο. Παράλληλα ο Blumlein παρουσίαζε ότι η διαφοροποίηση στην ένταση μεταξύ των σημάτων που τροφοδοτούν τα δύο ηχεία, είναι δυνατό να δημιουργήσει φασικές διαφορές μεταξύ των αυτιών. Το σύστημά του προοριζόταν για την δημιουργία ηχητικών εικόνων σε περιορισμένο εύρος γωνιών και γι' αυτό το λόγο ήταν καταλληλότερο για οικιακή χρήση. Αυτά όμως βρίσκονταν σε ερευνητικό επίπεδο και χρησιμοποιούνταν κυρίως στον κινηματογράφο. Τα πρώτα καταναλωτικά συστήματα Stereo εμφανίστηκαν την δεκαετία του 1950 και χρησιμοποιούσαν μεθόδους που έμοιαζαν με του Blumlein. Σε αυτά μονοφωνικά ηχογραφημένα υλικά, μετά από επεξεργασία με τεχνικές όπως Combfiltering και Bandsplitting, δημιουργούνταν η αίσθηση της στερεοφωνίας. Το 1958, οι Clark, Dutton και Vanderlyn της EMI, αναβίωσαν τις θεωρίες του Blumlein, και συσχέτισαν με μεγαλύτερη λεπτομέρεια την σχέση μεταξύ της επιθυμητής γωνίας προέλευσης της πηγής και της αντιληπτής από τον ακροατή κατά την αναπαραγωγή, αλλάζοντας την ένταση του σήματος μεταξύ των ηχείων. Επιπλέον ισχυρίστηκαν ότι ο Blumlein δεν έλαβε όλα τα πλεονεκτήματα της Binaural αναπαραγωγής, ειδικότερα το φαινόμενο της προπορείας 'precedence effect', αλλά χρησιμοποίησε μόνο ελάχιστα από αυτά που αξιολογεί το σύστημα ακρόασης του ανθρώπου. Περίπου 40 χρόνια πριν είχε όμως ήδη επιτευχθεί ηχογράφηση των ηχητικών πεδίο και αναπαραγωγή αυτών με μεγάλη ακρίβεια. Έτσι, από το 1920, η επιτυχία της Binaural μεθόδου αναπαραγωγής και καταγραφής των τρισδιάστατων ηχητικών πεδίων, ώθησε στην καταγραφή όλων των χωρικών στοιχείων που αξιολογούνται από τον άνθρωπο για την θέση της πηγής στο χώρο. Στην ηχογράφηση με την μέθοδο Binaural Stereo, δύο μικρόφωνα είναι τοποθετημένων στο άνοιγμα των αυτιών (ή και πιο μέσα, πιο κοντά στο τύμπανο), ενός πραγματικού ανθρώπου ή και ενός dummy Head. Ο όρος Binaural συνήθως περιγράφει ηχητικά γεγονότα που έχουν ηχογραφηθεί καταγράφοντας τις φασικές και διαφορές έντασης μεταξύ του ήχου που φτάνει στα αυτιά του ακροατή. Την δεκαετία 1950-1960 εμφανίστηκαν πολυκάναλα πρότυπα Stereo για τον κινηματογράφο, τα οποία έγιναν δημοφιλή και χρησιμοποιούσαν τρία κανάλια μπροστά και ένα περιφερειακό. Τέτοια συστήματα ήταν τα 3.0 Surround Stereo, Quadraphonic, 4.1 Channel Surround (Dolby Prologic), 5.1 Channel Surround (Dolby Prologic II) και 6.1 Channel Surround (Dolby Prologic Iix). Νωρίς το 1970, η Dolby Stereo παρουσίασε ένα τρόπο να μπορούν να κωδικοποιηθούν τέσσερα κανάλια ήχου surround σε δύο οπτικά κανάλια ενός φιλμ 35 mm. Αυτά τα πρότυπα χρησιμοποιούνται ακόμα και σήμερα, αλλά ο ήχος στο σινεμά κινείται σε περισσότερο προς την χρήση ψηφιακών καναλιών τα οποία ενσωματώνουν 5 ή και 7 κανάλια surround ήχου συν ένα κανάλι για τις πολύ χαμηλές συχνότητες. Τέτοια συστήματα κωδικοποίησης με 'Low bit rate' είναι τα Dolby Digital, Sony SDDS, και το Digital Theater System (DTS). Αντίστοιχα δημιουργήθηκαν ψηφιακά συστήματα στα οποία γίνεται η αναπαραγωγή και είναι τα 5.1 Surround Sound (Dolby Digital, DTS), 6.1 Channel Surround (Dolby Digital EX, DTS ES), 7.1 Surround Sound (Dolby Digital plus, DTS HD), 10.2 Channel Surround (THX αντίστοιχο του Imax) και 22.2 Channel Surround (Ultra High Definition Video). Ένα άλλο πρωτόκολλο που δημιουργήθηκε και παρουσίαζε πολλές δυνατότητες είναι το Ambisonic. Δημιουργήθηκε το 1970

(κυρίως από τον M.Gerzon και τον P B Fellgett) με σκοπό την πλήρη καταγραφή του ηχητικού πεδίου, ώστε να είναι δυνατή η ανασύνθεσή του με όλα τα 'χωρικά κριτήρια' από ένα σύστημα αναπαραγωγής που διαθέτει τουλάχιστον τέσσερα ηχεία. Βασίζεται κυρίως στη θεωρία του Blumlein, προσθέτοντας και την τρίτη διάσταση, αυτή του ύψους. Το Ambisonic έχει ένα μεγάλο δυναμικό αλλά δεν έχει χρησιμοποιηθεί εμπορικά ακόμα για διάφορους τεχνικούς λόγους. Η αξιοπιστία και οι δυνατότητες που παρουσιάζει αυτή η μέθοδος αποτέλεσε και κινητήριο έναυσμα για την επιλογή αυτής της μεθόδου ως την κύρια αυτής της εργασίας.

Γενικά ο κόσμος συνήθισε την ακρόαση του ήχου σε Surround στο σινεμά και πλέον και την έλευση του DVD, του Dolby Digital και του DTS γίνονται όλο και περισσότερες οικιακές εγκαταστάσεις συστημάτων μπορούν να προσομοιάσουν τον ήχο του σινεμά. Όπως φαίνεται από το 1881 υπάρχει μια τάση για την βελτιστοποίηση της καταγραφής και της αναπαραγωγής των ηχητικών πεδίων. Γενικά με την ανάπτυξη και διεύρυνση καταναλωτικά εφαρμογών 'Εικονικής πραγματικότητας' μέσω παιχνιδιών στους υπολογιστές και άλλων συστημάτων όπως οι τηλεσυνεδριάσεις ή ακόμα και η καλύτερη καταγραφή αναμνήσεων, έχει δημιουργηθεί η ανάγκη για την τροφοδότηση και άλλων αισθήσεων, εκτός της όρασης. Έτσι στην περίπτωση της ακοής απαιτούνται πιο ακριβή στοιχεία ώστε το σύστημα αντίληψης του ανθρώπου να μην το διαχωρίζει τα εικονικά ηχητικά πεδία από τα πραγματικά.



ΘΕΩΡΙΑ

Σε αυτό το κεφάλαιο περιλαμβάνονται βασικές θεωρητικές γνώσεις για την κατανόηση της λειτουργίας του συστήματος ακρόασης του ανθρώπου, των συστήματα και μεθόδων αναπαραγωγής και καταγραφής του ήχου, τα πλεονεκτήματα και μειονεκτήματα που έχουν αυτές, καθώς και μια αναλυτική αναφορά στη μέθοδο Ambisonic. Η αναλυτική αναφορά στη μέθοδο Ambisonic οφείλεται στο γεγονός ότι είναι και η μέθοδος στην οποία βασίστηκε η συγκεκριμένη εργασία.

Μια σημαντική διευκρίνιση που πρέπει να δοθεί πριν την εκκίνηση παρουσίασης των επόμενων κεφαλαίων είναι η διαφορά ανάμεσα στους όρους 'σύστημα' και 'μέθοδο'. Παραδείγματος χάρη, με τον όρο μέθοδο ηχοληψίας εννοούμε τον τρόπο με τον οποίο γίνεται η ηχοληψία. Έτσι στη 'μέθοδο' περιλαμβάνονται κανόνες και οδηγίες που αφορούν τη διάταξη των μικροφώνων, τα

πολικά τους διαγράμματα καθώς και την ενδεχόμενη επεξεργασία που πρέπει να γίνει σε αυτά τα σήματα. Με τον όρο 'σύστημα' ονομάζουμε την υλοποίηση μιας μεθόδου με την μορφή λογισμικού και hardware. Έτσι ένα σύστημα Ambisonic ηχογράφησης αποτελείται από το λογισμικό και το hardware το οποίο εφαρμόζει την μέθοδο ηχογράφησης Ambisonic. Στα παρακάτω κεφάλαια οποιοσδήποτε διαχωρισμός, ιδιότητα ή γενικότερη αναφορά γίνεται για τα συστήματα, αφορά και την αντίστοιχη μέθοδο, και αντίστροφα.

2.1. Τι είναι τρισδιάστατος ήχος και με ποιο τρόπο τον αντιλαμβάνεται ο άνθρωπος.

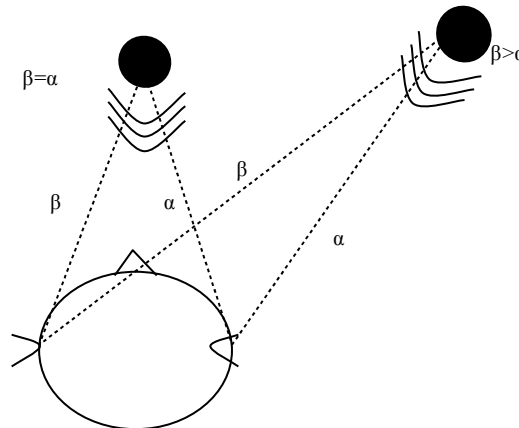
Όπως είναι γνωστό, ο άνθρωπος έχει την δυνατότητα εντοπισμού ηχητικών πηγών που βρίσκονται στον χώρο, εκτιμώντας την τρισδιάστατη θέση τους σε σχέση με τη δική του. Με τον όρο τρισδιάστατος ήχος εννοούμε τον ήχο που μπορεί να προέλθει από οποιαδήποτε αζιμούθια και κατακόρυφη γωνία (βλέπε παράρτημα 2), ή με άλλα λόγια τον ήχο που μπορεί να φτάσει στον ακροατή και από τις τρεις διαστάσεις. Στην καθημερινότητα του ο άνθρωπος λαμβάνει ήχους και από τις τρεις διαστάσεις. Το σύστημα ακοής του ανθρώπου έχει την ικανότητα να αντιλαμβάνεται την προέλευση των ήχων, αξιολογώντας ανεξάρτητα τα σήματα του ήχου που προσπίπτουν στο κάθε αυτί και τις διαφορές που έχουν αυτά μεταξύ τους.

2.1.1 Εντοπισμός της γωνίας προέλευσης του ήχου.

Κριτήρια που βασίζονται στην διαφορά χρόνου πρόσπτωσης και έντασης του ήχου μεταξύ των δύο αυτιών.

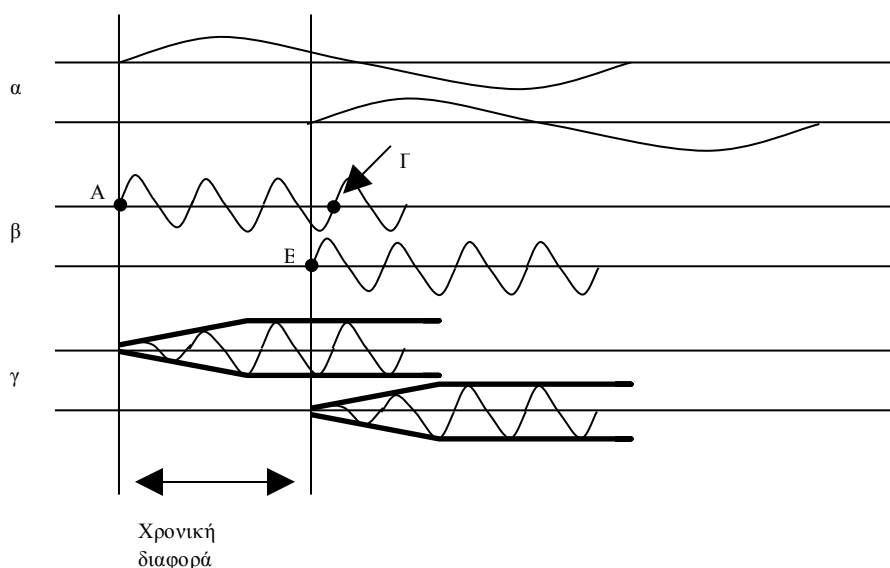
Πλευρική θέση των αυτιών (lateralization).

Λόγο της πλευρικής θέσης των αυτιών και της απόστασης μεταξύ τους, δημιουργείται μια διαφορά στην απόσταση της πηγής από κάθε αυτί (Εικόνα 1), που εξαρτάται από την αζιμούθια γωνία της πηγής σε σχέση με τον ακροατή. Αυτή η διαφορά απόστασης έχει σαν αποτέλεσμα την καθυστέρηση της άφιξης του ήχου σε ένα εκ των αυτιών. Αυτό αποτελεί ένα από τα φυσικά κριτήρια με τα οποία ο άνθρωπος αντιλαμβάνεται την αζιμούθια γωνία, και συναντάται στην βιβλιογραφία ως **Interaural Time Difference (ITD)** (βλέπε [2], [6], [7]). Ταυτόχρονα αυτή η διαφορά στην απόσταση έχει σαν αποτέλεσμα την σχετική μείωση της έντασης του ήχου που θα φτάνει στο ένα αυτί. Η σχετική ένταση του ήχου όμως διαφοροποιείται και λόγω της **ηχητικής επισκίασης**. Με τον όρο ηχητική επισκίαση εννοούμε το φιλτράρισμα που υφίσταται ο ήχος καθώς συναντάει κάποιο εμπόδιο στην διαδρομή του. Παρατηρώντας την εικόνα 1, μπορούμε να δούμε ότι στην περίπτωση της πηγής που βρίσκεται στις 0° αζιμούθια γωνία, η ηχητική επισκίαση θα ήταν η ίδια στα δύο αυτιά (ιδανικές συνθήκες σφαιρικού κεφαλιού, χωρίς ανωμαλίες). Όταν όμως η γωνία ξεφύγει από τις 0° (βλέπε εικόνα 1) τότε η διαδρομή του ήχου προς το ένα αυτί εμποδίζεται από μεγαλύτερο τμήμα του κεφαλιού, πράγμα που σημαίνει πως συχνότητες με μήκος κύματος μικρότερο από το μήκος το εμποδίου θα φιλτραριστούν και η ηχητική μείωση θα είναι ανάλογη της συχνότητας. Οι συχνότητες με μεγαλύτερο μήκος κύματος θα διαθλαστούν και έτσι δεν θα υπάρξει αισθητή διαφοροποίηση στην ένταση μεταξύ των δύο αυτιών (όσον αφορά το συγκεκριμένο φιλτράρισμα). Συνολικά αυτή η διαφοροποίηση της έντασης αποτελεί ένα από τα φυσικά κριτήρια με τα οποία ο άνθρωπος αντιλαμβάνεται την αζιμούθια γωνία της πηγής εκπομπής του ήχου, και στην βιβλιογραφία συναντάται ως **Interaural Intensity Differect (IID)** (βλέπε [2], [6], [7]).



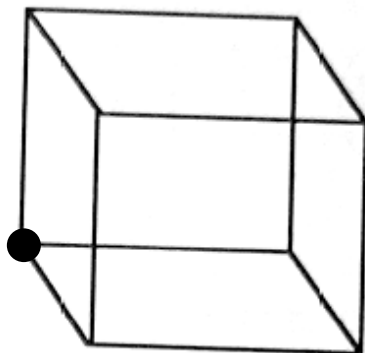
Εικόνα 1: Απεικόνιση ενός ανθρώπινου κεφαλιού και των δύο διαφορετικών ηχητικών πηγών. Ανάλογα με την αξιμότητα γωνία των πηγών υπάρχει διαφορά στην απόσταση μεταξύ της πηγής και του κάθε αυτιού.

Στην περίπτωση του ITD ο ανθρώπινος ωτικός μηχανισμός για την εύρεση της αξιμότητας γωνίας υπολογίζει την διαφορά φάσης μεταξύ του ήχου που λαμβάνουν τα δύο του αφτιά. Ο χρόνος που θα καθυστερήσει ο ήχος να φτάσει στο ένα αφτί, είναι ανάλογος της απόστασης που έχουν τα αφτιά μεταξύ τους, και κατ' επέκταση του μεγέθους του κεφαλιού. Όσο μεγαλύτερο είναι το κεφάλι του ανθρώπου, τόσο μεγαλύτερη θα είναι η απόσταση που έχουν τα αφτιά μεταξύ τους, άρα θα υπάρξει μεγαλύτερη χρονική διαφορά στην λήψη του ήχου, μεταξύ των δύο αφτιών. Ας υποθέσουμε ότι λόγω της απόστασης των αφτιών η χρονική διαφορά λήψης του ήχου θα είναι a sec. Για συχνότητες που η μισή περίοδος τους είναι μεγαλύτερη από τα $a/2$ sec, ο ανθρώπινος οργανισμός μπορεί εύκολα να προσδιορίσει την διαφορά φάσης (βλέπε Εικόνα 2, παράδειγμα α). Στην περίπτωση που η περίοδος της συχνότητας είναι μικρότερη από το $a/2$, το ανθρώπινο ωτικό σύστημα δεν μπορεί να καταλάβει βάση της διαφοράς φάσης ποίος από τους ήχους προηγείται, το Α του Β ή το Β του Γ (βλέπε Εικόνα 2, παράδειγμα β). Παρ' όλα αυτά ο ωτικός μηχανισμός καταφέρνει να υπολογίσει ποίος από τους δύο ήχους προηγείται, αξιοποιώντας το ξεκίνημα της περιβάλλουσας έντασης του ήχου. Πρέπει να σημειώσουμε ότι το ξεκίνημα της δυναμικής περιβάλλουσας αποτελείται από μία συχνότητα φορέα χαμηλή και την φέρουσα συχνότητα (η αρχική συχνότητα) που είναι υψηλή. Το ωτικό σύστημα με κάποιο τρόπο αναλύει την συνολική ένταση της περιβάλλουσας και στα δύο αφτιά για τις υψηλές συχνότητες, και μετράει τον χρόνο καθυστέρησης στην λήψη αυτών των δύο περιβάλλουσων (βλέπε Εικόνα 2, παράδειγμα γ). Αυτό το κριτήριο που χρησιμοποιεί ο ωτικός μηχανισμός το ονομάζουμε "ITD envelope" (βλέπε [2], [6], [7]).



Εικόνα 2: Απεικόνιση των φασικών διαφορών μεταξύ των δύο αυτιών, από έναν ήχο.

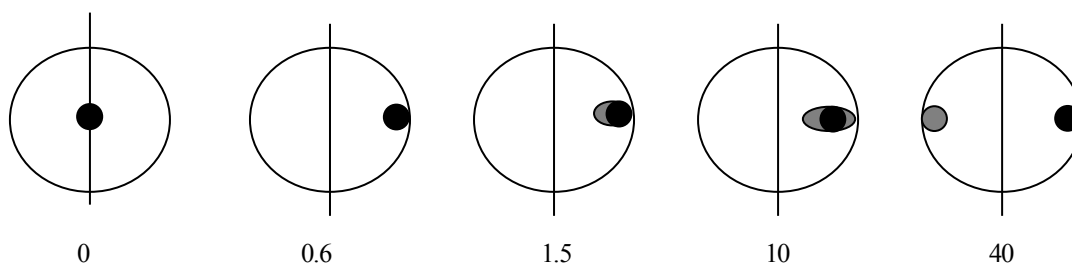
Χρησιμοποιώντας τα lateralization, μπορούμε να προσομοιάσουμε την αίσθηση που δημιουργεί μια πραγματική πηγή, και έτσι να ξεγελάσουμε το ωτικό σύστημα με μία εικονική πηγή. Εδώ πρέπει να αναφερθούμε σε κάποια χαρακτηριστικά της αντίληψης του ωτικού συστήματος. Άμα δρομολογήσουμε δύο ίδια ηχητικά σήματα και στα δύο αφτιά (είτε μέσω ακουστικών ή δύο ηχείων), τότε το ωτικό σύστημα δεν αντιλαμβάνεται δύο ξεχωριστές πηγές, αλλά μία. Η χωρική θέση που θα αντιληφθεί ότι προέρχεται ο ήχος θα εξαρτάται από τα ITD και IID. Άμα λοιπόν δεν υπάρχουν διαφορές ούτε χρονικές αλλά ούτε έντασης ανάμεσα στα δύο ηχητικά σήματα, το ωτικό σύστημα αντιλαμβάνεται ότι ο ήχος έρχεται από το κέντρο. Άμα αυξηθεί η τιμή ενός από τα παραπάνω κριτήρια, το ωτικό σύστημα αντιλαμβάνεται ότι ο ήχος είναι πιο κοντά στο αφτί με την μεγαλύτερη ένταση, ή αυτό που δεν παρουσιάζει χρονική καθυστέρηση. Όσο μεγαλώνει αυτή η διαφορά, τόσο τείνει ο ήχος να αποκλίνει από την διεύθυνση του κέντρου, και πλησιάζει την διεύθυνση του ενός αυτιού. Αυτό όμως συμβαίνει μέχρι μια μέγιστη τιμή των ITD και IID. Μετά από αυτή την τιμή η αίσθηση την διεύθυνσης προέλευσης του ήχου παραμένει η ίδια, στην διεύθυνση του ενός αυτιού. Η μέγιστη τιμή των ITD και IID εξαρτάται από το ηχητικό ερέθισμα, αλλά κάποιες προσεγγιστικές τιμές είναι για το ITD από 0.005 έως 1.5 msec και για το IID 1 μέχρι 10 dB (Sayers, 1964; Toole and Sayers, 1965). Τέλος, ένα τελευταίο χαρακτηριστικό στο οποίο πρέπει να αναφερθούμε είναι η αίσθηση που δημιουργείται σε μερικούς ότι ο ήχος, ενώ προέρχεται από την μπροστινή μεριά, στην πραγματικότητα προέρχεται από την πίσω. Το ίδιο μπορεί να συμβεί και για έναν ήχο ο οποίος δεν έχει εσω-ακουστικές διαφορές μεταξύ των αφτιών, μπορεί να θεωρηθεί πως ακούγεται από την πάνω μεριά, ενώ στην πραγματικότητα προέρχεται από την κάτω. Η εικόνα 3 απεικονίζει το οπτικό ισοδύναμο αυτής της αντίστροφης αίσθησης.



Εικόνα 3: The Necker cube, μια αναστροφή της αντίληψης της προοπτικής. Η τελεία μπορεί να βρίσκεται μέσα στον κύβο, στην πίσω, κάτω, αριστερή γωνία του ή εκτός αυτού στην μπροστά, κάτω, αριστερή γωνία. Αυτό συμβαίνει και στην περίπτωση της δημιουργούμενης αμφιβολίας εάν η προέλευση του ήχου είναι από μπροστά ή από πίσω.

Το φαινόμενο της προπορείας (Precedence ή Hass effect).

Κατά το φαινόμενο αυτό, αν λάβει ο άνθρωπος τον ίδιο ήχο με διαφορά χρόνου μέχρι ένα όριο, δεν τον αντιλαμβάνεται σαν ένα ξεχωριστό ήχο, αλλά σαν τον ίδιο, πχ με αντήχηση. Αυτός ο μηχανισμός βοηθάει τον άνθρωπο να αντιληφθεί την χωροτοποθέτηση του ήχου παρουσία αντηχήσεων. Ακολουθεί ένα πείραμα, στο οποίο χρησιμοποιώντας ένα σύστημα στερεοφωνικό (είτε στερεοφωνικά ακουστικά είτε δύο ηχεία) τροφοδοτούμε και τα δύο κανάλια με τον ίδιο ήχο, μόνο που εφαρμόζουμε ένα μεταβλητό delay στο αριστερό κανάλι. Στην Εικόνα 4 παρατηρούμε την χωρική θέση που αντιλαμβάνεται ο άνθρωπος, σε σχέση με την χρονική καθυστέρηση ανάμεσα στα δύο κανάλια.



Προσεγγιστικά ο χρόνος καθυστέρησης του αριστερού καναλιού σε msec

Εικόνα 4: Perceptual effect από την αύξηση του ITD από 0 έως 40 msec

Όπως φαίνεται και στην εικόνα 4, για καθυστερήσεις μεταξύ των 0-0.6 msec, ο ήχος μετατοπίζεται σταδιακά προς τα δεξιά στον οριζόντιο άξονα. Στο εύρος από 0.7 έως 1.5 msec η πηγή παραμένει στην δεξιά θέση, μόνο που έχει αρχίσει να φαίνεται μια επισκιασμένη περιοχή που συμβολίζει ότι φαίνεται να έχει μεγαλώσει το εύρος της πηγής. Από το 1.5 έως τα 10 msec, το εύρος της πηγής σταδιακά αυξάνεται, ενώ το κέντρο που συμβολίζει την κύρια θέση της πηγής μετατοπίζεται προς το κέντρο. Τέλος για χρόνους μεγαλύτερους των 10 msec, βάση της ανθρώπινης αντίληψης, συνεχίζεται η αύξηση του εύρους της πηγής, συνεχίζεται η μετατόπιση προς το κέντρο, αλλά μετά

από κάποιο ασαφές χρονικό όριο, που εξαρτάται από την πηγή ήχου, το διευρυμένο εύρος της πηγής θα διασπαστεί σε δύο ξεχωριστές πηγές ήχου (για περισσότερες πληροφορίες βλέπε [2], [6], [7]).

Κριτήρια που βασίζονται στην κίνηση του κεφαλιού ή στην κίνηση της πηγής του ήχου.

Κίνηση του κεφαλιού.

Ο άνθρωπος για να εντοπίσει κάποια ηχητική πηγή κινεί το κεφάλι του έτσι ώστε να ελαχιστοποιήσει τις εσω-ακουστικές διαφορές (ITD και IID). Από διάφορες μελέτες (βλέπε [2], [6], [7]) έχει φανεί ότι η κίνηση του κεφαλιού βοηθάει γενική στην προσέγγιση της θέσης του ήχου, και στην μείωση των φαινομένων αντιστροφής (έναν ήχο ενώ είναι στην μπροστινή μεριά τον αντιλαμβάνεται από πίσω). Η μείωση του φαινομένου της αντιστροφής μπορεί να εξηγηθεί με το ακόλουθο παράδειγμα:

Φανταστείτε ότι ένας ήχος προέρχεται από τις 30° αζιμούθια γωνία στο οριζόντιο επίπεδο. Η αντίστροφη θέση θα φανεί στις 150°. Όταν κινήσει το κεφάλι του προς τα δεξιά, θα αυξηθεί η ένταση που φτάνει ο ήχος στο αριστερό αφτί και θα μειωθεί αντίστοιχα η ένταση του ήχου που φτάνει στο δεξί. Ταυτόχρονα η χρονική καθυστέρηση που έφτανε ο ήχος στο αριστερό αφτί θα μειωθεί. Αντίθετα, αν ο ήχος πράγματι προερχόταν από τις 150°, κατά την κίνηση του κεφαλιού προς τα δεξιά, θα μείωνε την ένταση που φτάνει ο ήχος στο δεξί αυτί, θα αύξανε την ένταση που φτάνει στο αριστερό, και θα αύξανε την χρονική καθυστέρηση του αριστερού αφτιού. Έτσι θα αποσαφήνιζε ότι η χωρική θέση του ήχου είναι στις 30° μπροστά.

Ένα γεγονός που πρέπει να επισημάνουμε σε αυτό το σημείο είναι ότι κατά την ακρόαση από στερεοφωνική συστοιχία ηχείων, η κίνηση του κεφαλιού αντί να βοηθήσει στην εύρεση της εικονικής γωνίας προέλευσης του ήχου, θα οδηγούσε σε λανθασμένα συμπεράσματα, καθώς το στερεοφωνικό σύστημα έχει δημιουργηθεί θεωρώντας ως προϋπόθεση ότι ο προσανατολισμός του ανθρώπου θα είναι συγκεκριμένος.

Και στην περίπτωση όμως της ακρόασης από ακουστικά, η κίνηση του κεφαλιού δεν θα επιφέρει καμία αλλαγή στον ήχο που φτάνει στα αυτιά του ακροατή, άρα πάλι δεν μπορεί το ωτικό σύστημα να εκμεταλλευτεί αυτή του την δυνατότητα του στην εύρεση της διεύθυνσης των ηχητικών πηγών.

Κίνηση της πηγής του ήχου.

Ακριβώς όπως στην περίπτωση της κίνησης του κεφαλιού με μία σταθερή πηγή, έτσι και όταν το κεφάλι παραμένει ακίνητο και η πηγή κινείται, θα δημιουργήσει κάποιες δυναμικές αλλαγές με τις οποίες ο ανθρώπινος οργανισμός θα μπορέσει να αντιληφθεί την ακριβή θέση του ήχου. Ένας άλλος πολύ βασικός παράγοντας με τον οποίο το ωτικό σύστημα αντιλαμβάνεται την κίνηση της ηχητικής πηγής είναι η εμπειρία. Ένα παράδειγμα της χρήσης της εμπειρίας αποτελεί η δημιουργία της αίσθησης του περάσματος ενός αυτοκινήτου σε αγώνες. Αυτό μπορεί να επιτευχθεί με ένα μονοφωνικό ηχείο, το οποίο εκπέμπει ήχους συσχετισμένους με την εμπειρία του ανθρώπου. Το κύριο στοιχείο για την αίσθηση της κίνησης του ήχου είναι το Doppler effect.

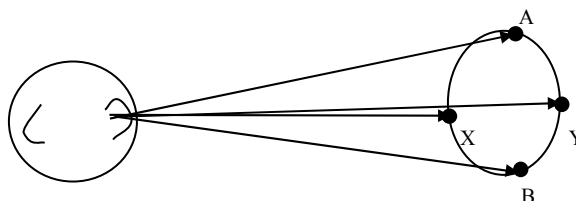
Ένα στοιχείο που προσμετράτε σε διάφορα πειράματα είναι η ελάχιστη ακουστή γωνία μετατόπισης (Minimum Audible Movement Angle, σε συντομογραφία **MAMA**) (βλέπε [2], [6], [7]). Ο άνθρωπος έχει την δυνατότητα να αντιληφθεί ακίνητες πηγές με διαφορά γωνίας γύρω στη μία μοίρα, κάτω από κατάλληλες συνθήκες. Η ελάχιστη αντιληπτή γωνία για κινούμενους ήχους

είναι της τάξης των 3 μοιρών κάτω από κατάλληλες συνθήκες, αλλά μπορεί να αυξηθεί συναρτήσει της έντασης της κίνησης, της χωρικής θέσης που γίνεται η κίνηση, αλλά και από τον τύπο της ηχητικής πηγής. Η MAMA παίρνει την ελάχιστη τιμή όταν οι ήχοι είναι ευρέως φάσματος. Αυτή η τιμή αυξάνεται για εντάσεις κίνησης εύρους από 2.8° έως 360° /sec, καθώς ο χρόνος που χρειάζεται το ωτικό σύστημα για να λάβει επιτυχημένα στιγμιότυπα κυμαίνονται μεταξύ των 150-300 msec (Grantham, 1986).

Κριτήρια που βασίζονται στις φασματικές τροποποιήσεις που προκαλούνται από την κόγχη.

Ασαφή κριτήρια ITD και IID.

Ας υποθέσουμε ένα κεφάλι το οποίο είναι σφαιρικό, που δεν έχει ασυμμετρίες, αγνοήσουμε τα χαρακτηριστικά του προσώπου, το οποίο δεν έχει κόγχες. Σε αυτή την ιδανική περίπτωση θα μπορούσαμε να πούμε ότι υπάρχουν σημεία στο χώρο τα οποία παρουσιάζουν τα ίδια ITD και IID (βλέπε εικόνα 5). Σημειώνω ότι αυτό το μοντέλο είναι τελείως θεωρητικό και δεν θα μπορούσε να υπάρξει σε ένα πραγματικό άτομο. Σε αυτά τα σημεία, άμα δεν λάβουμε υπόψη άλλα στοιχεία υπολογισμού της χωροτοποθέτησης της πηγής εκτός από τα ITD και IID, θα μπορούσε να υπάρξει μπέρδεμα του ωτικού συστήματος στον προσδιορισμό του σημείου που πραγματικά εκπέμπεται ο ήχος. Αυτά τα σημεία βρίσκονται στην επιφάνεια ενός κώνου ο οποίος ονομάζεται στη βιβλιογραφία "Cone of Confusion". Μετά από μελέτες που έγιναν στον εντοπισμό του τρόπου με τον οποίο το ωτικό σύστημα αποσαφηνίζει εάν ο ήχος προέρχεται από πάνω ή κάτω, από μπροστά ή από πίσω, κατέληξαν στο συμπέρασμα ότι **πρέπει να υπάρχει κάποιο φασματικό κριτήριο με το οποίο το σύστημα κάνει αυτόν τον εντοπισμό, και αυτό οφείλεται κυρίως στην κόγχη του αυτιού.**



Εικόνα 5: The cone of confusion. Παρομοιότητες τιμές των ITD και IID μπορούν να υπολογιστούν σε ένα σφαιρικό μοντέλο κεφαλιού, σε αντικρινά σημεία της επιφάνειας ενός κώνου. Σε αυτό το σχήμα με τα A-B απεικονίζεται η αμφιβολία για το ύψος μεταξύ αυτών των σημείων, ενώ με τα X-Y, η αμφιβολία για το μπροστά-πίσω αυτών των σημείων.

Οι Head Related Transfer Function.

Με τον όρο Head Related Transfer Function (εν συντομία **HRTF**) ορίζεται το συχνοτικό φίλτράρισμα του ήχου πριν φτάσει στο τύμπανο, και οφείλεται κυρίως στο εξωτερικό αφτί. Με τον όρο **Binaural HRTF** (βλέπε [2], [6], [7]) αναφερόμαστε σε HRTF και των δύο αυτιών, και μπορεί να οριστεί η ένταση του ήχου εξαρτώμενη από την συχνότητα και οι χρονικές καθυστερήσεις που οφείλονται στο πολύπλοκο σχήμα της κόγχης.

Οι πτυχές της κόγχης επιφέρουν στον ήχο μικρές χρονικές καθυστερήσεις της τάξης των 0-300 msec, οι οποίες τροποποιούν σημαντικά το φασματικό περιεχόμενο του ήχου που φτάνει στο αφτί, σε σχέση με αυτόν που θα λάμβανε ένα παντοκατευθυντικό μικρόφωνο στην ίδια

απόσταση. Η ασυμμετρία της κόγχης του αφτιού επιφέρει στον ήχο μικρές χρονικές καθυστερήσεις, συντονισμούς και διαθλάσεις του ήχου, που έχουν σαν αποτέλεσμα την αλλαγή του συχνοτικού του περιεχομένου σαν συνάρτηση της θέσης του της πηγής του ήχου σε σχέση με το κεφάλι.

Η χρήση των φασματικών τροποποιήσεων που επιφέρουν οι HRTF έχουν χαρακτηριστεί ως το συστατικό κλειδί για ένα σύστημα τρισδιάστατου ήχου, είτε για απευθείας μετρήσεις, είτε για μοντελοποίηση. Ο λόγος είναι ότι για την δημιουργία ενός ηχητικού χωρικού στοιχείου με μεγάλη ακρίβεια, πρέπει να τροποποιήσεις το φάσμα του ήχου με τέτοιο τρόπο ώστε όταν αυτό φτάσει στο τύμπανο να πλησιάζει όσο το δυνατό περισσότερο τον ήχο που θα έφτανε στο τύμπανο, κάτω από πραγματικές συνθήκες ακρόασης. Στην βιβλιογραφία ο όρος HRTF μπορεί να βρεθεί και ως Head Transfer Function (HTF), ως pinnae transform, ως Outer Ear Transfer Function (OETF) και ως Directional Transfer Function (DTF).

Ανάλογα με τα κριτήρια που θέτονται για μία εφαρμογή, μπορεί να είναι χρήσιμος ο επιπρόσθετος προσδιορισμός κάποιων χαρακτηριστικών του σώματος πέρα από την κόγχη οι οποίοι θα συνυπολογιστούν στην HRTF. Κάποιες μετρήσεις περιλαμβάνουν μόνο το την κόγχη και το κεφάλι, κάποιες άλλες συμπεριλαμβάνουν στις προηγούμενες το εσωτερικό κανάλι του αυτιού, τον κορμό του ανθρώπινου σώματος και τους ώμους, τα οποία επηρεάζουν λιγότερο το συνολικό φιλτράρισμα. Τέτοια χαρακτηριστικά μπορούν να χωριστούν σε αυτά που παραμένουν σταθερά ανεξάρτητα της διεύθυνσης του ήχου, και σε αυτά που αλλάζουν.

Σπουδαία χαρακτηριστικά των HRTF.

Έχουν γίνει διάφορα πειράματα σε ανθρώπους και έχει βρεθεί ότι οι HRTF ποικίλουν ανάλογα με το άτομο. Αυτό οφείλεται στο γεγονός ότι όλα τα άτομα δεν έχουν το ίδιο μέγεθος κόγχης αλλά ούτε και το ίδιο σχήμα. Το ιδανικό λοιπόν θα ήταν ανάλογα με τον χρήστη, να γινόταν χρήση της δικής του HRTF, πράγμα που δεν γίνεται στην υλοποίηση 3-D συστημάτων. Στις περισσότερες περιπτώσεις χρησιμοποιούνται γενικές HRTF (**Nonindividualized HRTF**) (βλέπε [2], [6], [7]). Αυτές μπορεί να προέρχονται από την μέση τιμή προσωπικών HRTF, πράγμα που μπορεί να μειώσει τα μέγιστα και ελάχιστα του φάσματος. Στις περισσότερες περιπτώσεις όμως χρησιμοποιούνται θεωρητικές HRTF – που έχουν υπολογιστεί είτε φυσικά είτε ψυχοακουστικά- οι οποίες γενικεύονται σε ένα μεγάλο αριθμό πληθυσμού.

Φασικά χαρακτηριστικά των HRTF.

Υπάρχουν ακόμα και διαφορές στις φάσεις που προκύπτουν από τις HRTF. Αυτό μπορεί να εξηγηθεί από το γεγονός ότι αφού υπάρχουν χρονικές καθυστερήσεις οι οποίες είναι ανάλογες με την συχνότητα, σίγουρα θα υπάρχουν και φασικές, πάλι ανάλογα με την συχνότητα. Η εξήγηση αυτού του φαινομένου είναι απλή. Καθυστερήση της τάξης της μίας περιόδου επιφέρει στο σήμα διαφορά στην φάση κατά 360° , ενώ καθυστέρηση κατά 0.5 της περιόδου επιφέρει διαφορά φάσης κατά 180° , κοκ. Η φασική απόκριση είναι πιθανός λιγότερο κρίσιμη σε σχέση με τις εσωακουστικές χρονικές διαφορές, για τον εντοπισμό της πηγής του ήχου. Δεν έχει αποσαφηνιστεί επίσης σε ποίο βαθμό είναι χρήσιμες οι καθυστερήσεις ανά συχνότητα σε σχέση με την συνολική καθυστέρηση σε όλες τις συχνότητες που προέρχεται από τις ITD.

Εντοπισμός βάση των HRTF.

Μετά από πειράματα που έχουν γίνει έχει βρεθεί ότι από ψυχοακουστικής άποψης, ότι ο κύριος ρόλος των HRTF είναι να αποσαφηνίσουν εάν ο ήχος προέρχεται από μπροστά ή από πίσω (σημεία A-B στον cone of confusion), ή το πάνω κάτω (σημεία X-Y στον cone of confusion)

(βλέπε εικόνα 5).

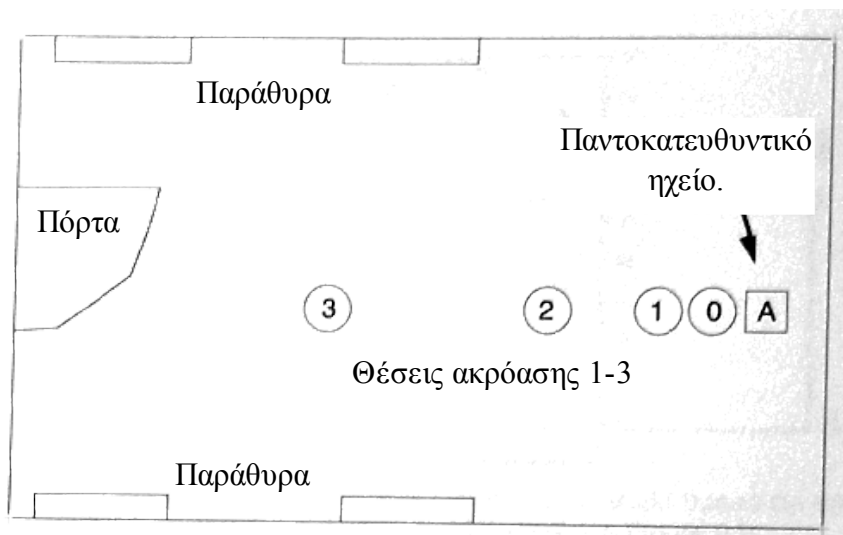
2.1.2 Υπολογισμός της απόστασης.

Στοιχεία έντασης και ηχηρότητας.

Το στοιχείο της έντασης ή της ηχηρότητας πιθανός να παίζουν ένα πιο σημαντικό ρόλο με άγνωστους ήχους παρά με γνωστούς. Η έκθεση σε μία συγκεκριμένη ηχητική πηγή σε διαφορετικές αποστάσεις επιτρέπει την ενσωμάτωση πολλών στοιχείων για την αντίληψη της απόστασης. Χωρίς όμως αυτή την έκθεση, σαν κριτήριο παραμένει μόνο η ηχηρότητα. Ένα παράδειγμα αποτελεί η σύγκριση της εμπειρίας της ακρόασης ήχων αμέσως πριν πάει να κοιμηθεί κάποιος, σε ένα άγνωστο με ένα γνωστό περιβάλλον. Σε ένα γνωστό περιβάλλον είναι γνωστή η απόσταση πχ, του ρολογιού που βρίσκεται στο δωμάτιο, και η απόσταση ενός αυτοκινήτου που περνάει. Αν και το αυτοκίνητο έχει μεγαλύτερη ένταση από το ρολόι, λόγω της οικειότητας με το περιβάλλον αναιρεί το στοιχείο της έντασης, καθώς ξέρει πως το αυτοκίνητο είναι πιο μακριά. Όταν όμως πηγαίνεις σε ένα άγνωστο μέρος, λόγω αυτής της έλλειψης οικειότητας, σε άγνωστους ήχους πιθανός να κριθεί η απόσταση βάση της κλίμακας έντασης. Μπορεί να υπολογιστεί όμως η κλίμακα έντασης;

Κάτω από ανηχοϊκές συνθήκες κάποιος μπορεί να υπολογίσει την ένταση του ήχου μιας παντοκατευθυντικής πηγής σε κάποιο σημείο βάση του νόμου του αντίστροφου τετραγώνου, ξέροντας την ένταση σε ένα συγκεκριμένο σημείο. Βάση των παραπάνω για κάθε διπλασιασμό της απόστασης θα έχουμε πτώση κατά 6 dB του ήχου. Στην περίπτωση που δεν είχαμε μια παντοκατευθυντική πηγή, αλλά μια γραμμική πηγή όπως ένας αυτοκινητόδρομος, τότε η πτώση της έντασης θα ήταν περίπου στα 3 dB για κάθε διπλασιασμό της απόστασης. Όμως υπάρχει μια σημαντική διαφορά ανάμεσα στην ένταση και αυτό που αντιλαμβάνεται ο άνθρωπος. Το αντίστοιχο μέγεθος που υπολογίζει την ένταση που αντιλαμβάνεται ο άνθρωπος ονομάζεται ηχηρότητα. Όμως κρίσεις όπως μισής ή διπλής ηχηρότητας συσχετίζονται καλύτερα με την sone scale παρά με το νόμο του αντίστροφου τετραγώνου. Τα sones είναι μια σκάλα που συσχετίζει ίσες ηχηρότητες για κάποια συχνότητα, με μία σκάλα έντασης. Διπλασιάζοντας τον αριθμό των sones επιτυγχάνεται διπλασιασμό της ηχηρότητας. Σε πειράματα που χρησιμοποιήθηκαν sones αντί για τον νόμο του αντίστροφου τετραγώνου, βρέθηκε ότι ο διπλασιασμός της απόστασης είχε σαν αποτέλεσμα την μείωση της ηχηρότητας κατά 10 dB. Περαιτέρω έρευνες έδειξαν ότι ή μείωση κατά 3 dB είναι μάλλον πολύ μικρή για να επιτευχθεί πιστικό αποτέλεσμα.

Όλα τα παραπάνω αναφέρονται σε ανηχοϊκές συνθήκες. Στην πραγματικότητα υπάρχει η αντήχηση και ταυτόχρονα άλλες πηγές. Οι παρακάτω εικόνες (εικόνες 6,) μπορούν να δείξουν πως η επιρροή της αντήχησης εμποδίζει την χρήση είτε του νόμου του αντίστροφου τετραγώνου, είτε την σκάλα ηχηρότητας. Η συνολική ένταση που φτάνει στον ακροατή από απευθείας και ανακλώμενο ήχο μπορεί να υπολογιστεί με την χρήση ενός προγράμματος προσομοίωσης δωματίων. Το αποτέλεσμα της υπολογισμού των συνολικών dB SPL φαίνεται στην εικόνα . Εδώ πρέπει να σημειώσουμε ότι τα συνολικά SPL δεν αλλάζουν περισσότερο από 3 dB μεταξύ του κοντινότερου και του μακρινότερου σημείου ακρόασης, παρά τον διπλασιασμό της απόστασης. Σε ένα αντηχητικό περιβάλλον η αναλογία έντασης μεταξύ της απευθείας και της ανακλώμενης ενέργειας αποτελεί βασικότερο κριτήριο για τον υπολογισμό της απόστασης από την κλίμακα της έντασης.



Εικόνα 6: Ένα μοντελοποιημένο δωμάτιο που χρησιμοποιήθηκε για τον υπολογισμό της έντασης μια πηγής σε διάφορα σημεία του. Η θέσεις του ακροατή 0-3 απέχουν από την πηγή 1, 2, 4, 8 feet αντίστοιχα. Το μοντελοποιημένο δωμάτιο αποτελείται από τέσσερα παράθυρα, από μια ξύλινη πόρτα, ένα χαλί πάνω στο συμπαγές δάπεδο, και κόντρα πλακέ.

Επίδραση της οικειότητας και της πρόβλεψης.

Το στοιχείο της απόστασης μπορεί να τροποποιηθεί σαν συνάρτηση της οικειότητας και της πρόβλεψης με την πηγή του ήχου, ειδικά με την ομιλία. Γι' αυτό το λόγο κάθε υλοποίηση των κριτηρίων της απόστασης σε ένα 3-D ηχητικό σύστημα, πιθανός να χρειαζόταν μία αποτίμηση της γνωστικής σχέσης με την ηχητική πηγή. Εάν η ηχητική πηγή είναι εντελώς συνθετική (πχ λευκός θόρυβος), τότε ο ακροατής θα χρειαστεί περισσότερο χρόνο για να αποκτήσει την οικειότητα με την παραμετρική αλλαγή της ηχηρότητας και άλλων στοιχείων που διαφοροποιούνται με τις διάφορες αποστάσεις. Εάν η ηχητική πηγή συσχετίζεται με μία συγκεκριμένη θέση από τις επαναλαμβανόμενες θέσεις, η προσομοίωση της απόστασης θα είναι ευκολότερη σε σχέση με αυτή που δεν είναι οικεία ή εμφανίζεται απρόσμενα. Για παράδειγμα είναι ευκολότερο να προσομοιάσεις έναν ψίθυρο έξι inches μακριά από το αφτί, παρά να τον προσομοιάσεις έξι μέτρα μακριά (βλέπε [2], [6], [7]).

Φασματικά και Binaural στοιχεία για την απόσταση.

Το φασματικό περιεχόμενο μιας ηχητικής πηγής σχετική με την θέση λήψης μπορεί να μεταβάλλεται σαν συνάρτηση της απόστασης. Αυτή η επιρροή περιλαμβάνει την επίδραση των ατμοσφαιρικών συνθηκών, την μοριακή απορρόφηση του αέρα, και την καμπυλότητα του κύματος. Από ψυχοακουστική άποψη, αυτά τα στοιχεία είναι σχετικά αδύναμα, σε σύγκριση με την ηχηρότητα, την οικειότητα και τα στοιχεία της αντήχησης. Παρόλα αυτά όλα αυτά τα στοιχεία μπορούν να αποβούν χρήσιμα για την δημιουργία ενός ωτικού μοντέλου με μεγάλη ακρίβεια. Στη σχέση 1 φαίνεται ο τύπος που δίνει την σχέση μείωσης της έντασης στις διάφορες συχνότητες συναρτήσει της απόστασης και άλλων παραγόντων είναι (μονάδα μέτρησης dB/meter):

$$a_{(f)} = 8.686f^2 \left\langle \left[1.84 * 10^{-11} \left(\frac{p_a}{p_r} \right)^{-1} \left(\frac{T}{T_r} \right)^{\frac{1}{2}} \right] + \left(\frac{T}{T_r} \right)^{\frac{-5}{2}} \{ 0.01275 [\exp(-\frac{2239.1}{T})] [\frac{f_{ro}}{f_{ro}^2 + f^2}] + 0.1068 [\exp(-\frac{3352.0}{T})] [\frac{f_{rN}}{f_{rN}^2 + f^2}] \} \right\rangle$$

Σχέση 1: Ο τύπος υπολογισμού της μείωσης της έντασης ανά συχνότητα συναρτήση της απόστασης και διαφόρων άλλων παραγόντων, και μετριέται σε dB/meter.

όπου:

f = η συχνότητα του ήχου στην οποία το ποσοστό εξασθένησης θα υπολογιστεί, σε Hz.

P_a = Περιβαλλοντική ατμοσφαιρική πίεση σε kPa.

P_r = Ατμοσφαιρική πίεση αναφοράς 101.325 kPa.

T = Ατμοσφαιρική θερμοκρασία σε K.

T_r = Θερμοκρασία αναφοράς 293.15 βαθμοί K.

$$f_{ro} = \frac{p_a}{p_r} \left[24 + \frac{(4.04 * 10^4 h)(0.02 + h)}{0.391 + h} \right] \quad [1]$$

$$f_{rN} = \left(\frac{p_a}{p_r} \right) \left(\frac{T}{T_r} \right)^{1/2} \left\langle 9 + 280(h) \exp \{ -4.170 \left[\left(\frac{T}{T_r} \right)^{-1/3} - 1 \right] \} \right\rangle \quad [2]$$

και για τις συναρτήσεις [1] και [2], το h είναι ισοδύναμο με το ποσοστό (της εκατό) εξαέρωσης των υγρών και υπολογίζεται ως εξής:

$$h = h_{rel} \left(\frac{p_{sat}}{p_r} \right) \left(\frac{p_a}{p_r} \right)^{-1}$$

όπου :

h_{rel} = Το ποσοστό (της εκατό) της υγρασίας

p_{sat} = (p_r)10^V

Στην παραπάνω εξίσωση για τον υπολογισμό της p_{sat}, η δύναμη V υπολογίζεται ως εξής:

$$V = 10.79586 \left[1 - \left(\frac{T_{01}}{T} \right) \right] - 5.02808 \log \left(\frac{T}{T_{01}} \right) + 1.50474 * 10^{-4} \{ 1 - 10^{-8.29692 \left[\left(\frac{T}{T_{01}} \right) - 1 \right]} \} + 0.42873 * 10^{-3} \{ -1 + 10^{4.76955 \left[1 - \left(\frac{T_{01}}{T} \right) \right]} \} - 2.2195983$$

όπου:

T₀₁ = Τριπλό σημείο ισοθερμικής θερμοκρασίας, 273,16 βαθμούς K.

Κάνοντας πράξη με την παραπάνω εξίσωση, βγαίνει το συμπέρασμα ότι όταν η πηγή απομακρύνεται από τον δέκτη, θα έχουμε αισθητή μείωση της έντασης των υψηλών συχνοτήτων, ενώ η μείωση θα είναι πολύ λιγότερη στις χαμηλές (βλέπε [19]).

Φασματικές αλλαγές και απόστασης της ηχητικής πηγής (μέτωπο κύματος).

Το ηχητικό μέτωπο μια πηγής, όταν φτάσει στα αυτιά του ακροατή είναι επίπεδο, αλλά μια πολύ κοντινή πηγή θα έχει καμπυλωτό μέτωπο. Μετατροπή του καμπύλου μετώπου του κύματος σε επίπεδο συμβαίνει καθώς η πηγή έχει μεγαλύτερη απόσταση από περίπου τα δύο μέτρα, σε σχέση με τον δέκτη. Αυτό έχει σαν αποτέλεσμα σε μια επιπρόσθετη έμφαση στις χαμηλές σε σχέση με τις υψηλές συχνότητες. Ένα αισθητό παράδειγμα για την εξήγηση αυτού του φαινομένου, είναι το "σκούρεμα" του ηχοχρώματος, καθώς μια πηγή πλησιάζει το αφτί του δέκτη. Αυτό το φαινόμενο συσχετίζεται με τις ισοακουστικές καμπύλες, οι οποίες δείχνουν μια ευαισθησία στις χαμηλές συχνότητες καθώς αυξάνεται η ηχητική πίεση.

Το στοιχείο του σκουρέματος του ηχοχρώματος για πολύ κοντινές πηγές συμβαίνει λιγότερο συχνά στις καθημερινές ακροάσεις σε σχέση με την μείωση των υψηλών συχνοτήτων που συμβαίνει καθώς μια ηχητική πηγή απομακρύνεται. Με την αύξηση της απόστασης, τα υψηλά στοιχεία ενός πολύπλοκου ήχου επηρεάζονται αισθητά από την υγρασία και την θερμοκρασία (βλέπε Σχέση 1).

Εντοπισμός του ηχητικού πεδίου μέσα στο κεφάλι του ακροατή (Inside the head localization).

Κατά την ακρόαση με ακουστικά είναι σύνηθες φαινόμενο να δημιουργηθεί η αίσθηση ότι ο εντοπισμός της ηχητικής πηγής γίνεται μέσα στο κεφάλι του ακροατή. Αυτό είναι πολύ σημαντικό, καθώς διαφορετικά η μεταφορά της εικονικής αίσθησης των τριών διαστάσεων μπορεί να επιτευχθεί πολύ καλύτερα σε ένα σύστημα ακουστικών, παρά μέσω ηχείων. Για την εξήγηση της ύπαρξης αυτού του φαινομένου ο Laws (1973) και άλλοι προσδιόρισαν σαν αιτίες διάφορες μη γραμμικές παραμορφώσεις που προκαλούνταν από διάφορα τμήματα της επικοινωνιακής αλυσίδας, και η χρήση free-field αντί για diffuse-field ισοροπημένα (equalized) ακουστικά.

Αντηχήσεις.

Οι ερευνητές συμφωνούν ότι η αντήχηση, φυσική ή τεχνητή, βελτιώνουν της εξωτερική του ήχου σε 3-D ακουστικά. Ας δούμε όμως ποια στοιχεία την αντήχησης μεταβάλλονται με την αύξηση της απόστασης.

Όταν κάποιος απομακρύνεται από μία πηγή σε ένα κλειστό χώρο, η ένταση του απευθείας ήχου θα μειωθεί, ενώ η ένταση των αντηχήσεων θα παραμείνει σταθερή. Οι μεταγενέστερες αντηχήσεις (late reverberation) πληροφορούν την αντίληψη για τον όγκο του χώρου που βρίσκεται μέσα η πηγή του ήχου. Η αναλογία έντασης μεταξύ του αντηχητικού σε σχέση με τον απευθείας ήχο αποτελεί κριτήριο της απόστασης.

Ο Von Békésy (1960) παρατήρησε ότι όταν άλλαζε την αναλογία μεταξύ του αντηχητικού και του απευθείας ήχου, ενώ η ηχηρότητα του ήχου παρέμενε σταθερή, άλλαζε η αίσθηση της απόστασης. Παρατήρησε ακόμα ότι με την αύξηση της αντήχησης αυξήθηκε και το φαινομενικό πλάτος της πηγής. Ένα ακόμα σημαντικό στοιχείο είναι ότι αυτός ο λόγος μεταξύ της έντασης του αντηχητικού και του απευθείας ήχου εξαρτάται σε μεγάλο βαθμό και από τον ίδιο τον χώρο, κάνοντας πιο αδύναμο αυτό το κριτήριο. Για παράδειγμα, σε ένα μικρό δωμάτιο, όπου έχει γίνει ακουστική μελέτη, αυτός ο λόγος θα ποικίλει ανάμεσα σε μικρότερα όρια απ' ότι σε ένα μεγάλο δωμάτιο, όπως ένα κλειστό γυμναστήριο.

Ο Sheeline (1982), όπως και ο Békésy (1960), βρήκαν ότι η αντήχηση προσδιορίζει ένα όριο για τον κρίση των αποστάσεων. Υπάρχει ένα ανώτατο όριο με το οποίο οι αντηχήσεις

μπορούν να μιξαριστούν με έναν ήχο πριν αυτός φτάσει στον "ακουστικό ορίζοντα" (auditory horizon").

Τέλος για την αντήχηση πρέπει να τονίσουμε ότι ενώ η παρουσία ενός αντηχητικού περιβάλλοντος μπορεί να δώσει την αίσθηση ότι το εικονικό ηχητικό πεδίο τείνει σε πραγματικό, μπορεί όμως ταυτόχρονα να μειώσει την ακρίβεια εντοπισμό, κάτω και από πραγματικές και εικονικές συνθήκες. Αυτό είναι πιθανόν να συμβεί σε καταστάσεις όπου η ένταση και ο χρόνος των πρώτων αντηχήσεων είναι τόσοι, ώστε η ατάκα της περιβάλλουσας έντασης αλλοιώνεται (δεν ακούγεται τόσο καθαρά) στο χρόνο (βλέπε [7],[18],[19]).

2.2. Συστήματα αναπαραγωγής Surround ήχου.

Με τον όρο Surround ονομάζονται τα συστήματα τα οποία μπορούν να δημιουργούν κατάλληλες ψευδαισθήσεις στον άνθρωπο ότι προέρχονται από διάφορες κατευθύνσεις της περιφέρειας του. Αυτά τα συστήματα θα χωριστούν σε δύο κατηγορίες βάση της ιδιότητάς τους να είναι ιεραρχικά (hierarchical). Με τον όρο ιεραρχικό ονομάζεται ένα σύστημα που αφού έχει κωδικοποιήσει ένα ηχητικό πεδίο μπορεί να το ανασυνθέσει ανεξαρτήτως του αριθμού των ηχείων που περιλαμβάνει το σύστημα αναπαραγωγής του ήχου. Τα συστήματα κατηγοριοποιημένα λοιπόν είναι (βλέπε [1]):

1. Συστήματα που δεν είναι ιεραρχικά :

- Dolby Digital – AC3 (Dolby Labs, 2004)
- DTS (Kramer, N.D.)
- Meridian Lossless Packaging¹ (De Lancie, 1998)
- Binaural
- Transaural

2. Ιεραρχικά συστήματα:

- Ambisonics
- Wavefield Synthesis
- Ambiphonics

Τα τρία πρώτα συστήματα της πρώτης κατηγορίας δεν έχουν την δυνατότητα δημιουργίας ηχητικών εικόνων και στις 360⁰ στον οριζόντιο επίπεδο, ενώ ταυτόχρονα δεν έχουν την δυνατότητα δημιουργίας ηχητικών εικόνων που βρίσκονται στο κατακόρυφο επίπεδο. Μπορούν να δημιουργήσουν ακριβείς ηχητικές εικόνες κυρίως στο μπροστινό πεδίο (με έναν περιορισμό στις γωνίες) ενώ μπορούν να δώσουν την αίσθηση ότι κάποιος ήχος προέρχεται γενικά από το πίσω αριστερό ή δεξιό πεδίο. Αντίθετα τα συστήματα Binaural και Transaural της πρώτης κατηγορίας έχει την δυνατότητα να δημιουργεί ακριβής ηχητικές εικόνες και στο οριζόντιο και το κατακόρυφο επίπεδο. Κάποια από τα συστήματα της δεύτερης κατηγορίας έχουν την δυνατότητα δημιουργίας ηχητικών εικόνων και στις τρεις διαστάσεις. Αυτό φυσικά προϋποθέτει την χρήση κατάλληλου συστήματος αναπαραγωγής και διάταξης των ηχείων. Έτσι το σύστημα ονομάζεται **παντοφωνικό (pantophonic system)** (βλέπε [1],[6],[12],[20]) στην περίπτωση που η αναπαραγωγή του ηχητικού πεδίου γίνεται μόνο στο οριζόντιο επίπεδο ενώ αν γίνεται και στις τρεις διαστάσεις, το σύστημα ονομάζεται **περιφωνικό (periphonic system)** (βλέπε [10]). Τα συστήματα που μπορούν να είναι

¹ Το συγκεκριμένο σύστημα θα έχει την δυνατότητα μελλοντικά να ενσωματώσει ιεραρχικές διαδικασίες.

περιφωνικά ανήκουν στην κατηγορία των συστημάτων τρισδιάστατου ήχου (3D Sound Systems).

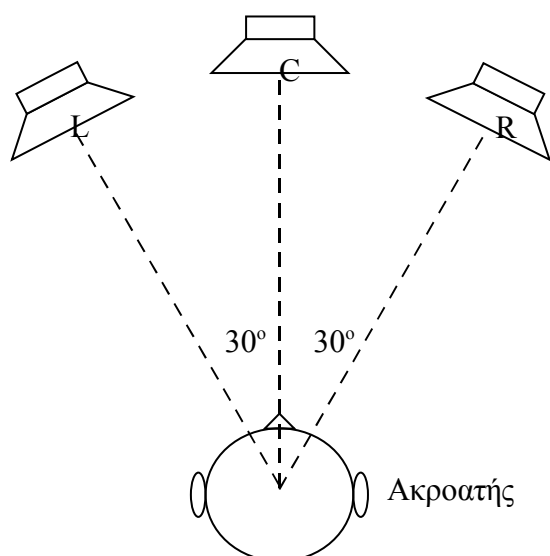
Παρακάτω αναφέρονται περιληπτικά τα παραπάνω συστήματα. Η μέθοδος Ambisonic θα αναλυθεί στο επόμενο κεφάλαιο καθώς οι αλγόριθμοι καταγραφής και αναπαραγωγής του ηχητικού πεδίου βασίστηκαν σε αυτό.

2.2.1 Τρία κανάλια στερεοφωνικού ήχου (3-0)

Το συγκεκριμένο πρωτόκολλο χρησιμοποιείται σπάνια, αν και αποτελεί την βάση για όλα τα επόμενα πολυκάναλα συστήματα, αυτός είναι και ο λόγος που αναφέρεται σε αυτή την παράγραφο αν και δεν αποτελεί σύστημα surround αναπαραγωγής. Αναπτύχθηκε από τον Steinberg και εμφανίστηκε το 1930. Για την υλοποίησή του απαιτεί την χρήση τριών καναλιών, τα οποία θα χρησιμοποιηθούν σαν αριστερό (L), σαν δεξί (R), και σαν κεντρικό (C), και η διάταξη των ηχείων φαίνεται στην εικόνα 7. Τα τρία κανάλια στερεοφωνικού ήχου χρησιμοποιήθηκαν κυρίως στον κινηματογράφο, καθώς έπρεπε να καλυφθεί φαρδύτερος χώρος ακρόασης, και για τον λόγο ότι οι Widescreen έτειναν να έχουν σαν αποτέλεσμα την μεγαλύτερη απόσταση μεταξύ των ηχείων (L) και (R). Τα δύο κανάλια στερεοφωνικού ήχου έγιναν πρότυπο για τα καταναλωτικά προϊόντα για λόγους ευκολίας και οικονομίας, καθώς και για το ότι ήταν ευκολότερος ο διαχωρισμός του αναλογικού δίσκου για δύο κανάλια αντί για τρία.

Τα πλεονεκτήματα του τρικάναλου στερεοφωνικού ήχου είναι πολλά. Το πρώτο είναι ότι δίνει την δυνατότητα μίας πιο φαρδιάς ηχητικής σκηνής, καθώς το κεντρικό ηχείο προσδιορίζει χωρικά συνεχώς το κέντρο, και έτσι υπάρχει η δυνατότητα τοποθέτησης των εξωτερικών ηχείων σε γωνίες έως και ($\pm 45^\circ$). Δεύτερον, παρέχεται η δυνατότητα ενός ευρύτερου πεδίου ακρόασης, καθώς ο ήχος δεν εγκλωβίζεται τόσο εύκολα στο κοντινότερο ηχείο, όπως συμβαίνει στην περίπτωση του δικάναλου στερεοφωνικού ήχου, καθώς υπάρχει το κεντρικό ηχείο. Τρίτον, ο ήχος δεν υφίσταται τόση μεγάλη επιρροή στο ηχόχρωμα του όταν αυτός τοποθετείται στο κέντρο, καθώς η αναπαραγωγή του γίνεται από πραγματική πηγή (C), και όχι από συνδυασμό των (L-R). Τέλος πρέπει να πούμε ότι σε εφαρμογές ήχου για εικόνα, το κεντρικό ηχείο χρησιμοποιείται και για την αναπαραγωγή των διαλόγων. Έτσι ακούγονται πιο καθαρά, και επικεντρώνουν το ενδιαφέρον του ακροατή στην οθόνη (κέντρο).

Ένα πρακτικό πρόβλημα του τρικάναλου στερεοφωνικού ήχου είναι η τοποθέτηση του κεντρικού ηχείου. Στο κινηματογράφο αυτό δεν αποτελεί πρόβλημα, καθώς μπορεί να τοποθετηθεί πίσω από μια διάφανη ακουστικά οθόνη. Στο καταναλωτικό περιβάλλον, σε ένα στούντιο ή σε ένα τηλεοπτικό περιβάλλον, το κεντρικό ηχείο πρέπει να τοποθετηθεί εκεί που θα υπάρχει ένα παράθυρο, μία τηλεόραση ή ένα monitor. Αυτό έχει σαν συνέπεια την τοποθέτηση του κεντρικού ηχείου πάνω ή κάτω από αυτά που αναφέρθηκαν παραπάνω, ή την σμίκρυνση του, σε σχέση με τα υπόλοιπα (βλέπε [1],[6],[12],[20]).



Εικόνα 7: Απεικόνιση της διάταξης των ηχείων για την αναπαραγωγή του (3-0) Stereo. Τα τρία ηχεία ισαπέχουν από τον ακροατή, και η γωνία μεταξύ των εξωτερικών είναι 60° , σύμφωνα με τα πρότυπα διαμόρφωσης του ITU για συμβατότητα με την δικάναλη αναπαραγωγή.

2.2.2 Surround τεσσάρων καναλιών (3-1stereo)

Αυτό το σύστημα έχει τα ίδια μειονεκτήματα με το (3-0 stereo). Έκτος από τα τρία μπροστινά κανάλια έχει προστεθεί και ένα περιφερειακό, το οποίο τροφοδοτεί ένα ηχείο στο πίσω μέρος του κοινού, ή περισσότερα στο πίσω και πλευρικό τμήμα του ακροατή (μονοφωνική τροφοδοσία, καθώς όλα τροφοδοτούνται από το ίδιο σήμα). Ο ρόλος του είναι να δώσει ένα μεγαλύτερο βαθμό ακουστικής εμπλοκής του κοινού στην οπτικοακουστική εμπειρία του έργου, εξασφαλίζοντας ένα κανάλι για περιφερειακά εφε. Δημιουργήθηκε το 1950 από την 20th Century Fox με προορισμό για τον κινηματογράφο, με σκοπό τον αποτελεσματικό ανταγωνισμό στο νέο τηλεοπτικό περιβάλλον.

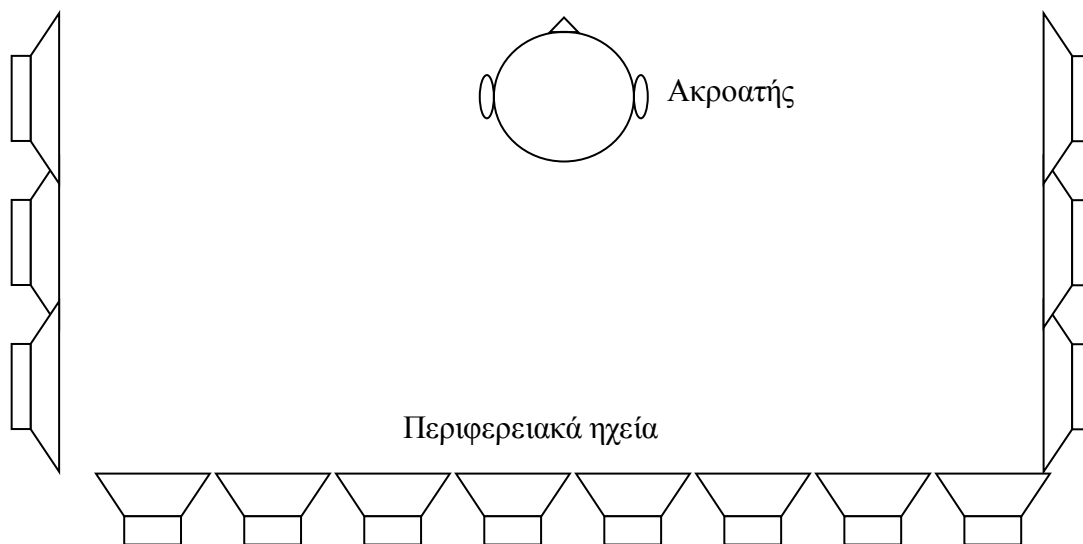
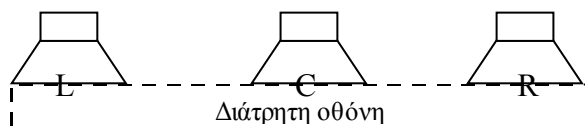
Τέλος πρέπει να πούμε ότι δεν υπήρχε πρόθεση να δημιουργηθεί η αίσθηση της χωροτοποθέτησης ήχων έως τις 360° , πράγμα που θα ήταν και ανέφικτο, καθώς έχουμε ένα κανάλι που τροφοδοτεί περισσότερα από ένα ηχεία (βλέπε [1],[6],[12],[20]).

Διαμόρφωση των ηχείων.

Η διαμόρφωση των ηχείων φαίνεται στην εικόνα 8. Σε αυτό το πρωτόκολλο γίνεται τροφοδοσία περισσότερων του ενός περιφερειακών ηχείων από το surround κανάλι, με σκοπό της διεύρυνσης του πεδίου ακρόασης. Αυτό είχε την τάση να δημιουργεί μία πιο διάχυτη και κατανομημένη αναπαραγωγή αυτών των εφέ σημάτων. Κάποιες φορές γινόταν και ηλεκτρονικός αποσυντονισμός του ηχητικού σήματος που τροφοδοτούσε το κάθε περιφερειακό ηχείο, ώστε το πεδίο που δημιουργούν να είναι διάχυτο, και έτσι να μην γίνεται αντιληπτό από τον ακροατή ότι το εφε προέρχεται από το ηχείο που βρίσκεται πιο κοντά του, ή να αντιλαμβάνεται τον ήχο μέσα στο κεφάλι του.

Στις καταναλωτικές υλοποιήσεις αυτού του πρωτοκόλλου, το μονοφωνικό surround

κανάλι τροφοδοτεί δύο ηχεία, τοποθετημένα όπως στο πρωτόκολλο (3-2) που θα αναλυθεί παρακάτω. Πρέπει τέλος να πούμε ότι σε αυτή την περίπτωση γίνεται μείωση της ηχητικής έντασης του σήματος που τροφοδοτεί τα surround ηχεία, έτσι ώστε λόγω της ακουστικής άθροισης των δύο ηχείων, να μην υπάρχει κακή ισορροπία στην ένταση μεταξύ του μπροστινού και του πίσω ηχητικού πεδίου (βλέπε [1],[6],[12],[20]).



Εικόνα 8: Το μονοφωνικό S (Surround) κανάλι τροφοδοτεί μια παράταξη από περιφερειακά ηχεία, έτσι ώστε να κάνει πιο διάχυτη τον ήχο των περιφερειακών εφε. Στη περίπτωση της διαμόρφωσης για καταναλωτικό περιβάλλον, γίνεται αναπαραγωγή από δύο ηχεία, με τεχνητό αποσυντονισμό, ή από δίπολα ηχεία, ώστε να προσομοιωθεί η διάχυτη αίσθηση του κινηματογράφου.

Περιορισμοί της τετρακάναλης αναπαραγωγής.

Ο κύριος περιορισμός της τετρακάναλης αναπαραγωγής είναι ο περιορισμός του ενός μόνο surround καναλιού. Παρά το γεγονός ότι η αναπαραγωγή του surround καναλιού γίνεται από περισσότερα από ένα ηχεία, ο άνθρωπος για να αντιληφθεί το περιβάλλον χρειάζεται δύο ξεχωριστά ασυσχέτιστα σήματα και από τις δύο πλευρές. Ψυχοακουστικές μελέτες έχουν αποδείξει ότι για την δημιουργία της αίσθησης του χώρου και για αίσθηση της χωροτοποθέτησης των εφε από το εφε, χρειάζονται τουλάχιστον δύο surround κανάλια.

2.2.3 5.1 κανάλια surround (3-2 stereo).

Το τετρακάναλο σύστημα αναπαραγωγής είχε το μειονέκτημα ότι περιελάμβανε ένα μονοφωνικό κανάλι για το surround. Έτσι το πρωτόκολλο 5.1 ήρθε για να δώσει λύση σε αυτό το πρόβλημα. Αυτό το σύστημα είναι προσανατολισμένο προς τα μπροστά, όπου όπως τα προαναφερθέντα συστήματα, δημιουργούν την εικόνα του ήχου βασισμένα στο τρικάναλο στέρεο, ενώ τα δύο περιφερειακά ηχεία χρησιμεύουν για την δημιουργία της αίσθησης του περιβάλλοντα χώρου, και των εφε. Αυτό σημαίνει ότι ο προορισμός αυτού του πρωτοκόλλου δεν ήταν να δώσει την δυνατότητα της χωροτοποθέτησης ήχων στις 360°, αλλά για να εμπλουτίσει ένα τρικάναλο σύστημα με εφε ή χώρο από τα δύο πίσω ηχεία. Γι' αυτό και πολλά πρότυπα επιμένουν να το ονομάζουν (3-2 stereo) αντί για 5.1 surround.

Ο διαχωρισμός μεταξύ του μπροστά και του πίσω αυτού του πρωτοκόλλου είναι ένας τομέας που πολλοί χρήστες του αγνοούν. Στη περίπτωση του δικάναλου ήχου είναι σχετικά εύκολη η μοντελοποίηση και η θεωρητική προσέγγιση για διανύσματα εντοπισμού των ήχων, σε οποιαδήποτε γωνία ανάμεσα στα ηχεία. Δεν είναι όμως το ίδιο εύκολη η μοντελοποίηση για πέντε κανάλια ήχου, καθώς τα ηχεία βρίσκονται σε διαφορετικές γωνίες μεταξύ τους, και η γωνία μεταξύ των πίσω (surround) ηχείων είναι ιδιαίτερα μεγάλη. Είναι, όμως, εφικτή η χρήση διαφορετικής σχέσης έντασης και φάσης ανάμεσα στα πέντε ηχεία για την προσομοίωση της διαφορετικής γωνίας προέλευσης του ήχου, όπως στα συστήματα Ambisonics. Η διαφορετική γωνία των ηχείων, όμως, θα οδηγήσει στη μη σταθερά αξιόπιστη δημιουργία της εικονικής θέσης προέλευσης του ήχου, ανάλογα με τον τομέα προέλευσής του. Συνίσταται λοιπόν, για αυτούς που δεν έχουν πρόσβαση σε πολύπλοκους νόμους χωροτοποθέτησης ή σε ψυχοακουστικά μοντέλα τροφοδότησης των πέντε καναλιών ανάλογα με την εικονική θέση του ήχου, να χρησιμοποιούν το συγκεκριμένο πρωτόκολλο όπως αυτό προοριζόταν για το κινηματογράφο, δηλαδή χρησιμοποιώντας τα μπροστινά κανάλια για την χωροτοποθέτηση των ήχων, και τα δύο πίσω για εφε και την δημιουργία της αίσθησης του χώρου. Ακόμα όμως και στη δεύτερη περίπτωση, είναι δυνατόν να δημιουργηθεί αρκετά πετυχημένη ψευδαίσθηση του χώρου, και να εξασφαλιστεί αρκετά καλός εντοπισμός της εικονικής θέσης προέλευσης των ήχων.

Τέλος, όσο αναφορά τον όρο .1 που υπάρχει σε αυτό το πρωτόκολλο, αφορά ένα επιπλέον κανάλι που υπάρχει LFE και αφορά εφε για τις χαμηλές συχνότητες. Ο συμβολισμός .1 προέρχεται από το περιορισμένο συχνοτικό εύρος αυτού του καναλιού σε σχέση με τα υπόλοιπα. Η διεθνής πρότυπη ονομασία για το surround 5.1 είναι (3-2-1 stereo) όπου το τελευταίο ψηφίο υποδηλώνει τον αριθμό των LFE καναλιών (βλέπε [1],[6],[12],[20]).

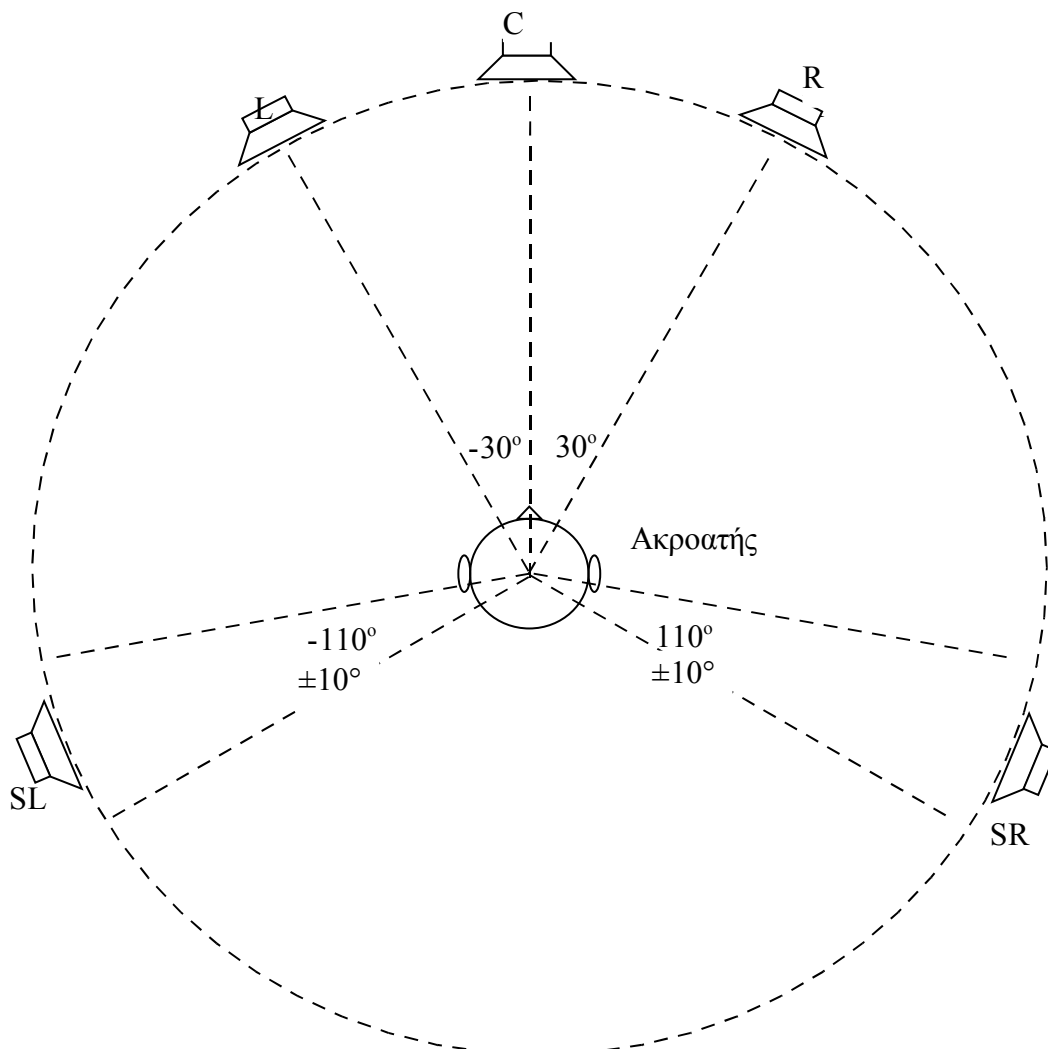
Παγκόσμια πρότυπα και σχετική διάταξη ηχείων.

Η σχετική διάταξη των ηχείων και η ρύθμιση των καναλιών περιγράφεται από τα πρότυπα ITU-R BS.775 (εικόνα 9). Το αριστερό και δεξί κανάλι τοποθετούνται στις $\pm 30^\circ$ ώστε να υπάρχει συμβατότητα με τα προηγούμενα στερεοφωνικά πρωτόκολλα. Για πολλά χρόνια αυτή η προσπάθεια για διατήρηση της συμβατότητας με τα προηγούμενα στερεοφωνικά πρότυπα, έχει σαν αποτέλεσμα να στενέψει η μπροστινή ηχητική σκηνή, ενώ λόγω του κεντρικού ηχείου, θα μπορούσε να είναι πιο ευρεία.

Τα surround ηχεία είναι τοποθετημένα στις 110° , και ο ρόλος τους είναι να εξασφαλίσουν την δυνατότητα της χωροτοποθέτησης των εφε πίσω από τον ακροατή και να διασφαλίσουν την ακουστική ενέργεια που χρειάζεται για την δημιουργία της αίσθησης της ηχητικής περιβολής του ακροατή. Στο πρότυπο 5.1 δεν υπάρχει κανονικά κανένα ηχείο ακριβώς πίσω από τον ακροατή. Έτσι η Dolby πρότεινε ένα νέο πρωτόκολλο που το ονόμασε EX και

περιλαμβάνει ένα επιπρόσθετο ηχείο στο κεντρικό πίσω μέρος του ακροατή. Το ITU πρότυπο επιτρέπει την προσθήκη επιπλέον ηχείων για την κάλυψη του περιβάλλοντος χώρου, ακριβώς όπως και στο πρότυπο (3-1 stereo) που έχει προαναφερθεί. Στη περίπτωση αυτή, τα επιπρόσθετα ηχεία θα πρέπει να τοποθετηθούν στις $\pm 60^\circ$ και $\pm 150^\circ$.

Όλα τα ηχεία πρέπει να είναι όμοια, ώστε να επιτευχθεί η ηχητική ομοιομορφία και ποιότητα γύρω από τον ακροατή. Υπάρχουν κάποιες συζητήσεις για την χρήση δίπολων ηχείων στις surround θέσεις. Αυτά τα ηχεία έχουν ένα πολικό διάγραμμα που μοιάζει με την εικόνα οκτώ, και για να επιτευχθεί πιο διάχυτη αίσθηση των εφε, προσανατολίζουν το μηδενικό σημείο της εικόνα οκτώ στη διεύθυνση του ακροατή. Ο ακροατής, λοιπόν, λαμβάνει περισσότερο διάχυτο ήχο παρά απευθείας, και του δίνεται η αίσθηση ενός πιο ευρύχωρου περιβάλλοντος πεδίου, που προσομοιάζει περισσότερο την αίσθηση που δίνεται σε μία κινηματογραφική αίθουσα. Παρ' όλα αυτά, η χρήση των δίπολων ηχείων, καθιστά δυσκολότερη την ακριβή χωροτοποθέτηση ηχητικών δρώντων στο πίσω μέρος του ακροατή (βλέπε [1],[6],[12],[20]).



Εικόνα 9: Απεικόνιση της σχετικής διάταξης των ηχείων για το πρότυπο Surround 5.1 (3-2-1 stereo) σύμφωνα με το πρότυπο ITU-R BS.775, το οποίο υποστηρίζει δύο ξεχωριστά Surround κανάλια τα οποία τροφοδοτούν ένα ή περισσότερα ηχεία ανά κανάλι.

Το LFE κανάλι και η χρήση των Μονάδα αναπαραγωγής χαμηλών συχνοτήτων (subwoofer).

Το LFE είναι ένα επιπλέον κανάλι υπο-μπάσων (sub-bass), με ένα ανώτερο συχνοτικά όρια στα 120 Hz. Προορίζεται για ηχητικά εφε για χαμηλές συχνότητες που χρειάζονται μεγαλύτερο Headroom και μεγαλύτερα επίπεδα ηχητικής πίεσης απ' ότι τα υπόλοιπα κανάλια. Δεν προορίζεται για να μεταφέρει κύρια συστατικά από τις εντάσεις των υπόλοιπων καναλιών, αλλά για την μεταφορά ειδικών εφε, που στην περίπτωση των εφαρμογών ήχου σε εικόνα, μπορεί να είναι εκρήξεις ή άλλοι υπόκωφοι θόρυβοι.

Σε καταναλωτικά συστήματα αναπαραγωγής του ήχου η αναπαραγωγή του LFE καναλιού είναι προαιρετική. Γι' αυτό το λόγο, οι μίξεις πρέπει να εξασφαλίζουν την ικανοποιητική ηχητική αναδόμηση του υλικού, ακόμα και στην περίπτωση που δεν αναπαράγεται το LFE.

Στην αναπαραγωγή στον κινηματογράφο, γίνεται προ-ενίσχυση του LFE καναλιού κατά 10 dB σε σχέση με τα υπόλοιπα κανάλια. Αυτό επιτυγχάνεται με ενίσχυση του σήματος του καναλιού που θα αναπαραχθεί, και όχι της έντασης ηχογράφησης του.

Περιορισμοί της αναπαραγωγής των 5.1 καναλιών.

Ο κύριος περιορισμός της αναπαραγωγής των 5.1 καναλιών, είναι ότι ο αρχικός προορισμός του δεν ήταν για να υποστηρίξει την ακριβή δημιουργία ηχητικών ειδώλων στις 360°. Ένας άλλος περιορισμός είναι ότι η ηχητική εικόνα στο μπροστινό μέρος είναι στενότερη απ' ότι θα ήταν μα δεν ήταν απαραίτητη η συμβατότητα με το πρότυπο (2-0) stereo. Τρίτος περιορισμός είναι ότι το κεντρικό ηχείο αποδεικνύεται προβληματικό στην ισορροπία της αναπαραγωγής στερεοφωνικής μουσικής. Ο λόγος είναι ότι νόμοι χωροτοποθέτησης και τεχνικές ηχογράφησης με συμπτωτικά μικρόφωνα, δεν έχουν βελτιστοποιηθεί ακόμα για αναπαραγωγή από τρία ηχεία, και οι είδη υπάρχοντες αναφέρονται στην αναπαραγωγή από δύο. Η απλή γεφύρωση του κεντρικού ηχείου ανάμεσα στο αριστερό και δεξί κανάλι, έχει σαν αποτέλεσμα την στένωση της μπροστινής ηχητικής εικόνας, σε σχέση με αυτήν που θα δημιουργώταν από την αναπαραγωγή από το πρωτόκολλο (2-0) stereo του ίδιου υλικού. Ένας τέταρτος περιορισμός είναι ότι η θέση των ηχείων LS και RS είναι συμβιβαστική, πράγμα που οδηγεί στην δημιουργία μια δυναμικής τρύπας της ηχητικής εικόνας στο πίσω μέρος του ακροατή, και προβλήματα στην εύρεση μια φυσικής θέσης για την τοποθέτηση τους σε πρακτικά δωμάτια (βλέπε [1],[6],[12],[20]).

Ένταση σημάτων στο πρωτόκολλο 5.1 surround.

Σε περιβάλλοντα ήχου για ταινίες, ο κανόνας είναι να γίνεται προενίσχυση της έντασης ηχογράφησης για τα surround κανάλια κατά 3 dB σε σχέση με την ένταση ηχογράφησης των μπροστινών. Αυτό γίνεται για να αντισταθμιστεί η εξασθένηση κατά 3 dB, για λόγους ακουστικής ευθυγράμμισης, που λαμβάνει χώρα κατά την αναπαραγωγή σε θέατρα. Είναι σημαντικό να γίνει αντιληπτή αυτή η ασυμφωνία ανάμεσα στις πρακτικές, καθώς είναι κανόνας στην μίξη της μουσικής και στα τηλεοπτικά, να γίνεται ευθυγράμμιση των καναλιών, για ίση ένταση και στο μέσο εγγραφής αλλά και στο μέσο ακουστικής αναπαραγωγής. Μετατροπή από master ταινιών σε καταναλωτικά ή τηλεοπτικά μέσα, προϋποθέτουν την μεταβολή της προενίσχυσης των surround καναλιών (βλέπε [1],[6],[12],[20]).

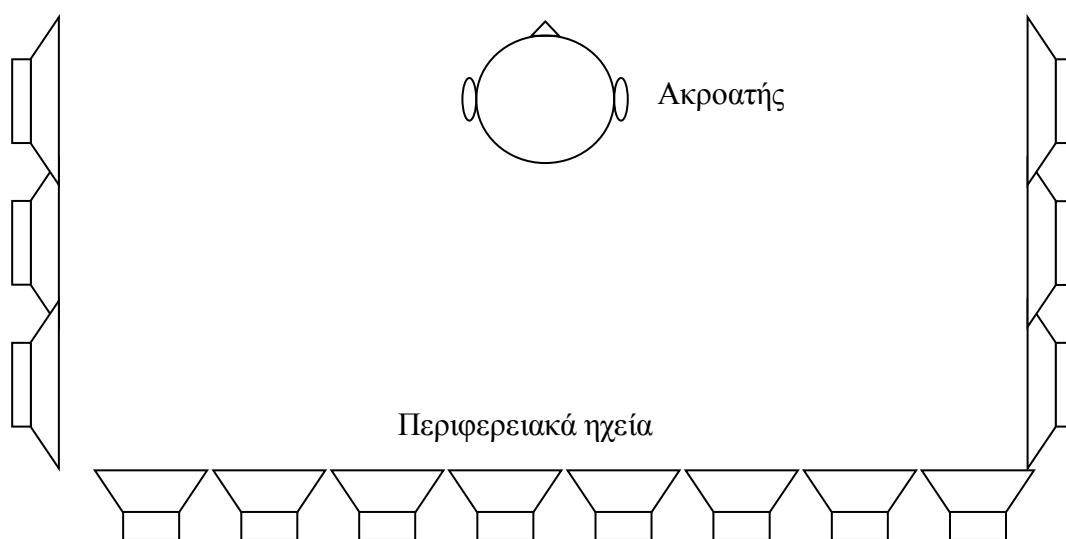
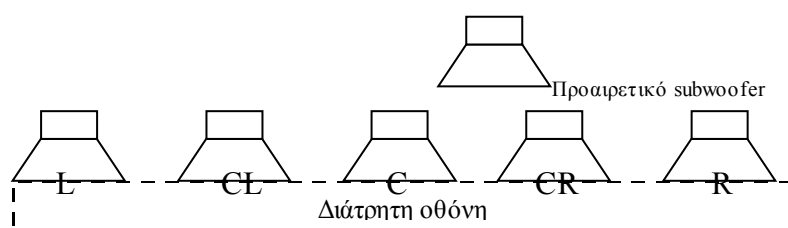
2.2.4 Πρωτόκολλο 7.1 καναλιών surround.

Το συγκεκριμένο πρωτόκολλο προέρχεται από το widescreen πρωτόκολλο για τον κινηματογράφο, προσθέτοντας στη διαμόρφωση του 5.1 surround δύο επιπλέον καναλιών, το LC και RC. Η σχετική διάταξη του συγκεκριμένου πρωτοκόλλου απεικονίζεται στην εικόνα 10. Αυτό το πρωτόκολλο δεν προοριζόταν αρχικά για καταναλωτικές εφαρμογές, αλλά για μεγάλες αίθουσες όπου λόγω του μεγάλου φάρδους της οθόνης χρειάζονταν επιπρόσθετα ηχεία για να καλύψουν επαρκώς όλες τις γωνίες ανάμεσα στα ηχεία, για όλες τις θέσεις του ακροατηρίου στην αίθουσα. Τα συστήματα της sony SDDS είναι ένας κοινός κάτοχος υλοποίησης αυτού του πρωτοκόλλου, καθώς είναι το γνήσιο πρότυπο 70 mm Dolby stereo.

Η Lexicon και η Meridian επίσης υλοποίησαν μια τεχνοτροπία με την οποία υποστήριζαν επτά κανάλια στους καταναλωτικούς αποκωδικοποιητές surround. Η τοποθέτηση που προτείνουν όμως διαφέρει από αυτή των εφαρμογών για τον κινηματογράφο. Τα επιπρόσθετα ηχεία χρησιμοποιούνται για να εξασφαλίσουν μία πιο φαρδιά πλευρική και μπροστινή σύνθεση, και να επιτρέψει στα surround ηχεία να τοποθετηθούν περισσότερο προς το πίσω μέρος, σε σχέση με την σύνθεση του 5.1 surround προτύπου (βλέπε [1],[6],[12],[20]).

2.2.5 10.2 κανάλια surround.

Ο Tomilson Holman έχει κάνει μια αξιοσημείωτη προσπάθεια προώθησης ενός συστήματος 10.2 καναλιών surround, ως το σύστημα αναπαραγωγής που θα επικρατήσει μελλοντικά για την ηχητική χωρική αναπαράσταση. Δεν έχει όμως υιοθετηθεί ως πρότυπο ακόμα. Στη βασική σύνθεση για 5.1 κανάλια έχει προσθέσει δύο πιο φαρδιά τοποθετημένα ηχεία μπροστά-πλευρικά, καθώς και ένα στο πίσω κεντρικό σημείο, για την κάλυψη της ακουστικής 'τρύπας' που διαθέτει το πρότυπο. Επιπλέον έχει προσθέσει δύο κανάλια 'ύψους' και ένα επιπλέον κανάλι LFE. Το δεύτερο LFE κανάλι έχει προοριστεί για να εξασφαλίσει πλευρικό διαχωρισμό των δύο ασυσχέτιστων περιεχομένων των χαμηλών συχνοτήτων και από τις δύο περιοχές της ακουστικής περιοχής, όπως προτείνεται από τον Griesinger για τον εμπλουτισμό της ευρυχωρίας των χαμηλών συχνοτήτων (βλέπε [1],[6],[12],[20]).



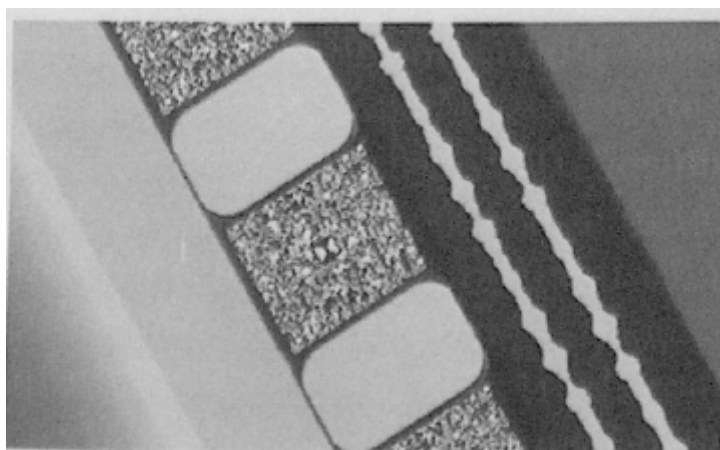
Εικόνα 10: Απεικόνιση του εμπλουτισμού της μπροστινής εικόνας σε μεγάλες αίθουσες από κάποια ηχητικά πρωτόκολλα για κινηματογράφο, με την προσθήκη δύο επιπλέον κανάλια στο μπροστινό τμήμα, σαν Center Left και Center Right.

2.2.6 Ψηφιακά πρωτόκολλα μεταφοράς πολυκάναλου ήχου

Dolby Digital

Η Dolby Digital ή AC3 κωδικοποιείται με στόχο την μεταφορά του 5.1 surround πρωτοκόλλου στον κινηματογράφο ή στο σπίτι, χωρίς να είναι απαραίτητη η χρήση αναλογικής μαθηματικής κωδικοποίησης. Ήταν αναμενόμενη να αντικαταστήσει το Dolby Stereo καθώς γίνεται σταδιακή αντικατάσταση των αναλογικών συστημάτων από τα ψηφιακά. Βασίζεται σε μια ψηφιακή, low-bit-rate κωδικοποίηση και αποκωδικοποίηση, η οποία δίνει την δυνατότητα μεταφοράς πολλών καναλιών surround μίξης, χωρίς τα προβλήματα διαχωρισμού και άλλων που υπάρχουν στα μαθηματικώς κωδικοποιημένα surround. Η Dolby Digital μπορεί να κωδικοποιήσει κανάλια βασισμένα στο ITU-πρότυπο 3-2-1 surround πρωτόκολλο των ηχείων και πρέπει να διακριθεί από τέτοια παγκόσμια πρότυπα καθώς είναι κυρίως μια μέθοδος αναπαράστασης κωδικοποιημένων σημάτων. Στη πραγματικότητα ο αλγόριθμος κωδικοποίησης AC-3 μπορεί να χρησιμοποιηθεί για ένα μεγάλο εύρος διαφορετικών διαμορφώσεων ηχητικών σημάτων και ποικιλία bit rate από 32 kbit/s για ένα μόνο κανάλι, έως και 640 kbit/s για surround σήματα.

Χρησιμοποιείται ευρέως για την διανομή των ψηφιακών ηχητικών καναλιών σε ένα φιλμ για ταινία 35 mm, όπου τα δεδομένα είναι αποθηκευμένα οπτικά στον χώρο ανάμεσα στις πλαϊνές τρύπες του φιλμ, όπως φαίνεται στο εικόνα 11. Με αυτό τον τρόπο, τα αναλογικά κανάλια του ήχου μπορούν να παραμείνουν στην κανονική τους θέση, παραπλεύρως της εικόνας, για λόγους συμβατότητας. Σε αυτό το πρότυπο έχει συνδυαστεί μία Dolby-SR κωδικοποιημένη αναλογικά Dolby Stereo μίξη, και το συνδυασμένο πρωτόκολλο που ονομάζεται Dolby SR-D. Το Dolby Digital χρησιμοποιείται επίσης για surround ήχο σε παρουσιάσεις DVD βίντεο, και σε ιδιαίτερες ψηφιακές τηλεοπτικές εφαρμογές.



Εικόνα 11: Τα δεδομένα του Dolby Digital είναι αποθηκευμένα οπτικά στον χώρο ανάμεσα στις τρύπες της ταινίας ενός φιλμ 35 mm

Ο έλεγχος της Dolby Digital κωδικοποίησης μπορεί να επιτευχθεί με την χρήση εφαρμογών λογισμικού την τροποποίηση πολλών παραμέτρων της διαδικασίας κωδικοποίησης. Το Dolby Digital μπορεί να λειτουργήσει σε δειγματοληψίες 32, 44.1 ή 48 kHz, και το LFE κανάλι δειγματοληπτείται στα 240 Hz (λόγο της περιορισμένου συχνοτικού του εύρους). Μια αλλαγή φάσης 90° συνήθως εφαρμόζεται και στο κάθε surround κανάλι, κατά την διάρκεια της κωδικοποίησης, το οποίο προφανώς βελτιώνει την μπροστά/πίσω χωροτοποθέτηση και μειώνει το crosstalk ανάμεσα στο κεντρικό και τα surround κανάλια κατά την διάρκεια την αποκωδικοποίησης από ένα Dolby Surround. Γι' αυτό το λόγο είναι σημαντική ή ακρόαση της εγγραφής κατά την διάρκεια της διαδικασίας κωδικοποίησης/αποκωδικοποίησης, για να διασφαλιστεί ότι αυτή η αλλαγή φάσης δεν επηρεάζει τις χωρικές προθέσεις του δημιουργού.

Εκτός από την αναπαράσταση των surround καναλιών στην σύνθετη ψηφιακή μορφή, η Dolby Digital περιείχε μια πληθώρα λειτουργικών χαρακτηριστικών τα οποία εμπλουτίζουν την ευελιξία του συστήματος και βοηθάει στην προσαρμογή της αναπαραγωγής για διάφορες καταναλωτικές καταστάσεις. Αυτό περιλαμβάνει dialogue normalization και την επιλογή να περιλαμβάνονται πληροφορίες δυναμικού εύρους, παράλληλα με τις ηχητικές πληροφορίες, για χρήση σε περιβάλλοντα όπου λόγω του θορύβου βάθους, δεν είναι δυνατή η ακρόαση όλου του δυναμικού εύρους του αρχικού υλικού. Επίσης μπορούν να περιλαμβάνονται πληροφορίες, παράλληλα με της πληροφορίες ήχου, με στόχο να μπορεί ο αποκωδικοποιητής να ξαναμιξάρει το υλικό για surround ήχο, ώστε να αναπαραχθεί μόνο από δύο κανάλια. Σαν κανόνας, οι πληροφορίες του Dolby Digital είναι αποθηκευμένες ή μεταδίδονται με τον μέγιστο αριθμό καναλιών που χρειάζονται για να παρουσιαστεί το τελικό υλικό, και κάθε συμβατό mixdown δημιουργείται στον

αποκωδικοποιητή. Αυτό διαφέρει από κάποια άλλα συστήματα όπου η μίξη για δύο κανάλια μεταφέρονται ταυτόχρονα με τις ηχητικές πληροφορίες του surround υλικού.

Η ένδειξη dialnorm μπορεί να χρησιμοποιηθεί σε τηλεοπτικά και άλλα υλικά για να διασφαλίσουν ότι η ένταση του ήχου θα παραμείνει περίπου σταθερή από πρόγραμμα σε πρόγραμμα. Έχει υποθεθεί ότι η ένταση των διαλόγων είναι ο κύριος παράγοντας που ελέγχει την ακουστική ένταση που χρησιμοποιείται στην σπίτια των ανθρώπων, και δεν θέλουν συνεχώς να την αλλάζουν, καθώς αλλάζουν τα προγράμματα (πχ από διαφημιστικό σε ειδήσεις). Η ένταση του dialnorm είναι ο μέσος όρος της έντασης των διαλόγων καθ' όλη την διάρκεια του προγράμματος, σε σύγκριση με την μέγιστη ένταση η οποία θα ήταν πιθανή, μετρήσιμη χρησιμοποιώντας ένα A τύπου L_{EQ} . Έτσι, για παράδειγμα, εάν ο μέσος όρος των διαλόγων κατά την διάρκεια του προγράμματος ήταν 70 dBA, και τα SPL που αντιστοιχούν στην μέγιστη πιθανή ένταση ήταν 100 dBA, η Dialnorm επιλογή θα ήταν -30 dB (βλέπε [1],[6],[12],[20]).

2.2.7 DTS

Το DTS (Digital Theater Systems) σύστημα είναι ένα ακόμη ψηφιακό πρωτόκολλο κωδικοποίησης, το οποίο μπορεί να μεταφέρει surround ήχο σε καταναλωτικές ή επαγγελματικές εφαρμογές, χρησιμοποιώντας μία low bit rate τεχνική κωδικοποίησης για την μείωση του data rate της ηχητικής πληροφορίας. Το σύστημα DTS μπορεί να υποστηρίξει ένα μεγάλο αριθμό bit rate από 32 kbits/ μέχρι και 4.096 Mbit/s με υποστήριξη μέχρι οκτώ κυρίων καναλιών και συχνότητες δειγματοληψίας μέχρι τα 192 kHz. Variable bit rate και lossless κωδικοποίηση είναι επίσης προαιρετική. Προβλέπονται επίσης στο σύστημα Downmixing και επιλογές ρύθμισης του δυναμικού εύρους.

Δεδομένα DTS βρίσκονται σε κάποιες παρουσιάσεις ταινιών και καταλαμβάνουν διαφορετική περιοχή στο φιλμ απ' ότι τα δεδομένα Dolby Digital και SDDS. Παρ' όλα αυτά είναι πιθανό να υπάρχουν σε μία εκτύπωση ταινίας σε μία multi-format έκδοση, και τα τρία ψηφιακά ηχητικά πρωτόκολλα συν τα αναλογικά Dolby Stereo κανάλια ταυτόχρονα στο ίδιο φιλμ, κάνοντάς το καθολικά συμβατό. DTS χρησιμοποιείται επίσης και σε έναν αριθμό από παρουσιάσεις surround CD, και είναι προαιρετικό στο DVD, απαιτώντας έναν ειδικό αποκωδικοποιητή για να αναπαράγει τα σήματα πληροφορίας από τον δίσκο (βλέπε [1],[6],[12],[20]).

SDDS

Το πρότυπο SDDS (Sony Dynamic Digital Sound) είναι το τρίτο από τα κύρια ανταγωνιστικά πρωτόκολλα για ψηφιακό ήχο για ταινίες. Χρησιμοποιώντας το ATRAC σύστημα μείωσης δεδομένων της Sony, και αυτό κωδικοποιεί τα δεδομένα με σημαντική εξοικονόμηση bit rate σε σύγκριση με το γνήσιο PCM (για πακετάρισμα 5:1). Το σύστημα SDDS υποστηρίζει 7.1 κανάλια αντί για 5.1, εξασφαλίζοντας λεπτομερή θέση κάλυψης της μπροστινής σκηνής. Δεν είναι σύνηθες να χρησιμοποιηθούν δεδομένα SDDS σε τίποτα άλλο εκτός από παρουσιάσεις εκτύπωσης ταινιών, παρόλο που θα μπορούσε να περιλαμβάνονται σε ένα DVD σαν ένα ιδιοκτησιακό πρωτόκολλο εάν χρειαστεί. Καταναλωτικοί αποκωδικοποιητές δεν είναι ακόμη διαθέσιμοι (βλέπε [1],[6],[12],[20]).

MPEG surround modes.

Το πρότυπο MPEG (Moving Pictures Expert Group) είναι χρησιμοποιούνται ευρέως για low bit rate αναπαραστάσεις σημάτων ήχου και εικόνας σε πολυμεσικές και άλλες εφαρμογές. Ενώ το MPEG-1 αποτελεί ένα δικάναλο πρωτόκολλο, MPEG-2 επεκτάθηκε σε πολυκαναλικές πληροφορίες. Υπάρχουν δύο εκδόσεις του MPEG-2, εκ των οποίων η μία αναπτύχθηκε με σκοπό

την συμβατότητα με τους αποκωδικοποιητές MPEG-1, και το άλλο, γνωστό ως MPEG-2 AAC (Advanced Audio Coding), που δεν είναι συμβατό με προηγούμενα πρωτόκολλα. Το πρότυπο MPEG-4 επίσης περιλαμβάνει κλιμακωτές επιλογές για πολυκάναλη κωδικοποίηση. Αυτά τα πρότυπα, περιγράφονται με λεπτομέρεια στα ISO 11172-3, 13818-3, 13818-7 και 14496 για αυτούς που θέλουν να καταλάβουν πως λειτουργούν αυτά, με λεπτομέρεια.

Η έκδοση MPEG-2 BC (Backwards Compatible) λειτουργεί με την κωδικοποίηση ενός μαθηματικά κωδικοποιημένου downmix, των surround καναλιών και του κεντρικού στα κανάλια left/right, ενός συμβατού με το MPEG-1 σχεδίου δομής. Αυτό μπορεί να αποκωδικοποιηθεί από ένα συμβατικό MPEG-1 αποκωδικοποιητή. Ένα τμήμα επέκτασης προστέθηκε στο τέλος του κουτιού, περιλαμβάνοντας μόνο τα κανάλια C, LS, RS. Κατά την αποκωδικοποίηση σε έναν MPEG-2 surround αποκωδικοποιητή, τα τρία πρόσθετα στοιχεία μπορούν να αναδομηθούν ξανά από τα σήματα L_0 και R_0 για την δημιουργία των γνήσιων πέντε καναλιών. Το κύριο πρόβλημα με τον MPEG-2 BC είναι (α) ότι το downmix γινόταν στον κωδικοποιητή, και έτσι δεν μπορούσε να αλλαχθεί στον αποκωδικοποιητή, (β) το data rate ήταν αξιοσημείωτα μεγάλο λόγω της ηθελημένης συμβατότητας με τα προηγούμενα πρότυπα, πράγμα που δεν θα συνέβαινε αν αυτό δεν ήταν θεμιτό. Συνεπώς το bit rate που χρειάζεται για την μεταφορά 5.1 καναλιών surround με λογική ποιότητα είναι της τάξης των 600-900 kbit/s.

Το MPEG-2 AAC, από την άλλη, είναι ένας πιο πολύπλοκος αλγόριθμος ο οποίος κωδικοποιεί τον πολυκάναλο ήχο για να δημιουργήσει ένα μονό bit stream ο οποίος αναπαριστά όλα τα κανάλια, με μία μορφή η οποία δεν είναι ικανή να αποκωδικοποιηθεί από έναν MPEG-1 αποκωδικοποιητή. Έχοντας παραμερίσει την απαίτηση της συμβατότητας με προηγούμενα πρωτόκολλα, το bit rate μπορεί τώρα να βελτιστοποιηθεί κωδικοποιώντας τα κανάλια σαν ένα γκρουπ, και χρησιμοποιώντας το πλεονέκτημα των εσω-καναλικών επαναλήψεων, εάν χρειαστεί. Η λειτουργία του τώρα μοιάζει περισσότερο με αυτήν του Dolby Digital, και τα bit rate που απαιτούνται για αποδεκτή ηχητική ποιότητα είναι παρόμοια. Το σύστημα MPEG-2 AAC έχει αναπτυχθεί με συνεισφορά πολλών διαφορετικών κατασκευαστών.

Ο αλγόριθμος MPEG surround δεν έχει υλοποιηθεί ακόμα ευρέως στις τηλεοπτικές, ταινίες και καταναλωτικές εφαρμογές. Παρόλο που το MPEG-2 BC προοριζόταν για χρήση σε εφαρμογές DVD στις χώρες που ανήκουν στο Region 2 (κυρίως Ευρώπη), αυτό δεν πραγματοποιήθηκε για τον λόγο της επικράτησης του Dolby Digital. Το πρότυπο δικάναλου MPEG, σαν το MPEG-1, Layer 3 (γνωστό ως .MP3 πρότυπο), από την άλλη, έχει ευρέως υιοθετηθεί για καταναλωτικούς σκοπούς (βλέπε [1],[6],[12],[20]).

Meridian Lossless Packing (MLP).

Η Meridian Lossless Packing (MLP) είναι μία τεχνική μείωσης των δεδομένων χωρίς απώλειες, για πολυκάναλο ήχο, έχοντας την άδεια από την Meridian Audio μέσω των Dolby Labs. Έχει προδιαγραφεί για πρωτόκολλα DVD-Audio σαν ένας τρόπος μείωσης του data rate που χρειάζεται για υψηλής ποιότητας εγγραφές, χωρίς καμία επιρροή στην ποιότητα (Με άλλα λόγια γίνεται αναδόμηση ακριβώς των αρχικών bit, το οποίο δεν συμβαίνει στις περιπτώσεις των Dolby Digital και MPEG). Χρησιμοποιώντας αυτή την τεχνική μπορεί να εξασφαλιστεί ένας μεγάλος χρόνος αναπαραγωγής στον δίσκο, ενώ παράλληλα εξασφαλίζεται μεγάλη ποιότητα μουσικής (συχνότητες δειγματοληψίας μέχρι και 192 kHz και ανάλυση ανάμεσα στα 16 και 24 bits). Τέλος πρέπει να πούμε ότι υποστηρίζονται μέχρι 6 κανάλια surround ήχου.

Το MLP δίνει την δυνατότητα μηχανικούς mastering να δημιουργήσουν ένα πολύπλοκο downmix (για δικάναλη αναπαραγωγή) του πολυκάναλου υλικού, το οποίο καταλαμβάνει λίγο περισσότερο χώρο στο δίσκο, λόγω της εκμετάλλευσης βιομηχανικά των ομοιοτήτων ανάμεσα σε αυτό το υλικό και την πολυκάναλη έκδοση κατά την διάρκεια της lossless κωδικοποίησης. Αυτό το

downmix μπορεί να έχει χαρακτηριστικά τα οποία διαφοροποιούνται κατά την διάρκεια του προγράμματος, και βασίζονται πλήρως στον καλλιτεχνικό έλεγχο του μηχανικού.

Υπάρχουν τεχνοτροπίες του MLP οι οποίες δεν έχουν παρουσιαστεί ακόμα. Για παράδειγμα, το σύστημα είναι επεκτάσιμο στην υποστήριξη αξιοσημείωτου αριθμού καναλιών, και έχει μια επιλογή να ενσωματώνει ιεραρχικά διαδικασίες κωδικοποίησης όπως το Ambisonics, όπου το συνθέτουν στοιχεία του ηχητικού πεδίου και όχι τα σήματα τροφοδοσίας για τα ηχεία. Αυτό μπορεί να χρησιμοποιηθεί μελλοντικά με σκοπό να ξεπεραστούν οι περιορισμοί στο όριο των ηχείων που μπορούν να τροφοδοτηθούν σε κάποια πρωτόκολλα, για την μεταφορά surround ήχου σε καταναλωτικές εφαρμογές (βλέπε 17).

2.2.8 Αμφιωτική ακουστική (Binaural).

Η μέθοδος Binaural αποτελεί μια από τις πιο ακριβείς μεθόδους ηχοληψίας και αναπαραγωγής τρισδιάστατων ηχητικών πεδίων. Η αρχή λειτουργίας της βασίζεται στην ηχογράφηση του ήχου που θα έφτανε μέσα στα αυτιά του ακροατή² ώστε κατά την αναπαραγωγή (μέσω ακουστικών) ο ακροατής να λάμβανε ακριβώς τα ίδια ηχητικά σήματα με αυτά που θα λάμβανε αν βρισκόταν στη θέση της ηχογράφησης του αρχικού ηχητικού πεδίου. Για την ηχογράφηση του ηχητικού πεδίου χρειάζονται μόλις δύο κανάλια εγγραφής, ενώ η αναπαραγωγή γίνεται από ακουστικά. Κατά την ηχογράφηση γίνεται χρήση dummy head στα αυτιά του οποίου έχουν τοποθετηθεί τα μικρόφωνα. Αυτά τα κεφάλια είναι κατάλληλα φτιαγμένα έτσι ώστε στις υποδοχές των μικροφώνων που βρίσκονται στα αυτιά αυτών, να δημιουργούνται όμοιες επιρροές στα ηχητικά σήματα που φτάνουν στο εσωτερικό του αυτιού με αυτές που θα δημιουργούνταν αν ο ακροατής βρισκόταν στη θέση της ηχογράφησης. Έτσι είναι κατασκευασμένες από κατάλληλο υλικό το οποίο να διαθλά και να ανακλά τον ήχο με παρόμοιο τρόπο με αυτόν του ανθρώπινου δέρματος, καθώς και έχουν διαμορφωμένες θέσεις για την τοποθέτηση κογχών στις θέσεις των αυτιών. Στην ηχογράφηση λοιπόν γίνεται η καταγραφή των ITD, IID και το φιλτράρισμα του κεφαλιού καθώς και των κογχών (βλέπε κεφ. 2.1.1 και 2.1.2). Ένα από τα προβλήματα που υφίστανται κατά την ηχογράφηση των ηχητικών πεδίων είναι η ηχογράφηση της επιρροής που έχει το εσωτερικό κανάλι του αυτιού. Όταν λοιπόν τοποθετηθούν τα μικρόφωνα στο εσωτερικό του αυτιού καταγράφεται και η συχνότητα συντονισμού του εσωτερικού καναλιού αυτού, που είναι γύρο στα 3 kHz. Όταν γίνει λοιπόν η αναπαραγωγή μέσω των ακουστικών σε κάποιον άνθρωπο, ο ίδιος ήχος θα ξαναπεράσει από το εσωτερικό κανάλι του αυτιού του με αποτέλεσμα να λαμβάνει τον ήχο με διπλό συντονισμό γύρο από τα 3 kHz. Αυτό σημαίνει ότι θα πρέπει να γίνει κάποια ισοστάθμιση στο ήδη ηχογραφημένο υλικό για να εξαλειφθεί αυτός ο παράγοντας. Ένα άλλο πρόβλημα που παρουσιάζεται κατά την ηχογράφηση των ηχητικών πεδίων είναι η διαφορετικότητα την οποία παρουσιάζουν οι κόγχες στους διάφορους ανθρώπους. Αυτό σημαίνει ότι σε κάθε άνθρωπο ο προσπίπτον ήχος θα υποβάλλεται σε τελείως διαφορετικό φιλτράρισμα. Έτσι κατά την αναπαραγωγή του ήδη ηχογραφημένου υλικού θα είναι σαν να ακούει ένας άνθρωπος με διαφορετικές κόγχες.

Έκτος όμως από την ηχογράφηση του ηχητικού πεδίου μπορεί να γίνει και σύνθεση αυτού. Ο τρόπος σύνθεσης γίνεται με την χρήση των HTRF δεδομένων. Ξέροντας λοιπόν την γωνίες πρόσπτωσης του ήχου (την αζιμούθια και κατακόρυφη γωνία) και κάνοντας συνέλιξη (convolution) στο χρόνο μπορεί να συντεθεί ο ήχος που θα έφτανε στα αυτιά του ανθρώπου. Με αυτό τον τρόπο δεν εξαλείφονται τα προβλήματα που υπάρχουν με την μέθοδο της ηχογράφησης

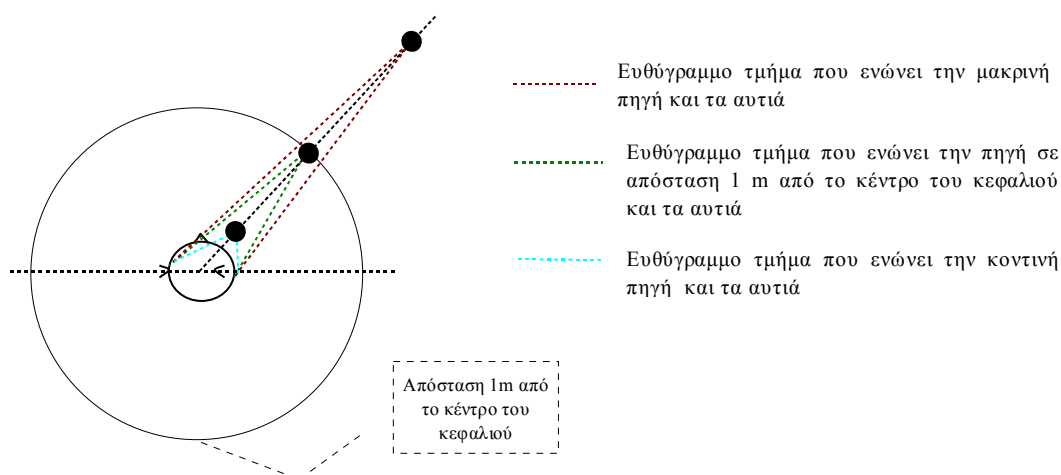
2 Αυτό σημαίνει ότι καταγράφεται και οποιαδήποτε επιρροή στον ήχο από την κογχή του αυτιού, που αποτελεί ένα από τα κριτήρια αξιολόγησης της χωρικής θέσης του ήχου για το σύστημα ακοής του ανθρώπου (βλέπε κεφ. 2.1.1.3.2)

του ηχητικού πεδίου αλλά υπάρχουν και επιπρόσθετα που προκύπτουν από τις HRTF που θα χρησιμοποιηθούν. Τα προβλήματα αυτά προκύπτουν στην προσομοίωση των αντηχήσεων ενός διαφορετικού χώρου και στην προσομοίωση της απόστασης. Ας δούμε όμως πρώτα με ποιο τρόπο δημιουργούνται οι HRTF.

Η δημιουργία των HRTF προκύπτει από την ηχογράφηση των κρουστικών αποκρούσεων ενός dummy head ή ενός πραγματικού ανθρώπου σε όλες τις αζιμούθιες και κατακόρυφες γωνίες (συνήθως με βήμα 5 μοιρών) με βάση το κέντρο του κεφαλιού του ανθρώπου. Η ηχογράφηση γίνεται σε συγκεκριμένη απόσταση (συνήθως ένα μέτρο) από το dummy head (ή τον άνθρωπο) και πραγματοποιούνται σε ανηχοϊκό θάλαμο. Ο λόγος χρήσης του ανηχοϊκού θαλάμου είναι η ευελιξία χρήσης αυτών των δειγμάτων για την προσθήκη αργότερα των αντηχήσεων του χώρου που επιθυμεί ο συνθέτης του ηχητικού πεδίου. Αν αντίθετα είχε γίνει ηχογράφηση σε ένα κοινό χώρο, θα γινόταν ηχογράφηση και των αντηχήσεων του. Αυτό σημαίνει ότι δεν θα είναι δυνατή η χρήση διαφορετικής προσομοίωσης χώρου από τον συνθέτη του ηχητικού πεδίου, καθώς δεν υπάρχει η δυνατότητα αφαίρεσης των ηχογραφημένων αντηχήσεων από αυτά τα δείγματα.

Ένα πρόβλημα που δημιουργείται με τις HRTF είναι στην προσομοίωση της απόστασης. Καθώς οι HRTF καταγράφονται σε συγκεκριμένη απόσταση, καταγράφονται ταυτόχρονα και τα αντίστοιχα IID, ITD και το φιλτράρισμα στο σήμα λόγο του κεφαλιού και της κόγχης του αυτιού. Τι γίνεται όμως άμα ο συνθέτης θέλει να προσομοιώσει ένα ήχο που βρίσκεται σε διαφορετική απόσταση; Όπως φαίνεται στην εικόνα 12 ανάλογα με την απόσταση της πηγής αλλάζει και η γωνία πρόσπτωσης του ήχου της στο κάθε αυτί αλλά και διαφορά της απόστασης μεταξύ της πηγής και του κάθε αυτιού. Αυτό σημαίνει ότι λανθασμένα θα χρησιμοποιούσαμε τις ίδιες HRTF που αφορούν και τα δύο αυτιά για την προσομοίωση διαφορετικών αποστάσεων. Θα έπρεπε λοιπόν είτε να γίνει η καταγραφή των HRTF ξεχωριστά για κάθε αυτί, με κέντρο το ίδιο, είτε να γίνει κατάλληλη επεξεργασία των ήδη υπάρχοντων HRTF ώστε να μπορεί να προκύψει το ίδιο αποτέλεσμα. Επιπλέον λόγο της μεταβολής της διαφοράς της απόστασης μεταξύ των αυτιών και της πηγής ανάλογα με την απόσταση, πρέπει να μεταβληθεί ο ITD και το IID μεταξύ των αυτιών. Αυτό θα μπορούσε να επιτευχθεί πριν την συνέλιξη (convolution) με τις HRTF, με χρήση χρονοκαθυστέρησης (delay) και κατάλληλου φιλτραρίσματος³ των σημάτων.

3 Το οποίο θα προσομοιώνει την συχνοτική επιρροή που δέχονται τα ηχητικά σήματα ανάλογα με την απόσταση διάδοσής στον αέρα.



Εικόνα 12: Απεικόνιση της διαφοράς στη γωνία πρόσπτωσης και της απόστασης με το κάθε αυτί που έχει μια πηγή που βρίσκεται στην ίδια γωνία σε σχέση με το κέντρο του κεφαλιού αλλά σε διαφορετική απόσταση.

Ένα πρόβλημα που μπορεί να προκύψει κατά την Binaural αναπαραγωγή με σήματα όχι κατάλληλα επεξεργασμένα (πχ απλή δρομολόγηση σημάτων στα ακουστικά χωρίς την χρήση της συνέληξης με τις HRTF), είναι το ηχητικό πεδίο να λαμβάνει χώρα μέσα στο κεφάλι του ακροατή (IHL, Inside the Head Locatedness). Αυτό οφείλεται στο γεγονός ότι σε αυτά τα σήματα δεν έχουν επέλθει όλες οι επιρροές που θα συνέβαιναν στην πραγματικότητα, που οφείλονται στο φιλτράρισμα του αέρα λόγω της απόστασης, φιλτράρισμα από το ίδιο το κεφάλι και τα αυτιά, ITD, IID κ.α. . Ακόμα όμως και αν γίνει μόνο συνέληξη με τις HRTF ενός ήχου μπορεί να παρουσιαστεί το ίδιο πρόβλημα. Ο λόγος είναι ότι ένα από τα στοιχεία που αποτιμά το σύστημα ακοής του ανθρώπου είναι ο λόγος μεταξύ του απευθείας ήχου και των αντηχήσεων. Αυτός ο λόγος είναι πολύ μικρός (αν όχι ανύπαρκτος) στις περιπτώσεις των HRTF καθώς δημιουργούνται σε ανηχοϊκούς θαλάμους. Το ίδιο δεν θα συνέβαινε με χρήση HRTF οι οποίες έχουν δημιουργηθεί σε “κανονικούς” χώρους.

Με τις HRTF μπορούμε λοιπόν να προσομοιωθούν όλα τα κριτήρια που χρειάζεται το ανθρώπινο σύστημα ακοής για την αξιολόγηση κυρίως της γωνίας πρόσπτωσης και μερικώς την απόσταση του ηχητικού πεδίου. Ένα λοιπόν από τα στοιχεία που δεν έχουν προσομοιωθεί είναι οι αντηχήσεις. Ας δούμε όμως με ποιο τρόπο μπορεί να γίνει η προσομοίωση χώρου κατά την Binaural αναπαραγωγή. Όπως προαναφέρθηκε, για να είναι δυνατή η προσομοίωση οποιαδήποτε χώρου από τον συνθέτη θα πρέπει να χρησιμοποιηθούν HRTF οι οποίες έχουν ηχογραφηθεί σε ανηχοϊκό θάλαμο. Για την δημιουργία ηχητικών τοπίων υψηλής ποιότητας θα πρέπει η κάθε μία από τις αντηχήσεις να αντιμετωπιστεί σαν διαφορετικό ηχητικό πεδίο. Έτσι θα πρέπει, εκτός από την συνέληξη που θα γίνει στον κύριο ήχο με την HRTF που αντιστοιχεί στις γωνίες πρόσπτωσης του ήχου, το ίδιο να γίνει και με κάθε αντήχηση. Το τελικό σήμα που θα δρομολογηθεί σε κάθε ακουστικό θα αποτελείται από την άθροιση όλων των συνέληξεων (βλέπε [1],[6],[12],[20]).

Περίληψη.

Τα **πλεονεκτήματα** που παρουσιάζει αυτή η μέθοδος αναπαραγωγής είναι:

- Έχει την δυνατότητα αναπαραγωγής ηχητικών πεδίων και στις τρεις διαστάσεις.
- Χρειάζεται μόνο δύο κανάλια για την εγγραφή και αναπαραγωγή του ηχητικού πεδίου.
- Είναι από τις πιο ακριβείς μεθόδους αναπαραγωγής ηχητικών πεδίων, συμπεριλαμβάνοντας

κριτήρια αξιολόγησης της απόστασης και της γωνίας πρόσπτωσης.

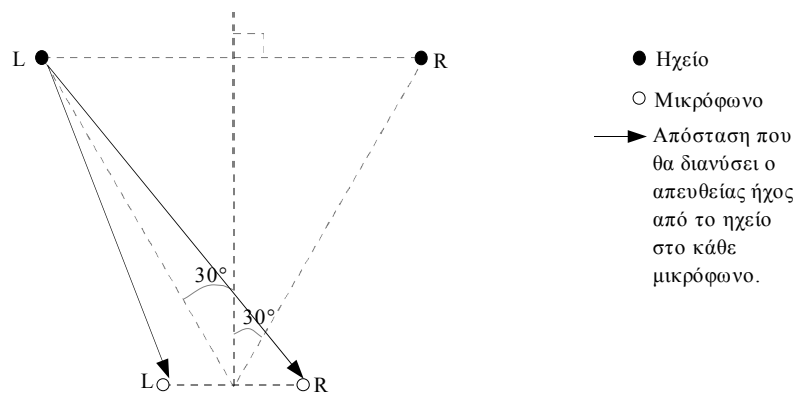
- Είναι δυνατή η ηχογράφιση των ηχητικών πεδίων με την χρήση Dummy Head.

Τα **μειονεκτήματα** που παρουσιάζει είναι:

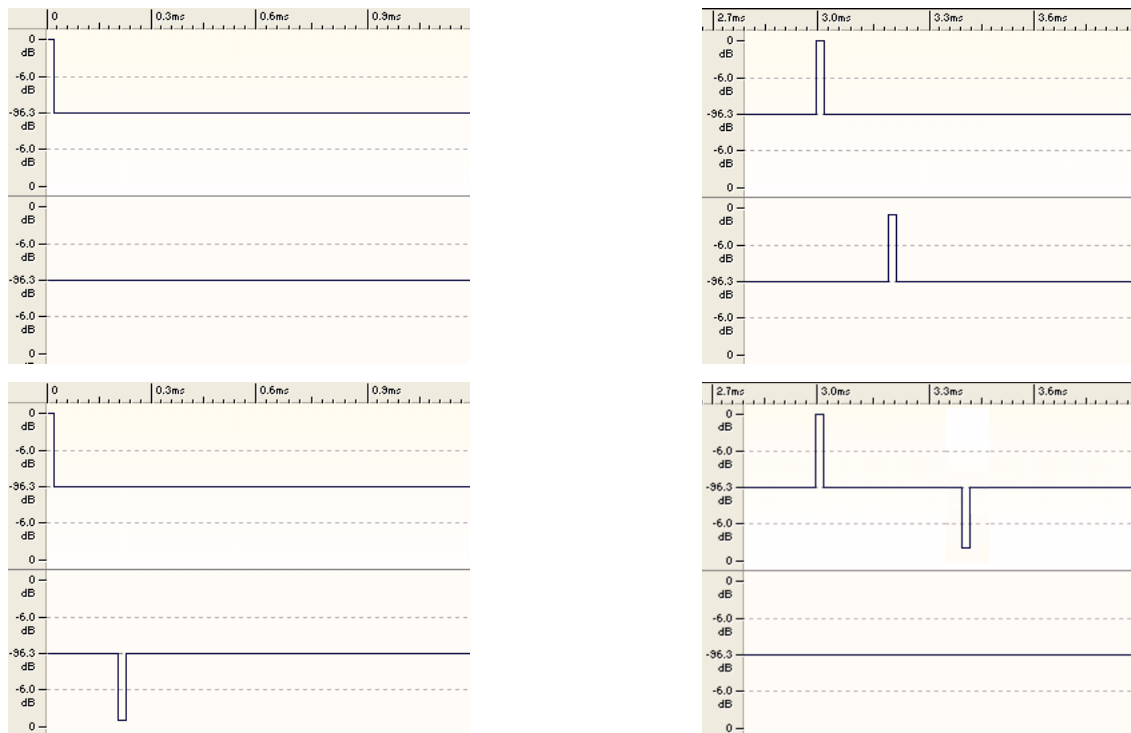
- Με την χρήση όχι ατομικών HRTF δεν γίνεται βέλτιστη αναπαραγωγή, δημιουργώντας διάφορα ανεπιθύμητα φαινόμενα όπως λανθασμένη γωνία προέλευσης του ήχου στον “cone of confusion”.
- Ο ακροατής πρέπει να φοράει ακουστικά.
- Δεν αλλάζει ο ήχος στα αυτιά του ακροατή κατά την κίνηση του κεφαλιού του (τροποποιήσεις στον ήχο θα βοηθούσαν στην καλύτερη εύρεση της προέλευσης του ήχου).
- Ένα ηχογραφημένο υλικό με την μέθοδο Binaural δεν μπορεί να χρησιμοποιηθεί σε άλλη μέθοδο αναπαραγωγής πολυκάναλου ήχου.

2.2.9 Αμφιωτική ακουστική με χρήση μεγαφώνων (Transaural).

Όπως προαναφέρθηκε στο κεφάλαιο 2.2.6 η αναπαραγωγή της μεθόδου Binaural γίνεται αποκλειστικά από ακουστικά. Η μέθοδος Transaural είναι η μέθοδος αναπαραγωγής των ηχητικών πεδίων που έχουν καταγραφή ή συντεθεί βάση της μεθόδου Binaural από δύο ηχεία. Αυτή η μέθοδος πρωτοπαρουσιάστηκε την δεκαετία του 1960 από τους Atal, Hill και Schroeder (Atal, 1966). Κατά την Binaural αναπαραγωγή έχουμε δύο ηχητικά σήματα (L, R) που προορίζονται να διοχετευθούν αντίστοιχα στα αυτιά του ακροατή μέσω των ακουστικών. Κατά την Transaural αναπαραγωγή είναι επιθυμητό αυτά τα σήματα να διοχετευθούν αντίστοιχα στα αυτιά του ακροατή μέσω των ηχείων. Υπάρχει όμως μια βασική διαφορά σε αυτούς τους δύο τρόπους μετάδοσης των σημάτων. Στην περίπτωση της αναπαραγωγής μέσω ακουστικών το κάθε αυτί θα λάβει μόνο το σήμα που τους αντιστοιχεί. Αντίθετα, στην αναπαραγωγή μέσω ηχείων το κάθε αυτί θα λάβει και τα δύο σήματα, εκ των οποίων το σήμα που λανθασμένα λαμβάνεται από το εκάστοτε αυτί να παρουσιάζει μια χρονική καθυστέρηση. Ας υποθέσουμε λοιπόν πως δεν υπάρχουν επιρροές στα δύο σήματα από το κεφάλι του ακροατή και ότι η αναπαραγωγή γίνεται σε ελεύθερο πεδίο (Free Field). Προσομοιάζουμε τα αυτιά του ανθρώπου με δύο μικρόφωνα τοποθετημένα σε αντίστοιχη απόσταση μεταξύ τους (περίπου 17 cm). Τα ηχεία ισαπέχουν από την μεσοκάθετο και το μέσω του διανύσματος που σχηματίζουν τα δύο μικρόφωνα και σχηματίζουν γωνία 60° μεταξύ τους (βλέπε εικόνα 13). Όπως φαίνεται στην εικόνα 13, όταν παίζει μόνο το ένα ηχείο ο ήχος θα φτάσει και στα δύο μικρόφωνα, μόνο που η απόσταση που θα διανύσει ο ήχος που εκπέμπεται από το αριστερό (L) ηχείο στο δεξί (R) μικρόφωνο είναι μικρότερη από αυτή από το αριστερό ηχείο στο αριστερό μικρόφωνο. Αυτό έχει σαν αποτέλεσμα την χρονική καθυστέρηση και την μείωση της έντασης του ήχου που θα φτάσει στο δεξί ηχείο σε σχέση με αυτή που θα φτάσει στο αριστερό. Επιθυμητό όμως θα ήταν ότι παίζει το αριστερό ηχείο να φτάνει μόνο στο αριστερό μικρόφωνο, και όταν παίζει το δεξί να φτάνει μόνο στο δεξί. Για να είναι εφικτό αυτό θα πρέπει να εκπεμφθεί ένας ανάστροφος (διαφορά φάσης 180°) και μειωμένος σε ένταση ήχος από το δεξί ηχείο ώστε όταν αυτός φτάσει στο αριστερό μικρόφωνο να δημιουργηθεί καταστροφική συμβολή. Όμως ομοίως ο ήχος θα εκπέμφθηκε από το αριστερό ηχείο θα φτάσει και στο δεξί μικρόφωνο, άρα πρέπει πάλι να ακολουθηθεί η προηγούμενη διαδικασία, και αυτό θα εξακολουθήσει επάπειρο.



Εικόνα 13: Αναπαράσταση της διάταξης ενός συστήματος Transaural και σημείωση της διαδρομής που θα διανύσει ο ήχος που παίζει ένα ηχείο μέχρι τα μικρόφωνα (προσομοίωση των αυτιών του ακροατή).



Εικόνα 14: Αριστερά φαίνονται τα σήματα που τροφοδότησαν τα ηχεία, και δεξιά τα σήματα που θα λάβουν τα μικρόφωνα. Στο πάνω μέρος απεικονίζεται τα σήματα όταν παίζει μόνο το ένα ηχείο, ενώ κάτω τα σήματα αν παίζουν και τα δύο ηχεία έχοντας εφαρμοστεί το ίδιο σήμα με χρονική καθυστέρηση και μείωση της έντασης στο άλλο για την δημιουργία μιας καταστροφικής συμβολής.

Το αποτέλεσμα αυτής της ακολουθίας που περιγράφηκε είναι ότι για την δημιουργία αυτών των καταστροφικών συμβολών θα χρειαζόταν:

- όσον αφορά το σήμα που θα εκπεμφθεί από το αριστερό ηχείο, το αρχικό σήμα αυτού να επαναληφθεί σε κάποια ισαπέχοντα χρονικά διαστήματα με σταθερά μειούμενη ένταση.
- Όσον αφορά το σήμα που θα εκπεμφθεί από το δεξί ηχείο, το αρχικό σήμα του αριστερού

να επαναληφθεί με αντίστροφη φάση σε κάποια ισαπέχοντα χρονικά διαστήματα με σταθερά μειούμενη ένταση.

Φυσικά η παραπάνω περιγραφή αφορούσε αναπαραγωγή ενός ήχου μόνο από το αριστερό ηχείο. Στην πραγματικότητα, ταυτόχρονα θα παίζει και το δεξί ηχείο και θα είναι επιθυμητή η εμφάνιση όσον αυτό αναπαράγει μόνο στο δεξί μικρόφωνο. Αυτό σημαίνει ότι θα η ανάλογη διαδικασία θα πρέπει να ακολουθηθεί και για την αναπαραγωγή από το δεξί ηχείο.

Όσα έχουν αναφερθεί μέχρι στιγμής είναι επαρκή για την κατανόηση της λειτουργίας της transaural αναπαραγωγής αγνοώντας την επίρεια του κεφαλιού του ανθρώπου και θεωρώντας την αναπαραγωγή σε ελεύθερο πεδίο. Δεν έχει αναφερθεί καθόλου η συνάρτηση υπολογισμού των σημάτων που πρέπει να τροφοδοτηθούν τα ηχεία ώστε να υπάρχει η βέλτιστη αναπαραγωγή. Τα δεδομένα που υπάρχουν είναι οι κρουστικές αποκρίσεις που έχει το κάθε ηχείο σε κάθε μικρόφωνο καθώς και ότι το επιθυμητό σήμα που πρέπει να φτάσει στο αριστερό ηχείο είναι το αριστερό σήμα του Binaural και αντίστοιχα στο δεξί το δεξί αυτού (βλέπε εικόνα 15). Τα σήματα που φτάνουν λοιπόν στα μικρόφωνα είναι:

$$\begin{bmatrix} \text{Mic}_1 \\ \text{Mic}_2 \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

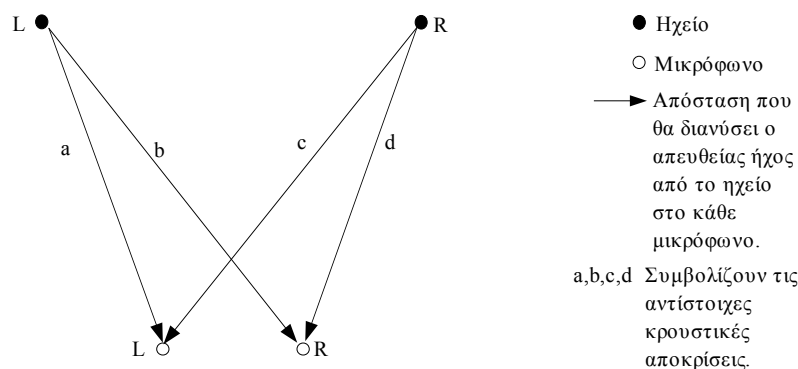
όπου :

- $\text{Mic}_1, \text{Mic}_2$: Το σήμα που εισέρχεται στο δεξί και αριστερό μικρόφωνο αντίστοιχα.
- a, b, c, d : Οι κρουστικές αποκρίσεις των ηχείων στο κάθε μικρόφωνο (βλέπε εικόνα 15).
- V_1, V_2 : Τα σήματα του αριστερού και δεξιού ηχείου αντίστοιχα.

Εμείς θέλουμε να υπολογίσουμε τα σήματα που πρέπει να τροφοδοτηθούν τα ηχεία άρα πρέπει να απομονώσουμε τον πίνακα που περιέχει τα ηχεία, πολλαπλασιάζοντας με τον αντίστροφο του άλλου και έτσι έχουμε:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \frac{1}{ad - cb} \begin{bmatrix} d & -c \\ -b & a \end{bmatrix} \begin{bmatrix} \text{Mic}_1 \\ \text{Mic}_2 \end{bmatrix}$$

Σχέση 2:



Εικόνα 15: Απεικόνιση των τέσσερις διαφορετικών κρουστικές αποκρίσεων κατά την Transaural αναπαραγωγή.

Ένα βασικό μειονέκτημα αυτής την μεθόδου αναπαραγωγής είναι ότι υπάρχει ένα συγκεκριμένο σημείο για την βέλτιστη αναπαραγωγή. Σε σημεία που ουσιαστικά δεν ισαπέχουν τα αυτιά του ακροατή από τα ηχεία υπάρχει η δημιουργία comb filter, που οφείλεται στο γεγονός ότι υπάρχει επανάληψη του σήματος σε συγκεκριμένες χρονικές αποκλίσεις.

Περίληψη.

Συνοψίζοντας, αυτή η μέθοδος παρουσιάζει τα παρόμοια πλεονεκτήματα και μειονεκτήματα με την Binaural αναπαραγωγή. Έτσι τα **πλεονεκτήματα** είναι:

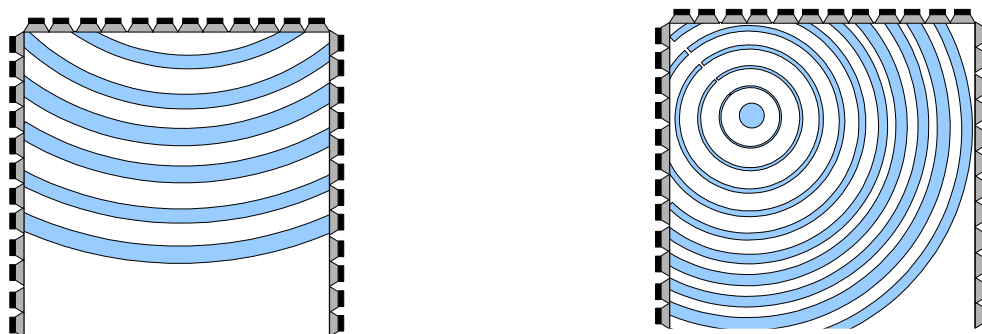
- Έχει την δυνατότητα αναπαραγωγής ηχητικών πεδίων και στις τρεις διαστάσεις.
- Χρειάζεται μόνο δύο κανάλια για την εγγραφή και αναπαραγωγή του ηχητικού πεδίου.
- Είναι από τις πιο ακριβείς μεθόδους αναπαραγωγής ηχητικών πεδίων, συμπεριλαμβάνοντας κριτήρια αξιολόγησης της απόστασης και της γωνίας πρόσπτωσης.
- Γίνεται η αναπαραγωγή από ηχεία.
- Είναι δυνατή η ηχογράφηση των ηχητικών πεδίων με την χρήση Dummy Head.

Τα **μειονεκτήματα** που παρουσιάζει είναι:

- Με την χρήση όχι ατομικών HRTF δεν γίνεται βέλτιστη αναπαραγωγή, δημιουργώντας διάφορα ανεπιθύμητα φαινόμενα όπως λανθασμένη γωνία προέλευσης του ήχου στον “cone of confusion”.
- Ο ακροατής πρέπει να βρίσκεται σε ένα συγκεκριμένο σημείο για να έχει βέλτιστη αναπαραγωγή.
- Δεν αλλάζει ο ήχος στα αυτιά του ακροατή κατά την κίνηση του κεφαλιού του (τροποποιήσεις στον ήχο θα βοηθούσαν στην καλύτερη εύρεση της προέλευσης του ήχου).
- Αυτό υλικό δεν μπορεί να χρησιμοποιηθεί σε άλλη μέθοδο αναπαραγωγής πολυκάναλου Ήχου.

2.2.10 Wave Field Synthesis (WFS), Wave Field Analysis (WFA) .

Η Wavefield Synthesis αποτελεί μια από τις πιο ακριβείς μεθόδους σύνθεσης και αναπαραγωγής ήχου στο οριζόντιο επίπεδο σε ευρύτερο πεδίο ακρόασης. Πρωτοδοκιμάστηκε από τα Bell Labs (Rumsey and McCormick), το 1994. Με αυτή την μέθοδο μπορούν να προσομοιωθούν διάφορες ηχητικές πηγές στον χώρο εντός και εκτός της επιφάνειας που καταλαμβάνει η διάταξη των ηχείων (βλέπε Εικόνα 16). Τα ηχεία είναι τοποθετημένα στις τρεις πλευρές ενός ορθογώνιου παραλληλογράμμου σε γραμμική διάταξη (line-array). Παρουσιάζει ίδια ακουστική ποιότητα σε ένα ευρύτερο πεδίο ακρόασης, που είναι και ένα από τα χαρακτηριστικά που δεν διαθέτουν τα περισσότερα από τα συστήματα πολυκάναλης αναπαραγωγής, καθώς έχουν ένα συγκεκριμένο σημείο που γίνεται βέλτιστη αναπαραγωγή (sweet spot). Το βασικότερο χαρακτηριστικό αυτής της μεθόδου αναπαραγωγής είναι ότι η εικονική πηγή φαίνεται να καταλαμβάνει συγκεκριμένη θέση στο χώρο . Ο κάθε ακροατής λοιπόν, ανάλογα με την θέση του στο ακροατήριο, αντιλαμβάνεται την γωνία προέλευσης και την απόσταση της πηγής, σαν αυτή να βρισκόταν πραγματικά στη θέση αυτή. Η μέθοδος αυτή βασίζεται στην μέθοδο Huygen (Huygen's principle). Σύμφωνα με αυτή, η διάδοση ενός κύματος σε ένα μέσο μπορεί να ανασυντεθεί μέσω της άθροισης των επιμέρους συμβολών των πηγών, που είναι τοποθετημένες κατά μήκος του μετώπου του κύματος αυτού ([1],[15],[16]). Για την δημιουργία λοιπόν του ιδίου αποτελέσματος που θα δημιουργούσε μια πηγή σε άπειρη απόσταση (θεωρητικά επίπεδο μέτωπο κύματος), θα μπορούσε να χρησιμοποιηθεί μια συστοιχία ηχείων σε γραμμική διάταξη. Ομοίως θα μπορούσε να



Εικόνα 16: Σε αυτή την εικόνα φαίνεται μια συστοιχία ηχείων σε γραμμική διάταξη (line array) τοποθετημένα στις τρεις πλευρές ενός τετραγώνου, καθώς και τα μέτωπα των κυμάτων που θα δημιουργούσαν οι εικονικές ηχητικές πηγές. Στα αριστερά φαίνεται μια συνθετική ηχητική πηγή εκτός των ορίων της διάταξης των ηχείων ενώ δεξιά μία ηχητική πηγή εντός αυτών.

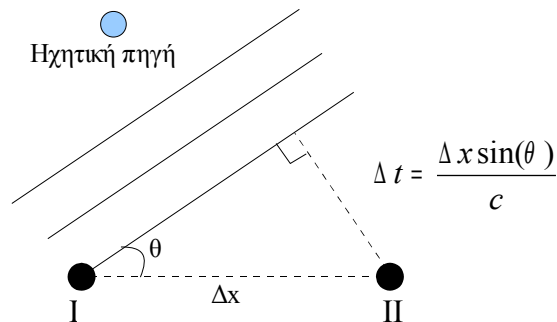
δημιουργηθεί μια κοντινότερη πηγή, της οποίας το μέτωπο κύματος θα ήταν καμπυλωτό με μία διάταξη ηχείων σε καμπύλη. Όμως με αυτή την μέθοδο σύνθεση και ανάλυσης μπορεί να συντεθεί οποιοδήποτε μέτωπο κύματος (καμπύλο ή γραμμικό) με μία διάταξη ηχείων σε γραμμική διάταξη, της οποίας τα σήματα που τροφοδοτούν τα ηχεία θα λάβουν τα αντίστοιχα σήματα με κατάλληλη καθυστέρηση (Βλέπε [15]).

Όμως αυτή η μέθοδος δεν είναι τόσο ιδανική. Έχει μια μέγιστη συχνότητα αναπαραγωγής που ονομάζεται 'Spatial Aliasing Frequency' (Berkhout et al., 1992) η οποία εξαρτάται από την απόσταση των ηχείων την ταχύτητα του ήχου και την γωνία εκπομπής. Με την Σχέση 3 (Verheijen et al., 1995) είναι δυνατός ο υπολογισμός αυτής της συχνότητας. Στην εικόνα 17 δίνεται ένα παράδειγμα για τον λόγο ύπαρξης αυτής της ανώτατης συχνότητας. Απεικονίζεται

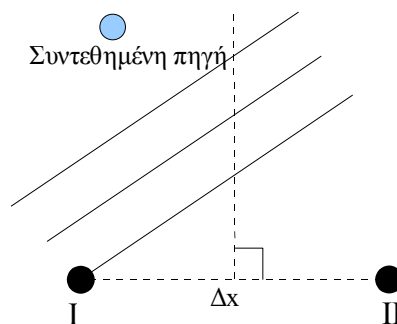
$$f_{Nyq} = \frac{c}{2\Delta x \sin(\theta)}$$

Σχέση 3: Σχέση υπολογισμού της μέγιστης συχνότητας αναπαραγωγής 'Spatial Aliasing Frequency', όπου (c) είναι η ταχύτητα του ήχου, Δx η απόσταση μεταξύ των ηχείων και θ είναι η γωνία εκπομπής του ήχου.

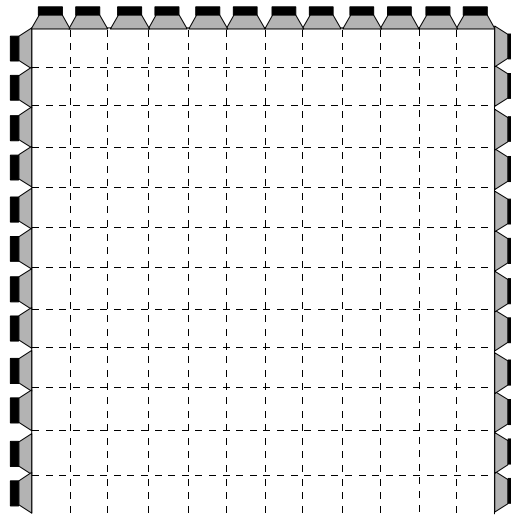
μια πηγή σε μεγάλη απόσταση της οποίας τα μέτωπα των ηχητικών κυμάτων είναι επίπεδα. Όπως φαίνεται στην εικόνα το μέτωπο του κύματος θα φτάσει στο δεύτερο ηχείο μετά από κάποιο χρόνο. Συχνότητες λοιπόν που θα έχουν την μίση περίοδο από αυτό τον χρόνο θα δημιουργήσουν καταστροφικές συμβολές κατά την ανασύνθεση του πεδίου, λόγω της καθυστέρησης του σήματος μεταξύ των ηχείων (βλέπε Εικόνες 18,19). Αυτό οφείλεται στο ότι η εκπομπή των ηχείων δεν είναι ιδανική, δηλαδή προς μία διεύθυνση μόνο, αλλά υπάρχει διάθλαση του κύματος. Τρόποι βελτιστοποίησης αυτής της μεθόδου αναπαραγωγής είναι η χρήση πιο κατευθυντικών ηχείων ή το φιλτράρισμα των σημάτων τροφοδότησης ηχείων (Βλέπε [1]).



Εικόνα 17: Απεικόνιση των επίπεδων μετώπων των κυμάτων μιας ηχητικής πηγής σε μεγάλη απόσταση, και η διαφορά χρόνου που θα φτάσει το μέτωπο από το ηχείο I στο ηχείο II.



Εικόνα 18: Αποτέλεσμα της σύνθεσης των μετώπων των κυμάτων από τα σήματα που έχουν ληφθεί από μικρόφωνα που βρίσκονται στις θέσεις των ηχείων. Η μέγιστη συχνότητα 'Spatial Aliasing Frequency' του ήχου που εκπέμπεται από τα ηχεία και βρίσκονται στην μεσοκάθετο της απόστασης μεταξύ των ηχείων, δεν θα είναι ακουστή καθώς θα δημιουργούν καταστροφικές συμβολές λόγω της χρονικής καθυστέρησης που θα έχει η εκπομπή του ήχου από το κάθε ηχείο.



Εικόνα 19: Απεικόνιση του πλέγματος καταστροφικών συμβολών στις που δημιουργείται στον χώρο λόγω της απόστασης μεταξύ των ηχείων. Έχει υποθεθεί ότι σε όλα τα ηχεία υπάρχει η ίδια γωνία πρόπτωσης των επιπέδων κυμάτων.

Περίληψη.

Η Wave Field Synthesis και Wave Field Analysis αποτελεί την πιο ακριβή μέθοδο αναπαραγωγής και σύνθεσης ηχητικών πεδίων στο οριζόντιο επίπεδο, όσον αφορά την αναπαραγωγή σε ευρύτερο πεδίο ακρόασης. Περιέχει όμως αρκετά **πλεονεκτήματα** και μειονεκτήματα. Τα **πλεονεκτήματα** αυτής της μεθόδου λοιπόν είναι:

- Είναι μια μέθοδος με την οποία η αναπαραγωγή του ηχητικού πεδίου γίνεται σε ευρύτερο πεδίο ακρόασης, με ίδια ποιότητα ακρόασης, αντίθετα με τις περισσότερες μεθόδους αναπαραγωγής πολυκάναλου ήχου, οι οποίες έχουν ένα σημείο βέλτιστης αναπαραγωγής (sweet spot).
- Είναι η μόνη μέθοδος με καλή προσομοίωση της απόστασης της εικονικής πηγής. Αυτό συμβαίνει για τον λόγο ότι η θέση της συντεθειμένης πηγής καταλαμβάνει συγκεκριμένη θέση στο χώρο, σαν να υπήρχε πραγματικά.

Σημείωση: Μια διαφοροποίηση αυτής της μεθόδου σε σχέση με αυτές που αναφέρονται στην αναπαραγωγή του ηχητικού πεδίου μέσω πολυκάναλης αναπαραγωγής, είναι ότι δεν γίνεται σύνθεση της ίδιας της πηγής αλλά των μετώπων των κυμάτων αυτής.

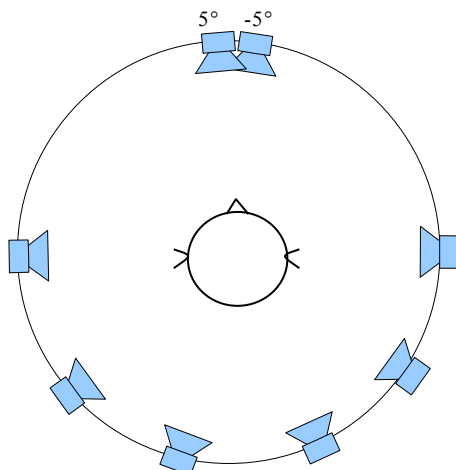
Τα **μειονεκτήματα** αυτής της μεθόδου είναι:

- Χρειάζεται μεγάλο αριθμό καναλιών αναπαραγωγής για καλή ακρόαση στο οριζόντιο επίπεδο.
- Δεν είναι έως τώρα δυνατή η αναπαραγωγή ηχητικών πεδίων και στις τρεις διαστάσεις.
- Υπάρχει ανώτατη συχνότητα αναπαραγωγής 'Spatial Aliasing Frequency' η οποία εξαρτάται από διάφορους παράγοντες όπως η απόσταση μεταξύ των ηχείων, η γωνία εκπομπής της εικονικής πηγής κ.α.

- Υπάρχει αδυναμία πραγματικής ηχογράφησης ηχητικού πεδίου. Έτσι λειτουργεί κυρίως μόνο ως μέθοδος σύνθεσης ηχητικών πεδίων.
- Χρειάζεται μεγάλος αριθμός καναλιών καταγραφής και μεγάλη υπολογιστική ισχύ για την τροφοδότηση των ηχείων με τα κατάλληλα ηχητικά σήματα.

2.2.11 Ambiphonics

Αυτό η μέθοδος δημιουργήθηκε από το Glasgow (2001). Δεν ανήκει στις γενικευμένες μεθόδους αναπαραγωγής τρισδιάστατων ηχητικών πεδίων, αλλά προορίζεται μόνο για την αναπαραγωγή ηχητικών πεδίων ηχογραφημένων από συναυλιακούς χώρους. Αποτελεί ένα υβρίδιο μεταξύ της Binaural/Transaural αναπαραγωγής συνδυασμένη με έναν σωστότερο ψυχοακουστικά αλγόριθμο δημιουργίας αντηχήσεων, έτσι ώστε να ξεγελιέται το σύστημα ακοής του ανθρώπου και να του δημιουργείται η ψευδαίσθηση ότι βρίσκεται σε μια πραγματική αίθουσα. Αυτή η μέθοδος βασίζεται στην καλή αναπαραγωγή του ηχητικού πεδίου που βρίσκεται στην μπροστινή διεύθυνση του ακροατή, ενώ οι αντηχήσεις πρόκειται να συντεθούν στην συνέχεια με την χρήση του αλγόριθμου των αντηχήσεων. Έτσι κατά την ηχογράφηση του ηχητικού μπροστινού ηχητικού πεδίου μπορεί να χρησιμοποιηθεί ένα μικρόφωνο dummy head χωρίς κόγχες ή οποιαδήποτε άλλη τεχνική στερεοφωνικής ηχογράφησης. Πρέπει όμως να περιοριστούν όσο το δυνατό όλες οι ανακλάσεις που φτάνουν στο μικρόφωνο από άλλες διευθύνσεις εκτός από την μπροστινή. Η διάταξη των ηχείων κατά την αναπαραγωγή είναι δύο στην μπροστινή διεύθυνση στις $\pm 5^\circ$ ενώ δεν υπάρχει συγκεκριμένη διάταξη για τα πίσω ηχεία. Η αναπαραγωγή από τα μπροστινά ηχεία γίνεται με χρήση των καταστροφικών συμβολών, όμοια με αυτή της Transaural αναπαραγωγής. Αυτό σημαίνει ότι αν κάθονται οι ακροατές σε μία ευθεία θα έχουν καλή αναπαραγωγή. Το κάθε ένα από τα πίσω ηχεία θα προσομοιώσει μία από τις πρώτες αντηχήσεις μαζί με μία ουρά (ασυσχέτιστες) οι οποίες προσομοιώνουν το διάχυτο πεδίο. Το σήμα που θα τις τροφοδοτήσει θα είναι ένα σύνεληξη των σημάτων της μπροστινής διεύθυνσης και της κρουστικής απόκρουσης που θα είχαν οι αντηχήσεις από αυτή την διεύθυνση (Στις αντηχήσεις δεν θα υπάρχει καθόλου αρχικό -dry- σήμα). Έτσι για να βρεθεί η κατάλληλη θέση των πίσω ηχείων στην οποία η αναπαραγωγή είναι βέλτιστη θα πρέπει να γίνουν δοκιμαστικές αναπαραγωγές (βλέπε [1],[6],[12],[20]). Το πλεονέκτημα αυτής της μεθόδου είναι ότι έχουμε μια καλή εικόνα του ηχητικού τοπίου από την μπροστινή διεύθυνση, ενώ για τις αντηχήσεις, καθώς χρησιμοποιούνται πραγματικές πηγές τοποθετημένες στον χώρο, το σύστημα ακοής του ανθρώπου λαμβάνει όλα τα απαραίτητα κριτήρια για την χωροτοποθέτηση αυτών (κυρίως κριτήρια διεύθυνσης).



Εικόνα 20: Μια τυπική διάταξη των ηχείων κατά την Ambisonic αναπαραγωγή. Όπως φαίνεται μόνο τα μπροστινά ηχεία έχουν συγκεκριμένη θέση στις (5°, -5°).

2.3. Ambisonic

Το Ambisonic είναι μια μέθοδος αποσύνθεσης και αναπαραγωγής ενός τρισδιάστατου ηχητικού πεδίου. Ένα σύστημα Ambisonic αναλαμβάνει να αποσυνθέσει ένα ηχητικό πεδίο σε σφαιρικές αρμονικές (Encoder B-Format Ambisonic) και στη συνέχεια να το ανασυνθέσει μέσω ενός συστήματος αναπαραγωγής ήχου, υπολογίζοντας τους συντελεστές των σφαιρικών αρμονικών που θα πρέπει να τροφοδοτηθεί το κάθε ηχείο του συστήματος (Decoder B-Format Ambisonic). Ανήκει στα Ιεραρχικά (hierarchical) συστήματα καθώς μπορεί να ανασυνθέσει ένα ήδη κωδικοποιημένο ηχητικό πεδίο (B-Format) ανεξαρτήτως με τον αριθμό των ηχείων του συστήματος αναπαραγωγής. Αρκεί η χρήση κατάλληλου αποκωδικοποιητή, ώστε να γίνει σωστή ανασύνθεση του ηχητικού πεδίου⁴.

Ανάλογα με τον αποκωδικοποιητή και το σύστημα αναπαραγωγής (διάταξη ηχείων) μπορεί να ανασυσταθεί ολόκληρο το ηχητικό πεδίο, δηλαδή και στις τρεις διαστάσεις, ή ένα μέρος αυτού, δηλαδή μόνο το οριζόντιο επίπεδο. Γι' αυτό το λόγο, το σύστημα αυτό μπορεί να είναι είτε **παντοφωνικό (pantophonic system)**, στην περίπτωση που η αναπαραγωγή του ηχητικού πεδίου γίνεται μόνο στο οριζόντιο επίπεδο, είτε **περιφωνικό (periphonic system)** στην περίπτωση που η διάταξη των ηχείων είναι κατάλληλη για την αναπαραγωγή ηχητικών πεδίων και στις τρεις διαστάσεις (βλέπε [10]).

2.3.1 Φυσική ερμηνεία της μεθόδου Ambisonic.

Η μέθοδος Ambisonic βασίζεται στην αποσύνθεση και ανασύνθεση του ηχητικού πεδίου βάση των σφαιρικών αρμονικών (βλέπε [28]). Η αποσύνθεση του ηχητικού πεδίου γίνεται σε ένα σημείο που θεωρείται και το κέντρο του ηχητικού πεδίου, ενώ η ανασύνθεσή του γίνεται από μια διάταξη ηχείων σε κάποια απόσταση από το σημείο ακρόασης, τέτοια ώστε τα μέτωπα των κυμάτων αυτών που φτάνουν στο σημείο ακρόασης να θεωρούνται επίπεδα. Ας δούμε όμως τι είναι οι σφαιρικές αρμονικές και από που προκύπτουν .

Σημείωση: Στην απόδειξη που ακολουθεί υπάρχει μια αναντιστοιχία στα σύμβολα 'θ' και 'φ' σε σχέση με την υπόλοιπη εργασία που βασίζεται στο παράρτημα 2. Εδώ το σύμβολο 'φ'

⁴ Βλέπε σελ 97 [1]

συμβολίζει την αζιμούθια γωνία ενώ το 'θ' την γωνία ανύψωσης, η οποία μετριέται από το θετικό του άξονα z και όχι από το θετικό του x. Για να επιτευχθεί η αντιστοιχία με την υπόλοιπη εργασία αρκεί η αντικατάσταση του 'φ' με το 'θ' και του 'θ' με 'φ+π/2'. Αυτή η διαφοροποίηση στην εργασία έγινε για λόγους συμφωνίας με τους συμβολισμούς που χρησιμοποιούνται στην βιβλιογραφία σε κάθε περίπτωση.

Για να μπορέσουμε να εξηγήσουμε τι είναι οι σφαιρικές αρμονικές θα πρέπει πρώτα να αναφερθούμε στην εξίσωση που περιγράφει το κύμα στον χώρο και της λύσεις αυτής. Η εξίσωση του κύματος στο χώρο φαίνεται στη Σχέση 4.

$$\frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = \nabla^2 p$$

Σχέση 4:

Όπου το ∇^2 αποτελεί την ταυτότητα του Laplace (Laplacian operator) για τις τρεις διαστάσεις. Η ταυτότητα αυτή μορφοποιείται ανάλογα με το σύστημα συντεταγμένων που αναφερόμαστε. Σε αυτή την περίπτωση πρόκειται να χρησιμοποιήσουμε σαν σύστημα συντεταγμένων τις σφαιρικές αρμονικές. Η ταυτότητα Laplace σε αυτή την περίπτωση παίρνει την μορφή που φαίνεται στη Σχέση 5.

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

Σχέση 5:

Αντικαθιστώντας την Σχέση 5 στη Σχέση 4 παίρνουμε την Σχέση 6.

$$\frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial p}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial p}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 p}{\partial \phi^2}$$

Σχέση 6:

Αν θέσουμε ότι η p ισούται με την παραγοντοποιημένη ποσότητα $p=T(t)\Phi(\varphi)R(r)P(\theta)$ και αντικαταστήσουμε στη Σχέση 6 δημιουργείται η Σχέση 7.

$$\frac{RP\Phi}{c^2} \frac{\partial^2 T}{\partial t^2} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 P\Phi T \frac{\partial R}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta R\Phi T \frac{\partial P}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} RPT \frac{\partial^2 \Phi}{\partial \phi^2} \Leftrightarrow$$

$$\frac{RP\Phi}{c^2} \frac{\partial^2 T}{\partial t^2} = \frac{2P\Phi T}{r} \frac{\partial R}{\partial r} + P\Phi T \frac{\partial^2 R}{\partial r^2} + \frac{R\Phi T}{r^2 \sin \theta} \cos \theta \frac{\partial P}{\partial \theta} + \frac{R\Phi T}{r^2} \frac{\partial^2 P}{\partial \theta^2} + \frac{RPT}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2}$$

Σχέση 7:

Κάνοντας πράξεις στη Σχέση 7 διαχωρίζουμε όλες τις μεταβλητές που μεταβάλλονται με το T, και έτσι παίρνουμε την πρώτη σχέση της Σχέση 8. Παίρνοντας την μερική παράγωγο ως προς T φαίνεται ότι το δεξιό μέρος της ισότητας θα ισούται με το μηδέν. Ολοκληρώνοντας στη συνέχεια φαίνεται ότι το αριστερό μέρος της αρχικής ισότητας θα ισούται με μια σταθερή ποσότητα (βλέπε

Σχέση 8 την δεύτερη ισότητα).

$$\frac{1}{Tc^2} \frac{\partial^2 T}{\partial t^2} = \frac{2}{rR} \frac{\partial R}{\partial r} + \frac{1}{R} \frac{\partial^2 R}{\partial r^2} + \frac{1}{r^2 \sin^2 \theta P} \cos^2 \theta \frac{\partial P}{\partial \theta} + \frac{1}{r^2 P} \frac{\partial^2 P}{\partial \theta^2} + \frac{1}{r^2 \Phi \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2}$$

$$\frac{1}{Tc^2} \frac{\partial^2 T}{\partial t^2} = \omega^2$$

Σχέση 8:

Ομοίως εργαζόμαστε για τις για τις υπόλοιπες συναρτήσεις, και η διαδικασία αυτή περιγράφεται στις Σχέση 9-Σχέση 12.

$$\frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2} = \frac{r^2 \sin^2 \theta}{c^2 T} \frac{\partial^2 T}{\partial t^2} - \frac{2r \sin^2 \theta}{R} \frac{\partial R}{\partial r} - \frac{\sin^2 \theta r^2}{R} \frac{\partial^2 R}{\partial r^2} - \frac{\sin \theta \cos \theta}{P} \frac{\partial P}{\partial \theta} - \frac{\sin^2 \theta}{P} \frac{\partial^2 P}{\partial \theta^2}$$

$$\frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2} = n^2$$

Σχέση 9:

$$\frac{2r}{R} \frac{\partial R}{\partial r} + \frac{r^2}{R} \frac{\partial^2 R}{\partial r^2} = \frac{r^2}{c^2 T} \frac{\partial^2 T}{\partial t^2} - \frac{1}{P \sin \theta} \cos \theta \frac{\partial P}{\partial \theta} - \frac{1}{P} \frac{\partial^2 P}{\partial \theta^2} - \frac{1}{\Phi \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2}$$

$$\frac{\partial \left(\frac{2r}{R} \frac{\partial R}{\partial r} + \frac{r^2}{R} \frac{\partial^2 R}{\partial r^2} \right)}{\partial r} = \frac{2r}{c^2 T} \frac{\partial^2 T}{\partial t^2}$$

$$\frac{1}{2r} \frac{\partial \left(\frac{2r}{R} \frac{\partial R}{\partial r} + \frac{r^2}{R} \frac{\partial^2 R}{\partial r^2} \right)}{\partial r} = \frac{1}{c^2 T} \frac{\partial^2 T}{\partial t^2}$$

$$\frac{1}{2r} \frac{\partial \left(\frac{2r}{R} \frac{\partial R}{\partial r} + \frac{r^2}{R} \frac{\partial^2 R}{\partial r^2} \right)}{\partial r} = k^2$$

$$\frac{\partial \left(\frac{2r}{R} \frac{\partial R}{\partial r} + \frac{r^2}{R} \frac{\partial^2 R}{\partial r^2} \right)}{\partial r} = 2rk^2$$

$$2r \frac{\partial R}{\partial r} + r^2 \frac{\partial^2 R}{\partial r^2} = (r^2 k^2 - C) R$$

Σχέση 10:

$$\frac{1}{P \sin \theta} \cos \theta \frac{\partial P}{\partial \theta} + \frac{1}{P} \frac{\partial^2 P}{\partial \theta^2} = \frac{r^2}{c^2 T} \frac{\partial^2 T}{\partial t^2} - \frac{1}{\Phi \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2} - \frac{2r}{R} \frac{\partial R}{\partial r} - \frac{r^2}{R} \frac{\partial^2 R}{\partial r^2}$$

$$\frac{\partial \left(\frac{1}{P \sin \theta} \cos \theta \frac{\partial P}{\partial \theta} + \frac{1}{P} \frac{\partial^2 P}{\partial \theta^2} \right)}{\partial \theta} = - \frac{2 \cos \theta}{\Phi \sin^3 \theta} \frac{\partial^2 \Phi}{\partial \phi^2}$$

$$\frac{\sin^3 \theta}{2 \cos \theta} \frac{\partial \left(\frac{1}{P \sin \theta} \cos \theta \frac{\partial P}{\partial \theta} + \frac{1}{P} \frac{\partial^2 P}{\partial \theta^2} \right)}{\partial \theta} = - \frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2}$$

Σχέση 11:

$$\frac{\sin^3 \theta}{2 \cos \theta} \frac{\partial \left(\frac{1}{P \sin \theta} \cos \theta \frac{\partial P}{\partial \theta} + \frac{1}{P} \frac{\partial^2 P}{\partial \theta^2} \right)}{\partial \theta} = n^2$$

$$\frac{\partial \left(\frac{1}{P \sin \theta} \cos \theta \frac{\partial P}{\partial \theta} + \frac{1}{P} \frac{\partial^2 P}{\partial \theta^2} \right)}{\partial \theta} = \frac{2n^2}{\sin^3 \theta} \cos \theta$$

$$\frac{1}{\sin \theta} \cos \theta \frac{\partial P}{\partial \theta} + \frac{\partial^2 P}{\partial \theta^2} = \left(\frac{n^2}{\sin^2 \theta} - C \right) P$$

Σχέση 12: Συνέχεια από Σχέση 11.

Παρατηρώντας τις διαφορικές εξισώσεις που αφορούν την μεταβολή βάσης του 'θ' και του 'φ', φαίνεται ότι αν βρεθούν οι λύσεις αυτών των διαφορικών εξισώσεων, λύση θα αποτελεί και το άθροισμα των λύσεων με οποιοδήποτε βάρος. Η λύση της Σχέση 9 είναι $\sin(n\theta)$ ή $\cos(n\theta)$, ενώ η λύση της Σχέση 12 είναι η πολυώνυμο Legendre (ή συνεταιρισμένη συνάρτηση Legendre, Associated Legendre function) τάξης 'n' και βαθμού 'm', η οποία περιγράφεται στη Σχέση 13.

$$P_m^n(\eta) = (1 - \eta^2)^{n/2} \frac{d^n P_m(\eta)}{d\eta^n}$$

Σχέση 13:

όπου $\eta = \cos \theta$ και:

$$P_m(\eta) = \frac{1}{2^m m!} \frac{d^m}{d\eta^m} (\eta^2 - 1)^m$$

Σχέση 14:

Η γενική λύση της εξίσωσης για τα φ,θ δίνεται από το γινόμενο των λύσεων των διαφορικών εξισώσεων για τις συναρτήσεις P και Φ (σύμφωνα με την παραδοχή που έχουμε πάρει $\rho = T(t)\Phi(\varphi)R(r)P(\theta)$). Αυτό το γινόμενο αποτελεί τις σφαιρικές αρμονικές, οι οποίες περιγράφονται στη Σχέση 15 και έχουν την ιδιότητα που φαίνεται στη Σχέση 16.

$$Y_m(\theta) = Y_{m0}^1(\theta, \phi) = P_m(\eta), \quad Y_{m0}^{-1}(\theta, \phi) = 0$$

$$Y_{mn}^1(\theta, \phi) = \cos(n\phi) P_m^n(\eta)$$

$$Y_{mn}^{-1}(\theta, \phi) = \sin(n\phi) P_m^n(\eta)$$

Σχέση 15:

$$\int_0^{2\pi} \int_0^\pi Y_{mn}^\sigma(\theta, \phi) Y_{lk}^\tau(\theta, \phi) \sin\theta d\theta d\phi = \frac{2\pi}{2m+1} \frac{(m+n)!}{(m-n)!} \delta_{ml} \delta_{nk} \delta_{\sigma\tau}$$

Σχέση 16:

όπου 'σ,τ' παίρνουν τιμές (-1,1), $\delta_{ml}=0$ αν $m \neq l$ ενώ αντίθετα $\delta_{ml}=1$ ενώ το n παίρνει τιμές στο διάστημα $[0,m]$.

Λόγω της Σχέση 16 και ότι στη συγκεκριμένη διαφορική εξίσωση το άθροισμα με διαφορετικά βάρη των λύσεων της αποτελεί λύση της, μπορεί οποιαδήποτε συνάρτηση που περιγράφει την εξίσωση της επιφανείας της σφαίρας (εξαρτάται από τα ϕ και $\cos\theta$) να αποτελεί λύση της, και να περιγραφεί με ένα άθροισμα απείρων σφαιρικών αρμονικών με κατάλληλα βάρη, όπως φαίνεται στη Σχέση 17.

$$F(\theta, \phi) = \sum_{m=0}^{\infty} \sum_{n=0}^{m} \sum_{\sigma=-1}^1 B_{mn}^\sigma Y_{mn}^\sigma(\theta, \phi)$$

Σχέση 17:

όπου B_{mn}^σ οι συντελεστές των αρμονικών, οι οποίοι περιγράφονται από την Σχέση 18.

$$B_{mn}^\sigma = \frac{2m+1}{2\pi} \frac{(m-n)!}{(m+n)!} \int_0^{2\pi} \int_0^\pi F(\theta, \phi) Y_{lk}^\tau(\theta, \phi) \sin\theta d\theta d\phi$$

Σχέση 18:

η οποία προκύπτει από την Σχέση 17, αν πολλαπλασιαστούν και τα δύο μέλη με $\sin\theta$ και Y_{lk}^τ , και στη συνέχεια πάρουμε το ολοκλήρωμα ως προς 'φ' και 'θ' για τα διαστήματα $[0,2\pi]$ και $[0,\pi]$ αντίστοιχα. Το ολοκλήρωμα που προκύπτει παίρνει τιμές διάφορες του μηδενός μόνο για $m=l$, $n=k$, $\tau=\sigma$, άρα και προκύπτει η Σχέση 18. Στον πίνακα (8) στο παράρτημα 1, απεικονίζονται τύποι καθώς και το πολικό διάγραμμα των σφαιρικών αρμονικών μέχρι την τρίτη τάξη. Στα τρισδιάστατα σχέδια των πολικών διαγραμμάτων με κόκκινο συμβολίζονται οι γωνίες που η λήψη του ήχου γίνεται συμφασική, ενώ με μπλε, οι γωνίες που η λήψη γίνεται με αναστροφή φάσης κατά 180° .

Τέλος η γενική λύση της Σχέση 4 φαίνεται στη Σχέση 19.

$$p\left(\vec{r}\right) = \sum_{m=0}^{\infty} j^m j_m(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{mn}^\sigma Y_{mn}^\sigma(\theta, \phi) + \sum_{m=0}^{\infty} j^m h_m^-(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} A_{mn}^\sigma Y_{mn}^\sigma(\theta, \phi)$$

Σχέση 19:

όπου B_{mn}^σ και A_{mn}^σ οι συντελεστές των σφαιρικών αρμονικών για το κύμα που συγκλίνει και αυτό που αποκλίνει από το κέντρο, η $j_m(kr)$ η σφαιρική συνάρτηση Bessel για το συγκλίνον κύμα, ενώ η h_m^- η σφαιρική συνάρτηση Hankel για το αποκλίνον κύμα. Το 'k' είναι ο κυματάρθμος και ισούται με ' $2\pi f/c$ '.

2.3.2 Ambisonic Format .

Υπάρχουν 5 διαφορετικές τυποποιήσεις για τα σήματα ambisonic οι οποίες είναι:

- A-Format
- B-Format
- D-Format
- UHJ-Format ή C-Format
- G-Format

Η κάθε τυποποίηση από αυτές αντιστοιχεί σε ένα στάδιο της μεθόδου Ambisonic. Έτσι η τυποποίηση A-Format αντιστοιχεί στο στάδιο της ηχοληψίας, το B-Format στο στάδιο της αποσύνθεσης του ηχητικού πεδίου, ενώ τα D,HJ,G-Format στο στάδιο της ανασύνθεσης του ηχητικού πεδίου (βλέπε [14]).

A-Format ,ηχογράφηση ηχητικού πεδίου.

Το A-Format αφορά το μέρος της ηχοληψίας του τρισδιάστατου ηχητικού πεδίου, βάση της μεθόδου Ambisonics. Αποτελείται από τα τέσσερα σήματα που εξάγονται από ένα μικρόφωνο Soundfield, τα οποία είναι (βλέπε [13]):

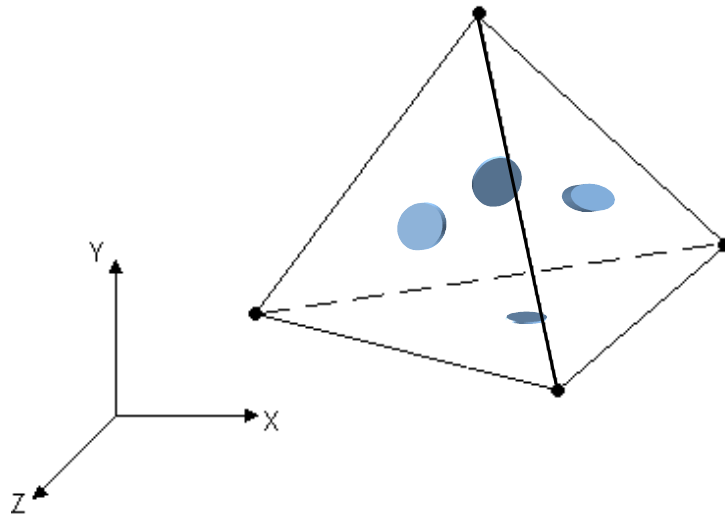
FUL: Front Up Left

BUR: Front Up Right

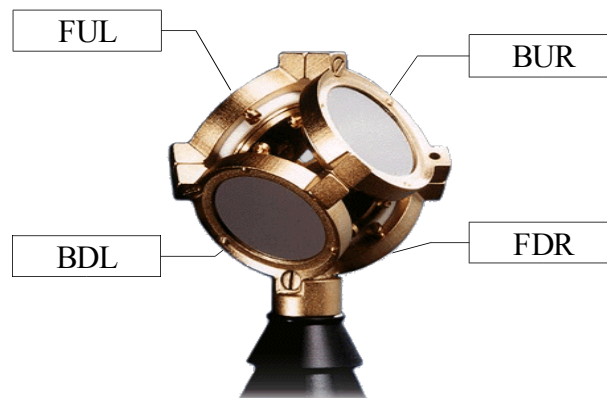
BDL: Back Down Left

FDR: Front Down Right

Αυτά είναι τα σήματα που λαμβάνουν οι τέσσερις κάψες ενός Soundfield μικροφώνου (βλ. Εικόνα 22). Αυτές βρίσκονται στις πλευρές ενός τετράεδρου (βλέπε Εικόνα 21). Το πολικό διάγραμμα που έχουν αυτές οι κάψες είναι υπο-καρδιά (sub-cardioid) . Τα σήματα αυτά θα εξισορροπηθούν για να αντισταθμιστεί η απόσταση που έχουν από το κέντρο του ηχητικού πεδίου (η του τετράεδρου) (βλέπε [11]). Αυτά τα σήματα θα τροφοδοτηθούν στην συνέχεια σε έναν κωδικοποιητή B-Format, για να δημιουργηθούν τα σήματα των σφαιρικών αρμονικών 0^{ης} και 1^{ης} τάξης. Οι λήψη των σημάτων των σφαιρικών αρμονικών θα μπορούσαν να επιτευχθούν με την χρήση ενός παντοκατευθυντικού (omni) μικροφώνου για την δημιουργία του σήματος (W) ενώ για τα σήματα (X), (Y), (Z) θα μπορούσε να γίνει λήψη με τρία διπολικά (Figure of eight) μικρόφωνα προσανατολισμένα προς τα θετικά των αξόνων X, Y, Z του συστήματος συντεταγμένων αντίστοιχα. Με αυτό τον τρόπο θα παίρναμε κατευθείαν το B-Format χωρίς να χρειαστεί κωδικοποίηση. Δεν είναι ανέφικτο όμως αυτά τα μικρόφωνα να είναι συμπτωτικά, γι' αυτό και δημιουργούνται μικρόφωνα όπως τα Soundfield, για την λήψη του ηχητικού πεδίου. Για την καλύτερη ηχοληψία του ηχητικού πεδίου θα ήταν απαραίτητη η λήψη σφαιρικών αρμονικών μεγαλύτερης τάξης. Τέτοια μικρόφωνα δεν έχουν υλοποιηθεί μέχρι στιγμής εκτός από μία θεωρητική προσεγγίσεις του Philip Cotterell στο PHD thesis για ένα μικρόφωνο δεύτερης τάξης, και της France Telecom που δουλεύει σε ένα πρότυπο μικρόφωνο τέταρτης τάξης, που δεν υπάρχει όμως καμία εγγύηση για την εμπορική του κυκλοφορία (βλ. [14]). Στις εικόνες (22) και (23) φαίνονται δύο μικρόφωνα soundfield.



Εικόνα 21: Απεικόνιση ενός τετράεδρου όπου με μπλε κυλίνδρους έχουν συμβολιστεί οι τέσσερις κάψες ενός Soundfield μικροφώνου.



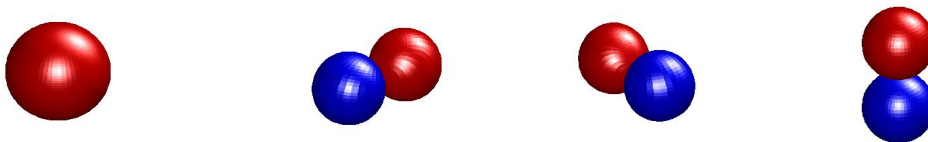
Εικόνα 22: Μικρόφωνο soundfield MK V. Στα πλαίσια απεικονίζονται οι ονομασίες των σημάτων που βγαίνουν από το μικρόφωνο και η αντιστοιχία τους με τις κάψες.



Εικόνα 23: Μικρόφωνο AGM MR2.

B-Format.

Το B-Format αποτελεί το τμήμα της αποσύνθεσης του ηχητικού πεδίου σε σφαιρικές αρμονικές για μια μέθοδο Ambisonic. Αποτελείται από τέσσερα σήματα (βλ. Εικόνα 24). Αυτά είναι τα (W), (X), (Y), (Z). Το (W) συμβολίζει την σφαιρική αρμονική 0^{ης} τάξης του ηχητικού πεδίου. Αυτό το σήμα ταυτίζεται με αυτό που θα λάμβανε ένα μικρόφωνο παντοκατευθυντικό (omni) αν βρισκόταν στο κέντρο του ηχητικού πεδίου. Τα (X), (Y), (Z) συμβολίζουν τις σφαιρικές αρμονικές 1^{ης} τάξης, και ταυτίζονται με τα σήματα που θα λάμβαναν τρία δικατευθυντικά (figure of eight) μικρόφωνα προσανατολισμένα στις 0^ο αζιμούθια γωνία και 0^ο γωνία ανύψωσης, 90^ο αζιμούθια γωνία και 0^ο γωνία ανύψωσης και 0^ο αζιμούθια γωνία και 90^ο γωνία ανύψωσης αντίστοιχα αν βρίσκονταν στο κέντρο του ηχητικού πεδίου. Το B-Format μπορεί να προκύψει είτε από την κωδικοποίηση ενός σήματος A-Format, είτε συνθετικά, υπολογίζοντας τις σφαιρικές αρμονικές ενός εικονικού ηχητικού πεδίου. Συνθετικά υπάρχει η δυνατότητα δημιουργίας εικονικών ηχητικών πεδίων μεγαλύτερης τάξης. Σε αυτή την περίπτωση θα υπήρχαν και τα σήματα μεγαλύτερης τάξης σφαιρικών αρμονικών όπως (S), (T), (U), (V) κτλ. (βλ. πίνακα 8).



Εικόνα 24: Απεικόνιση των πολικών διαγραμμάτων των σφαιρικών αρμονικών 0ης και 1ης τάξης. Το πρώτο εξ' αυτών συμβολίζει το σήμα W ενώ τα υπόλοιπα τα σήματα X, Y, Z αντίστοιχα.

D-Format

Το D-Format αφορά το τμήμα της μεθόδου Ambisonic της τροφοδότησης των ηχείων του συστήματος αναπαραγωγής για την ανασύνθεση του ηχητικού πεδίου. Αυτός η τυποποίηση χρησιμοποιείται για ιστορικούς λόγους, και δεν μπορεί να συγκεκριμενοποιηθεί ως προς τον αριθμό των σημάτων, καθώς αναφέρεται γενικά στα σήματα που θα τροφοδοτήσουν τα ηχεία, που ο αριθμός τους όμως ποικίλει ανάλογα με το σύστημα αναπαραγωγής. Έτσι για συγκεκριμένες διατάξεις και αριθμό ηχείων έχουν δημιουργηθεί άλλες τυποποιήσεις όπως το UHJ-Format και το G-Format.

UHJ-Format ή C-Format.

Το UHJ-Format (Universal HJ Format) ή αλλιώς C-Format αφορά το τμήμα της μεθόδου Ambisonic για την τροφοδότηση των ηχείων. Αναφέρεται σε συστήματα αναπαραγωγής με δύο ή τρία ηχεία. Αυτό το πρότυπο δημιουργήθηκε για να μπορούν ακροατές που δεν πρόκειται να ακούσουν Surround να μπορούν να ακούσουν την πληροφορία Ambisonic από δύο ή τρία κανάλια, ώστε να ακούν την πληροφορία του ηχητικού πεδίου από ένα Mono ή Stereo (η ανασύνθεση του ηχητικού πεδίου δεν μπορεί να γίνει από Mono, ενώ μπορεί να γίνει από Stereo). Η χρησιμότητα αυτού του πρωτοκόλλου έγκειται στην αναδρομική συμβατότητα με προηγούμενα συστήματα αναπαραγωγής που αποτελεί πλεονέκτημα για κάθε νέο καταναλωτικό σύστημα. Η

πλήρης μορφή αυτής της τυποποίησης αποτελείται από τέσσερα κανάλια με ονομασίες 'Q', 'T', 'R' και 'L'. Η χρήση της πλήρους μορφής αυτής της τυποποίησης μπορεί να δημιουργήσει ακριβής ηχητικές εικόνες και στις τρεις διαστάσεις όταν η αναπαραγωγή γίνεται από ένα Stereo. Το σήμα 'Q' περιέχει την πληροφορία του ηχητικού πεδίου που αφορά το ύψος. Η παράλειψή του αφαιρεί από το σύστημα αναπαραγωγής την δυνατότητα δημιουργίας ηχητικών πεδίων και στις τρεις διαστάσεις και περιορίζεται στη δημιουργία ηχητικών πεδίου στο οριζόντιο επίπεδο (βλέπε [14]). Το L αντιστοιχεί στο αριστερό κανάλι ενός συμβατού Stereo, ενώ το R στο δεξί. Το T επιτρέπει την πιο ακριβή οριζόντια αποκωδικοποίηση. Η αναλογία στα σήματα του B-format τα οποία είναι συνδυασμένα έτσι ώστε να δημιουργηθεί ένα C-Format σήμα, είναι προσεκτικά βελτιστοποιηθεί για την καλύτερη συμβατότητα με συμβατικά πρωτόκολλα Stereo και Mono. Αν οριστεί το Σ (όμοιο με το M στο MS stereo) σαν L+R και το Δ (όμοιο με το S στο MS stereo) σαν L-R, τότε:

$$\Sigma = 0.9397W + 0.1856X$$

$$\Delta = j(-0.3420W + 0.5099X) + 0.6555Y$$

$$T = j(-0.1432W + 0.6512X) - 0.7071Y$$

$$Q = 0.9772Z$$

Σχέση 20: Σχέσεις μετατροπής των σφαιρικών συντεταγμένων στο C-Format.

όπου j (ή $\sqrt{-1}$) αναπαριστάει μία πρόοδο της φάσης κατά 90° .

Το δεύτερο, τρίτο και τέταρτο κανάλι ενός C-Format σήματος μπορούν να χρησιμοποιηθούν με τον βαθμό ανάλυσης της κατεύθυνσης που είναι επιθυμητή, με διαθέσιμα δύο και μισό κανάλια επιλογής, όπου το τρίτο κανάλι (T) είναι περιορισμένου συχνοτικού εύρους. Για συμβατότητα με το Stereo χρησιμοποιούνται μόνο τα L και R (το L και R είναι αντίστοιχα $0.5(\Sigma + \Delta)$ και $0.5(\Sigma - \Delta)$).

G-Format.

Το G-Format αφορά το τμήμα της μεθόδου Ambisonic για την τροφοδότηση των ηχείων, ώστε να ανασυσταθεί το ηχητικό πεδίο. Προορίζεται για την τροφοδότηση των ηχείων ενός σύστημα αναπαραγωγής που συμφωνεί με το πρότυπο του ITU "5.1 Surround sound". Οι προδιαγραφές αυτής της τυποποίησης δεν έχουν οριστεί γι' αυτό και δεν έχει χρησιμοποιηθεί ευρέως.

Μέσα όπως το DVD και το SACD έχουν την δυνατότητα μεταφοράς δύο επιπλέον καναλιών ήχου. Την ώρα της αποκωδικοποίησης του B-Format για την δημιουργία του G-Format, υπάρχει η δυνατότητα της ταυτόχρονης κωδικοποίησης και του UHJ-Format. Με αυτό τον τρόπο είναι η εφικτή η μεταφορά από αυτά τα μέσα του G και του UHJ-Format, ώστε να μπορούν οι ακροατές να ακούσουν Surround, Stereo ή Mono. Αυτή η τυποποίηση ονομάστηκε G+2-Format (βλέπε [14]).

2.3.3 Πλεονεκτήματα.

- Είναι δυνατή η ηχογράφηση τρισδιάστατων ηχητικών πεδίων.
- Είναι ιεραρχική μέθοδος, δηλαδή αφού έχει γίνει κωδικοποίηση του ηχητικού πεδίου είναι δυνατή η αναπαραγωγή του ασχέτως με τον αριθμό των ηχείων που θα χρησιμοποιηθεί (φυσικά χρειάζεται κατάλληλος αποκωδικοποιητής).

- Τα κωδικοποιημένα σήματα με κατάλληλη επεξεργασία μπορούν να αναπαραχθούν και από άλλα συστήματα όπως 5.1 surround, Binaural, Transaural και έχει συμβατότητα με μονοφωνική και στερεοφωνική αναπαραγωγή.
- Στα συστήματα πρώτης τάξης χρειάζονται μόλις 4 κανάλια για την ηχογράφιση και καταγραφή των ηχητικών πεδίων, ενώ για την ηχογράφιση πεδίο στο οριζόντιο επίπεδο χρειάζονται μόλις 3 κανάλια.
- Υπάρχει διατήρηση του ηχητικού πεδίου κατά την περιστροφή του κεφαλιού.

2.3.4 Μειονεκτήματα.

- Υπάρχει συγκεκριμένο σημείο στο οποίο γίνεται βέλτιστη αναπαραγωγή.
- Για καλύτερη αναπαραγωγή του ηχητικού πεδίου πρέπει να γίνει ηχογράφιση και μεγαλύτερης τάξης σφαιρικών αρμονικών του ηχητικού πεδίου, πράγμα που δεν είναι εφικτό μέχρι στιγμής.
- Δεν υπάρχουν όλα τα δεδομένα για την αίσθηση της απόστασης, όπως διαφορά στις γωνίες πρόσπτωσης στα αυτιά.
- Κατά την ηχογράφιση στα σήματα του A-Format πρέπει να γίνει ισοστάθμιση στις χαμηλές συχνότητες λόγω της δημιουργίας του proximity effect όταν το ηχητικό πεδίο βρίσκεται κοντά στα μικρόφωνα (Για τα σήματα X, Y, Z).

2.3.5 Κωδικοποιητής (Encoder) Ambisonic B Format .

Ο κωδικοποιητής αναλαμβάνει να επεξεργαστεί κάποια σήματα στην είσοδό του και να εξάγει κάποια άλλα τα οποία θα είναι το αποτέλεσμα της επεξεργασίας και θα έχουν κάποια συγκεκριμένα χαρακτηριστικά (βλέπε [5]). Τα σήματα εξόδου θα αποτελούν το αποσυντεθειμένο ηχητικό πεδίο βάση των σφαιρικών αρμονικών. Ο αριθμός των σημάτων εισόδου και εξόδου ποικίλει ανάλογα με την τάξη του κωδικοποιητή. Επίσης, τα σήματα που πρέπει να εισαχθούν στον κωδικοποιητή διαφοροποιούνται στη περίπτωση ενός πραγματικού ηχητικού πεδίου και ενός εικονικού που πρόκειται να συντεθεί. Ένας κωδικοποιητής Ambisonic B-Format πρώτης τάξης μπορεί να λαμβάνει στην είσοδό του είτε τα τέσσερα σήματα από ένα μικρόφωνο Soundfield (υπάρχον ηχητικό πεδίο), είτε τον μονοφωνικό ήχο που θέλουμε να χωροτοποθετήσουμε μαζί με την πληροφορία της αξιμούθιας και κατακόρυφης γωνίας της χωρικής θέσης αυτού, σε σχέση με την κέντρο αναπαραγωγής⁵ (εικονικό ηχητικό πεδίο που πρόκειται να συντεθεί). Στην περίπτωση ενός κωδικοποιητή πρώτης τάξης τα σήματα που θα εξαχθούν από αυτόν είναι τέσσερα (W, X, Y, Z), που αντιπροσωπεύουν τις σφαιρικές αρμονικές 0^{ης} και 1^{ης} τάξης του ηχητικού πεδίου (βλέπε παράρτημα 1 πίνακας 8).

Encoder B-Format από A-Format (υπάρχον ηχητικό πεδίο).

Το A-Format αποτελείται από τέσσερα σήματα (FUL,FDR,BDL,BUR) τα οποία προέρχονται από κάψες με υποκαρδιοειδές πολικό διάγραμμα. Ας θεωρήσουμε λοιπόν ότι οι κάψες των μικροφώνων είναι συμπτωτικές. Η εξίσωση που δίνει την εξίσωση του μικροφώνου είναι (βλέπε παράρτημα 3):

⁵ Η πληροφορία της αξιμούθιας και κατακόρυφης γωνίας ορίζει την θέση πάνω σε στην επιφάνεια μια σφαίρας από την οποία θα θέλαμε να βρίσκεται η ηχητική πηγή που θα παρήγαγε τον συγκεκριμένο ήχο.

$$M_{pollar}(\theta, \phi) = C + (1 - C)\cos(\phi)\cos(\Phi)\sin(\theta)\sin(\Theta) + (1 - C)\cos(\phi)\cos(\Phi)\cos(\theta)\cos(\Theta) + \dots \\ \dots(1 - C)\sin(\phi)\sin(\Phi)$$

Σχέση 21:

Αντικαθιστώντας τις συναρτήσεις της 'Σχέση 21' στην 'Σχέση 48' έχουμε :

$$S_{mic} = \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi) * C + \dots \\ \dots \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi) * (1 - C)\cos(\phi)\cos(\Phi)\sin(\theta)\sin(\Theta) + \dots \\ \dots \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi) * (1 - C)\cos(\phi)\cos(\Phi)\cos(\theta)\cos(\Theta) + \dots \\ \dots \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi) * (1 - C)\sin(\phi)\sin(\Phi)$$

Σχέση 22:

Για διευκόλυνση των πράξεων θέτουμε τις παρακάτω σχέσεις:

$$a(\theta, \phi) = \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi) \\ b(\theta, \phi) = \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi)\sin(\theta)\cos(\phi) \\ e(\theta, \phi) = \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi)\cos(\theta)\cos(\phi) \\ g(\theta, \phi) = \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi)\sin(\phi)$$

Σχέση 23:

Αντικαθιστώντας την 'Σχέση 23' στη 'Σχέση 22' έχουμε:

$$S_{mic} = Ca(\theta, \phi) + (1-C)\cos(\Phi)\sin(\Theta)b(\theta, \phi) + \dots$$

$$\dots(1-C)\cos(\Phi)\cos(\Theta)e(\theta, \phi) + (1-C)\sin(\Phi)g(\theta, \phi)$$

Σχέση 24:

Σύμφωνα με τη 'Σχέση 24', τις γωνίες των μικροφώνων και την κατευθυντικότητα τους (subcardioid, $C=3/4$) που έχουν υπολογιστεί στο "Παράρτημα 6", οι εξισώσεις λήψης των τεσσάρων μικροφώνων σε ένα Soundfield μικρόφωνο είναι:

$$FUL = \frac{3}{4}a(\theta, \phi) + \frac{1}{4}\cos(35.2642)\sin(45)b(\theta, \phi) + \dots$$

$$\dots\frac{1}{4}\cos(35.2642)\cos(45)e(\theta, \phi) + \frac{1}{4}\sin(35.2642)g(\theta, \phi)$$

$$BDL = \frac{3}{4}a(\theta, \phi) + \frac{1}{4}\cos(-35.2642)\sin(135)b(\theta, \phi) + \dots$$

$$\dots\frac{1}{4}\cos(-35.2642)\cos(135)e(\theta, \phi) + \frac{1}{4}\sin(-35.2642)g(\theta, \phi)$$

$$BUR = \frac{3}{4}a(\theta, \phi) + \frac{1}{4}\cos(35.2642)\sin(-135)b(\theta, \phi) + \dots$$

$$\dots\frac{1}{4}\cos(35.2642)\cos(-135)e(\theta, \phi) + \frac{1}{4}\sin(35.2642)g(\theta, \phi)$$

$$FDR = \frac{3}{4}a(\theta, \phi) + \frac{1}{4}\cos(-35.2642)\sin(-45)b(\theta, \phi) + \dots$$

$$\dots\frac{1}{4}\cos(-35.2642)\cos(-45)e(\theta, \phi) + \frac{1}{4}\sin(-35.2642)g(\theta, \phi)$$

Σχέση 25:

Αντίστοιχα οι εξισώσεις που αναπαριστούν τα σήματα των σφαιρικών αρμονικών 1^{ης} και 0^{ης} τάξης είναι:

$$W = a$$

$$X = e$$

$$Y = b$$

$$Z = g$$

Σχέση 26:

Παρατηρώντας την 'Σχέση 25' και κάνοντας πράξει μπορούν να προκύψουν τα σήματα των σφαιρικών αρμονικών της 'Σχέση 26'. Έτσι έχουμε:

$$W = \frac{1}{3}[FUL + BDL + BUR + FDR] = 0.3333[FUL + BDL + BUR + FDR]$$

$$X = \frac{1}{\cos(35.2642)\cos(45)}[FUL - BDL + FDR - BUR] = 1.7320[FUL - BDL + FDR - BUR]$$

$$Y = \frac{1}{\cos(35.2642)\sin(45)}[FUL + BDL - FDR - BUR] = 1.7320[FUL + BDL - FDR - BUR]$$

$$Z = \frac{1}{\sin(35.2642)}[FUL - BDL - FDR + BUR] = 1.7321[FUL - BDL - FDR + BUR]$$

Σχέση 27:

Από την 'Σχέση 27' αποδείχθηκε ότι οι σφαιρικές αρμονικές μπορούν να προκύψουν από απλές προσθαιρέσεις των σημάτων των που λαμβάνουν οι κάψες ενός Soundfield Μικροφώνου. Στην πραγματικότητα αυτό δεν είναι μια τόσο απλή διαδικασία καθώς τα μικρόφωνα δεν είναι συμπτωτικά αλλά σχεδόν συμπτωτικά. Τα προβλήματα λοιπόν δημιουργούνται κυρίως λόγω της διαφοράς απόστασης μεταξύ των καψών. Η διαφορά απόστασης έχει σαν αποτέλεσμα την χρονική διαφορά λήψης των σημάτων. Δημιουργούνται έτσι φασικές διαφορές και στη διαδικασία μετατροπής του σήματος A-Format σε B-Format (προσθαιρέσεις μεταξύ των σημάτων) υπάρχουν ανεπιθύμητες καταστροφικές και ενισχυτικές συμβολές (Comb-filter). Ένα δεύτερο πρόβλημα που δημιουργεί αυτή η απόσταση είναι ότι υπάρχει διαφορά στις γωνίες πρόσπτωσης του ήχου, άρα παραμόρφωση του ηχητικού πεδίου. Πρόβλημα αποτελεί και το γεγονός ότι τα μικρόφωνα, όσο ποιοτικά και να είναι, ποτέ δεν είναι ιδανικά. Το πολικό τους διάγραμμα θα αλλάξει ανάλογα με την συχνότητα και συνήθως όσο πιο υψηλή είναι η συχνότητα, τόσο πιο κατευθυντική είναι η λήψη τους. Αυτό έχει σαν αποτέλεσμα την λανθασμένη λήψη του ηχητικού πεδίου κυρίως στις υψηλές συχνότητες. Τέλος για πηγές ήχου που βρίσκονται σε απόσταση μικρότερη του μέτρου από τα μικρόφωνα, υπάρχει μια ενίσχυση της έντασης που λαμβάνουν στις χαμηλές συχνότητες λόγω του proximity effect και έχει σαν αποτέλεσμα την λανθασμένη λήψη του ηχητικού πεδίου φασματικά. Τα περισσότερα από αυτά τα προβλήματα μπορούν να λυθούν με την χρήση της συσκευής που συνοδεύει τα Soundfield μικρόφωνα. Με αυτή τη συσκευή ο χρήστης μπορεί να βαθμονομήσει και να ισοσταθμίσει τις λήψεις των τεσσάρων σημάτων ώστε να ληφθεί όσο πιο σωστά γίνεται το ηχητικό πεδίο. Στη Σχέση 28 φαίνεται ο τρόπος μετατροπής του A-Format σήματος σε B-Format, όπου με τις συναρτήσεις F_x συμβολίζονται οι συναρτήσεις όλων των επεξεργασιών που γίνονται στη συσκευή.

$$W = F_0(\omega) \times (FUL + BDL + FDR + BUR)$$

$$X = F_1(\omega) \times (FUL - BDL + FDR - BUR)$$

$$Y = F_1(\omega) \times (FUL + BDL - FDR - BUR)$$

$$Z = F_1(\omega) \times (FUL - BDL - FDR + BUR)$$

Σχέση 28: Ο ολοκληρωμένος τύπος μετατροπής του A σε B format, όπου F_0 και F_1 είναι συναρτήσεις ισοστάθμισης που εξαρτώνται από την συχνότητα του σήματος.

Σε αυτή την εργασία για την δημιουργία της μετατροπής του A-Format σήματος σε B-Format θα θεωρηθεί ότι οι κάψες των μικροφώνων είναι ιδανικές και τοποθετημένες συμπτωτικά.

Encoder B-Format από εικονικό ηχητικό πεδίο (σύνθεση).

Σε αυτή την περίπτωση μπορούμε να γίνει μια εικονική ηχογράφιση του ηχητικού πεδίου. Δεν χρειάζεται να γίνει προσομοίωση των σημάτων που θα λάμβανε ένα Soundfield μικρόφωνο. Αντίθετα μπορούμε να προσομοιώσουμε τα τρία δικατευθυντικά μικρόφωνα (Figure-8) προσανατολισμένα στα θετικά των τριών αξόνων συντεταγμένων που θα λάμβανε την ένταση του ηχητικού πεδίου, και ένα παντοκατευθυντικό (omni) μικρόφωνο που θα λαμβάνει την πίεση του πεδίου. Τα τρία αυτά μικρόφωνα θα τεθούν συμπτωτικά στο κέντρο του ηχητικού πεδίου. Με αυτόν τον τρόπο λαμβάνουμε τις σφαιρικές αρμονικές 0^{th} και 1^{st} τάξης, άρα δημιουργείται απευθείας το B-Format. Στη Σχέση 29 φαίνονται οι μαθηματικοί τύποι που θα έδιναν τα σήματα που θα λάμβαναν αυτά τα μικρόφωνα.

$$W=1/(2^{0.5})$$

$$X=\cos(\theta)*\cos(\varphi)$$

$$Y=\sin(\theta)*\cos(\varphi)$$

$$Z=\sin(\varphi)$$

Σχέση 29: Μαθηματικοί τύποι που αναπαριστούν την απόκριση των εικονικών μικροφώνων. Τα W,X,Y,Z αναπαριστούν τα σήματα του B-Format. Βλέπε σχέδιο 1 για τις γωνίες.

Ο πολλαπλασιασμός της έντασης του W με $1/(2^{0.5})$ (ή αλλιώς μείωση της έντασης του 3 dB) γίνεται για συμβατότητα με παλιότερα συστήματα όπου κατά την ηχογράφισης ηχητικών πεδίων όπου κατά την ηχογράφιση υπήρχε αυτή η διαφορά στην ένταση μεταξύ του σήματος W και των X,Y,Z, σύμφωνα με το άρθρο [5]. Στην πράξη γίνεται για την εκμετάλλευση του δυναμικού εύρους που έχει ένα σύστημα ηχογράφισης, καθώς το σήμα 'W' έχει εμφανώς μεγαλύτερη ένταση από τα σήματα 'X', 'Y' και 'Z'. Κατά την αποκωδικοποίηση γίνεται επανάκτηση της έντασης του W πολλαπλασιάζοντας την έντασή του με $2^{0.5}$ (ή αλλιώς αύξηση του σήματος κατά 3 dB).

2.3.6 Ambisonic αποκωδικοποιητής (Decoder).

Ο αποκωδικοποιητής αναλαμβάνει να επεξεργαστεί τα σήματα του B-Format κατάλληλα και να τα διανέμει στα ηχεία, έτσι ώστε να ανασυντεθεί το ηχητικό πεδίο (βλέπε [25]). Σκοπός της επεξεργασίας είναι να δημιουργηθεί στον ακροατή η αίσθηση του αρχικού ηχητικού πεδίου. Ο άνθρωπος αντιλαμβάνεται την γωνία προέλευσης του ήχου αξιολογώντας κάποια στοιχεία κατευθυντικότητας (βλέπε κεφ. 2). Έτσι δεν είναι σκοπός της επεξεργασίας να ανασυνθέσει ακριβώς το ίδιο ηχητικό πεδίο με το αρχικό, αλλά να τροφοδοτήσει τον ακροατή με αυτά τα στοιχεία προέλευσης του ήχου, αξιολογώντας τα οποία να του δημιουργηθεί η ίδια αίσθηση προέλευσης του ήχου, με αυτή που θα του δημιουργόταν αν βρισκόταν στο αρχικό ηχητικό πεδίο.

Για να είναι εφικτή η αξιολόγηση της ακρίβειας ενός αναπαραγόμενου ηχητικού πεδίου θα πρέπει να γίνει ένας συσχετισμός της γωνίας εντοπισμού που γίνεται αντιληπτή από το ανθρώπινο σύστημα ακοής, με την πραγματική γωνία εντοπισμού του εικονικού ηχητικού πεδίου (ή πραγματικού εφόσον πρόκειται για ηχογραφημένο). Υπάρχουν δύο ψυχοακουστικές θεωρίες για τον υπολογισμό της αντιληπτής γωνίας εντοπισμού του ήχου, μια που βασίζεται στη διαφορά φάσης του ήχου που προσπίπτει στα αυτιά, και αυτή που βασίζεται στην διαφορά έντασης.

Η πρώτη ψυχοακουστική θεωρία που χρησιμοποιείται βασίζεται στις διαφορές φάσης⁶ μεταξύ των σημάτων που φτάνουν στα αυτιά αναφέρονται κυρίως για συχνότητες κάτω των 700 Hz. Βάση αυτής το διάνυσμα εντοπισμού του αντιληπτού ήχου προσδιορίζεται με το πηλίκο της συνολικής ενίσχυσης του διανύσματος της τάσης του αναπαραγόμενου ήχου προς την ενίσχυση της ακουστικής πίεσης στο σημείο του ακροατή. Ας υποθέσουμε λοιπόν ότι έχουμε 'i' ηχεία τοποθετημένα σε ίσες αποστάσεις από τον ακροατή στο οριζόντιο επίπεδο, σε αξιμούθιες γωνίες 'θ_i', τα οποία προσδίδουν ενίσχυση στον ήχο κατά 'g_i' αντίστοιχα (υποθέτουμε ότι η ενίσχυση είναι πραγματικός αριθμός για ευκολία). Η ακουστική ενίσχυση της πίεσης στη θέση του ακροατή είναι το άθροισμα των ενισχύσεων που προσδίδει το κάθε ηχείο, και ο τύπος υπολογισμού του φαίνεται στη Σχέση 30.

⁶ διαφορετικά μπορούν να βρεθούν και σαν διαφορά χρόνου, ITD

$$P = \sum_i g_i$$

Σχέση 30:

Η αύξηση της τάσης του σήματος είναι η πρόσθεση των διανυσμάτων αύξησης της πίεσης του κάθε ηχείου 'g'. Το κάθε ένα από αυτά έχει προσανατολισμό 'θ_i' στο οριζόντιο επίπεδο

(διάνυσμα προσανατολισμένο προς το αντίστοιχο ηχείο κάθε φορά). Το διάνυσμα τάσης θα μπορούσε να αναλυθεί σε συνιστώσες στον 'X' και 'Y' (βλέπε Σχέση 31).

$$V_x = \sum_i g_i \cos(\theta_i)$$

$$V_y = \sum_i g_i \sin(\theta_i)$$

Σχέση 31:

Διαιρώντας⁷ λοιπόν αυτό το διάνυσμα της τάσης με την συνολική πίεση 'P', προκύπτει το διάνυσμα της τάσης 'r_v' (ή αλλιώς Velocity localization vector) και προσανατολισμένο προς την γωνία 'θ_v' (βλέπε Σχέση 32).

$$r_v \cos(\theta_v) = V_x / P$$

$$r_v \sin(\theta_v) = V_y / P$$

Σχέση 32

Η γωνία 'θ_v' ονομάζεται 'αζιμούθια γωνία εντοπισμού του διανύσματος τάσης' ('velocity vector localization azimuth', ή αλλιώς 'Makita localization azimuth'). Το μέτρο του διανύσματος 'r_v' θα είναι μεγαλύτερο ή ίσο με το μηδέν⁸. Στην περίπτωση που δεν είχαμε αναπαραγωγή ηχητικού πεδίου από μια συστοιχία ηχείων, αλλά από μια πραγματική πηγή, το μέτρο του 'r_v' θα ήταν ίσο με μονάδα και η γωνία 'θ_v' θα ταυτιζόταν με την πραγματική γωνία προέλευσης της πηγής (αυτό προκύπτει από αντικατάσταση στις σχέσεις 13-15).

Η δεύτερη ψυχοακουστική θεωρία αφορά την διαφορά έντασης του ήχου που προσπίπτει στα αυτιά (IID). Αυτή αποτελεί κριτήριο εντοπισμού του ήχου για το εύρος των συχνοτήτων 700 Hz με 4kHz, και για συχνότητες κάτω από τα 700 Hz για ακροατές που δεν βρίσκονται στο κέντρο (Hot spot). Βάση αυτής το διάνυσμα εντοπισμού του αντιληπτού ήχου μπορεί να υπολογιστεί με τον λόγο της αύξησης της έντασης όλων των ηχείων, προς την αύξηση της ακουστικής έντασης του αναπαραγόμενου ήχου. Θεωρώντας την ίδια διάταξη ηχείων που αναφέρθηκε και στην προηγούμενη ψυχοακουστική θεωρία, η συνολική ακουστική ένταση του αναπαραγόμενου ήχου στο σημείο του ακροατή υπολογίζεται από τη Σχέση 33.

$$E = \sum_i g_i^2$$

Σχέση 33:

7 Η διαίρεση γίνεται για ανεξαρτητοποίηση του μέτρου του 'διανύσματος εντοπισμού τάσης' από την ένταση του ήχου. Γίνεται λοιπόν μια αναγωγή στη μονάδα.

8 Πάντα το μέτρο ενός διανύσματος είναι μεγαλύτερο από το μηδέν, απλά γίνεται η διευκρίνιση για να μην γίνει εσφαλμένος υπολογισμός της γωνίας, θεωρώντας ότι το μέτρο είναι αρνητικό.

Αντίστοιχα η αύξηση της ηχητικής έντασης υπολογίζεται από το άθροισμα των διανυσμάτων έντασης του κάθε ηχείου, που έχει διεύθυνση και φορά από τον ακροατή προς το αντίστοιχο ηχείο. Οι συνιστώσες αυτού φαίνονται στη Σχέση 34.

$$E_x = \sum_1^j g_i^2 \cos(\theta_i)$$

$$E_y = \sum_1^j g_i^2 \sin(\theta_i)$$

Σχέση 34:

Διαιρώντας⁹ λοιπόν την διάνυσμα αύξησης της έντασης όλων των ηχείων με την συνολική αύξηση της έντασης του αναπαραγόμενου ήχου παίρνουμε το 'διάνυσμα εντοπισμού βάσης της έντασης' (Energy Localization vector) το οποίο συμβολίζεται ως 'r_E'. Το μέτρο αυτού μπορεί να πάρει τιμές μεγαλύτερες ή ίσες του μηδενός¹⁰ και έχει διεύθυνση 'θ_E'.

$$r_E \cos(\theta_E) = E_x / E$$

$$r_E \sin(\theta_E) = E_y / E$$

Σχέση 35:

Η γωνία 'θ_E' ονομάζεται και 'γωνία εντοπισμού του διανύσματος έντασης' (ή αλλιώς 'Energy vector localization azimuth'). Και σε αυτή την περίπτωση άμα είχαμε πραγματική πηγή που προερχόταν ο ήχος, το 'r_E' θα έπαιρνε την τιμή 1. Έχει βρεθεί ότι ανεξάρτητα με τη διάταξη ή των αριθμό των ηχείων που πρόκειται να χρησιμοποιηθεί, οι δύο γωνίες εντοπισμού, βάση των δύο ψυχοακουστικών μεθόδων, πρέπει ιδανικά να είναι ίσες για οποιαδήποτε προσομοίωση γωνίας εντοπισμού (βλέπε Σχέση 36), τουλάχιστον για ακουστές συχνότητες μέχρι τα 3.5 ή 4 kHz.

$$\theta_V = \theta_E$$

Σχέση 36:

Αυτό διασφαλίζει ότι με χρήση διαφορετικού μηχανισμού εντοπισμού της προέλευσης του ήχου το ακουστικό σύστημα θα αντιλαμβάνονται την ίδια γωνία προέλευσης σε ένα ευρύ συχνοτικό εύρος, και ιδιαίτερα σε συχνοτικά εύρη όπου το σύστημα ακοής αποτιμά και τους δύο μηχανισμούς για τον εντοπισμό. Έχει βρεθεί λοιπόν ότι όταν ισχύει η Σχέση 36 η ψευδαίσθηση της προέλευσης της ηχητικής εικόνας είναι πιο ακριβής και ξεκάθαρη, σε σχέση με περιπτώσεις όπου αυτές οι γωνίες διαφέρουν ουσιαστικά.

Το γεγονός ότι η γωνία προέλευσης της ηχητικής πηγής είναι ακριβής και ξεκάθαρη δεν εξασφαλίζει και την σταθερότητα της ηχητικής πηγής πχ σε κινήσεις του κεφαλιού. Για την εξασφάλιση της σταθερότητας της ηχητικής εικόνας πρέπει να ληφθούν υπόψη περισσότερες προϋποθέσεις που αφορούν τα 'r_E' και 'r_V'. Έτσι για οικιακά συστήματα αναπαραγωγής surround ήχου είναι απαραίτητη η διασφάλιση ότι το 'r_V' θα παίρνει την τιμή '1' τυπικά για τις συχνότητες κάτω από τα 400 Hz, για ακρόαση στο κέντρο της διάταξης των ηχείων. Αντίστοιχα για τις τιμές

9 Η διαίρεση γίνεται για ανεξαρτητοποίηση του μέτρου του 'διανύσματος έντασης εντοπισμού' από την ένταση του ήχου. Γίνεται λοιπόν μια αναγωγή στη μονάδα.

10 Πάντα το μέτρο ενός διανύσματος είναι μεγαλύτερο από το μηδέν, απλά γίνεται η διευκρίνιση για να μην γίνει εσφαλμένος υπολογισμός της γωνίας, θεωρώντας ότι το μέτρο είναι αρνητικό.

συχνοτήτων πάνω από τα 400 Hz είναι επιθυμητό οι τιμές του r_E να παίρνει την μέγιστη δυνατή τιμή. Συνήθως γίνονται κάποιοι συμβιβασμοί στην σχεδίαση του decoder, καθώς είναι αδύνατη η ταυτόχρονη ταύτιση της μεγιστοποίησης του r_E για όλες τις αναπαραγόμενες διευθύνσεις. Ο συμβιβασμός που γίνεται είναι η σχεδίαση έτσι ώστε τα r_E να παίρνουν τις μέγιστες τιμές (για συχνότητες πάνω από τα 400 Hz) κυρίως στο μπροστινό πεδίο, προσέχοντας όμως να μην γίνει υπερβολικά ασταθές το πλευρικό και το πίσω ηχητικό πεδίο. Ένα άλλο ανεπιθύμητο φαινόμενο είναι η διαφορά $1-r_E$ να είναι μεγάλη σε μικρές αλλαγές της γωνίας, καθώς θα παρουσιάζονται γειτονικοί ήχοι με λιγότερη ή περισσότερη σταθερότητα. Ας δούμε λοιπόν τα χαρακτηριστικά που πρέπει να διαθέτει ένας αποκωδικοποιητής για να ονομάζεται Ambisonic.

Για να ονομάζεται ένας αποκωδικοποιητής Ambisonic, κατά την αναπαραγωγή ηχητικών πεδίων στις 360°, πρέπει να πληρεί κάποιες προϋποθέσεις που αφορούν το ηχητικό πεδίο που λαμβάνει ένας ακροατής που κάθεται στο κέντρο του συστήματος αναπαραγωγής. Σύμφωνα με τον M.Gerzon στο [25] αυτές είναι:

- Η γωνία του energy και velocity vector του αναπαραγόμενου πεδίου πρέπει να είναι ίσες και αμετάβλητες ανεξαρτήτως της συχνότητας, για τουλάχιστον την συχνοτική περιοχή μέχρι τα 4 kHz.
- Για τις συχνότητες κάτω από των 400 Hz πρέπει το διάνυσμα εντοπισμού της τάσης r_V (velocity localization vector) να ισούται με την μονάδα για όλες τις αναπαραγόμενες αζιμουθιακές γωνίες.
- Για τις μεσαίες και υψηλές συχνότητες ανάμεσα στα 700 Hz και τα 4 kHz το ένταση του διανύσματος της έντασης r_E (Energy localization Vector) πρέπει να είναι όσο το δυνατόν πιο κοντά στη μέγιστη τιμή του για όσο το δυνατό μεγαλύτερο μέρος του ηχητικού πεδίου των 360°.

Η επίτευξη των παραπάνω προϋποθέσεων μπορεί να πραγματοποιηθεί σε μεγάλο βαθμό με την χρήση συστημάτων αναπαραγωγής με διάταξη ηχείων με αυξημένη συμμετρία, όπως κανονικά N-γωνια, τετράγωνα ή διατάξεις με αντιδιαμετρικά ζεύγη ηχείων. Παρ' αυτά στο [25] ο M. Gerzon παρουσιάζει συστήματα αναπαραγωγής Ambisonic που παρουσιάζουν λιγότερη συμμετρία.

Σύμφωνα με τον M.Gerzon στο [25] υπάρχουν δύο εμπειρικοί κανόνες για την σχεδίαση αποκωδικοποιητών υψηλών συχνοτήτων. Σύμφωνα με τον πρώτο κανόνα, όταν το r_E παίρνει την μέγιστη τιμή, έχει βρεθεί ότι το r_E ισούται με το r_V , έτσι ώστε οι καλύτερες σχεδιάσεις κωδικοποιητών για τις υψηλές συχνότητες ήταν όταν το r_V ποικίλει ανάλογα με την αζιμούθια γωνία κωδικοποίησης, με γενικό μοτίβο ίδιο με αυτό του r_E , όσο αυτό είναι δυνατό. Γενικά το r_E ποικίλει με την αζιμούθια γωνία με μεγαλύτερο αρμονικά στοιχεία σε σχέση με το r_V , και έτσι σπάνια αυτά τα δύο έχουν ακριβώς την ίδια τιμή για όλες τις γωνίες. Ένα καλό όμως ταίριασμα αυτών των δύο γενικά δίνει τα καλύτερα δυνατά αποτελέσματα στις υψηλές συχνότητες.

Κατά τον δεύτερο κανόνα, σύμφωνα με τον M.Gerzon στο [25], η αύξηση της συνολικής εκπεμπόμενης ενέργειας E τείνει να είναι μέγιστη όταν το r_E είναι ελάχιστο, και το αντίθετο. Αυτό δυστυχώς είναι σε τέτοιο βαθμό ώστε να δίνεται έμφαση στις συμβολές των διευθύνσεων που δεν είναι τόσα καλά εντοπισμένες. Κατά την σχεδίαση κωδικοποιητών στους οποίους έχει επιλεγεί να παίρνεται το μέγιστο r_E στην μπροστινή διεύθυνση, καλό είναι να μετατραπούν οι κωδικοποιημένες γωνίες του B-Format με τέτοιο τρόπο ώστε να επικρατεί η 'κυριαρχία' της μπροστινής διεύθυνσης ώστε να αυξηθεί η αντίστοιχη ένταση των μπροστινών καναλιών και έτσι να αντισταθμιστεί αυτή η απώλεια ενίσχυσης. Γενικά τέτοιες τροποποιήσεις θα μετατοπίσουν τις αντιληπτές γωνίες προέλευσης σε σχέση με αυτές του κωδικοποιημένου ηχητικού

πεδίου, και πλέων δεν θα υπάρχει ταύτιση αυτών. Στην πράξη τέτοιες τροποποιήσεις δεν θα είναι τόσο μεγάλες που να δημιουργήσουν τόσο μεγάλες παραμορφώσεις του αρχικού ηχητικού πεδίου.

Για την υλοποίηση του αποκωδικοποιητή Ambisonic πρόκειται να υπολογιστούν οι συντελεστές των σφαιρικών αρμονικών ώστε να τροφοδοτηθεί κατάλληλα η διάταξη των ηχείων που πρόκειται να γίνει η αναπαραγωγή του ηχητικού πεδίου. Όπως έχει προαναφερθεί σε αυτή την παράγραφο πρόκειται να υπολογιστούν συντελεστές ώστε να γίνεται σωστή αναπαραγωγή του ηχητικού πεδίου κάτω από τα 400 Hz (σύμφωνα με την ψυχοακουστική θεωρία της διαφοράς φάσης) και άλλοι για αναπαραγωγή πάνω από τα 400 Hz (σύμφωνα με την θεωρία της διαφοράς έντασης). Αυτό σημαίνει ότι πρόκειται να χρησιμοποιηθούν δύο φίλτρα που θα διαχωρίσουν τα κωδικοποιημένα σήματα B-Format σε δύο συχνοτικά εύρη. Σύμφωνα με τον M.Gerzon στο [25], θα πρέπει να γίνει αντιστάθμιση της διαφοράς φάσης που ενδεχομένως να υπάρχει λόγω της επιρροής των φίλτρων, καθώς κάθε λάθος στη φάση πρόκειται να μεγεθυνθεί λόγω του μεγάλου γωνιακού εύρους (360°) διανομής αυτών των σημάτων. Ας δούμε όμως ποια είναι η συνάρτηση υπολογισμού του σήματος που θα λάβει το κάθε ηχείο.

Στην περίπτωση που όλα τα ηχεία ισαπέχουν από τον ακροατή, η γενική εξίσωση υπολογισμού του σήματος που φτάνει στο ακροατή προκύπτει από μια πρόσθεση με διαφορετικά βάρη των σφαιρικών αρμονικών. Ταυτόχρονα, καθένα από τα σήματα (XYZ) λαμβάνουν μέρος σε αυτή την πρόσθεση βάση της προβολής αυτών στο διάνυσμα με κατεύθυνση προς το ηχείο από την θέση του ακροατή. Η εξίσωση λοιπόν για τον υπολογισμό του σήματος που πρέπει να τροφοδοτηθεί το κάθε ηχείο φαίνεται στη Σχέση 37.

$$g_i = \alpha * W + \beta * X * \cos(\theta) * \cos(\varphi) + \gamma * Y * \sin(\theta) * \cos(\varphi) + \delta * Z * \sin(\varphi)$$

Σχέση 37: Γενική σχέση υπολογισμού του σήματος που πρέπει να λάβει το κάθε ηχείο της διάταξης αναπαραγωγής για να ανασυνθέσει το ηχητικό πεδίο.

όπου τα 'α,β,γ,δ' μεταβάλλονται ανάλογα με το εύρος των συχνοτήτων, πχ υψηλές συχνότητες και χαμηλές.

Σύμφωνα με τον Farina στο [5], όταν η διάταξη των ηχείων που χρησιμοποιείται είναι κανονικό N-γωνο, η συνάρτηση υπολογισμού του σήματος που θα λάβει κάθε ηχείο είναι:

$$g_i = (G_1/2) * W + (G_2/2) * (X * \cos(\theta) * \cos(\varphi) + Y * \sin(\theta) * \cos(\varphi) + Z * \sin(\varphi))$$

Σχέση 38: Σχέση υπολογισμού του σήματος που θα λάβει κάθε ηχείο αν η διάταξή τους είναι κανονικό N-γωνο σύμφωνα με τον Farina στο [5].

όπου τα G_1 και G_2 είναι διαφορετικά για τις υψηλές και χαμηλές συχνότητες, και μεταβάλλονται ανάλογα με τον αριθμό των ηχείων του συστήματος αναπαραγωγή. Στην πίνακα (1) φαίνονται διάφοροι τύποι υπολογισμού των G_1 και G_2 από διάφορους συγγραφείς, για την περίπτωση της αναπαραγωγής από διάταξη ηχείων σε κανονικό N-γωνο, στο οριζόντιο επίπεδο (βλέπε [5]). Στη τελευταία στήλη του πίνακα φαίνεται το 'Γ' που είναι ο λόγος του G_2 προς το G_1 .

Έκδοση	Όνομα	Συγγραφέας	G_1	G_2	$\Gamma=G_2/G_1$
a)	Για αίθουσες συναυλιών για όλες τις συχνότητες	D.Malham	$\sqrt{\frac{8}{3N}}$	$\sqrt{\frac{8}{3N}}$	1
b)	Για στούντιο, για τις υψηλές συχνότητες	M.Gerzon	$\sqrt{\frac{8}{4N}}$	$\sqrt{\frac{8}{2N}}$	$\sqrt{2}$
c)	Για στούντιο, για τις χαμηλές συχνότητες	M.Gerzon	$\sqrt{\frac{8}{6N}}$	$\sqrt{\frac{16}{3N}}$	2
d)	Για στούντιο, για τις πολύ χαμηλές συχνότητες	J. M. Jot	$\sqrt{\frac{8}{2N^2}}$	$\sqrt{\frac{16}{N^2}}$	2

Πίνακας 1: Πίνακας υπολογισμού των G_1 και G_2 για να αναπαραγωγή σε οριζόντιο επίπεδο από διάταξη ηχείων σε κανονικό N-γωνο. Ο πίνακας είναι προέρχεται από το [5].

Μια άλλη αναφορά στο [5] που αφορά τον υπολογισμό των G_1 και G_2 αναφέρεται στην U.S. pat. N. 3,997,725 του M.Gerzon που αναφέρεται στην υλοποίηση ενός κωδικοποιητή και για τις τρεις διαστάσεις.

Συχνότητες	G_1	G_2	$\Gamma=G_2/G_1$
>500	$\sqrt{2}$	$\sqrt{2}$	1
<500	1	$\sqrt{3}$	$\sqrt{3}$

Πίνακας 2: Αυτός ο πίνακας αναφέρεται στον υπολογισμό των G_1 και G_2 σύμφωνα με την πατέντα του m.Gerzon (U.S. pat. N. 3,997,725).

Μια σημαντική παρατήρηση που πρέπει να προσθέσουμε σε αυτό το σημείο είναι ότι αν η τιμή του 'Γ' είναι ίση με την μονάδα έχει σαν αποτέλεσμα να μην εκπέμπεται ήχος από τα ηχεία που είναι τοποθετημένα σε αντίθετη διεύθυνση από την προέλευση της πηγής και αυτό έχει σαν αποτέλεσμα να δημιουργείται πιο σταθερή αίσθηση του surround σε μεγαλύτερο εύρος ακρόασης.

Μια εναλλακτική εξίσωση για τον υπολογισμό του σήματος που πρέπει να δρομολογηθεί στο κάθε ηχείο φαίνεται στη Σχέση 39 (Farina et al.,2001)¹¹. Με την εξίσωση ουσιαστικά προσομοιάζεται ένα μικρόφωνο προσανατολισμένο προς αυτό, και το σήμα που λαμβάνει αποτελεί το σήμα που θα τροφοδοτήσει το ηχείο. Αυτό που μπορεί να τροποποιήσει ο χρήστης είναι το πολικό διάγραμμα του εικονικού μικροφώνου, αλλάζοντας τον δείκτη κατευθυντικότητας 'd'. Έτσι, δίνοντας κατάλληλες τιμές στο 'd' μπορεί να υπολογίσει τα σήματα που θα τροφοδοτήσουν τα ηχεία για τις υψηλές και χαμηλές συχνότητες.

$$g_i = ((2-d)/2) * W + (d/2) * (X * \cos(\theta) * \cos(\varphi) + Y * \sin(\theta) * \cos(\varphi) + Z * \sin(\varphi))$$

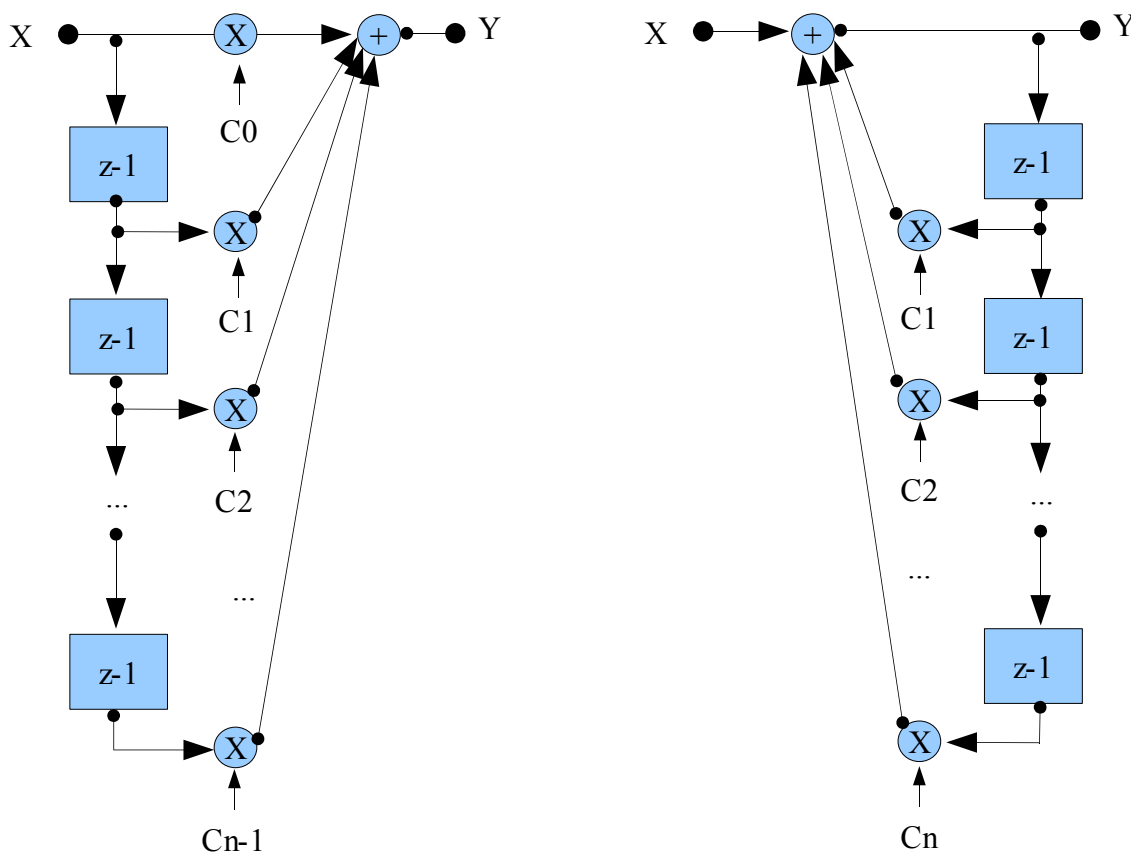
Σχέση 39: Συνάρτηση υπολογισμού του σήματος που θα λάβει ένα ηχείο χρησιμοποιώντας εικονικά μικρόφωνα στη θέση του ακροατή προσανατολισμένο προς το εκάστοτε ηχείο (Farina et al., 2001).

11 Και σε αυτή την περίπτωση θεωρείται ότι η διάταξη των ηχείων είναι κανονικό N-γωνο.

Ένα σύστημα αναπαραγωγής Ambisonic μπορεί να χρησιμοποιεί και μια διάταξη ηχείων λιγότερο συμμετρική. Για την περίπτωση που τα ηχεία δεν είναι τοποθετημένα σε κανονικό N-γωνο, αλλά παρουσιάζουν συμμετρία ως προς τον X άξονα (συμμετρία αριστερά-δεξιά) έχει αναφερθεί ο M.Gerzon στο [25].

2.3.7 Κατηγορίες ψηφιακών φίλτρων.

Ένα ψηφιακό φίλτρο μπορεί να υλοποιηθεί είτε σαν IIR είτε σαν FIR. Η διαφορά ανάμεσα σε αυτούς τους δύο τύπους φίλτρων είναι ο χρόνος της κρουστικής απόκρισης (impulse response), καθώς και ο τρόπος υλοποίησης. Στην περίπτωση των φίλτρων IIR, ο χρόνος της κρουστικής απόκρισης είναι άπειρος, ενώ στην περίπτωση των FIR περιορισμένος. Αυτό οφείλεται στον τρόπο υλοποίησης αυτών των φίλτρων. Στην περίπτωση των FIR φίλτρων, το σήμα εξόδου προκύπτει από το σήμα εισόδου¹² με διαδοχικές καθυστερήσεις και διαφορετικό βάρος κάθε φορά. Στα IIR φίλτρα αντίθετα το σήμα εξόδου προκύπτει από αυτό της εισόδου έχοντας προσθέσει σε αυτό διαδοχικές καθυστερήσεις του σήματος εξόδου¹³ με διαφορετικό βάρος (ανάδραση του σήματος εξόδου στην είσοδο). Στην εικόνα (25) φαίνεται το διάγραμμα ροής ενός FIR και ενός IIR φίλτρου.



Εικόνα 25: Στα αριστερά απεικονίζετε ένα φίλτρο FIR τάξης n ενώ στα δεξιά ένα φίλτρο IIR τάξης n . Με τα σύμβολα X και Y συμβολίζεται η είσοδος και η έξοδος του φίλτρου αντίστοιχα, με το σύμβολο z^{-1} η καθυστέρηση του σήματος κατά ένα $sample$ και με τους όρους C_n οι συντελεστές του φίλτρου.

12 Πολλαπλασιασμένα με διαφορετικό συντελεστή κάθε φορά.

13 Πολλαπλασιασμένα με διαφορετικό συντελεστή κάθε φορά.

Τα πλεονεκτήματα των φίλτρων FIR έναντι των IIR είναι τα εξής:

1. Είναι δυνατή η εξάλειψη της παραμόρφωσης της φάσης (επιλογή των συντελεστών του φίλτρου ώστε να παρουσιάζει γραμμική απόκριση φάσης)
2. Εγγυημένη σταθερότητα ακόμα και στην περίπτωση που οι συντελεστές έχουν κβαντοποιηθεί.

Το βασικότερο μειονέκτημα αυτών των φίλτρων είναι ότι καταναλώνουν μεγάλη υπολογιστική ισχύ (βλέπε [8], [9]).

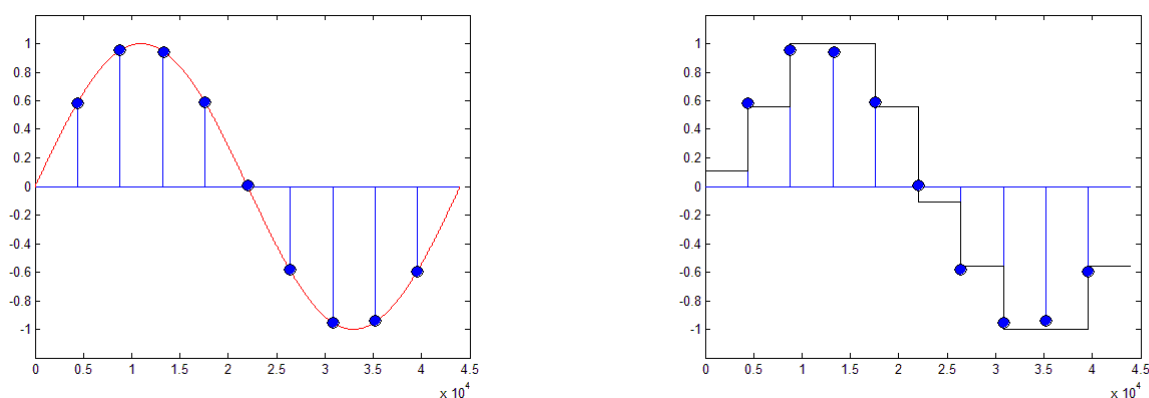
2.3.8 Κβαντισμός.

Τα αναλογικά σήματα είναι συνεχή στον χρόνο και μπορούν να πάρουν οποιαδήποτε τιμή έντασης. Τα ψηφιακά αντίθετα έχουν συγκεκριμένες τιμές ανά χρονικά διαστήματα (1/ (συχνότητα δειγματοληψίας)) ενώ οι τιμές της έντασής τους μπορεί να λάβει συγκεκριμένες τιμές (ανάλογα με το bit rate). Η τάση ενός αναλογικού σήματος μπορεί να παίρνει οποιαδήποτε τιμή σε κάποιο διάστημα, όπως πχ. [-1,1]. Στο ίδιο διάστημα ένα ψηφιακό σήμα μπορεί να πάρει διακριτές τιμές (βλέπε [21], [22], [23], [24]). Αυτές, κατά την υλοποίηση ενός midtread κβαντιστή, υπολογίζονται από την Σχέση 40.

$$Q = -1 + k \frac{1 - (-1)}{2^N - 1} \quad \text{όπου:} \quad k=0,1,\dots,2^N - 1$$

Σχέση 40:

Κβαντοποίηση, λοιπόν, είναι η διαδικασία αντιστοίχησης των πραγματικών τιμών έντασης του σήματος στις κοντινότερες δυνατές τιμές που έχουν οριστεί βάση του bit rate. Στην εικόνα (26) φαίνεται ένα σήμα διακριτού χρόνου που έχει προκύψει από την δειγματοληψία (μπλε κουκκίδες) ενός αναλογικού (κόκκινη γραμμή). Με την μαύρη γραμμή φαίνεται το σήμα που θα προκύψει μετά την κβαντοποίηση (υπάρχει διατήρηση της τιμής για το χρονικά διάστημα του 1/sample rate λόγω της διαδικασίας sample hold που διαθέτει).



Εικόνα 26: Αριστερά φαίνεται με κόκκινη γραμμή το αρχικό αναλογικό σήμα ενώ και στις δύο εικόνες με μπλε κουκκίδες απεικονίζονται οι στιγμιαίες τιμές του σήματος βάση της δειγματοληψίας που έχει γίνει. Στη δεξιά εικόνα φαίνεται με μαύρη γραμμή η τιμές του σήματος μετά την κβαντοποίηση (midtread).

Ένας κβαντιστής μπορεί να υλοποιηθεί με δύο τρόπους, είτε ως midtread είτε σαν midriser. Και οι δύο αυτοί στρογγυλοποιούν την πραγματική τιμή του σήματος εισόδου στην κοντινότερη δυνατή τιμή που μπορεί να πάρει βάση του bit rate που έχει επιλεγεί. Η διαφορά τους είναι ότι ο midtread

παράγει και μηδενικές τιμές στην έξοδο του ενώ ο midriser όχι. Στη Σχέση 42 φαίνονται οι τύποι υλοποίησης του καθενός εξ αυτών. Η ελάχιστη ένταση του σήματος που χρειάζεται να προστεθεί στο σήμα για να πάει στο επόμενο bit συμβολίζεται με Δ και δίνεται από την Σχέση 41. Στην εικόνα (27) φαίνεται η αντιστοίχιση των τιμών της εισόδου (signal) με την έξοδο του κβαντιστή (Q).

$$\Delta = \frac{2S_{\max}}{2^N - 1}$$

Σχέση 41:

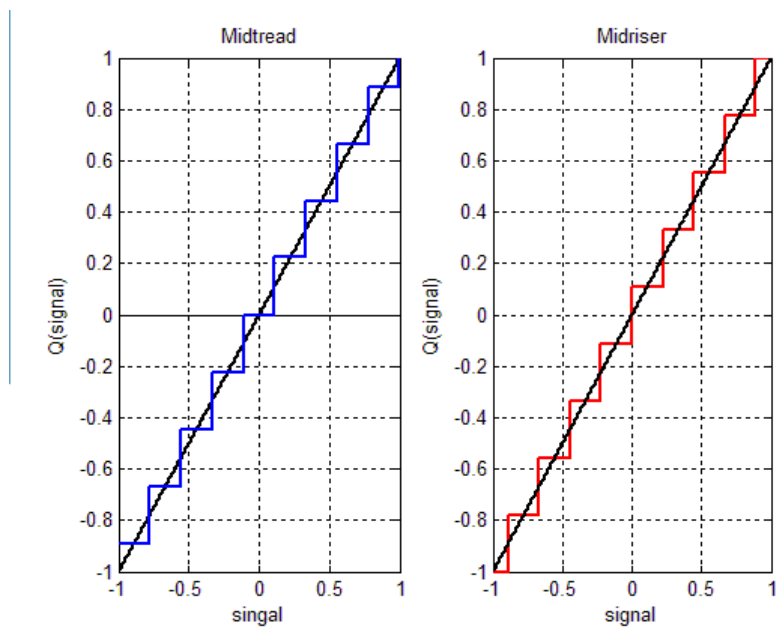
$$Q(\text{signal}) = \Delta \left\lfloor \frac{\text{signal}}{\Delta} + \frac{1}{2} \right\rfloor$$

$$Q(\text{signal}) = \Delta \left\lceil \frac{\text{signal}}{\Delta} \right\rceil + \frac{\Delta}{2}$$

Σχέση 42: Η πάνω σχέση υλοποιεί έναν midread κβαντιστή ενώ η κάτω έναν midriser

όπου :

' S_{\max} ' και ' $-S_{\max}$ ' η μέγιστη και ελάχιστη τιμή που θέλουμε να παίρνει ο κβαντιστής και $\lfloor \cdot \rfloor$ συμβολισμός για την στογγυλοποίηση προς τα κάτω στην κοντινότερη ακέραια τιμή.



Εικόνα 27: Απεικόνιση της αντιστοίχισης τιμών του σήματος με τις διακριτές τιμές που μπορεί να πάρει ένας Midread και ένας Midriser κβαντιστής αντίστοιχα.

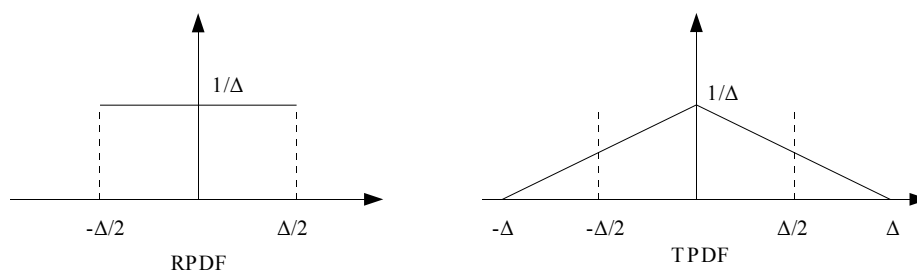
Γενικά η κβαντοποίηση είναι μια μη γραμμική διαδικασία καθώς μετά την κβαντοποίηση είναι αδύνατο να υπολογιστούν οι ακριβείς αρχικές τιμές του σήματος εισόδου. Επιπλέον κατά την κβαντοποίηση προκύπτει ένα σφάλμα κβαντισμού είναι ακουστό και μετρήσιμο. Αυτό εμφανίζεται σαν πρόσθετος θόρυβος στο αρχικό σήμα και η έντασή του αυξάνει αντιστρόφως ανάλογα με το bit rate, και έχει σημαντική επίπτωση στην ποιότητα του σήματος. Αυτή η υποβάθμιση της ποιότητας του σήματος είναι πιο εμφανής για τις τιμές πλάτους του αρχικού σήματος μικρότερες του βήματος κβαντισμού ' Δ '. Για την μερική βελτίωση της ποιότητας του σήματος μπορεί να προστεθεί στο αρχικό σήμα ένας θόρυβος μικρού πλάτους (πλάτος μεγαλύτερο του Δ) και συγκεκριμένων στατιστικών χαρακτηριστικών ο οποίος ονομάζεται dither (βλέπε [21], [22], [23], [24]). Τα πλεονεκτήματα χρήσης του dither είναι:

- καθιστά ανεξάρτητο το σφάλμα κβαντισμού από το σήμα εισόδου
- Αφαιρεί την αρμονική παραμόρφωση που θα εμφανιζόταν στο σήμα εξόδου η οποία εμφανιζόταν όταν το σήμα εισόδου είχε πολύ μικρό πλάτος, και δίνει στον θόρυβο κβαντισμού χαρακτηριστικά λευκού θορύβου.
- Αυξάνει κατά υποκειμενικό τρόπο το δυναμικό τρόπο το δυναμικό εύρος του κβαντιστή.

Το βασικό μειονέκτημα του dither είναι η μείωση του λόγου σήματος προς θόρυβο, ο οποίος λόγος μειώνεται με την αύξηση του πλάτους του dither.

Το Dither πρέπει να χρησιμοποιείται σε όλες τις επεξεργασίες του ψηφιακού σήματος όπου χρειάζεται επανακβαντισμός, όπως για παράδειγμα στην μεταβολή της στάθμης του σήματος ή στο ψηφιακό φιλτράρισμα του σήματος σε μια ψηφιακή συσκευή.

Υπάρχουν διάφορα είδη dither των οποίων η διαφορά τους έγκειται στη συνάρτηση πυκνότητας πιθανότητας του θορύβου (Probability Density Function, PDF) και στο φασματικό τους περιεχόμενο. Στα ψηφιακά συστήματα συγκεκριμένα χρησιμοποιούνται γεννήτριες dither ομοιόμορφης κατανομής (Rectangular PDF, RPDF) πλάτους 1 LSB (Δ) p-p και τριγωνικής κατανομής (Triangular PDF, TPDF) πλάτους 2LSB (2Δ) p-p. Οι μορφές των συναρτήσεων πυκνότητας αυτών φαίνεται στην εικόνα (28) (βλέπε [21], [22], [23], [24]).



Εικόνα 28: Στα αριστερά φαίνεται η συνάρτηση πυκνότητας μιας τετραγωνικής PDF (RPDF) ενώ στα δεξιά μιας τριγωνικής PDF (TPDF) με πλάτη $\Delta/2$ και 2Δ αντίστοιχα.

Σύμφωνα με το [21], από τα είδη dither που φαίνονται στην εικόνα 28, αυτό με συνάρτηση πυκνότητας TPDF παρουσιάζει καλύτερη στατιστική συμπεριφορά και ελαχιστοποιεί την συσχέτιση του θορύβου κβαντισμού με το σήμα. Το μειονέκτημα αυτής είναι ότι μειώνει το SNR κατά 6 dB.

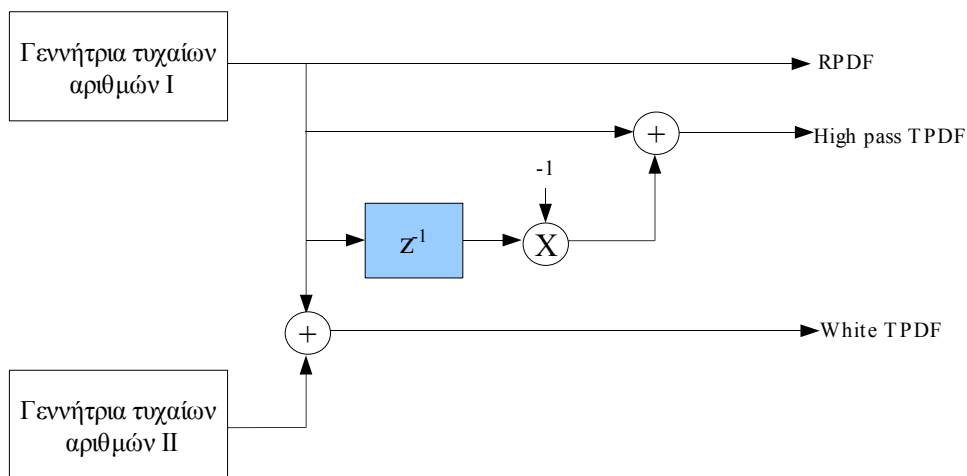
Η υλοποίηση ενός θορύβου dither γίνεται με την χρήση γεννήτριας παραγωγής τυχαίων αριθμών. Μια ακολουθία που μπορεί να δώσει μια σειρά ψευδοτυχαίων αριθμών για την δημιουργία dither

ομοιόμορφης κατανομής (RPDF) δίνεται από την Σχέση 43.

$$X_i = (aX_{i-1} + \gamma) \bmod B$$

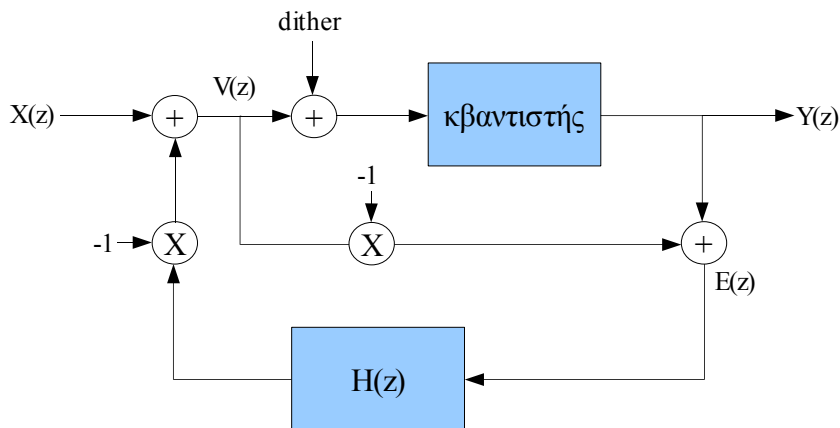
Σχέση 43:

όπου a, B, γ, X_0 ακέραιοι αριθμοί και $a > 0, \gamma < B$ και X_0 η αρχική τιμή που θα πάρει η ακολουθία των αριθμών. Η δημιουργία μιας ακολουθίας TPDF μπορεί να δημιουργηθεί με το άθροισμα δύο ανεξάρτητων RPDF. Επίσης, αν σε μία RPDF δημιουργηθεί μια ανάδραση έτσι ώστε η έξοδός της να προκύπτει από την πρόσθεση του τελευταίου με το προηγούμενο δείγμα, δημιουργείται μια TPDF με υψηλοδιαβατά χαρακτηριστικά. Στην εικόνα 29 φαίνεται το μπλοκ διάγραμμα για την δημιουργία TPDF ακολουθιών.



Εικόνα 29: Block διάγραμμα για την δημιουργία TPDF και High pass TPDF (παρουσιάζει υψηλοδιαβατά χαρακτηριστικά).

Όπως προαναφέρθηκε το dither είναι ένας επιπρόσθετος θόρυβος που εισάγεται στο σήμα εισόδου για την μείωση της αίσθησης του θορύβου που προκύπτει λόγω του κβαντισμού. Στις χαμηλές εντάσεις του σήματος εισόδου όμως ο θόρυβος κβαντισμού μπορεί να περάσει το ελάχιστο κατώφλι ακοής και έτσι να γίνει ακουστός. Το συχνотικό εύρος που γίνεται πιο αισθητό αυτό το φαινόμενο είναι στις υψηλές συχνότητες πάνω από τα 4 kHz. Για τον περιορισμό αυτού του φαινομένου χρησιμοποιούνται τεχνικές μορφοποίησης του θορύβου (Noise-Shaping) με τις οποίες διαμορφώνεται συχνотικά ο θόρυβος κβαντισμού. Η μορφοποίηση αυτή πραγματοποιείται με χρήση ανατροφοδοτούμενων ψηφιακών φίλτρων που λειτουργούν ως κλάδοι ανατροφοδότησης γύρω από έναν κβαντιστή. Στην εικόνα 30 φαίνεται το block διάγραμμα ενός κβαντιστή με λειτουργία μορφοποίησης θορύβου, ενώ στη Σχέση 44 φαίνονται οι εξισώσεις υπολογισμού του σήματος εξόδου του αυτού.



Εικόνα 30: Μπλοκ διάγραμμα κβαντιστή με χρήση μορφοποίησης θορύβου.

$$Y(z) = X(z) - E(z) * H(z) + E(z) = X(z) + (1 - H(z)) * E(z)$$

Σχέση 44: Συνάρτηση υπολογισμού του σήματος εξόδου ενός κβαντιστή που χρησιμοποιεί (Noise-Shaping).

όπου

- $H(z)$ η συνάρτηση μεταφοράς του ψηφιακού φίλτρου.
- $X(z)$ το σήμα εισόδου του φίλτρου.
- $Y(z)$ το σήμα εξόδου της μονάδας κβαντοποίησης με χρήση Noise-Shaping.
- $E(z)$ το σφάλμα κβαντισμού.

Τα πρώτα φίλτρα που εμφανίστηκαν είχαν συνάρτηση μεταφοράς με την μορφή που φαίνεται στη Σχέση 45.

$$1 - H(z) = (1 - Z^{-1})^n$$

Σχέση 45: Εξίσωση των πρώτων φίλτρων που εμφανίστηκαν για την συνάρτηση μεταφοράς της τεχνικής Noise-Shape.

όπου 'n' η τάξη του φίλτρου.

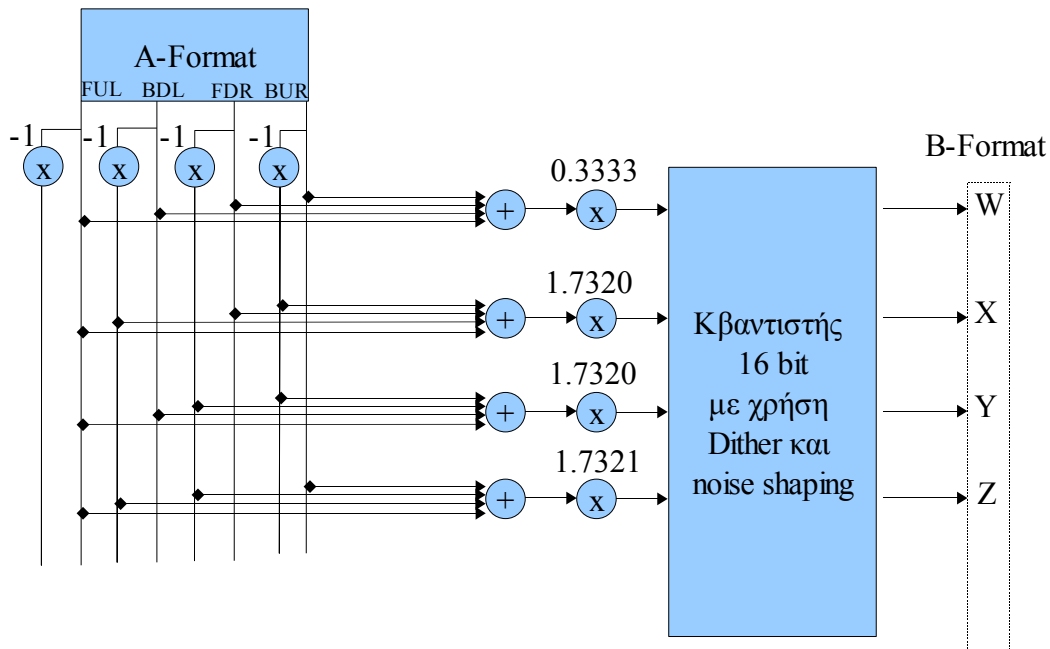


ΥΛΟΠΟΙΗΣΗ

Η υλοποίηση του αλγόριθμου κωδικοποίησης και αποκωδικοποίησης των σημάτων για την δημιουργία του B-Format (encoder) και την δημιουργία του D-Format (decoder) θα πραγματοποιηθούν σε περιβάλλον Matlab.

3.1. Encoder B-Format.

3.1.1 Encoder B-Format από A-Format (υπάρχον ηχητικό πεδίο).



Εικόνα 31: Μπλοκ διάγραμμα για την κωδικοποίηση του σήματος A-Format σε B-Format Ambisonic.

Αρχικά στην εικόνα 31 φαίνεται το μπλοκ διάγραμμα ενός κωδικοποιητή B Format από A-Format Ambisonic. Αυτός ο αλγόριθμος πρόκειται να υλοποιηθεί σε περιβάλλον Matlab. Όπως φαίνεται στην εικόνα 31 και έχει αποδειχθεί στο κεφάλαιο 2.3.5.1 τα σήματα του B-Format μπορούν να προκύψουν από απλές προσθαφαιρέσεις των τεσσάρων σημάτων που λαμβάνει ένα Soundfield μικρόφωνο. Οι συντελεστές που πολλαπλασιάζονται τα σήματα προκύπτουν από την Σχέση 27. Η μονάδα κβαντοποίησης πρόκειται για απαραίτητο τμήμα αυτής του αλγορίθμου καθώς είναι υποχρεωτικός ο επανακβαντισμός των σημάτων καθώς έχουν πολλαπλασιαστεί με έναν συντελεστή. Στο Παράρτημα 9' υπάρχει ο κώδικας που τον υλοποιεί. Για την υλοποίηση του κβαντιστή έχουν χρησιμοποιηθεί δύο συναρτήσεις, μια που αναλαμβάνει να κάνει την κβαντοποίηση (Quantiser.m, βλέπε παράρτημα 7) η οποία στη συνέχεια χρησιμοποιεί μια δεύτερη που αναλαμβάνει την δημιουργία του dither (MakeDither.m, βλέπε παράρτημα 8). Η συνάρτηση Quantiser.m δίνει την δυνατότητα στον χρήστη να επιλέξει τον τρόπο κβαντοποίησης (Midrise, Midtreat), αν θα χρησιμοποιηθεί dither και αν ναι τι είδους θα είναι αυτό (RPDF, TPDF, Highpass TPDF). Επίσης μπορεί να επιλεγεί αν θα χρησιμοποιηθεί τεχνική noise shape και να οριστεί το εύρος εντάσεων του κβαντιστή και του αριθμού των διακριτών τιμών που θα έχει αυτός. Η συνάρτηση MakeDither.m δίνει την δυνατότητα στον χρήστη να δημιουργήσει θόρυβο dither επιλέγοντας τον τύπο αυτού (RPDF, TPDF, Highpass TPDF) καθώς και του μέγιστης έντασης αυτού (ορισμός της έντασης που αντιστοιχεί το Δ ή αλλιώς LSB). Και οι δύο αυτές συναρτήσεις έχουν το πλεονέκτημα ότι μπορούν να διαχειριστούν αρχεία που περιλαμβάνουν περισσότερα από ένα κανάλια.

Τεκμηρίωση κώδικα Matlab για κωδικοποιητή A-Format ambisonic σε B-Format.

Σε αυτό το κεφάλαιο θα γίνει τεκμηρίωση του κώδικα που βρίσκεται στο παράρτημα 9. Στην πρώτη γραμμή του κώδικα ορίζεται η συνάρτηση 'encodeabformat' η οποία εξάγει ένα πίνακα που αποτελείται από τέσσερις στήλες και τόσες γραμμές όσες είναι και το κάθε ένα από τα σήματα εισόδου. Σαν ορίσματα σε αυτή την συνάρτηση εισάγονται τα τέσσερα σήματα που εξάγονται από ένα μικρόφωνο Soundfield (FUL, BDL, BUR, FDR). Τα ορίσματα θα μπορούσαν να είναι περισσότερα ώστε ο χρήστης να μπορεί να επιλέξει τον τρόπο κβαντοποίησης, αν θα έχει dither και τι τύπου θα είναι αυτό κ.α. . Αποφασίστηκε όμως πως θα επαρκούσε σε αυτή την εφαρμογή η κβαντοποίηση να είναι τύπου Midtread, με ανάλυση 16 bit, μέγιστη ένταση '1', με χρήση dither τύπου TPDF και χρήση τεχνικής noiseshape τρίτης τάξης. Η χρήση TPDF dither έγινε γιατί σύμφωνα με την θεωρία παρουσιάζει καλύτερη αίσθηση του σήματος σε σχέση με τον θόρυβο κβαντισμού. Η επιλογή του Noiseshape τρίτης τάξης έγινε για μείωση της έντασης του θορύβου κβαντισμού ανάλογα με την συχνότητα, ώστε να μην γίνεται ακουστός στις χαμηλές εντάσεις του σήματος. Στις γραμμές 20-30 δίνονται αρχικές τιμές που αφορούν τον τρόπο που θα γίνει η κβαντοποίηση καθώς και ο καθορισμός όλων των πινάκων που πρόκειται να χρησιμοποιηθούν. Στις γραμμές 33-36 γίνεται η εφαρμογή της Σχέση 27 για την δημιουργία των τεσσάρων σημάτων του B-Format (W,X,Y,Z). Στη συνέχεια στη γραμμή 38 ενσωματώνονται όλα τα σήματα (W,X,Y,Z) σε ένα πίνακα (ALL), όπου έχει τον ίδιο αριθμό γραμμών με αυτά, ενώ στη πρώτη του στήλη έχει το σήμα 'W', στη δεύτερη το σήμα 'X', στη τρίτη το σήμα 'Y' και τέλος στη τέταρτη το 'Z'. Τέλος στην γραμμή 40 γίνεται χρήση της συνάρτησης 'Quantiser' με την οποία γίνεται κβαντοποίηση του τετρακάναλου σήματος (πίνακα) 'ALL' ώστε να δημιουργηθεί το τετρακάναλο σήμα εξόδου της συνάρτησης.

Τεκμηρίωση κώδικα Matlab για την συνάρτηση κβαντοποίησης.

Ο κώδικας της συνάρτησης κβαντοποίησης με το όνομα Quantiser βρίσκεται στο παράρτημα 7. Το πλεονέκτημα αυτής της συνάρτησης είναι ότι ο χρήστης μπορεί να επιλέξει την μέθοδο που επιθυμεί να ακολουθηθεί για να πραγματοποιηθεί η κβαντοποίηση, αν θα χρησιμοποιηθεί dither, τι τύπου θα είναι αυτό, αν θα χρησιμοποιηθεί Noise shape, ποιά θα είναι η μέγιστη τιμή που επιθυμεί να παίρνει ο κβαντιστής και τέλος να ορίσει και το bit rate. Επίσης πλεονέκτημα αυτής της συνάρτησης είναι ότι μπορεί να επεξεργαστεί και σήματα που περιλαμβάνουν περισσότερα από ένα κανάλια. Εξάγει ένα πίνακα με τόσες γραμμές και στήλες¹⁴ όσες έχει και ο πίνακας του σήματος που εισάγεται για να κβαντιστεί. Τα ορίσματα εισόδου του είναι έξι. Στο όρισμα 'InSignal' εισάγεται το σήμα που πρόκειται να κβαντοποιηθεί, το 'N' ορίζει τον αριθμός των διακριτών τιμών που μπορεί να πάρει ο κβαντοποιητής (2^N), το 'Smax' ορίζει το διάστημα μεταξύ του οποίου θα οριστούν οι διακριτές τιμές του κβαντιστή [-Smax,Smax], το 'MidriserFlag' αν ο κβαντοποιητής θα είναι Midriser ή Midtread (1,0 αντίστοιχα), το 'NSFlag' αν θα χρησιμοποιηθεί τεχνική Noise shape με δυνατές τιμές (0,2,3)¹⁵ και τέλος με το όρισμα 'DitherFlag' ορίζει αν πρόκειται να χρησιμοποιηθεί dither και αν ναι τον τύπο αυτού. Οι τιμές που μπορεί να πάρει το 'DitherFlag' είναι '0', '1', '2', '3' εκ των οποίων στη πρώτη δεν θα χρησιμοποιηθεί dither ενώ στις υπόλοιπες θα είναι τύπου RPDF, TPDF, Highpass TPDF αντίστοιχα.

Στην πρώτη γραμμή ορίζεται η συνάρτηση. Στις γραμμές 32-24 γίνεται ένας έλεγχος για

¹⁴ αυτές αντιστοιχούν στα κανάλια του σήματος εισόδου

¹⁵ Με την τιμή '0' δεν πρόκειται να χρησιμοποιηθεί Noiseshape ενώ με τις τιμές '2' και '3' θα χρησιμοποιηθεί, και θα είναι 2^{15} και 3^{15} τάξης αντίστοιχα.

τις τιμές που έχει δώσει ο χρήστης στο όρισμα 'NSFlag' για να διαπιστωθεί μήπως έχει δώσει κάποια τιμή εκτός του πεδίου ορισμού της. Με την χρήση της συνάρτησης 'or' μέσα σε μια 'or' επιτυγχάνεται η δημιουργία μιας 'or' που δέχεται τρεις εισόδους. Έτσι αν έστω και μία από αυτές τις τρεις τιμές εισόδου πάρει την τιμή '1', η έξοδος θα είναι '1'. Διαφορετικά η έξοδος της 'or' θα ισούται με το μηδέν. Με την αφαίρεση που κάνω της μονάδας από την 'or' επιτυγχάνω ακριβώς τις αντίθετες τιμές εξόδου, άρα αν έστω και ένα από τα σήματα είναι μονάδα η έξοδος θα ισούται με το μηδέν. Αν το όρισμα της συνάρτησης 'if' είναι '0' τότε θεωρείται ψευδής και προχωράει απευθείας στο 'else' ή στο 'end' αυτής, διαφορετικά εκτελεί τις σχέσεις που βρίσκονται μεταξύ του 'if' και του 'else' και στη συνέχεια πάει στο 'end'. Έτσι, αν η τιμή που έχει δώσει ο χρήστης στο NSFlag δεν είναι μέσα στα επιτρεπόμενα όρια, η συνάρτηση 'if' βγαίνει αληθής και έτσι η ακολουθία της συνάρτησης 'Quantiser' διακόπτεται βγάζοντας 'error' στον χρήστη, διαφορετικά πηγαίνει κατευθείαν στο 'end' και συνεχίζεται ο αλγόριθμος. Στις γραμμές 36-37 ορίζονται απλώς ο πίνακας εξόδου 'OutSignal' και το 'dither'. Στην γραμμή 40 ορίζω τις μεταβλητές 'gr' και 'st' οι οποίες αντιστοιχούν στον αριθμό των γραμμών και τον αριθμό των στηλών του σήματος εισόδου. Είναι χρήσιμες αυτές οι μεταβλητές γιατί προσδιορίζουν το πλήθος των στοιχείων που θα υπολογιστούν για το σήμα εξόδου (γραμμές) ενώ οι στήλες μου προσδιορίζουν από πόσες στήλες (κανάλια) θα αποτελείται ο πίνακας εξόδου. Στη γραμμή 46 υλοποιείται η Σχέση 41. Στη γραμμή 48 ξεκινάει ένας έλεγχος για την τιμή της μεταβλητής MidriserFlag που ορίζει τον τρόπο που θα γίνει κβαντοποίηση. Στην περίπτωση που η τιμή του Midriser είναι '0' εκτελείται ο κώδικας από την γραμμή 49-130. Αν η τιμή του Midriser είναι '1' εκτελείται ο κώδικας από την γραμμή 134-217, διαφορετικά η συνάρτηση 'Quantiser' βγάζει 'error' και τερματίζεται. Η μόνη διαφορά που παρουσιάζει ο κώδικας που βρίσκεται μεταξύ των γραμμών 49-130 σε σχέση με αυτών στις γραμμές 134-217 είναι η εξίσωση κβαντισμού. Στη συνέχεια θα αναλυθεί ο κώδικας 49-130, ενώ από τον κώδικα 134-217 θα αναλυθούν μόνο οι διαφορές που παρουσιάζει. Στις γραμμές 52-59 γίνεται έλεγχος για την τιμή του ορίσματος DitherFlag. Στην περίπτωση που πάρει έχει την τιμή '0' δεν χρησιμοποιείται dither και αυτό που γίνεται στις γραμμές 53-54 είναι η προσθήκη στο αρχικό σήμα ενός μηδενικού σήματος, άρα το αρχικό παραμένει το ίδιο. Στη περίπτωση που το DitherFlag έχει κάποια από τις τιμές (1,2,3) τότε καλείται η συνάρτηση 'MakeDither' για να δημιουργήσει κατάλληλο θόρυβο dither (ανάλογα με την τιμή του 'DitherFlag') και να τον προσθέσει στη γραμμή 57 στο αρχικό σήμα. Διαφορετικά βγάζει 'error' και διακόπτει την ροή της συνάρτησης 'Quantiser'. Από την γραμμή 65 μέχρι την 130 γίνεται η διαδικασία της κβαντοποίησης ανάλογα με την τιμή του 'NSFlag'. Έτσι στις γραμμές 66-75 γίνεται η κβαντοποίηση χωρίς την χρήση Noise shape. Στη γραμμή 66 γίνεται υλοποίηση της Σχέση 42. Στις γραμμές 73-75 γίνεται έλεγχος για τις τιμές των δειγμάτων που έχουν ξεπεράσει σε απόλυτη τιμή την μέγιστη τιμή που μπορεί να πάρει ο κβαντιστής, και δίνουν σε αυτές την μέγιστη ή την ελάχιστη τιμή αυτού ανάλογα. Στην περίπτωση που βρεθούν δείγματα που ξεπερνούν σε απόλυτη τιμή τη μέγιστη επιτρεπτή ένταση κβαντισμού υπάρχει ψαλιδισμός του σήματος. Στην υλοποίηση ενός Midtread η ελάχιστη και μέγιστη επιτρεπτή τιμή εξόδου είναι $[-S_{max} + \Delta/2, S_{max} - \Delta/2]$. Στις γραμμές 77-99 γίνεται η υλοποίηση με τεχνική Noise shape δεύτερης τάξης. Εδώ υλοποιείται το βλοκ διάγραμμα της εικόνας 30. Η έξοδος της συνάρτησης $H(z)$ σε αυτή την περίπτωση είναι η μεταβλητή 'w' που υπάρχει στην γραμμή 84 (Οι τιμές των συντελεστών αυτής έχουν παρθεί από το αντίστοιχο εργαλείο του κ. Α. Φλώρου 'QuantizeAudio.m'). Τα 'f1' και 'f2' είναι δύο διαδοχικές θέσεις μνήμης που καταγράφουν την τιμή που είχε ο θόρυβος κβαντισμού κατά το τελευταίο και προτελευταίο δείγμα, και μέσω της συνάρτησης 'w' γίνεται ανάδραση (σήμα από την έξοδο επιστρέφει στην είσοδο) για να προστεθεί στο επόμενο δείγμα εισόδου. Αφού οι πράξεις γίνονται σε κάθε δείγμα ξεχωριστά, η συνάρτηση αυτή θα μπορούσε να υλοποιηθεί με την δημιουργία ενός βρόγχου, με χρήση της συνάρτησης 'for' στον οποίον σε κάθε ανάδραση του θα εισήγαγε σε όλες τις συναρτήσεις το επόμενο δείγμα

εισόδου. Ο αριθμός των δειγμάτων εισόδου ισούται όμως με τον αριθμό των γραμμών του πίνακα εισόδου, και υπάρχει αντιστοιχία ότι η τιμή της πρώτης γραμμής του πίνακα αντιστοιχεί στο πρώτο δείγμα, η δεύτερη στο δεύτερο και ούτω καθ' εξής. Έτσι αρκεί στην συνάρτηση μεταβολής της μεταβλητής να τίθεται κάθε φορά η επόμενη θέση του πίνακα, ξεκινώντας από την τιμή '1' και καταλήγοντας στην μέγιστη θέση του πίνακα (βλέπε γραμμή 81, και συνολικός βρόγχος ανάδρασης γραμμές 81-100). Οι τιμές των 'f1' και 'f2' όταν ξεκινάει η διαδικασία πρέπει να έχουν την τιμή '0' καθώς δεν υπάρχει προηγούμενο δείγμα. Η συνάρτηση Quantiser, όπως προαναφέρθηκε, μπορεί να δεχθεί και δείγματα με περισσότερα του ενός κανάλια, που ορίζονται με τις στήλες. Άρα για να υλοποιηθεί η προηγούμενη διαδικασία και στα υπόλοιπα κανάλια θα πρέπει να δημιουργηθεί αντίστοιχα ένας βρόγχος ανάδρασης για τις στήλες (γραμμές 78-101). Κάθε φορά που θα αλλάζει όμως η στήλη (διαφορετικό κανάλι) θα πρέπει να δίνονται οι αρχικές τιμές στις μεταβλητές 'f1' και 'f2' (βλέπε γραμμές 79,80). Στη γραμμή 85 είναι η συνάρτηση δημιουργίας ενός κβαντιστή Midtread όπου σαν σήμα εισόδου μπαίνει η συνάρτηση που φαίνεται στη Σχέση 44. Στις γραμμές 91-95 γίνεται έλεγχος για τιμές εξόδου κβαντιστή που έχουν ξεπεράσει τα επιτρεπτά τα όρια τιμών αυτού και τους δίδεται η κατάλληλη τιμή. Στις γραμμές 97-99 γίνεται μια ολισθητική εγγραφή των προηγούμενων και της νέας τιμής δειγμάτων στις διαδοχικές μνήμες 'f1' και 'f2'. Στις γραμμές 104-128 έχουμε ακριβώς την ίδια διαδικασία με τις γραμμές 78-101, και αυτό που αλλάζει είναι η συνάρτηση της μεταβλητής 'w' και το ότι πλέον υπάρχουν τρεις διαδοχικές θέσεις μνήμης. Όπως έχει προαναφερθεί ο κώδικας των γραμμών 133-215 είναι παρόμοιος με αυτών στις γραμμές 48-130, με μόνες διαφορές την αλλαγή της συνάρτησης κβαντισμού που έχει αντικατασταθεί από την Midriser της Σχέση 42 (βλέπε πχ γραμμή 151), και το ότι πλέον η μέγιστη και ελάχιστη τιμή κβαντισμού είναι 'Smax' και '-Smax' αντίστοιχα. Τέλος στη γραμμή 216 είναι το error που θα βγάλει η συνάρτηση σε περίπτωση που οι τιμές του MidriserFlag δεν είναι '0' ή '1'.

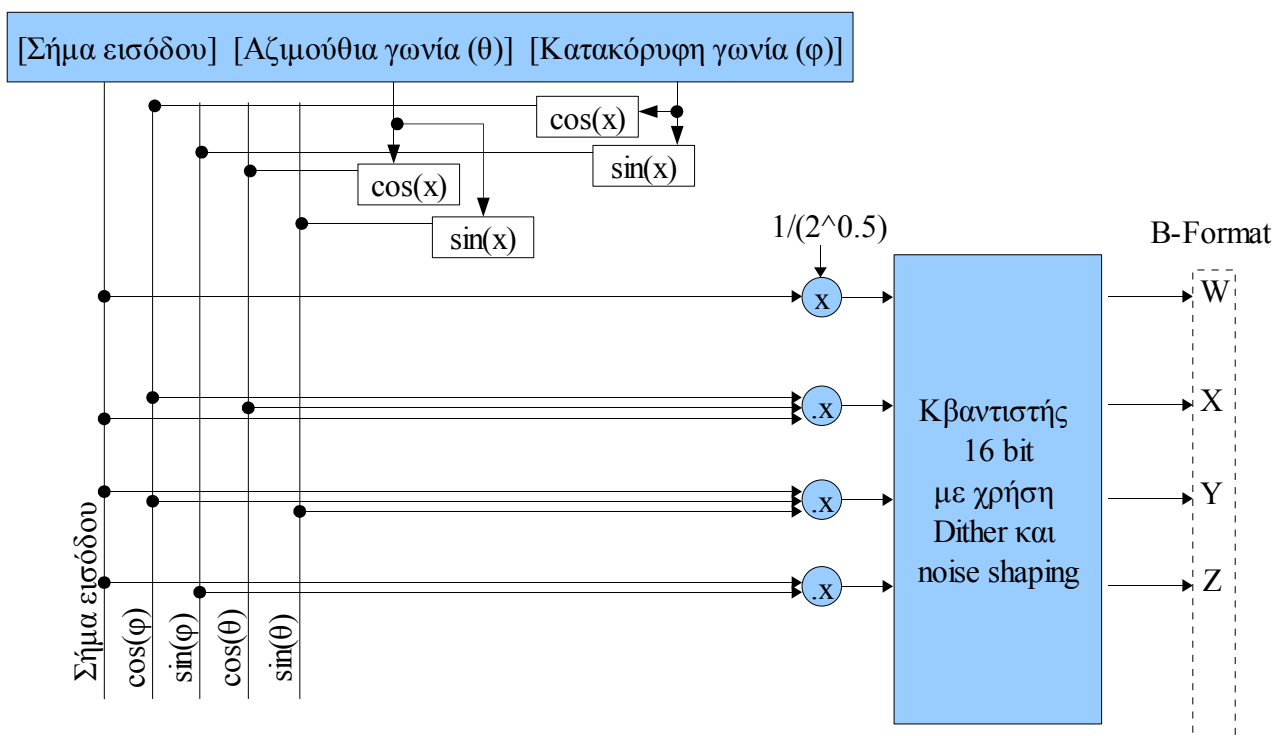
Τεκμηρίωση κώδικα Matlab για την συνάρτηση δημιουργίας θορύβου dither.

Ο κώδικας της συνάρτησης MakeDither βρίσκεται στο παράρτημα 8. Η συνάρτηση μπορεί να δημιουργεί θόρυβο Dither τύπου RPDF, TPDF, Highpass TPDF. Έχει τέσσερα ορίσματα (dithtype, min, gr, st), εκ των οποίων με το πρώτο ο χρήστης μπορεί να επιλέξει τον τύπο του θορύβου dither, με δυνατές επιλογές (1,2,3) οι οποίες αντιστοιχούν σε τύπους (RPDF, TPDF, Highpass TPDF). Τα (min, gr, st) αντιστοιχούν στην ελάχιστη τιμή κβαντισμού (LSB ή αλλιώς Δ), στο μήκος του θορύβου dither (γραμμές) και τέλος στο πόσα κανάλια dither θα έχει το σήμα εξόδου (στήλες). Στην πρώτη γραμμή ορίζεται η συνάρτηση. Στις 25-35 δίνονται αρχικές τιμές κάποιων σταθερών καθώς και ορίζεται ένας κενός πίνακας ο 'dith' ο οποίος στο τέλος θα αποτελέσει το σήμα εξόδου. Στις γραμμές 40-49 γίνεται ένας έλεγχος για την τιμή του 'dithtype' και αν έχει την τιμή '1' πραγματοποιείται ο κώδικας που περιλαμβάνεται σε αυτές τις γραμμές. Ο τύπος dither που πρόκειται να δημιουργηθεί είναι RPDF. Η εξίσωση που φαίνεται στη γραμμή 45 αντιστοιχεί στη Σχέση 43. Οι τιμές εξόδου αυτής της συνάρτησης κυμαίνονται στο διάστημα $[0, m-1]$ ¹⁶. Οι τιμές που θέλουμε να παίρνει όμως ένας RPDF είναι μεταξύ του $[-\min/2, \min/2]$. Αφαιρώντας λοιπόν από τις τιμές της συνάρτησης της γραμμής 45 το $(m-2)/2$, οι τιμές αυτής θα κυμαίνονται μεταξύ των $[-m+2, m-2]$. Διαιρώντας λοιπόν αυτό το διάστημα με το 'm-2' το διάστημα τιμών θα είναι $[-1/2, 1/2]$. Πολλαπλασιάζοντας στη συνέχεια με το 'min' το διάστημα των τιμών θα

16 Αυτό οφείλεται στο γεγονός ότι το πηλίκο μιας διαίρεσης, όταν ο διαιρετέος και ο διαιρέτης ακέραιοι, είναι πάντα ακέραιοι. Ο διαιρέτης σε αυτή την συνάρτηση είναι ακέραιος και η αρχική τιμή του διαιρετέου επίσης. Επειδή όλες οι επόμενες τιμές που παίρνει ο διαιρετέος είναι αποτέλεσμα της προηγούμενης πολλαπλασιασμένη με ακέραιο και αφαιρώντας πάλι ακέραιο, θα είναι πάντα ακέραιες. Άρα η μεγαλύτερη τιμή υπολοίπου που μπορεί να βγει από αυτή την συνάρτηση είναι η τιμή του διαιρετέου μείον ένα.

είναι το $[-\min/2, \min/2]$. Η αναγωγή στο επιθυμητό διάστημα τιμών γίνεται στην γραμμή 46 όπου ορίζεται και κάθε χρονική στιγμή η τιμή του dither. Όπως φαίνεται στη Σχέση 43 για να υλοποιηθεί χρειάζεται να χρησιμοποιείται η τιμή του προηγούμενου δείγματος. Ένας τρόπος για να υλοποιηθεί αυτό είναι με την δημιουργία ενός βρόγχου με χρήση της συνάρτησης 'for'. Έχουν δημιουργηθεί λοιπόν δύο 'for' ένα για τις τιμές των γραμμών, και ένα για τις γραμμές των στηλών. Ο ορισμός κάποιων τιμών εκτός των βρόγχων 'for' (βλέπε γραμμές 41,42) έγινε για να υπάρχουν λιγότερες πράξεις άρα να γίνονται γρηγορότερα οι υπολογισμοί. Δεν έχει επιλεγθεί επίσης η επαναφορά των αρχικών τιμών κάθε φορά που αλλάζει στήλη, ώστε να μην είναι συσχετισμένος ο θόρυβος μεταξύ των καναλιών. Ο βρόγχος στις γραμμές 53-65 είναι παρόμοιος με αυτόν στις γραμμές 40-49, με μόνη διαφορά ότι υπάρχουν δύο διαφορετικές γεννήτριες RPDF και οι έξοδοι τους προστίθενται. Αυτό έχει σαν αποτέλεσμα την δημιουργία ενός θορύβου TPDF (βλέπε εικόνα 29). Στη γραμμή 62 θα χρειαζόταν μια διαίρεση της τιμής με το δύο, ώστε να πάρουμε το μέσο όρο των δύο, η οποία έχει πραγματοποιηθεί στη γραμμή 54, και ο λόγος είναι για να γίνεται μια λιγότερη πράξη κάθε φορά που γίνεται ανάδραση στον βρόγχο. Τέλος ο βρόγχος στις γραμμές 69-80 είναι και αυτός παρόμοιος με αυτόν των γραμμών 53-65, με μόνη διαφορά αντί να χρησιμοποιεί μια δεύτερη γεννήτρια RPDF χρησιμοποιεί μια καθυστέρηση αυτής κατά ένα δείγμα την οποία την αφαιρεί από την επόμενη. Το αποτέλεσμα είναι η δημιουργία μια Highpass TPDF (βλέπε εικόνα 29).

3.1.2 Encoder B-Format από εικονικό ηχητικό πεδίο (σύνθεση).



Εικόνα 32: Μπλοκ διάγραμμα ενός αλγορίθμου σύνθεσης B-Format Ambisonics. Με το '*' δηλώνεται ο πολλαπλασιασμός του κάθε στοιχείου του ενός πίνακα με το αντίστοιχο του άλλου.

Στην εικόνα 32 φαίνεται το μπλοκ διάγραμμα ενός κωδικοποιητή B-Format Ambisonics εικονικού ηχητικού πεδίου. Όπως φαίνεται τα σήματα (X,Y,Z) προκύπτουν από το γινόμενο ημίτονων και συνημίτονων της αζιμούθιας και κατακόρυφης γωνίας, ενώ το 'W' προκύπτει με το γινόμενο του σήματος με το $1/(2^{0.5})$ (βλέπε Σχέση 29). Στο μπλοκ διάγραμμα έχει

χρησιμοποιηθεί ο συμβολισμός της πράξης '*' που συμβολίζει τον πολλαπλασιασμό του κάθε στοιχείου του ενός πίνακα με το αντίστοιχο του άλλου (συμβολισμός πράξης στο Matlab). Αυτό έχει γίνει για να είναι εφικτό ο συγκεκριμένος ήχος να μην έχει στατική αλλά δυναμική θέση στο χώρο. Έτσι οι γωνίες 'φ' και 'θ' μπορούν να είναι πίνακες μεταβολής αυτών των γωνιών στο χρόνο. Μέτα τον πολλαπλασιασμό τα σήματα πρέπει να επανακβαντιστούν, και για να επιτευχθεί αυτό έχουν χρησιμοποιηθεί τα εργαλεία για το Matlab 'Quantiser.m' και 'Makedither.m' των οποίων η λειτουργία έχει αναλυθεί στις παραγράφους 3.1.1.2 και 3.1.1.3 αντίστοιχα. Ο αλγόριθμος υλοποίησης του encoder B-Format Ambisonics εικονικού ηχητικού πεδίου βρίσκεται στο παράρτημα 10.

Τεκμηρίωση κώδικα Matlab για την συνάρτηση κωδικοποίησης εικονικού ηχητικού πεδίου σε B-Format Ambisonics.

Ο κώδικας για τον encoder B-Format Ambisonics εικονικού ηχητικού πεδίου βρίσκεται στο παράρτημα 10. Η συνάρτηση δημιουργεί ένα τετρακάναλο σήμα B-Format (W,X,Y,Z) για ένα εικονικό ηχητικό πεδίο που αποτελείται από τον ήχο του σήματος εισόδου. Τα ορίσματα της συνάρτησης υλοποίησης είναι τρία (InSignal, Azimuth, Elevation), τα οποία αντιστοιχούν στον πίνακα του σήματος εισόδου, στο αυτόν με την αζιμούθια γωνία που έχει ο ήχος κάθε χρονική στιγμή και αυτόν με την κατακόρυφη γωνία αυτού. Το 'InSignal' πρέπει να είναι ένα σήμα μονοφωνικό, σε αντίθετη περίπτωση ο κώδικας θα λάβει το πρώτο κανάλι αυτού. Ένας άλλος περιορισμός αυτού είναι ότι το μέγεθός του πρέπει να είναι τουλάχιστον 200 δείγματα. Οι πίνακες 'Azimuth' και 'Elevation' καθορίζουν από πόσα δείγματα θα είναι το σήμα εξόδου. Η σύνταξη που πρέπει να έχουν τα δεδομένα σε αυτούς του πίνακες είναι ['1^η τιμή γωνίας' 'αριθμός δείγματος 2^{ης} γωνίας' '2^η τιμή γωνίας' 'αριθμός δείγματος 3^{ης} γωνίας']. Λόγο της σύνταξης ο μικρότερος αριθμός στοιχείων που μπορεί να έχει ο πίνακας είναι '3' και το μέγεθός του μπορεί να είναι μόνο μονός αριθμός. Αυτό που θα δημιουργήσει στη συνέχεια είναι ένας πίνακας του οποίου οι τιμές θα ξεκινάνε από την τιμή '1^η τιμή γωνίας' και γραμμικά θα φτάνουν στην τιμή '2^η τιμή γωνίας' στην θέση 'αριθμός δείγματος 2^{ης} γωνίας' του πίνακα. Στη συνέχεια από την τιμή '2^η τιμή γωνίας' θα μεταβεί γραμμικά στην τιμή '3^η τιμή γωνίας' φτάνοντας σε αυτή στη θέση του πίνακα 'αριθμός δείγματος 3^{ης} γωνίας' και ούτω καθ' εξής. Είναι σαφές λοιπόν ότι το μέγεθος του πίνακα των γωνιών που θα δημιουργηθεί θα διαθέτει τόσες θέσεις όσες εκφράζει το προτελευταίο στοιχείο του πίνακα 'Azimuth' ή 'Elevation'. Το μέγεθος των πινάκων με τις γωνίες είναι αυτό που θα καθορίσει και το μέγεθος του σήματος εξόδου της συνάρτησης. Αυτό σημαίνει ότι η προτελευταία τιμή των δύο αυτών πινάκων ('Azimuth' και 'Elevation') θα πρέπει να είναι η ίδια. Ένας περιορισμός που υπάρχει από την συνάρτηση είναι ότι αυτή τη τιμή πρέπει να είναι μεγαλύτερη από το 200. Γενικά ο περιορισμός του 200 που έχει προαναφερθεί και στα τρία σήματα εισόδου οφείλεται στο μέγεθος του fade out και του crossfade που πρόκειται να χρησιμοποιηθεί αργότερα στον κώδικα. Σε περίπτωση που η προτελευταία τιμή των δύο αυτών πινάκων ('Azimuth' και 'Elevation') είναι μεγαλύτερη από το μέγεθος του σήματος εισόδου, θα γίνει επανάληψη αυτού μετά το τέλος του, με αλληλοεπικάλυψη και χρήση cross-fade σε εύρος 100 δειγμάτων ώστε να δημιουργηθεί ένα σήμα εισόδου με τον ίδιο αριθμό δειγμάτων. Στη τελευταία επανάληψη του σήματος εισόδου θα χρησιμοποιηθούν τόσα δείγματα αυτού, όσα τα δείγματα που περισσεύουν για να επιτευχθεί ο ίδιος αριθμός δειγμάτων. Τα τελευταία 100 δείγματα του νέου σήματος εισόδου θα γίνει fade out. Τέλος θα εφαρμοστεί η Σχέση 29 για την δημιουργία των τεσσάρων σημάτων του σήματος B-Format Ambisonic και στη συνέχεια θα γίνει κβαντοποίηση των τιμών.

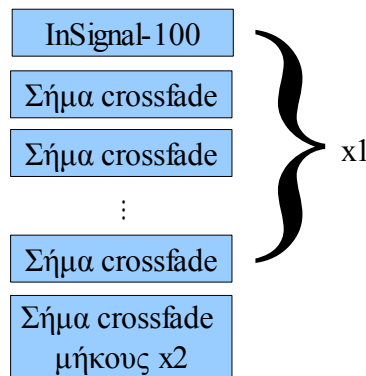
Στην πρώτη γραμμή ορίζεται η συνάρτηση με μία μεταβλητή εξόδου και 3 ορίσματα εισόδου. Στις γραμμές 31-48 δίνονται αρχικές τιμές και ορίζονται κάποιοι πίνακες που πρόκειται να

χρησιμοποιηθούν στη συνέχεια. Στη γραμμή 51 καθορίζεται ως σήμα εισόδου το πρώτο κανάλι (πρώτη στήλη) του ορίσματος InSignal. Στις γραμμές 52-53 ορίζονται ως επιθυμητές τιμές αυτές που βρίσκονται μόνο στην πρώτη γραμμή. Ο καθορισμός των επιθυμητών τιμών στις γραμμές 51-53 γίνεται για να αποφευχθούν ενδεχόμενα λάθη στη συνέχεια του κώδικα. Στις γραμμές 55-57 δημιουργούνται δύο πίνακες που πρόκειται να χρησιμοποιηθούν για την δημιουργία fade in και fade out. Στις γραμμές 59-61 ορίζονται τα μεγέθη των τριών ορισμάτων εισόδου ώστε να μπορούν να γίνουν έλεγχοι και διάφορες πράξεις στη συνέχεια. Στις γραμμές 66-101 γίνεται διάφοροι έλεγχοι που κατά σειρά αφορούν :

- γραμμές 66-68, έλεγχος αν ο αριθμός των στηλών των πινάκων εισόδου 'Azimuth' και 'Elevation' είναι μονός.
- γραμμές 69-71, έλεγχος αν ο αριθμός των στηλών των πινάκων εισόδου 'Azimuth' και 'Elevation' είναι μεγαλύτερος του 3 όπου είναι και η μικρότερη τιμή που μπορεί να έχουν αν υπάρχει ορθή σύνταξή τους.
- γραμμές 74-79, έλεγχος αν υπάρχει λάθος στη σύνταξη του πίνακα 'Azimuth' όσον αφορά την αύξηση που πρέπει να έχουν οι τιμές των 'αριθμός δείγματος n^{th} γωνίας ' καθώς μεγαλώνει η τιμή της γωνίας 'n'.
- γραμμές 81-87, έλεγχος αν υπάρχει λάθος στη σύνταξη του πίνακα 'Elevation' όσον αφορά την αύξηση που πρέπει να έχουν οι τιμές των 'αριθμός δείγματος n^{th} γωνίας ' καθώς μεγαλώνει η τιμή της γωνίας 'n'.
- γραμμές 89-91, έλεγχος αν η τιμή του προτελευταίου στοιχείου των πινάκων 'Azimuth' και 'Elevation' είναι ίδιο. Αυτό γίνεται γιατί αυτές καθορίζουν το μέγεθος του σήματος εξόδου, άρα πρέπει να είναι ίσες.
- γραμμές 93-95, έλεγχος αν η το μέγεθος του σήματος εισόδου καθώς και η τιμή του προτελευταίου δείγματος του πίνακα 'Azimuth' είναι μεγαλύτερη του 200. Γίνεται έλεγχος μόνο στον πίνακα 'Azimuth' καθώς αφού έχει προηγηθεί έλεγχος αν η προτελευταία τιμή αυτού ισούται με την αντίστοιχη του 'Elevation'.
- γραμμές 97-101, έλεγχος αν η τιμές των γωνιών του Elevation βρίσκονται μεταξύ των ορίων [-90,90].

Στις γραμμές 104-114 γίνεται η δημιουργία του πίνακα των αζιμούθιων γωνιών που αντιστοιχούν σε κάθε δείγμα (σε κάθε χρονική στιγμή), αποκωδικοποιώντας την σύνταξη του πίνακα 'Azimuth'. Δημιουργείται ένας βρόγχος 'for' για να παίρνονται διαδοχικά τιμές που αντιστοιχούν στις μονές θέσεις του πίνακα 'Azimuth' εκτός της τελευταίας του. Οι μονές τιμές του πίνακα 'Azimuth' αντιστοιχούν στη τιμή αζιμούθιας γωνίας που πρέπει να έχει ο πίνακας αζιμούθιων γωνιών σε συγκεκριμένες τιμές δειγμάτων. Για τον υπολογισμό των ενδιάμεσων τιμών έχει χρησιμοποιηθεί η σχέση (αρχική τιμή):[(τελική τιμή)-(αρχική τιμή)]/[(αριθμός δειγμάτων που μεσολαβούν)-1]: (τελική τιμή) (βλέπε γραμμή 109). Το 'w2' έχει χρησιμοποιηθεί σαν μονάδα μνήμης της θέσης του πίνακα που αντιστοιχεί στην 'αρχική τιμή' της εξίσωσης που προαναφέρθηκε, ώστε να βρεθεί ο αριθμός δειγμάτων που μεσολαβούν. Η αρχική τιμή του 'w2' είναι μηδέν και χρησιμοποιείται για να μην υπάρχει πρόβλημα όταν το 'i' παίρνει την τιμή '1'. Το 'r' χρησιμεύει για να γίνει η αφαίρεση του '1' που πρέπει να γίνει στο πρώτο δείγμα για να είναι σωστός ο αριθμός των δειγμάτων που μεσολαβούν. Μόλις ολοκληρωθεί μια φορά ο βρόγχος το 'r' παίρνει την τιμή μηδέν. Αυτό συμβαίνει γιατί στον πίνακα δεν πρέπει να υπάρχει επανάληψη της τελικής τιμής της τελευταίας θέσης με την επόμενη. Για παράδειγμα κατά την σύνταξη αν ο πίνακας 'Azimuth' ήταν [1 3 3 5 5] τότε βάση της εξίσωσης οι τιμές εξόδου θα ήταν [1 2 3 3 5]. Όπως φαίνεται στην 3 και τέταρτη

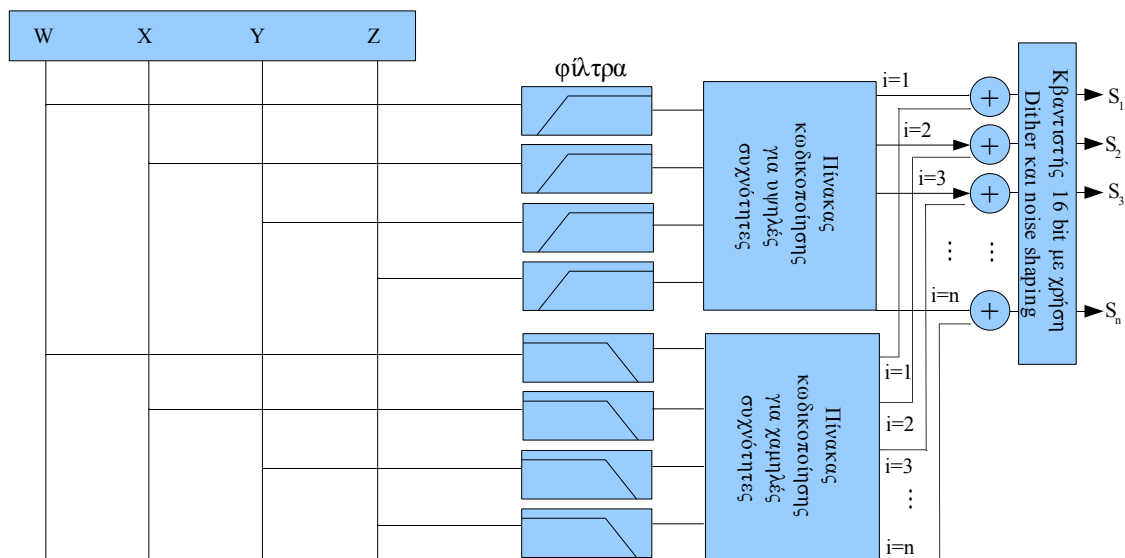
θέση γίνεται επανάληψη της τιμής. Για να υπάρχει συνέχεια και να μην δημιουργείται το παραπάνω σφάλμα υπολογίζεται ένα δείγμα παραπάνω (θέτοντας $i=0$), και στη συνέχεια αγνοείται η πρώτη τιμή που ουσιαστικά είναι η επανάληψη. Με αυτόν τον τρόπο ο πίνακας εξόδου θα ήταν [1 2 3 4 5]. Δεν θα πρέπει να αγνοείται η πρώτη τιμή μόνο για $i=1$ για αυτό τον λόγο και στη γραμμή 111 έχει γίνει ένα πολλαπλασιασμός του '-1' με το 'r'. Ένας δεύτερος πολλαπλασιασμός του 'r' με το -1 γίνεται στο τέλος της 111 καθώς για την τιμή $i=1$ το μέγεθος του πίνακα ισούται με 'Azimuth(2)'. Στην γραμμή 105 πραγματοποιείται ένας έλεγχος, όπου αν η αρχική τιμή της γωνίας ισούται με την τελική, να πραγματοποιείται η εξίσωση που βρίσκεται στην γραμμή 107. Αυτός ο έλεγχος είναι απαραίτητος καθώς αν $\alpha=\beta$ στην εξίσωση $\alpha:(\beta-\alpha)/(\gamma-1):\beta$ βγάζει στην έξοδο του μηδενικό πίνακα, ενώ στην πραγματικότητα βάση της σύνταξης θα έπρεπε ο πίνακας που θα σχηματιστεί να έχει γ στοιχεία που όλα θα ισούνται με το α . Τέλος ο σχηματισμός του συνολικού πίνακα γίνεται με την προσθήκη του καινούριου πίνακα που δημιουργείται για κάθε 'i' στον συνολικό (βλέπε γραμμή 111). Στη γραμμή 116 γίνεται μια αναστροφή των στηλών σε γραμμές του πίνακα 'azimhtable' καθώς σε μεταγενέστερες πράξεις θα πρέπει να είναι πίνακας στήλη. Στις γραμμές 118-131 υπάρχει υλοποίηση του ίδιου κώδικα που υπάρχει στις γραμμές 104-116 μόνο που τώρα αφορά την δημιουργία του πίνακα τιμών των κατακόρυφων γωνιών από τον πίνακα 'Elevation'. Στις γραμμές 134-153 δημιουργείται το συνολικό σήμα εισόδου που θα έχει το ίδιο μέγεθος με το μέγεθος του πίνακα των γωνιών. Αυτό χρειάζεται επειδή έγινε αρχική παραδοχή ότι αυτός που θα καθορίζει το μέγεθος του σήματος εξόδου είναι το μέγεθος του πίνακα γωνιών που πρόκειται να δημιουργηθεί. Εδώ διακρίνονται δύο περιπτώσεις. Η μία είναι το 'InSignal' να είναι μικρότερο ή ίσο με αυτό που τότε απλώς περιορίζεται το εύρος αυτού από την αρχική του τιμή σε αυτή του πίνακα γωνιών (βλέπε 134-136). Στη γραμμή 136 γίνεται και ένας πολλαπλασιασμός των τελευταίων 100 δειγμάτων του σήματος εισόδου με το 'fadeout' για να υπάρχει ομαλό σβήσιμο του σήματος. Η δεύτερη περίπτωση είναι το μέγεθος του σήματος 'InSignal' να είναι μικρότερο από την επιθυμητό. Σε αυτή την περίπτωση θα γίνεται επανάληψη του σήματος 'InSignal' όσες φορές είναι επιθυμητό με crossfade διάρκειας 100 δειγμάτων. Το κοινό στοιχείο σε κάθε επανάληψη θα ήταν ένα σήμα που θα ξεκίναγε με crossfade και θα τελείωνε πριν αρχίσει το επόμενο crossfade. Το μήκος που θα έχει αυτό το σήμα θα είναι όσο του 'InSignal' μείον 100. Το αρχικό σήμα σε αυτή την ακολουθία θα ήταν το 'InSignal' χωρίς τα τελευταία 100 δείγματα. Δημιουργούνται λοιπόν στις γραμμές 140-142 το σήμα που θα εμπεριέχει το crossfade, και το μειωμένο αρχικό σήμα. Στη γραμμή 138 βρίσκεται πόσες φορές χωράει ακέραιο το 'InSignal-100' στον συνολικό μέγεθος που πρέπει να έχει αυτό. Στη συνέχεια βρίσκεται το υπόλοιπο της διαίρεσης του συνολικού μεγέθους που πρέπει να έχει το σήμα εισόδου με το 'InSignal-100', το οποίο αντιστοιχεί στα επιπλέον δείγματα που χρειάζονται από ένα σήμα crossfade για να επιτευχθεί το επιθυμητό μέγεθος πίνακα (βλέπε γραμμή 139). Η μορφή που θα έχει πλέον το αρχικό σήμα φαίνεται στην εικόνα (33).



Εικόνα 33: Απεικόνιση του σχηματισμού του συνολικού σήματος εισόδου. Το $x1$ αντιστοιχεί σε πόσες φορές χωράει το μέγεθος του 'InSignal-100' στο συνολικό μέγεθος που πρέπει να έχει το σήμα εισόδου. Το $x2$ αντιστοιχεί στο υπόλοιπο της διαίρεσης του μεγέθους που πρέπει να έχει το σήμα εισόδου δια του 'InSignal'.

Ο σχηματισμός του συνολικού αρχικού σήματος γίνεται μεταξύ των γραμμών 143-152. Στις γραμμές 157-160 γίνεται η δημιουργία των σημάτων του B-Format Ambisonics σύμφωνα με την Σχέση 29. Στη γραμμή 164 δημιουργείται ένας πίνακας (σήμα) που περιέχει και τα τέσσερα σήματα (στην πρώτη στήλη το 'W', στη δεύτερη το 'X' στην τρίτη το 'Y' και στη τέταρτη το 'Z'). Τέλος στη γραμμή 166 γίνεται κβαντοποίηση αυτών των σημάτων.

3.2. Ambisonic αποκωδικοποιητής (Decoder).

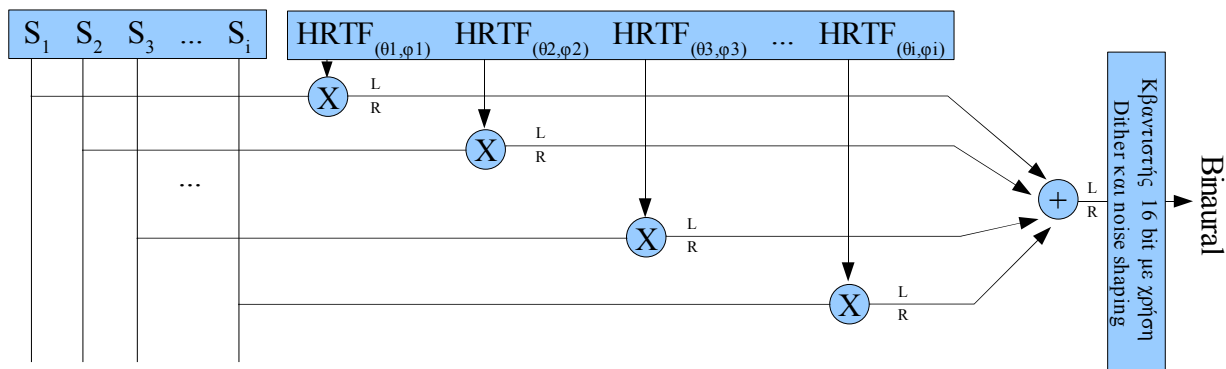


Εικόνα 34: Μπλοκ διάγραμμα αποκωδικοποιητή Ambisonic. Τα S_i αντιστοιχεί στο σήμα που θα τροφοδοτήσει το κάθε ηχείο, ενώ τα φίλτρα High και Low pass έχουν συχνότητα αποκοπής στα 400 Hz. Ανάλογα με την διάταξη των ηχείων που πρόκειται να χρησιμοποιηθεί αλλάζουν και οι πίνακες κωδικοποίησης για τις υψηλές και χαμηλές συχνότητες. Επίσης υπάρχει αντιστάθμιση των φάσεων μεταξύ των φίλτρων για να μην υπάρχουν παραμορφώσεις φάσης.

Στην εικόνα 34 φαίνεται το μπλοκ διάγραμμα ενός αποκωδικοποιητή Ambisonics πρώτης τάξης. Τα σήματα εισόδου του είναι τα τέσσερα σήματα του B-Format που αντιστοιχούν

στην σφαιρικές αρμονικές μηδενικής και πρώτης τάξης. Το κάθε ένα από αυτά τα σήματα δρομολογείται σε ένα crossover¹⁷ δύο περιοχών, το οποίο διαχωρίζει τα σήματα σε μία περιοχή κάτω και μία πάνω από τα 400 Hz (το crossover στην εικόνα 34 αποτελείται από ένα φίλτρο Highpass και ένα Lowpass των οποίων η συχνότητα αποκοπής είναι τα 400 Hz). Τα φίλτρα που τα αποτελούν είναι σχεδιασμένα έτσι ώστε να έχουν γραμμική απόκριση φάσης, και να παρουσιάζουν όσο το δυνατό λιγότερη διαφορά στην φάση μεταξύ τους. Στη συνέχεια, αυτά τα σήματα δρομολογούνται σε δύο συναρτήσεις που αναλαμβάνουν να προσθέσουν τα σήματα του B-Format με διαφορετικά βάρη, για τον υπολογισμό του σήματος που πρέπει να τροφοδοτηθεί σε κάθε ηχείο, όσον αναφορά τις υψηλές και χαμηλές συχνότητες. Στη συνέχεια προστίθενται αυτά τα δύο σήματα ώστε να ανασυσταθεί όλο το συχνοτικό εύρος του σήματος για το κάθε ηχείο.

Στον κωδικοποιητή έχει ενσωματωθεί και ένας αλγόριθμος κωδικοποίησης σε Binaural του αναπαραγόμενου, από την διάταξη ηχείων, ηχητικού πεδίου, του οποίου το μπλοκ διάγραμμα φαίνεται στην εικόνα 35.



Εικόνα 35: Μπλοκ διάγραμμα για την κωδικοποίηση σε Binaural του ηχητικού πεδίου που δημιουργείται από την Ambisonic αναπαραγωγή. Με το 'X' συμβολίζεται η συνέλιξη (convolution) στο χρόνο μεταξύ των δύο σημάτων.

3.2.1 Δημιουργία συνάρτηση υπολογισμού της κρουστικής απόκρισης για τα φίλτρα Highpass και Lowpass και τρόπος φιλτραρίσματος σήματος.

Τα φίλτρα που πρόκειται να χρησιμοποιηθούν είναι ένα Lowpass για την περιοχή των συχνοτήτων κάτω από τα 400 Hz και ένα φίλτρο Highpass για την περιοχή πάνω από τα 400 Hz. Ο σχεδιασμός του φίλτρου θα πραγματοποιηθεί στο Matlab με το εργαλείο fdatool. Το φίλτρο πρόκειται να υλοποιηθεί σαν FIR για να έχουμε γραμμική απόκριση φάσης. Τα χαρακτηριστικά που πρέπει να έχει το φίλτρο είναι :

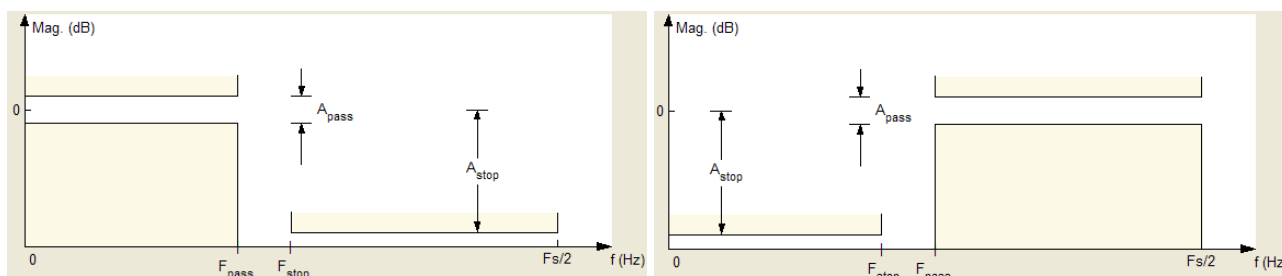
- Γραμμική φάση για την συχνοτική περιοχή Band pass, και ελαχιστοποίηση όσο γίνεται της διαφοράς φάσης μεταξύ των δύο φίλτρων.
- Μείωση της έντασης στο εύρος συχνοτήτων της περιοχής Band stop (περιοχή μετά ή πριν το Fstop ανάλογα με το αν το φίλτρο θα είναι lowpass ή highpass) στα 90 dB, ώστε το σήμα αυτών των περιοχών να βρίσκεται σχεδόν στο όριο του θορύβου.
- Στη συχνοτική περιοχή Pass band (περιοχή πριν ή μετά το Fstop ανάλογα με το αν

¹⁷ Διάταξη διαχωρισμού των σημάτων σε έναν αριθμό συχνοτικών περιοχών. Το συχνοτικό εύρος της κάθε περιοχής εξαρτάται από τον σχεδιασμό του.

το φίλτρο θα είναι Lowpass ή highpass) να έχει συχνотική απόκλιση που δεν θα ξεπερνάει τα ± 0.5 dB.

- Η συχνότητα αποκοπής και των δύο φίλτρων θα ήταν τα 400 Hz, όσο το δυνατό μικρότερο εύρος συχνотήτων για την μετάβαση από το συχνотικό εύρος Pass band στο Stop band.

Η τάξη του φίλτρου θα επιλεγεί ώστε να είναι η μικρότερη δυνατή ώστε να εκπληρώνονται τα χαρακτηριστικά που πρέπει να έχει. Στην εικόνα 36 φαίνεται πως αντιστοιχίζονται οι μεταβλητές που χρησιμοποιούνται στο `fdatool` του matlab με τα χαρακτηριστικά του φίλτρου. Για την σχεδίαση του lowpass φίλτρου, οι μεταβλητές ορίστηκαν με τις τιμές ($A_{pass}=0.5$ dB, $A_{stop}=90$ dB, $F_{pass}=400$ Hz, $F_{stop}=500$ Hz, και συχνότητα δειγματοληψίας $F_s=44100$). Αντίστοιχα για το Highpass φίλτρο, οι μεταβλητές ορίστηκαν με τις τιμές ($A_{pass}=0.5$ dB, $A_{stop}=90$ dB, $F_{pass}=500$ Hz, $F_{stop}=400$ Hz, και συχνότητα δειγματοληψίας $F_s=44100$).

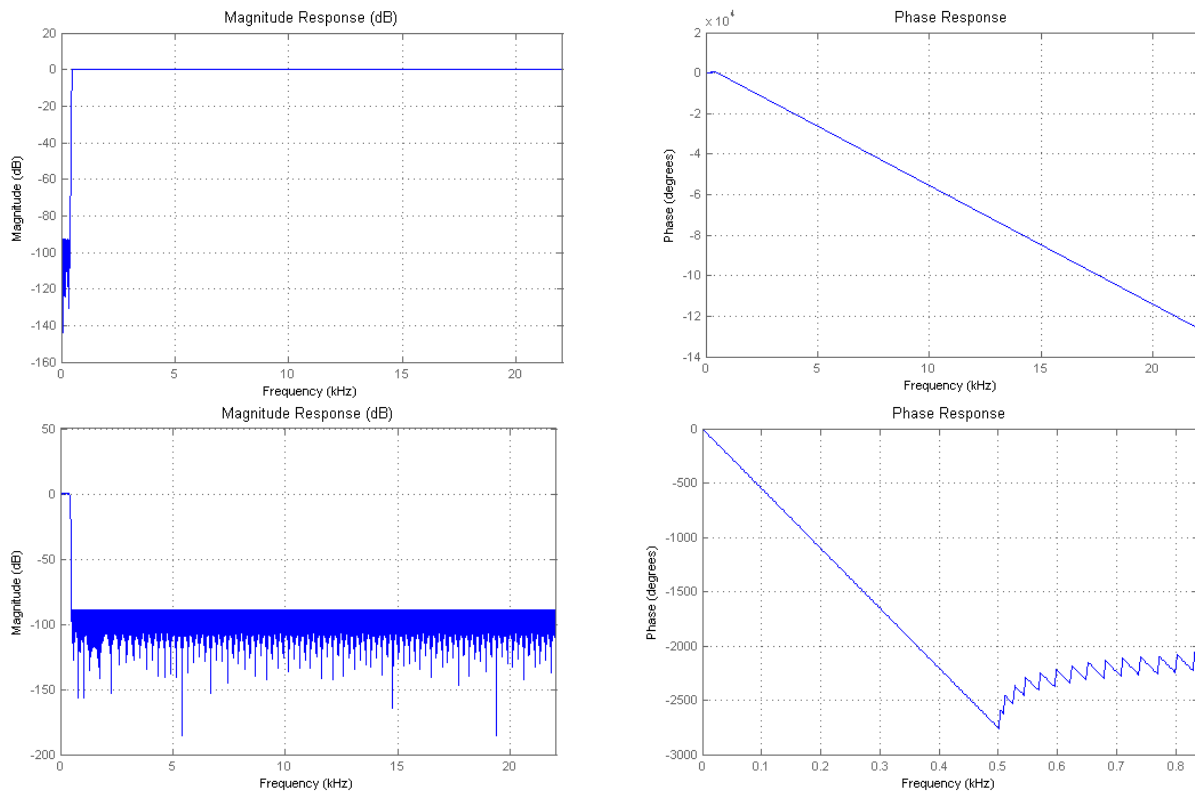


Εικόνα 36: Απεικόνιση του τρόπου με τον οποίο συμβολίζονται τα στοιχεία του φίλτρου στο Matlab, και πιο συγκεκριμένα στο `fdatool`.

Η συχνотική απόκλιση και η απόκλιση φάσης των δύο φίλτρων φαίνεται στην εικόνα 37.

Η καθυστέρηση φάσης και το 'Group delay' ενός φίλτρου FIR είναι ίσα στην συχνотική περιοχή που η απόκλιση φάσης είναι γραμμική. Για ένα φίλτρο FIR τάξης 'n' που παρουσιάζει γραμμική φάση, το 'Group delay' ισούται με $n/2$ και το σήμα που υπόκειται στο φιλτράρισμα πρόκειται να υποστεί καθυστέρηση ίση με $n/2$ δείγματα.

Από την στιγμή που υπολογίστηκαν οι συντελεστές υλοποίησης των FIR φίλτρων highpass και lowpass, για να εφαρμοστούν σε ένα σήμα αρκεί να γίνει συνέλιξη (convolution) αυτού με τους συντελεστές του φίλτρου. Στη συνέχεια πρέπει να αγνοηθούν τα πρώτα $n/2$ δείγματα για να γίνει η εξομάλυνση της φάσης. Όμως μετά τη συνέλιξη (convolution) το μέγεθος του σήματος που υπ'αίσθη το φιλτράρισμα αυξήθηκε κατά 'n' δείγματα. Αυτό σημαίνει ότι για να ανακτήσει το σήμα το αρχικό του μέγεθος, θα πρέπει να αγνοηθούν και τα τελευταία $n/2$ δείγματα.



Εικόνα 37: Σε αυτή την εικόνα αριστερά φαίνεται η συχνωτική απόκριση ενώ δεξιά η απόκριση φάσης για το κάθε φίλτρο. Οι πάνω εικόνες αφορούν το φίλτρο highpass ενώ κάτω το φίλτρο lowpass.

3.2.2 Τεκμηρίωση κώδικα Matlab για την δημιουργία του αποκωδικοποιητή B-Format.

Ο κώδικας υπάρχει στο Παράρτημα 11. Με τον συγκεκριμένο αποκωδικοποιητή, ο χρήστης μπορεί να αποκωδικοποιήσει ένα ηχητικό πεδίο από B-Format Ambisonic σε οποιοδήποτε αριθμό και διάταξη ηχείων επιθυμεί, αρκεί να διαθέτει τον κατάλληλο πίνακα συντελεστών (βάρη για την άθροιση των 'W,X,Y,Z'). Ταυτόχρονα το αναπαραγόμενο ηχητικό πεδίο από τα ηχεία μπορεί να το κωδικοποιήσει σε Binaural. Οι έξοδοι της συγκεκριμένης συνάρτησης είναι δύο (OutSignal,binaural) εκ των οποίων η πρώτη αντιστοιχεί σε ένα πολυκάναλο σήμα που το κάθε του κανάλι (στήλη) αντιστοιχεί στο σήμα που πρέπει να τροφοδοτηθεί σε κάθε ηχείο, ενώ το άλλο είναι ένα στερεοφωνικό σήμα που αντιστοιχεί στη binaural κωδικοποίηση του αναπαραγόμενου ηχητικού πεδίου. Τα ορίσματα εισόδου είναι τρία (InSignal,computebinauralflag,coefficientindex). Το 'InSignal' είναι ένα τετρακάναλο σήμα που αντιστοιχεί στο B-Format Ambisonic. Το 'computebinauralflag' παίρνει τιμές (1,0) οι οποίες καθορίζουν άμα θα γίνει ή όχι κωδικοποίηση του ηχητικού πεδίου σε Binaural. Το 'coefficientindex' αντιστοιχεί στον πίνακα με τις σφαιρικές συντεταγμένες του κάθε ηχείου και τους συντελεστές της πρόσθεσης των (W,X,Y,Z) για την δημιουργία του σήματος που θα τροφοδοτηθεί στο κάθε ηχείο. Η σύνταξη που πρέπει να έχει αυτός ο πίνακας είναι ο εξής: η κάθε γραμμή του αντιστοιχεί στα στοιχεία του κάθε ηχείου, η πρώτη στήλη στην αζιμούθια γωνία αυτού, η δεύτερη στην γωνία ανύψωσης, οι στήλες 3,4,5,6 στους συντελεστές των σημάτων (W,X,Y,Z) για τις υψηλές συχνότητες ενώ οι στήλες 7,8,9,10 στους συντελεστές των σημάτων (W,X,Y,Z) για τις χαμηλές συχνότητες. Από τα παραπάνω είναι

προφανές ότι ο αριθμός των γραμμών αυτού του πίνακα ταυτίζεται με τον αριθμό των ηχείων που πρόκειται να χρησιμοποιηθούν. Σε αυτή την εργασία οι συντελεστές των πινάκων θα χρησιμοποιηθούν από ήδη σχηματισμένα συστήματα είτε θα δημιουργηθούν εκ νέου. Οι περιορισμοί στην γωνία ανύψωσης ισχύουν όπως στους προηγούμενους κώδικες. Το πλεονέκτημα που παρουσιάζει ο συγκεκριμένος κώδικας είναι ότι μπορεί να αποκωδικοποιήσει το B-Format Ambisonic, ανεξάρτητα από την διάταξη των ηχείων που πρόκειται να χρησιμοποιηθεί, αρκεί ο χρήστης να διαθέτει τους πίνακες συντελεστών. Ταυτόχρονα μπορεί να γίνει ακρόαση του ηχητικού πεδίου που ανασυστήνεται σε Binaural. Στη συνέχεια ακολουθεί ανάλυση του κώδικα.

Στη πρώτη γραμμή ορίζεται η συνάρτηση με τις μεταβλητές εξόδου και τα ορίσματα εισόδου. Στις γραμμές 36-43 καθορίζονται οι αρχικές τιμές και οι πίνακες εξόδου. Στις γραμμές 46-47 ορίζονται τέσσερις μεταβλητές που αντιστοιχούν στα μεγέθη των ορισμάτων εισόδου (InSignal, coefficientindex). Στις γραμμές 54-68 γίνονται τρεις διαδοχικοί έλεγχοι, εκ των οποίων ο πρώτος αφορά κατά πόσον το σήμα εισόδου διαθέτει τα τέσσερα σήματα (W,X,Y,Z) που απαρτίζουν το B-Format Ambisonics. Ο δεύτερος ελέγχει αν ο πίνακας 'coefficientindex' διαθέτει 10 στήλες, που αντιστοιχούν στη σωστή σύνταξή του, ενώ ο τρίτος ελέγχει αν κάποια από τις γωνίες ανύψωσης του πίνακα 'coefficientindex' (τιμές δεύτερης στήλης) παίρνει τιμή εκτός των ορίων [-90,90]. Στις γραμμές 72-86 και 93-108 δημιουργούνται δύο φίλτρα (Highpass, Lowpass) με τις προδιαγραφές που έχουν οριστεί στην παράγραφο 3.2.1. Ο κώδικας αυτών δημιουργήθηκε με το εργαλείο fdatool του Matlab. Στο τέλος του κάθε φίλτρου γίνεται μια αναστροφή του πίνακα των συντελεστών του φίλτρου ώστε από 'πίνακας γραμμής', να γίνει 'πίνακας στήλης', για ευκολία στις πράξεις στη συνέχεια. Στις γραμμές 111-126 γίνεται φιλτράρισμα των σημάτων (W,X,Y,Z) ώστε να χωριστούν σε ένα υψηλό και χαμηλό συχνοτικό εύρος, ενώ στη συνέχεια μεταβάλλεται κατάλληλα το μέγεθος του πίνακα για να έχουμε ίδιο σήμα με το αρχικό και να γίνει αντιστοίχιση της φάσης (βλέπε παράγραφο 3.2.1). Στη γραμμή 129 δημιουργείται ένας πίνακας που περιέχει το κάθε ένα από αυτά τα σήματα σε διαφορετική στήλη με την διάταξη (Whigh, Xhigh, Yhigh, Zhigh, Wlow, Xlow, Ylow, Zlow). Στη γραμμή 131 βρίσκεται η μέγιστη απόλυτη τιμή αυτού του πίνακα, ώστε στην γραμμή 133 να γίνει Normalize του σήματος. Αυτό είναι απαραίτητο καθώς στο φίλτρο έχει δοθεί απόκριση συχνότητας ± 0.5 dB και αυτό σημαίνει ότι αν το σήμα, στην συχνοτική περιοχή Passband έχει μέγιστη τιμή ± 1 αυτό μπορεί να παραμορφωθεί. Στις γραμμές 136-146 γίνεται υλοποίηση της Σχέση 37 για να δημιουργηθούν τα σήματα που θα τροφοδοτήσουν το κάθε ηχείο. Ο βρόγχος με το 'for' έχει χρησιμοποιηθεί για να υπολογίζεται διαδοχικά το σήμα που θα τροφοδοτήσει το κάθε ηχείο, και με αυτό τον τρόπο να παίρνονται και οι αντίστοιχες τιμές από τον πίνακα 'coefficientindex'. Το τελικό σήμα που εξάγεται από τον βρόγχο 'for' αποτελείται από τόσα σήματα (κανάλια) όσα και τα ηχεία που περιγράφονται στον πίνακα 'coefficientindex' με τις γραμμές αυτού. Στις γραμμές 148-149 γίνεται ένα normalize όλων των σημάτων που προορίζονται για τα ηχεία, βάση της μέγιστης τιμής έντασης αυτών συνολικά. Στη γραμμή 152 γίνεται κβαντοποίηση του σήματος εξόδου. Στις γραμμές 158-229 γίνεται υπολογισμός του Binaural σήματος στην περίπτωση που το επιλέξει ο χρήστης. Στη γραμμή 158 γίνεται έλεγχος αν ο χρήστης επιθυμεί να γίνει υπολογισμός του σήματος Binaural. Στις γραμμές 159-161 γίνεται αναγωγή των αξιμούθιων γωνιών που έχει θέσει ο χρήστης στο εύρος [0-360]. Στις γραμμές 162-164 γίνεται έλεγχος αν οι γωνίες ανύψωσης είναι μικρότερες από τις -45° , όπου όταν αυτό είναι αληθές η συνάρτηση διακόπτεται και βγάζει 'error'. Αυτό συμβαίνει για τον λόγο ότι οι HRTF συναρτήσεις που διαθέτουμε σε αυτή την εργασία δεν υπάρχουν για τιμές γωνίας ανύψωσης μικρότερη των -40° . Ο έλεγχος για τις 45° γίνεται για τον λόγο ότι με την στρογγυλοποίηση που πρόκειται να γίνει στις 'μονάδες' των γωνιών, οι γωνίες μέχρι τις 45° θα στρογγυλοποιούνται στις 40° . Λόγω του ότι το βήμα στις αξιμούθιες γωνίες βάση των HRTF που διαθέτουμε διαφέρει, ανάλογα με την γωνία ανύψωσης, και το γεγονός ότι έχουμε HRTF με συγκεκριμένες γωνίες ανύψωσης, έχει

δημιουργηθεί ο κώδικας στις γραμμές 167-190 όπου στρογγυλοποιεί την αζιμούθια γωνία και την γωνία ανύψωσης των ηχείων στις κοντινότερες δυνατές τιμές γωνιών βάση των HRTF. Στη συνέχεια στις γραμμές 192-209 σχηματίζεται η διαδρομή 'path' στο οποίο βρίσκονται οι HRTF και φορτώνεται κάθε φορά η κατάλληλη. Η εισαγωγή των σημάτων HRTF γίνεται στις γραμμές 212-213, ενώ στις γραμμές 214 και 216 γίνεται συνέλιξη (convolution) μεταξύ του ηχείου και της κατάλληλης HRTF (βάση των σφαιρικών συντεταγμένων). Στις γραμμές 215 και 217 γίνεται άθροιση όλων των σημάτων που εκπέμπονται από τα ηχεία (μετά την συνέλιξη) για το κάθε σήμα του binaural (L,R). Τέλος στη γραμμή 223 γίνεται η δημιουργία ενός σήματος που αποτελείται από το (L,R), ενώ στις γραμμές 25-28 γίνεται normalize και κβαντοποίηση αυτού.

3.2.3 Τεκμηρίωση κώδικα Matlab για τον σχηματισμό του πίνακα συντελεστών για τη δημιουργία των σημάτων που θα τροφοδοτήσουν τα ηχεία, που θα χρησιμοποιηθούν στη συνέχεια στην αποκωδικοποίηση.

Το συγκεκριμένο εργαλείο μπορεί ο χρήστης να δημιουργήσει έναν πίνακα με τους συντελεστές των 'W' και των 'X,Y,Z' και τις γωνίες του κάθε ηχείου που θα χρησιμοποιηθούν στη συνέχεια στον αποκωδικοποιητή ώστε να δημιουργηθούν τα κατάλληλα σήματα που πρέπει να τροφοδοτηθούν στα ηχεία. Ο χρήστης μπορεί να επιλέξει ανάμεσα σε 4 διαφορετικούς τρόπους υπολογισμού των συντελεστών, εκ των οποίων οι δύο συνιστώνται από τον M. Gerzon, ένας από τον D. Malham και ένας από τον Farina. Ο χρήστης σε τρεις από τις τέσσερις περιπτώσεις μπορεί να επιλέξει τον αριθμό των ηχείων του συστήματος αναπαραγωγής, τα οποία βρίσκονται στις κορυφές ενός κανονικού n-γώνου, στο οριζόντιο επίπεδο. Τα ορίσματα εισόδου είναι δύο, εκ των οποίων με το πρώτο 'todoflag' ορίζεται η μέθοδος με την οποία ο χρήστης επιθυμεί να υπολογιστούν οι συντελεστές. Οι τιμές που μπορεί να θέσει είναι τέσσερις (1,2,3,4). Με την πρώτη τιμή (1) δημιουργείται ο πίνακας συντελεστών για ένα σύστημα ηχείων που αποτελείται από πέντε ηχεία τα οποία είναι διατεταγμένα στο οριζόντιο επίπεδο με γωνίες (0,45,130,-130,-45). Τον υπολογισμό των συντελεστών αυτού του συστήματος την έχει κάνει ο M. Gerzon (βλέπε 25). Με την τιμή '2' υπολογίζονται οι συντελεστές για ένα σύστημα με τόσα ηχεία όσα έχει ορίσει ο χρήστης (δεύτερο όρισμα εισόδου 'speakers', που οι τιμές του μπορούν να είναι μεγαλύτερες του 4), που προορίζεται να παίξει σε συναυλιακούς χώρους. Η μέθοδος υπολογισμού των σημάτων έχει δημιουργηθεί από τον D. Malham. Με την τιμή '3' υπολογίζονται οι συντελεστές για ένα σύστημα αναπαραγωγής που αποτελείται από τόσα ηχεία, όση η αριθμητική τιμή του δεύτερου ορίσματος εισόδου (speakers), και προορίζεται για αναπαραγωγή σε στουντιακούς χώρους. Ο τρόπος υπολογισμού των συντελεστών δημιουργήθηκε από τον M. Gerzon. Με τη τέταρτη τιμή γίνεται υπολογισμός των συντελεστών των ηχείων για ένα σύστημα που αποτελείται από τόσα ηχεία όσος ο αριθμός του δεύτερου ορίσματος. Ο γενικός τύπος για τον υπολογισμό τον έχει δημιουργήσει ο Farina, αλλά πιο συγκεκριμένα η επιλογή του 'd' γίνεται ερευνώντας την καταλληλότερη τιμή 'd' κατά την οποία υπάρχει μικρότερη απόκλιση από την μονάδα των 're' και 'rv' (και δευτερευόντως των r_v και r_e αντίστοιχα) αντίστοιχα, για τις υψηλές και χαμηλές συχνότητες για όλες τις αναπαραγόμενες γωνίες προέλευσης, ενώ ταυτόχρονα ισχύει $\theta_v = \theta_e = \theta$ για όλο το εύρος των συχνοτήτων. Τα ορίσματα εξόδου είναι τέσσερα (coefficientindextable, table_rv_re, dH, dL). Το πρώτο είναι ο πίνακας που περιέχει τους συντελεστές για τα ηχεία, και την αζιμούθια και κατακόρυφη γωνία αυτού. Η μορφή που έχει ο πίνακας εξόδου είναι [azimuth, elevation, WH, XH, YH, ZH, WL, XL, YL, ZL]. Το δεύτερο (table_rv_re) είναι ένας πίνακας που περιέχει για κάθε πραγματική γωνία αναπαραγωγής στο οριζόντιο επίπεδο, τις τιμές των r_v, r_e και των αντιληπτών γωνιών βάση του r_v και του r_e για τις υψηλές και χαμηλές συχνότητες. Η μορφή που

θα έχει ο πίνακας εξόδου είναι [azim rHv rHe VHazim EHazim rLv rLe VLazim Elazim]. Το τρίτο και τέταρτο όρισμα εξόδου είναι οι τιμές των 'd' για βέλτιστη αναπαραγωγή για τις υψηλές και χαμηλές συχνότητες, στην περίπτωση που στο πρώτο όρισμα εισόδου επιλεγεί το 4 (τέταρτη μέθοδος υπολογισμού των συντελεστών). Ο κώδικας για αυτό το εργαλείο βρίσκεται στο παράρτημα 12. Στη συνέχεια ακολουθεί η τεκμηρίωση του κώδικα.

Στη πρώτη γραμμή ορίζεται η συνάρτηση με τα ορίσματα εισόδου και εξόδου. Στις γραμμές 30-42 γίνεται έλεγχος αν είναι σωστός ο αριθμός των ορισμάτων εισόδου, ανάλογα με την αριθμητική τιμή του πρώτου ορίσματος εισόδου, καθορίζεται η τιμή που θα βρίσκεται το πρώτο ηχείο, καθώς και η γωνία που θα σχηματίζουν μεταξύ τους δύο διαδοχικά ηχεία. Στις γραμμές 46-49 γίνεται έλεγχος αν οι τιμή που έχει δώσει ο χρήστης στο πρώτο όρισμα 'todoflag' είναι μέσα στις επιτρεπτές τιμές. Στις γραμμές 52-75 δίνονται αρχικές τιμές για τις διάφορες μεταβλητές που πρόκειται να χρησιμοποιηθούν στη συνέχεια. Στις γραμμές 77-136 γίνεται η δημιουργία των συντελεστών και ο πίνακας αποτίμησης (δεύτερο όρισμα εξόδου) για την περίπτωση που ο χρήστης επιλέξει ότι το πρώτο όρισμα εισόδου 'todoflag' θα έχει την τιμή 1. Στις γραμμές 80-84 δημιουργείται ο πίνακας των συντελεστών για αυτή την περίπτωση (5 ηχεία σε γωνίες 0,45,130,-130,-45) του M. Gerzon, βλέπε 25). Στις γραμμές 86-136 δημιουργούνται τα αθροίσματα από τις Σχέση 30, Σχέση 31, Σχέση 33 και Σχέση 34, ώστε στη συνέχεια να μπορεί να συμπληρωθεί ο πίνακας της αποτίμησης του συστήματος. Σε αυτές τις εξισώσεις γίνεται η κωδικοποίηση και αποκωδικοποίηση ενός σήματος με πλάτος μονάδα, για το διάστημα γωνιών 0-345 με βήμα 15°. Για κάθε γωνία προστίθεται μια καινούρια γραμμή σε κάθε ένα από αυτούς του πίνακες. Για τις τρεις πρώτες περιπτώσεις, ο τελικός πίνακας με την αποτίμηση θα δημιουργηθεί στις γραμμές 193-256. Στις γραμμές 139-152 ορίζονται οι τιμές των G1 και G2 (βλέπε Σχέση 38 και πίνακα 1) που θα χρησιμοποιηθούν στη συνέχεια για να οριστεί ο πίνακας των συντελεστών. Στις γραμμές 155-187 πραγματοποιούνται τα αθροίσματα από τις από τις Σχέση 30, Σχέση 31, Σχέση 33 και Σχέση 34 όπως και στις γραμμές 86-136. Στις γραμμές 193-200 πραγματοποιούνται οι Σχέση 32 και Σχέση 35. Στις γραμμές 201-204 γίνεται άθροιση των τετραγώνων των Σχέση 32 και Σχέση 35 ώστε να μείνουν μόνο τα 'lv' και 're'. Στις γραμμές 205-252 γίνεται υπολογισμός της αντιληπτής γωνίας προέλευσης. Αυτό επιτυγχάνεται διαιρώντας τις σχέσεις των Σχέση 32 και Σχέση 35 μεταξύ τους στη κάθε μία, ώστε να προκύψει η εφαιπτομένη της γωνίας αντίληψης και στη συνέχεια παίρνω την αντίστροφη συνάρτηση της εφαιπτομένης για να βρω την γωνία. Επειδή όμως η γωνίες που θα προκύψουν είναι δύο με διαφορά 180°, γίνεται έλεγχος για το ποία είναι σωστή. Ένας άλλο πράγμα που πρέπει να αποφευχθεί είναι η διαίρεση με το μηδέν, όπου γίνεται έλεγχος στις γραμμές 209-211 (και αντίστοιχα σε επόμενες γραμμές για τις άλλες γωνίες) και υπολογίζεται αν η γωνία προέλευσης είναι 90° ή 270°. Στις γραμμές 253-256 δημιουργείται ο πίνακας αποτίμησης εφόσον έχει επιλεγεί ένας από τους τρεις πρώτους τρόπους υπολογισμού των συντελεστών (todoflag=1-3). Στις γραμμές 261-437 γίνεται η σύνθεση του πίνακα συντελεστών και της αποτίμησης για την τέταρτη περίπτωση. Αυτή περιγράφει μια γενικευμένη λύση που ο χρήστης πρέπει να επιλέξει κατάλληλο 'd' για τις ψηλές και χαμηλές συχνότητες ώστε το σύστημα να παίζει καλά σύμφωνα με τον Farina (βλέπε Σχέση 38). Το 'd' παίρνει τιμές στο διάστημα [1,2). Στην ουσία σε αυτό το κομμάτι του κώδικα γίνεται ένας έλεγχος για το ποια τιμή του 'd' είναι καλύτερη για τις ψηλές και χαμηλές συχνότητες του decoder. Στις γραμμές 261-376 γίνονται ουσιαστικά οι ίδιοι υπολογισμοί με τις προηγούμενες τρεις πρώτες περιπτώσεις (todoflag=1-3) μόνο που γίνονται για εύρος γωνιών 0°-359° με βήμα 1°. Στις γραμμές 379-429 γίνεται έλεγχος για το κατά πόσο η νέα τιμή του 'd' δημιουργεί αποκωδικοποιητή που παίζει καλύτερα στις χαμηλές και ψηλές συχνότητες. Η λογική είναι να ελεγχθεί αν ο πίνακας αποτίμησης για το τελευταίο 'd' κάθε χρονική στιγμή, παίζει καλύτερα σε περισσότερες γωνίες από ότι αυτός που έχει αποθηκευτεί ως καλύτερος. Στη περίπτωση που παίζει καλύτερα σε περισσότερες γωνίες, αποθηκεύεται αυτός ως καλύτερος. Στη

συνέχεια το 'd' παίρνει την επόμενη τιμή. Ο έλεγχος γίνεται για τις τιμές των ν_n , ν_e , αντιληπτή αξιμούθια γωνία για το ν_n και για το ν_e , ξεχωριστά για τις ψηλές και χαμηλές συχνότητες. Στις χαμηλές συχνότητες για να θεωρηθεί ότι ο κωδικοποιητής με το καινούριο 'd' παίζει καλύτερα από τον παλιό εξετάζεται πρώτα ποιος έχει ν_n μέγιστο σε περισσότερες γωνίες και δευτερευόντως η αξιμούθιες γωνίες για το ν_n και ν_e να είναι πιο κοντά στην πραγματική γωνία. Στην περίπτωση που βρεθούν δύο κωδικοποιητές με τον ίδιο αριθμό γωνιών που παίζει καλύτερα ο ένας από τον άλλο, γίνεται έλεγχος ποιανού το ν_e είναι μεγαλύτερο για περισσότερες πραγματικές γωνίες. Αντίστοιχα συμβαίνει με τις ψηλές συχνότητες απλώς αντικαθίστανται τα ν_n με ν_e , και το αντίστροφο στον έλεγχο. Στις γραμμές 431-436 δημιουργείται ο τελικός πίνακας αποτίμησης του αποκωδικοποιητή με την καλύτερη απόδοση και το εύρος γωνιών είναι 0° - 345° με βήμα 15° .



ΑΠΟΤΙΜΗΣΗ ΚΑΙ ΠΑΡΑΤΗΡΗΣΕΙΣ

Σε αυτό το μέρος θα γίνεται αξιολόγηση της απόδοσης των κωδικοποιητών, καθώς και παρατηρήσεις που αφορούν την συνολική δημιουργία και λειτουργία του συστήματος ambisonic. **Σημείωση: στους πίνακες αναγράφονται τιμές μέχρι τις 180° καθώς υπάρχει συμμετρία ως προς τον Y άξονα.**

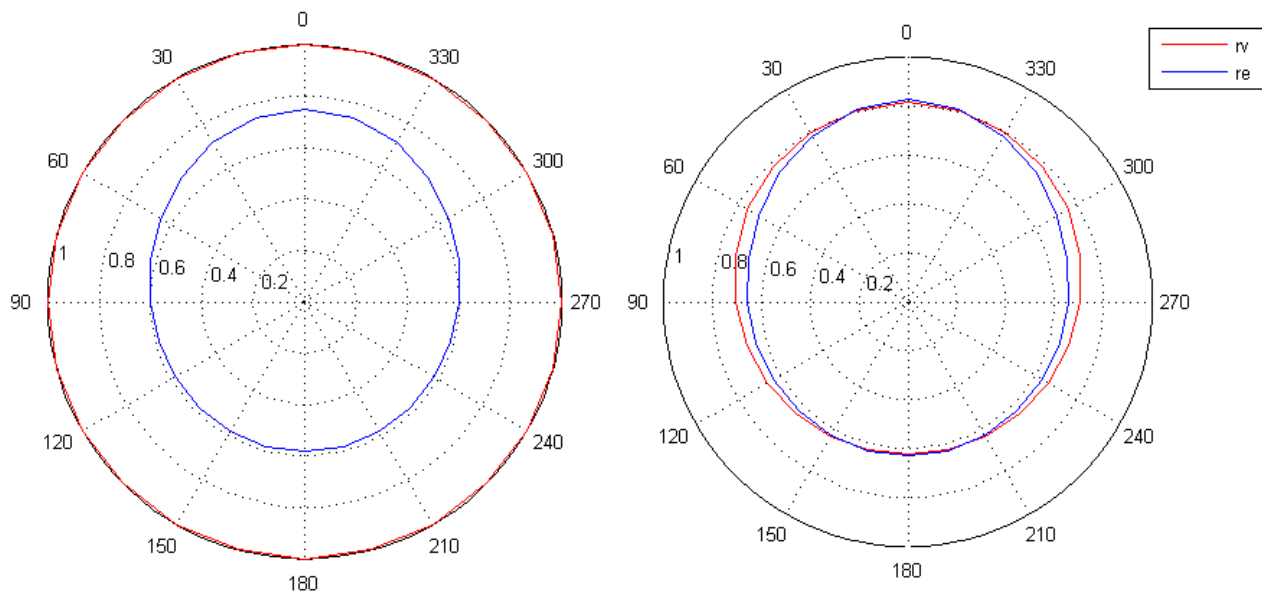
Στο πίνακα (3) φαίνεται η ψυχοακουστική αποτίμηση ενός αποκωδικοποιητή ambisonic που έχει σχεδιαστεί από τον Gerzon για αναπαραγωγή από 5 ηχεία στο οριζόντιο επίπεδο, με αζιμουθιες γωνίες αυτών (0,45,130,-130,45) μοίρες αντίστοιχα. Στην εικόνα 38 φαίνονται τα

πολικά διαγράμματα των 'rv' και 're' για της χαμηλές και ψηλές συχνότητες (αριστερά και δεξιά αντίστοιχα).

Πραγματική γωνία	rv για τις ψηλές συχνότητες	re για τις ψηλές συχνότητες	Αντιληπτή γωνία στις υψηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις ψηλές συχνότητες σε σχέση με το re	rv για τις χαμηλές συχνότητες	re για τις χαμηλές συχνότητες	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το re
0,0000	0,8158	0,8273	0,0000	0,0000	1,0001	0,7494	0,0000	0,0000
15,0000	0,8116	0,8148	15,0000	14,9990	1,0000	0,7396	15,0000	15,0010
30,0000	0,7994	0,7828	30,0010	30,0000	1,0000	0,7139	30,0000	30,0000
45,0000	0,7807	0,7431	45,0010	45,0010	1,0000	0,6808	45,0000	44,9990
60,0000	0,7576	0,7059	60,0010	60,0020	1,0000	0,6482	60,0000	59,9970
75,0000	0,7324	0,6761	75,0010	75,0020	1,0000	0,6212	75,0000	74,9960
90,0000	0,7071	0,6550	90,0000	90,0020	1,0000	0,6017	90,0000	89,9960
105,0000	0,6835	0,6416	105,0000	105,0000	1,0000	0,5896	105,0000	105,0000
120,0000	0,6629	0,6341	120,0000	120,0000	1,0000	0,5834	120,0000	120,0000
135,0000	0,6462	0,6306	135,0000	135,0000	1,0000	0,5814	135,0000	135,0000
150,0000	0,6339	0,6295	150,0000	150,0000	1,0000	0,5817	150,0000	150,0000
165,0000	0,6265	0,6294	165,0000	165,0000	1,0000	0,5827	165,0000	165,0000
180,0000	0,6240	0,6294	180,0000	180,0000	1,0000	0,5832	180,0000	180,0000

Πίνακας 3: Απεικόνιση των τιμών των 're' και 'rv' για τις υψηλές και χαμηλές συχνότητες καθώς και οι αντιληπτές γωνίες όσων αφορά τους δύο διαφορετικούς μηχανισμούς αντίληψης της προέλευσης του ήχου. Οι πίνακας αφορά το Ambisonic σύστημα που σχεδίασε ο Gerzon για 5 ηχεία στο οριζόντιο επίπεδο, με γωνίες (0,45,130,-130,-45).

Κατά την θεωρητική ψυχοακουστική αποτίμηση αυτού του κωδικοποιητή (σύμφωνα με την θεωρία του Gerzon που περιγράφεται στο [25]) για τις χαμηλές συχνότητες κάτω από τα 400 Hz αυτός ο αποκωδικοποιητής έχει σωστή αντιστοιχία ανάμεσα στις πραγματικές και αναπαραγόμενες γωνίες προέλευσης του ήχου, και παρουσιάζει μεγάλη σταθερότητα σε αυτές κατά την κίνηση του κεφαλιού του ακροατή. Όπως φαίνεται από το πολικό διάγραμμα 38 για αυτές τις συχνότητες υπάρχει μεγαλύτερη σταθερότητα στις γωνίες από το μπροστινό επίπεδο, και αυτό οφείλεται στο ότι η τιμή του 're' σε αυτή την περιοχή είναι μεγαλύτερο, ενώ το 'rv' έχει την τιμή 1. Στις ψηλές συχνότητες έχει μεγαλύτερη σταθερότητα στο μπροστινό επίπεδο καθώς η τιμή των 're' είναι μεγαλύτερη. Ταυτόχρονα υπάρχει ακρίβεια στην αναπαραγωγή σε όλες τις αζιμουθίες γωνίες. Ο συγκεκριμένος, καθώς η διάταξη των ηχείων δεν παρουσιάζει τόση συμμετρία (σε σχέση πχ με το κανονικό n=γωνο), δεν έχει την ίδια απόκριση καθώς αλλάζει η γωνία της συντεθειμένης πηγής.



Εικόνα 38: Πολικά διαγράμματα των 're' και 'rv' για τις υψηλές (δεξιά) και χαμηλές (αριστερά) συχνότητες του συστήματος Ambisonics που σχεδίασε ο Gerzon για 5 ηχεία στο οριζόντιο επίπεδο, με γωνίες (0,45,130,-130,-45). Το αριστερό πολικό διάγραμμα αφορά τις χαμηλές συχνότητες ενώ το δεξί τις ψηλές.

Αντίστοιχα στον πίνακα (4),(5) φαίνεται η ψυχοακουστική αποτίμηση ενός αποκωδικοποιητή Ambisonic σύμφωνα με τον D. Malham για αναπαραγωγή σε συναυλιακούς χώρους από διάταξη ηχείων σε κανονικό n-γωνο (n=4, n=5 αντίστοιχα) στο οριζόντιο επίπεδο.

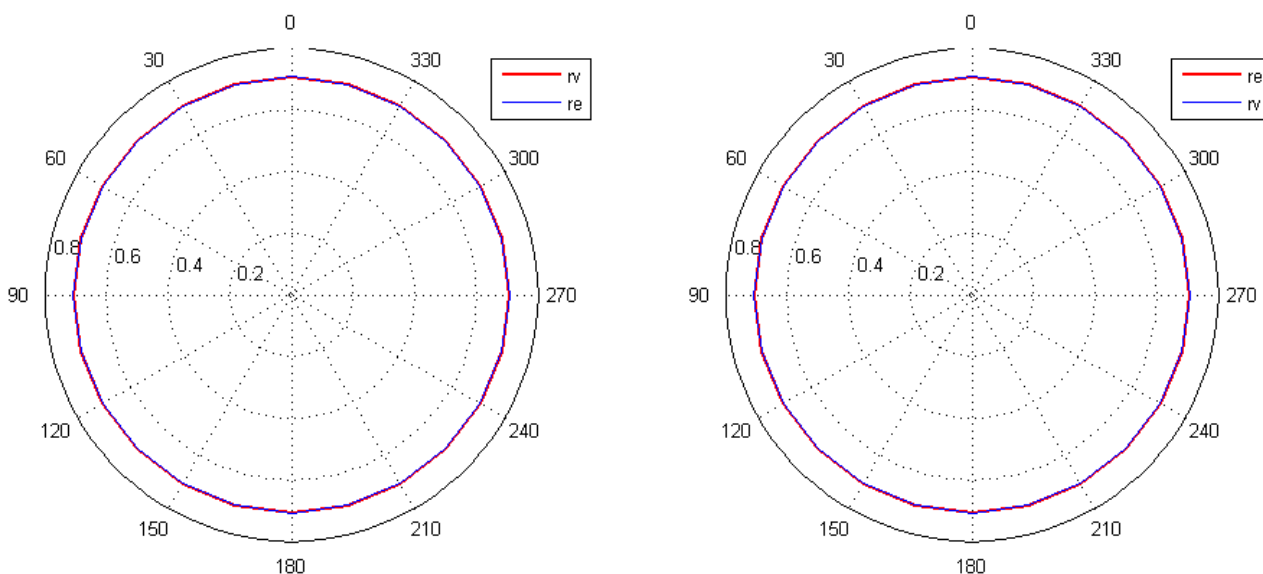
Πραγματική γωνία	rv για τις ψηλές συχνότητες	re για τις ψηλές συχνότητες	Αντιληπτή γωνία στις υψηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις ψηλές συχνότητες σε σχέση με το re	rv για τις χαμηλές συχνότητες	re για τις χαμηλές συχνότητες	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το re
0,0000	0,7071	0,7071	0,0000	0,0000	0,7071	0,7071	0,0000	0,0000
15,0000	0,7071	0,7071	15,0000	15,0000	0,7071	0,7071	15,0000	15,0000
30,0000	0,7071	0,7071	30,0000	30,0000	0,7071	0,7071	30,0000	30,0000
45,0000	0,7071	0,7071	45,0000	45,0000	0,7071	0,7071	45,0000	45,0000
60,0000	0,7071	0,7071	60,0000	60,0000	0,7071	0,7071	60,0000	60,0000
75,0000	0,7071	0,7071	75,0000	75,0000	0,7071	0,7071	75,0000	75,0000
90,0000	0,7071	0,7071	90,0000	90,0000	0,7071	0,7071	90,0000	90,0000
105,0000	0,7071	0,7071	105,0000	105,0000	0,7071	0,7071	105,0000	105,0000
120,0000	0,7071	0,7071	120,0000	120,0000	0,7071	0,7071	120,0000	120,0000
135,0000	0,7071	0,7071	135,0000	135,0000	0,7071	0,7071	135,0000	135,0000
150,0000	0,7071	0,7071	150,0000	150,0000	0,7071	0,7071	150,0000	150,0000
165,0000	0,7071	0,7071	165,0000	165,0000	0,7071	0,7071	165,0000	165,0000
180,0000	0,7071	0,7071	180,0000	180,0000	0,7071	0,7071	180,0000	180,0000

Πίνακας 4: Απεικόνιση των τιμών των rv και re, καθώς και των αντιληπτών γωνιών προέλευσης του ήχου για έναν αποκωδικοποιητή Ambisonic από τον D. Malham για αναπαραγωγή στο οριζόντιο επίπεδο σε συναυλιακούς χώρους. Ο πίνακας αναφέρεται για αναπαραγωγή από 4 ηχεία.

Πραγματική γωνία	rv για τις ψηλές συχνότητες	re για τις ψηλές συχνοτήτες	Αντιληπτή γωνία στις υψηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις ψηλές συχνότητες σε σχέση με το re	rv για τις χαμηλές συχνότητες	re για τις χαμηλές συχνότητες	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το re
0,0000	0,7071	0,7071	0,0000	0,0000	0,7071	0,7071	0,0000	0,0000
15,0000	0,7071	0,7071	15,0000	15,0000	0,7071	0,7071	15,0000	15,0000
30,0000	0,7071	0,7071	30,0000	30,0000	0,7071	0,7071	30,0000	30,0000
45,0000	0,7071	0,7071	45,0000	45,0000	0,7071	0,7071	45,0000	45,0000
60,0000	0,7071	0,7071	60,0000	60,0000	0,7071	0,7071	60,0000	60,0000
75,0000	0,7071	0,7071	75,0000	75,0000	0,7071	0,7071	75,0000	75,0000
90,0000	0,7071	0,7071	90,0000	90,0000	0,7071	0,7071	90,0000	90,0000
105,0000	0,7071	0,7071	105,0000	105,0000	0,7071	0,7071	105,0000	105,0000
120,0000	0,7071	0,7071	120,0000	120,0000	0,7071	0,7071	120,0000	120,0000
135,0000	0,7071	0,7071	135,0000	135,0000	0,7071	0,7071	135,0000	135,0000
150,0000	0,7071	0,7071	150,0000	150,0000	0,7071	0,7071	150,0000	150,0000
165,0000	0,7071	0,7071	165,0000	165,0000	0,7071	0,7071	165,0000	165,0000
180,0000	0,7071	0,7071	180,0000	180,0000	0,7071	0,7071	180,0000	180,0000

Πίνακας 5: Σε αυτό τον πίνακα φαίνονται οι τιμές των rv και re, καθώς και των αντιληπτών γωνιών προέλευσης του ήχου για έναν αποκωδικοποιητή Ambisonic από τον D. Malham για αναπαραγωγή στο οριζόντιο επίπεδο σε συναυλιακούς χώρους. Ο πίνακας αναφέρεται για αναπαραγωγή από 5 ηχεία.

Παρατηρήθηκε ότι ανεξάρτητα από τον αριθμό των ηχείων, οι πίνακες αποτίμησης παραμένουν οι ίδιοι. Αναμενόμενο λοιπόν είναι και τα πολικά τους διαγράμματα για τα re και rv να παραμένουν τα ίδια. Αυτά φαίνονται στην εικόνα 39, στις χαμηλές (αριστερά) και ψηλές (δεξιά) συχνότητες.



Εικόνα 39: Πολικά διαγράμματα των 'rv' και 're' για τον αποκωδικοποιητή του D. Malham για αναπαραγωγή σε συναυλιακούς χώρους από διάταξη ηχείων τοποθετημένα σε κανονικό n-γωνο στο οριζόντιο επίπεδο. Το αριστερό πολικό διάγραμμα αφορά τις χαμηλές συχνότητες ενώ το δεξί τις ψηλές.

Ο συγκεκριμένος αποκωδικοποιητής έχει καλή αντιστοιχία ανάμεσα στις γωνίες αναπαραγωγής και τις πραγματικές γωνίες. Η απόδοσή του είναι καλή για αναπαραγωγή σε μεγαλύτερο πεδίο ακρόασης, καθώς δεν παράγεται ήχος με αντίθετη φάση από τα ηχεία που βρίσκονται σε αντίθετη διεύθυνση από αυτή της προέλευσης του πραγματικού ήχου. Σύμφωνα με τις προδιαγραφές που αναφέρει ο Gerzon στο [25], ο συγκεκριμένος δεν μπορεί να θεωρηθεί σαν

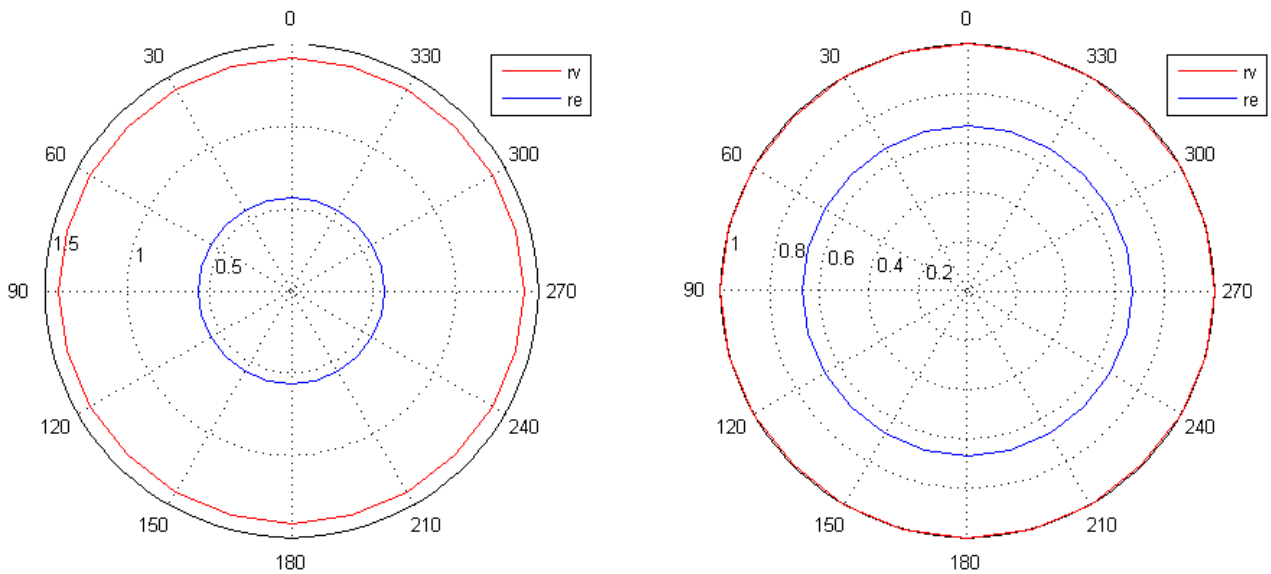
αποκωδικοποιητής Ambisonic καθώς δεν συμπληρώνει τις απαιτούμενες προϋποθέσεις, όπως για τις χαμηλές συχνότητες να ισχύει ότι $rv=1$. Ο συμβιβασμός όμως που έχει κάνει, πιστεύω πως είναι καλώς για αναπαραγωγή σε σφαιρικούς χώρους, όπου το πεδίο ακρόασης είναι ευρύτερο.

Στο πίνακα (6) φαίνεται η ψυχοακουστική αποτίμηση ενός αποκωδικοποιητή Ambisonic σύμφωνα με τον Gerzon για αναπαραγωγή σε σφαιρικούς χώρους από διάταξη ηχείων σε κανονικό n -γωνο στο οριζόντιο επίπεδο.

Πραγματική γωνία	rv για τις υψηλές συχνότητες	re για τις υψηλές συχνότητες	Αντιληπτή γωνία στις υψηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις υψηλές συχνότητες σε σχέση με το re	rv για τις χαμηλές συχνότητες	re για τις χαμηλές συχνότητες	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το re
0,0000	1,0000	0,6667	0,0000	0,0000	1,4142	0,5657	0,0000	0,0000
15,0000	1,0000	0,6667	15,0000	15,0000	1,4142	0,5657	15,0000	15,0000
30,0000	1,0000	0,6667	30,0000	30,0000	1,4142	0,5657	30,0000	30,0000
45,0000	1,0000	0,6667	45,0000	45,0000	1,4142	0,5657	45,0000	45,0000
60,0000	1,0000	0,6667	60,0000	60,0000	1,4142	0,5657	60,0000	60,0000
75,0000	1,0000	0,6667	75,0000	75,0000	1,4142	0,5657	75,0000	75,0000
90,0000	1,0000	0,6667	90,0000	90,0000	1,4142	0,5657	90,0000	90,0000
105,0000	1,0000	0,6667	105,0000	105,0000	1,4142	0,5657	105,0000	105,0000
120,0000	1,0000	0,6667	120,0000	120,0000	1,4142	0,5657	120,0000	120,0000
135,0000	1,0000	0,6667	135,0000	135,0000	1,4142	0,5657	135,0000	135,0000
150,0000	1,0000	0,6667	150,0000	150,0000	1,4142	0,5657	150,0000	150,0000
165,0000	1,0000	0,6667	165,0000	165,0000	1,4142	0,5657	165,0000	165,0000
180,0000	1,0000	0,6667	180,0000	180,0000	1,4142	0,5657	180,0000	180,0000

Πίνακας 6: Απεικόνιση των τιμών των rv και re , καθώς και των αντιληπτών γωνιών προέλευσης του ήχου για έναν αποκωδικοποιητή Ambisonic από τον Gerzon για αναπαραγωγή στο οριζόντιο επίπεδο σε σφαιρικούς χώρους.

Σε αυτή την περίπτωση ξαναπατηρήθηκε ότι ανεξάρτητα από τον αριθμό των ηχείων, οι πίνακες αποτίμησης παραμένουν οι ίδιοι. Στην εικόνα 40 φαίνονται τα πολικά διαγράμματα των re και rv του αποκωδικοποιητή που σχεδίασε ο Gerzon για αναπαραγωγή από διάταξη ηχείων σε κανονικό n -γωνο στο οριζόντιο επίπεδο.



Εικόνα 40: Πολικά διαγράμματα των 'rv' και 're' για των αποκωδικοποιητή του M. Gerzon για αναπαραγωγή σε σπουντιακούς χώρους από διάταξη ηχείων τοποθετημένα σε κανονικό n-γωνο στο οριζόντιο επίπεδο. Το αριστερό πολικό διάγραμμα αφορά τις χαμηλές συχνότητες ενώ το δεξί τις ψηλές.

Αυτός ο αποκωδικοποιητής παρουσιάζει περίεργη συμπεριφορά, καθώς το 'rv' παίρνει την τιμή 1 για την περίπτωση των υψηλών συχνοτήτων και μεγαλύτερο re για την περίπτωση των χαμηλών. Πιστεύω πως πρέπει να έχει γίνει κάποιο λάθος στη βιβλιογραφία.

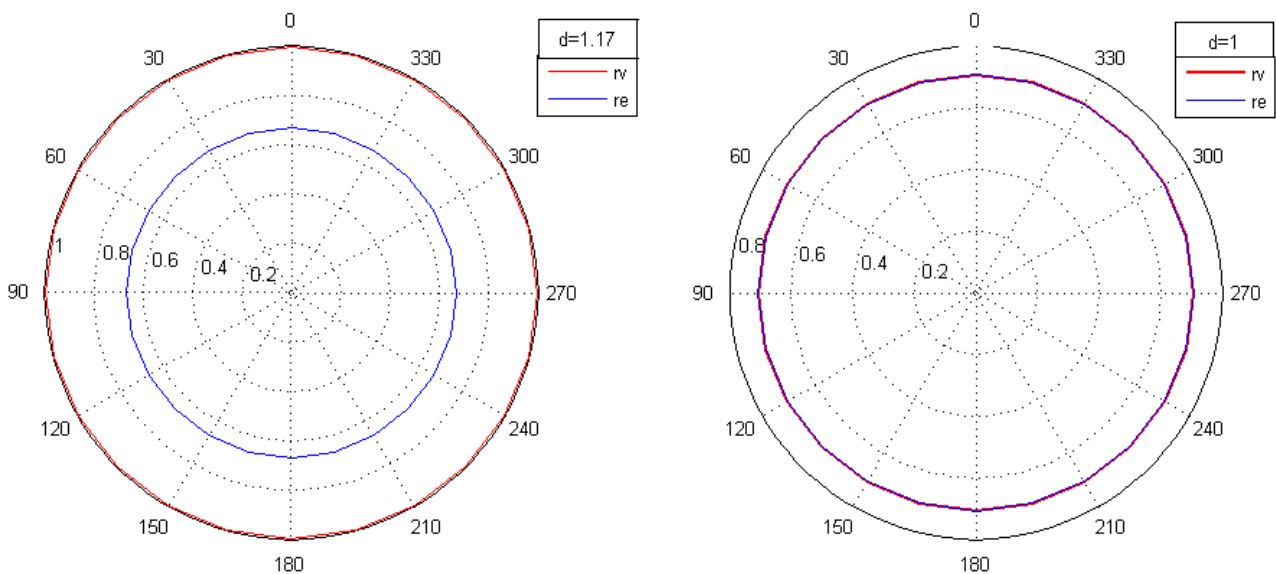
Πραγματική γωνία	rv για τις ψηλές συχνότητες	re για τις ψηλές συχνότητες	Αντιληπτή γωνία στις υψηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις ψηλές συχνότητες σε σχέση με το re	rv για τις χαμηλές συχνότητες	re για τις χαμηλές συχνότητες	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το rv	Αντιληπτή γωνία στις χαμηλές συχνότητες σε σχέση με το re
0,0000	0,7071	0,7071	0,0000	0,0000	0,9968	0,6674	0,0000	0,0000
15,0000	0,7071	0,7071	15,0000	15,0000	0,9968	0,6674	15,0000	15,0000
30,0000	0,7071	0,7071	30,0000	30,0000	0,9968	0,6674	30,0000	30,0000
45,0000	0,7071	0,7071	45,0000	45,0000	0,9968	0,6674	45,0000	45,0000
60,0000	0,7071	0,7071	60,0000	60,0000	0,9968	0,6674	60,0000	60,0000
75,0000	0,7071	0,7071	75,0000	75,0000	0,9968	0,6674	75,0000	75,0000
90,0000	0,7071	0,7071	90,0000	90,0000	0,9968	0,6674	90,0000	90,0000
105,0000	0,7071	0,7071	105,0000	105,0000	0,9968	0,6674	105,0000	105,0000
120,0000	0,7071	0,7071	120,0000	120,0000	0,9968	0,6674	120,0000	120,0000
135,0000	0,7071	0,7071	135,0000	135,0000	0,9968	0,6674	135,0000	135,0000
150,0000	0,7071	0,7071	150,0000	150,0000	0,9968	0,6674	150,0000	150,0000
165,0000	0,7071	0,7071	165,0000	165,0000	0,9968	0,6674	165,0000	165,0000
180,0000	0,7071	0,7071	180,0000	180,0000	0,9968	0,6674	180,0000	180,0000

Πίνακας 7: Απεικόνιση των τιμών των rv και re, καθώς και των αντιληπτών γωνιών προέλευσης του ήχου για έναν αποκωδικοποιητή Ambisonic σχεδιασμένο βάση των τύπων του Farina για αναπαραγωγή στο οριζόντιο επίπεδο σε οικιακούς χώρους. Η επιλογή του d έγινε από το αλγόριθμο υπολογισμού των συντελεστών για τα σήματα των ηχείων που σχεδιάστηκε για αυτή την εργασία. Το d για τις χαμηλές συχνότητες έχει την τιμή '1.17' ενώ για τις ψηλές '1'.

Στο πίνακα (7) φαίνεται η ψυχοακουστική αποτίμηση ενός αποκωδικοποιητή Ambisonic σύμφωνα με τον Farina για αναπαραγωγή σε οικιακούς χώρους από διάταξη ηχείων σε κανονικό n-γωνο στο οριζόντιο επίπεδο. Στον συγκεκριμένο, η επιλογή των 'd' για βέλτιστη λειτουργία στις ψηλές και χαμηλές συχνότητες επιλέχθηκε βάση διερεύνησης της βέλτιστης λειτουργίας στο πεδίο ορισμού

του d (d ανοίγει στο $[1,2)$), με ακρίβεια δεύτερου δεκαδικού, από τον αλγόριθμο υπολογισμού των συντελεστών για τα σήματα των ηχείων που δημιουργήθηκε για αυτή την εργασία. Βρέθηκε ότι ανεξάρτητα από τον αριθμό των ηχείων, για βέλτιστη αναπαραγωγή στις χαμηλές συχνότητες, το d πρέπει να ισούται με '1.17', ενώ για τις ψηλές συχνότητες με το '1'.

Τέλος, και σε αυτή την περίπτωση παρατηρήθηκε ότι ανεξάρτητα από τον αριθμό των ηχείων, οι πίνακες αποτίμησης παραμένουν οι ίδιοι. Στην εικόνα 41 φαίνονται τα πολικά διαγράμματα των 're' και 'rv' του αποκωδικοποιητή που σχεδιάστηκε βάση της θεωρίας του Farina για αναπαραγωγή από διάταξη ηχείων σε κανονικό n -γωνο στο οριζόντιο επίπεδο.



Εικόνα 41: Πολικά διαγράμματα των 'rv' και 're' για των αποκωδικοποιητή που σχεδιάστηκε βάση της θεωρίας του Farina για αναπαραγωγή σε οικιακούς χώρους από διάταξη ηχείων τοποθετημένα σε κανονικό n -γωνο στο οριζόντιο επίπεδο. Το αριστερό πολικό διάγραμμα αφορά τις χαμηλές συχνότητες ενώ το δεξί τις ψηλές.

Ο συγκεκριμένος αποκωδικοποιητής παρουσιάζει καλή αντιστοιχία στις πραγματικές και τις αντιληπτές γωνίες προέλευσης του ήχου για τις χαμηλές και ψηλές συχνότητες. Υπάρχει ίδια ποιότητα αναπαραγωγής ανεξάρτητα από την πραγματική γωνία προέλευσης, και σταθερότητα αυτής κατά την κίνηση του κεφαλιού του ακροατή. Ο συγκεκριμένος είναι σχεδιασμένος για να πληρεί όσο το δυνατό περισσότερες από τις προϋποθέσεις που έχει θέσει ο Gerzon.



ΣΥΜΠΕΡΑΣΜΑΤΑ

Σε αυτή την εργασία ασχοληθήκαμε με την ανασύνθεση εικονικών και πραγματικών ηχητικών πεδίων με την μέθοδο Ambisonics. Πράγματι αποτελεί ένα από τα πιο ακριβή συστήματα αναπαραγωγής ηχητικών πεδίων. Μπορεί να κάνει πιστή αναπαράσταση του ηχητικού πεδίου ανεξάρτητα από τον αριθμό των ηχείων που πρόκειται να χρησιμοποιηθούν. Το γεγονός επίσης ότι η αναπαραγωγή και αποτίμηση έγινε μέσω του συστήματος Binaural δίνει την δυνατότητα της αναπαραγωγής των ηχητικών πεδίων που είναι ηχογραφημένα με την μέθοδο Ambisonics, σε συστήματα Binaural ή Transaural. Ταυτόχρονα, η μέθοδος που ακολουθήθηκε για την αποτίμηση του συστήματος μπορεί να αποτελέσει βάση ποιοτικού ελέγχου για όλα τα συστήματα αναπαραγωγής τρισδιάστατου ή Surround ήχου, καθώς και μέθοδο τροποποίησης των κωδικοποιημένων ηχητικών πεδίων ώστε να μπορούν να αναπαραχθούν μέσω των συστημάτων Binaural και Transaural. Η υλοποίηση αυτής της εργασίας επιτεύχθηκε με την μελέτη της

θεωρητικής τεκμηρίωσης και φυσικής ερμηνείας της μεθόδου Ambisonics, αλλά και την μελέτη συγγραμμάτων ανθρώπων που έχουν δημιουργήσει τέτοια συστήματα αναπαραγωγής. Η υλοποίηση έγινε σε περιβάλλον Matlab. Η αποτίμηση του συστήματος έγινε με χρήση της μεθόδου Binaural, η οποία αποτελεί την πιο ακριβή μέθοδο αναπαραγωγής τρισδιάστατων ηχητικών πεδίων. Με την ίδια μεθοδολογία μπορούν να αποτιμηθούν και συστήματα Ambisonic μεγαλύτερης τάξης. Ένα από τα πιο βασικά συμπεράσματα που απορρέει από αυτή την εργασία είναι ότι για να επιτευχθεί βέλτιστη αναπαραγωγή στο μεγαλύτερο εύρος γωνιών στο οριζόντιο επίπεδο, θα πρέπει η διάταξη των ηχείων να βρίσκεται σε κανονικό N-γωνο. Σε αντίθετη περίπτωση θα υπάρχει βέλτιστη αναπαραγωγή σε συγκεκριμένες γωνίες στο οριζόντιο επίπεδο. Ένα δεύτερο συμπέρασμα είναι ότι για την βέλτιστη αναπαραγωγή σε συγκεκριμένες γωνίες πρέπει ο σχεδιαστής του συστήματος να έχει μεγάλη εμπειρία και πρέπει να κάνει αρκετές δοκιμές για να την δημιουργία κατάλληλου αποκωδικοποιητή που θα πληρεί της προϋποθέσεις.

Βιβλιογραφία

- [1] Bruce Wiggins, "An investigation into the real-time manipulation and control of three-dimensional sound field", University of Derby,2004
- [2] Ville Pulkki, "Spatial sound generation and perception by amplitude panning techniques", Helsinki University of Technology, Laboratory of acoustics and Audio Signal Processing,2001
- [3] P. Schillebeeckx, i. Paterson-Stephens, B. Wiggins, "Using Matlab/simulink as an implementation tool for Multi-Channel Surround Sound", Signal processing Applications Research Group, University of Derby,
- [4] Bruce Wiggins, I. Paterson-Stephens, Val Lowndes, S. Berry, "The design and optimization for surround sound decoders using heuristic methods", University of Derby,
- [5] A. Farina, E. Ugolotti, "Software Implementation of B-Format Encoding and Decoding", 104th AES Convention, Amsterdam, Preprint 4691,1998
- [6] William G.Gardner, "3-d audio using loudspeakers", Kluwer Academic Publishers,Boston,Dordrecht,London,1998
- [7] Durand R.Begault, "3D Sound for virtual reality and multimedia", A.P Professional,copuright by Academic Press Inc.,1994
- [8] Ken Steiglitz, "A Digital Signal Processing Primer", Addison-Wesley Publishing Company Inc.,1996
- [9] Monson H. Hayes, "Digital Signal Processing", Schaum's OUTlines,1999
- [10] Robert A.Humphrey, "Automatic Loudspeaker Location Detection for use in Ambisonic Systems", University of York,1996
- [11] Karin Carlsson, "Objective Localization in Ambisonic Surround-sound", Department of Speech, Music and Hearing, Royal Institute of Technology,
- [12] Mathieu Guillaume, Yves Grenier, "Theoretical Sound Field Analysis", Telecom Paris, Departement Traitement du Signal et de Image, Groupe Audio, Acoustic et Ondes,2006
- [13] Hugh Robjohns, "A brief History of microphones", Microphone Data Book,2001
- [14] Aristotel Digenis, "Towards An Ambisonic Audio Unit Plug-in Suite", University of York,2004
- [15] Jerome Daniel, Rozenn Nicol, Sebastien Moreau, "Further Investigation of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging", Audio Engineer Society, Convention Paper 5788,2003
- [16] Diemer de Vries, Marinus M.Boone, "Wave Field Synthesis and Analysis using Array Technology", Laboratory of Acoustical Imaging and Sound Control, Delft University of Technology ,1999
- [17] Francis Rumsey,Tim McCormick, "Sound And Recording An Introduction", Focal Press,Oxford,1997
- [18] Jens Blauert, "Spatial Hearing (The Psychophysics of Human Sound Localization)", MIT press, Cambridge, London, England,1997
- [19] Edward J. Rickley, Gregg G.Fleming, John A.Volpe, "Computing the Absorption of Sound by the Atmosphere and its Applicability to Aircraft Noise Certification",

- Office of Environment and Energy, Washington, D.C. 2059.,1998
- [20] James A.Moorer, "Extension of the method of Spacial Harmonics to three Dimensions", ,2000-2001
- [21] Α.Χ. Φλώρος, Ι.Ν. Μουρτζόπουλος, "Ψηφιακή Τεχνολογία του Ήχου", Τμήμα Τεχνολόγων Μηχανικών και Τεχνολογίας Υπολογιστών του Πανεπιστημίου Πατρών,2005
- [22] Α. Φλώρος, "Ψηφιακή Τεχνολογία Ήχου, Στοιχεία κβαντισμού ηχητικών σημάτων", ,
- [23] R.A. Wannamaker, "The Theory of Dithered Quantization", University of Waterloo,2003
- [24] R.A. Wannamaker, S.P. Lipshitz, J. Vanderkooy, J.N. Write, "A Theory of Non-Subtractive Dither", University of Waterloo,
- [25] M.A. Gerzon, G.J. Barton, "Ambisonic Decoders for HDTV", AES,
- [26] M.A. Gerzon, "General Metatheory of Auditory Localization", AES,
- [27] B. Wiggins, I. Paterson-Stephens, P.Schillebeeckx, "The analysis of multi-channel sound reproduction algorithms using HRTF data", University of Derby,
- [28] F. Rumsey, "Spatial Audio", Focal Press,2001
- [29] P.M. Morse, K.U. Ingard, "Theoretical Acoustics", Princeton University Press, 1986

Ευρετήριο λέξεων

3	M
3D Sound Systems.....26	Makita localization azimuth.....62
A	N
A-Format.....53	Noise-Shaping.....72
Associated Legendre function.....51	P
B	pantophonic system.....25, 47
B-Format.....55	periphonic system.....25, 47
C	Precedence.....17
C-Format.....55	S
D	Spatial Alising Frequency.....44
D-Format.....55	U
dither.....70	UHQ-Format.....55
E	V
Energy localization Vector.....64	velocity localization vector.....64
Energy Localization vector.....63	Velocity localization vector.....62
Energy vector localization azimuth.....63	velocity vector localization azimuth.....62
G	I
G-Format.....56	Ιεραρχικά.....47
G+2-Format.....56	μ
H	μέθοδο.....14
Hass effect.....17	π
hierarchical.....25, 47	παντοφωνικό.....25, 47
HRTF.....19	περιφωνικό.....25, 47
I	σ
Inside the Head Locatedness.....39	σύστημα.....14
Interaural Intensity Differect (IID).....15	σφαιρικές αρμονικές.....51
Interaural Time Difference (ITD).....14	τ
L	τεχνικές μορφοποίησης του θορύβου.....72
Laplacian operator.....48	

Συντομογραφίες.

ITU: International Telecommunication Union

IIR: Infinite Impulse Response (feedback filter)

FIR: Finite Impulse Response (feedforward filter)

DFT: Discreet Fourier Transform

GUI: Graphical User Interface

WFS: Wave Field Synthesis

WFA: Wave Field Analysis

IID: Interaural Intensity Difference

ITD: Interaural Time Difference

HRTF: Head Related Transfer Function

IHL: Inside the Head Locatedness

PDF: Probability Density Function

RPDF: Rectangular Probability Density Function

TPDF: Triangular Probability Density Function




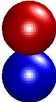
SNR: Signal to Noise Ratio







LSB: Less Significant Bit







ΠΑΡΑΡΤΗΜΑ.

Παράρτημα 1 Πίνακας 8. Σφαιρικές αρμονικές μέχρι την 3η τάξη.

Σε αυτό τον πίνακα φαίνονται οι σφαιρικές αρμονικές μέχρι 3^{ης} τάξης με τους τύπους και τα πολικά διαγράμματά τους. Με το κόκκινο χρώμα συμβολίζονται οι γωνίες που λαμβάνει το σήμα με κανονική φάση, ενώ με μπλε όταν λαμβάνει με ανεστραμμένη.

Τάξη σφαιρικών αρμονικών	m,n,σ	Ονομασία καναλιού	Συνάρτηση υλοποίησης	Πολικό διάγραμμα
0	0,0,1	W	1	
1	1,1,1	X	$\cos(\theta)\cos(\varphi)$	
1	1,1,-1	Y	$\sin(\theta)\cos(\varphi)$	
1	1,0,1	Z	$\sin(\varphi)$	

Τάξη σφαιρικών αρμονικών	m,n,σ	Ονομασία καναλιού	Συνάρτηση υλοποίησης	Πολικό διάγραμμα
2	2,0,1	R	$(3\sin^2(\varphi)-1)/2$	
2	2,1,1	S	$(3^{0.5}/2)\cos(\theta)\sin(2\varphi)$	
2	2,1,-1	T	$(3^{0.5}/2)\sin(\theta)\sin(2\varphi)$	
2	2,2,1	U	$(3^{0.5}/2)\cos(2\theta)\cos^2(\varphi)$	
2	2,2,-1	V	$(3^{0.5}/2)\sin(2\theta)\cos^2(\varphi)$	
3	3,0,1	K	$\sin(\varphi)(5\sin^2(\varphi)-2)/2$	

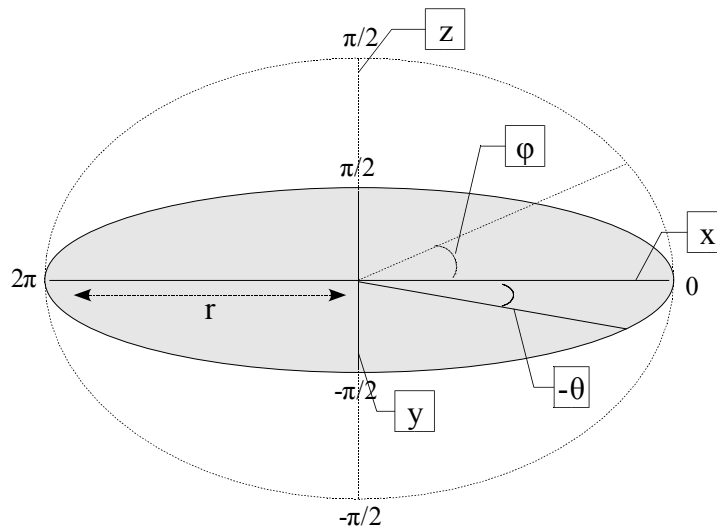
Τάξη σφαιρικών αρμονικών	m,n,σ	Ονομασία καναλιού	Συνάρτηση υλοποίησης	Πολικό διάγραμμα
3	3,1,1	L	$(3/8)0.5\cos(\theta)\cos(\varphi)(5\sin^2(\varphi)-1)$	
3	3,1,-1	M	$(3/8)0.5\sin(\theta)\cos(\varphi)(5\sin^2(\varphi)-1)$	
3	3,2,1	N	$(15^{0.5}/2)\cos(2\theta)\sin(\varphi)\cos^2(\varphi)$	
3	3,2,-1	O	$(15^{0.5}/2)\sin(2\theta)\sin(\varphi)\cos^2(\varphi)$	
3	3,3,1	P	$(5/8)^{0.5}\cos(3\theta)\cos^3(\varphi)$	
3	3,3,-1	Q	$(5/8)^{0.5}\sin(3\theta)\cos^3(\varphi)$	

Παράρτημα 2 Ορισμός και σύμβολα για σφαιρικές συντεταγμένες.

Μια θέση στο χώρο μπορεί να προσδιοριστεί μέσω των σφαιρικών συντεταγμένων με τρία διαφορετικά στοιχεία, την αζιμούθια γωνία, την κατακόρυφη γωνία και την απόσταση από την αρχή των αξόνων. Την αζιμούθια γωνία θα την συμβολίζω με 'θ', την κατακόρυφη με 'φ' και την απόσταση με 'r'. Οι τιμές που μπορούν να πάρουν οι γωνίες είναι:

$$0 \leq \theta \leq 2\pi \text{ και } -\frac{\pi}{2} \leq \phi \leq \frac{\pi}{2}$$

Σχέση 46:



Σχέδιο 1: Απεικόνιση του συστήματος συντεταγμένων με καρτεσιανές και σφαιρικές συντεταγμένες. Σημείωση: Οι γωνίες μετρώνται αντίθετα με την φορά των δεικτών του ρολογιού.

Παράρτημα 3 Η εξίσωση που δίνει το πολικό διάγραμμα οποιαδήποτε μικροφώνου στον χώρο.

Οποιοδήποτε πολικό διάγραμμα μικροφώνου που λαμβάνει στο κατακόρυφο επίπεδο με τον ίδιο τρόπο που λαμβάνει και στο οριζόντιο, μπορεί να δημιουργηθεί από την παρακάτω εξίσωση:

$$M_{pollar}(\theta, \phi) = C + (1 - C)[\cos(\theta) \cos(\Phi) \sin(\theta) \sin(\theta) + \cos(\theta) \cos(\Phi) \cos(\theta) \cos(\theta) + \sin(\theta) \sin(\Phi)]$$

Σχέση 47:

όπου :

- Θ: η αξιμούθια γωνία που σχηματίζει το μικρόφωνο αναφορικά με τις πολικές συντεταγμένες.
- Φ: η κατακόρυφη γωνία που σχηματίζει το μικρόφωνο αναφορικά με τις πολικές συντεταγμένες.
- θ: η αξιμούθια γωνία πρόσπτωσης του ήχου αναφορικά με τις πολικές συντεταγμένες.
- φ: η κατακόρυφη γωνία πρόσπτωσης του ήχου αναφορικά με τις πολικές συντεταγμένες.

Για την περίπτωση των καρδιοειδών μικροφώνων ισχύει ότι $C=0.5$, των παντοκατευθυντικών $C=1$, των δικατευθυντικών $C=0$, των υπερ-καρδιοειδών (hyper cardioid) $C=0.25$, των σούπερ-καρδιοειδών (super-cardioid) $C=(3/8)$ και των υπο-καρδιοειδών (sub-cardioid) $C=(3/4)$.

Παράρτημα 4 Συνάρτηση υπολογισμού του σήματος που λαμβάνει ένα μικρόφωνο.

Έστω ότι υπάρχει ένα ηχητικό πεδίο γύρω από την κάψα του μικροφώνου. Αυτό θα μπορούσε να εκφραστεί σαν μια συνάρτηση που μεταβάλλεται βάση της αζιμούθιας και κατακόρυφης γωνίας. Ας ονομάσουμε λοιπόν αυτή την συνάρτηση $f(\theta, \phi)$. Αυτό που θα λάμβανε το μικρόφωνο θα ήταν:

$$S_{mic} = \sum_{\theta=0}^{2\pi} \sum_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \phi) * M_{pollar}$$

Σχέση 48:

Παράρτημα 5 Κώδικας Matlab για σχεδιασμό σφαιρικών αρμονικών μέχρι 3^{ης} τάξης.

```
%plot 0, 1,2 order spherical harmonics
%number of nglcs to be measured
Angles=64;
%azimuthal and elevation angle
azimuth=0:2*pi/(Angles-1):2*pi;
elevation=-pi/2:pi/(Angles-1):pi/2;
%η λούπα με την οποία θα δημιουργηθούν οι πίνακες με τις
%θέσεις των σημείων στους άξονες X,Y,Z για τα σήματα των σφαιρικών
%αρμονικών και τα χρώματα αυτών.
a=1;
b=1;
for i=1:Angles-1;
    for j=1:Angles-1
        %δημιουργία πίνακα για το σήμα W
        r=1/sqrt(2);
        [WX(a,b),WY(a,b),WZ(a,b)]= ...
            sph2cart(azimuth(i),elevation(j),abs(r));
        [WX(a+1,b),WY(a+1,b),WZ(a+1,b)]= ...
            sph2cart(azimuth(i+1),elevation(j),abs(r));
        [WX(a+2,b),WY(a+2,b),WZ(a+2,b)]= ...
            sph2cart(azimuth(i+1),elevation(j+1),abs(r));
        [WX(a+3,b),WY(a+3,b),WZ(a+3,b)]= ...
            sph2cart(azimuth(i),elevation(j+1),abs(r));
        [WX(a+4,b),WY(a+4,b),WZ(a+4,b)]= ...
            sph2cart(azimuth(i),elevation(j),abs(r));
        WC(:,b)=[1;1;1;0];
        %δημιουργία πίνακα για το σήμα X
        r=cos(azimuth(i))*cos(elevation(j));
        [XX(a,b),XY(a,b),XZ(a,b)]= ...
            sph2cart(azimuth(i),elevation(j),abs(r));
        r=cos(azimuth(i+1))*cos(elevation(j));
        [XX(a+1,b),XY(a+1,b),XZ(a+1,b)]= ...
            sph2cart(azimuth(i+1),elevation(j),abs(r));
        r=cos(azimuth(i+1))*cos(elevation(j+1));
        [XX(a+2,b),XY(a+2,b),XZ(a+2,b)]= ...
            sph2cart(azimuth(i+1),elevation(j+1),abs(r));
        r=cos(azimuth(i))*cos(elevation(j+1));
        [XX(a+3,b),XY(a+3,b),XZ(a+3,b)]= ...
            sph2cart(azimuth(i),elevation(j+1),abs(r));
        r=cos(azimuth(i))*cos(elevation(j));
```

```
[XX(a+4,b),XY(a+4,b),XZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    XC(:,b)=[1;1;1;0];
else
    XC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα y
r=sin(azimuth(i))*cos(elevation(j));
[YX(a,b),YY(a,b),YZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=sin(azimuth(i+1))*cos(elevation(j));
[YX(a+1,b),YY(a+1,b),YZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=sin(azimuth(i+1))*cos(elevation(j+1));
[YX(a+2,b),YY(a+2,b),YZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=sin(azimuth(i))*cos(elevation(j+1));
[YX(a+3,b),YY(a+3,b),YZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=sin(azimuth(i))*cos(elevation(j));
[YX(a+4,b),YY(a+4,b),YZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    YC(:,b)=[1;1;1;0];
else
    YC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα Z
r=sin(elevation(j));
[ZX(a,b),ZY(a,b),ZZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=sin(elevation(j));
[ZX(a+1,b),ZY(a+1,b),ZZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=sin(elevation(j+1));
[ZX(a+2,b),ZY(a+2,b),ZZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=sin(elevation(j+1));
```

```

[ZX(a+3,b),ZY(a+3,b),ZZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=sin(elevation(j));
[ZX(a+4,b),ZY(a+4,b),ZZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    ZC(:,b)=[1;1;1;0];
else
    ZC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα R
r=1.5*((sin(elevation(j)))^2)-0.5;
[RX(a,b),RY(a,b),RZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=1.5*((sin(elevation(j)))^2)-0.5;
[RX(a+1,b),RY(a+1,b),RZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=1.5*((sin(elevation(j+1)))^2)-0.5;
[RX(a+2,b),RY(a+2,b),RZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=1.5*((sin(elevation(j+1)))^2)-0.5;
[RX(a+3,b),RY(a+3,b),RZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=1.5*((sin(elevation(j)))^2)-0.5;
[RX(a+4,b),RY(a+4,b),RZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    RC(:,b)=[1;1;1;0];
else
    RC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα S
r=cos(azimuth(i))*sin(2*elevation(j))*(3^0.5)/2;
[SX(a,b),SY(a,b),SZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=cos(azimuth(i+1))*sin(2*elevation(j))*(3^0.5)/2;
[SX(a+1,b),SY(a+1,b),SZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=cos(azimuth(i+1))*sin(2*elevation(j+1))*(3^0.5)/2;
[SX(a+2,b),SY(a+2,b),SZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=cos(azimuth(i))*sin(2*elevation(j+1))*(3^0.5)/2;
[SX(a+3,b),SY(a+3,b),SZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=cos(azimuth(i))*sin(2*elevation(j))*(3^0.5)/2;
[SX(a+4,b),SY(a+4,b),SZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    SC(:,b)=[1;1;1;0];
else
    SC(:,b)=[0;0;0;0];

```

```

end
%δημιουργία πίνακα για το σήμα T
r=sin(azimuth(i))*sin(2*elevation(j))*(3^0.5)/2;
[TX(a,b),TY(a,b),TZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=sin(azimuth(i+1))*sin(2*elevation(j))*(3^0.5)/2;
[TX(a+1,b),TY(a+1,b),TZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=sin(azimuth(i+1))*sin(2*elevation(j+1))*(3^0.5)/2;
[TX(a+2,b),TY(a+2,b),TZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=sin(azimuth(i))*sin(2*elevation(j+1))*(3^0.5)/2;
[TX(a+3,b),TY(a+3,b),TZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=sin(azimuth(i))*sin(2*elevation(j))*(3^0.5)/2;
[TX(a+4,b),TY(a+4,b),TZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    TC(:,b)=[1;1;1;0];
else
    TC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα U
r=(3^0.5)/2*cos(2*azimuth(i))*(cos(elevation(j)))^2;
[UX(a,b),UY(a,b),UZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=(3^0.5)/2*cos(2*azimuth(i+1))*(cos(elevation(j)))^2;
[UX(a+1,b),UY(a+1,b),UZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=(3^0.5)/2*cos(2*azimuth(i+1))*(cos(elevation(j+1)))^2;
[UX(a+2,b),UY(a+2,b),UZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=(3^0.5)/2*cos(2*azimuth(i))*(cos(elevation(j+1)))^2;
[UX(a+3,b),UY(a+3,b),UZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=(3^0.5)/2*cos(2*azimuth(i))*(cos(elevation(j)))^2;
[UX(a+4,b),UY(a+4,b),UZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    UC(:,b)=[1;1;1;0];
else
    UC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα V
r=(3^0.5)/2*sin(2*azimuth(i))*(cos(elevation(j)))^2;
[VX(a,b),VY(a,b),VZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=(3^0.5)/2*sin(2*azimuth(i+1))*(cos(elevation(j)))^2;
[VX(a+1,b),VY(a+1,b),VZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=(3^0.5)/2*sin(2*azimuth(i+1))*(cos(elevation(j+1)))^2;

```

```

[VX(a+2,b),VY(a+2,b),VZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=(3^0.5)/2*sin(2*azimuth(i))*(cos(elevation(j+1)))^2;
[VX(a+3,b),VY(a+3,b),VZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=(3^0.5)/2*sin(2*azimuth(i))*(cos(elevation(j)))^2;
[VX(a+4,b),VY(a+4,b),VZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    VC(:,b)=[1;1;1;0];
else
    VC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα K
r=sin(elevation(j))*(5*(sin(elevation(j))^2)-3)/2;
[KX(a,b),KY(a,b),KZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=sin(elevation(j))*(5*(sin(elevation(j))^2)-3)/2;
[KX(a+1,b),KY(a+1,b),KZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=sin(elevation(j+1))*(5*(sin(elevation(j+1))^2)-3)/2;
[KX(a+2,b),KY(a+2,b),KZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=sin(elevation(j+1))*(5*(sin(elevation(j+1))^2)-3)/2;
[KX(a+3,b),KY(a+3,b),KZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=sin(elevation(j))*(5*(sin(elevation(j))^2)-3)/2;
[KX(a+4,b),KY(a+4,b),KZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    KC(:,b)=[1;1;1;0];
else
    KC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα L
r=cos(azimuth(i))*cos(elevation(j))*...
    (5*(sin(elevation(j))^2)-1)*((3/8)^0.5);
[LX(a,b),LY(a,b),LZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=cos(azimuth(i+1))*cos(elevation(j))*...
    (5*(sin(elevation(j))^2)-1)*((3/8)^0.5);
[LX(a+1,b),LY(a+1,b),LZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=cos(azimuth(i+1))*cos(elevation(j+1))*...
    (5*(sin(elevation(j+1))^2)-1)*((3/8)^0.5);
[LX(a+2,b),LY(a+2,b),LZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=cos(azimuth(i))*cos(elevation(j+1))*...
    (5*(sin(elevation(j+1))^2)-1)*((3/8)^0.5);
[LX(a+3,b),LY(a+3,b),LZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));

```

```

r=cos(azimuth(i))*cos(elevation(j))*...
    (5*(sin(elevation(j))^2)-1)*((3/8)^0.5);
[LX(a+4,b),LY(a+4,b),LZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    LC(:,b)=[1;1;1;0];
else
    LC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα M
r=sin(azimuth(i))*cos(elevation(j))*...
    (5*(sin(elevation(j))^2)-1)*((3/8)^0.5);
[MX(a,b),MY(a,b),MZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=sin(azimuth(i+1))*cos(elevation(j))*...
    (5*(sin(elevation(j))^2)-1)*((3/8)^0.5);
[MX(a+1,b),MY(a+1,b),MZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=sin(azimuth(i+1))*cos(elevation(j+1))*...
    (5*(sin(elevation(j+1))^2)-1)*((3/8)^0.5);
[MX(a+2,b),MY(a+2,b),MZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=sin(azimuth(i))*cos(elevation(j+1))*...
    (5*(sin(elevation(j+1))^2)-1)*((3/8)^0.5);
[MX(a+3,b),MY(a+3,b),MZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=sin(azimuth(i))*cos(elevation(j))*...
    (5*(sin(elevation(j))^2)-1)*((3/8)^0.5);
[MX(a+4,b),MY(a+4,b),MZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    MC(:,b)=[1;1;1;0];
else
    MC(:,b)=[0;0;0;0];
end
%δημιουργία πίνακα για το σήμα N
r=cos(2*azimuth(i))*sin(elevation(j))*...
    (cos(elevation(j))^2)*(15^0.5)/2;
[NX(a,b),NY(a,b),NZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=cos(2*azimuth(i+1))*sin(elevation(j))*...
    (cos(elevation(j))^2)*(15^0.5)/2;
[NX(a+1,b),NY(a+1,b),NZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=cos(2*azimuth(i+1))*sin(elevation(j+1))*...
    (cos(elevation(j+1))^2)*(15^0.5)/2;
[NX(a+2,b),NY(a+2,b),NZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=cos(2*azimuth(i))*sin(elevation(j+1))*...
    (cos(elevation(j+1))^2)*(15^0.5)/2;
[NX(a+3,b),NY(a+3,b),NZ(a+3,b)]= ...

```

```

    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=cos(2*azimuth(i))*sin(elevation(j))*...
    (cos(elevation(j))^2*(15^0.5)/2;
[NX(a+4,b),NY(a+4,b),NZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    NC(:,b)=[1;1;1;1;0];
else
    NC(:,b)=[0;0;0;0;0];
end
%δημιουργία πίνακα για το σημείο
r=sin(2*azimuth(i))*sin(elevation(j))*...
    (cos(elevation(j))^2*(15^0.5)/2;
[OX(a,b),OY(a,b),OZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=sin(2*azimuth(i+1))*sin(elevation(j))*...
    (cos(elevation(j))^2*(15^0.5)/2;
[OX(a+1,b),OY(a+1,b),OZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=sin(2*azimuth(i+1))*sin(elevation(j+1))*...
    (cos(elevation(j+1))^2*(15^0.5)/2;
[OX(a+2,b),OY(a+2,b),OZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=sin(2*azimuth(i))*sin(elevation(j+1))*...
    (cos(elevation(j+1))^2*(15^0.5)/2;
[OX(a+3,b),OY(a+3,b),OZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=sin(2*azimuth(i))*sin(elevation(j))*...
    (cos(elevation(j))^2*(15^0.5)/2;
[OX(a+4,b),OY(a+4,b),OZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    OC(:,b)=[1;1;1;1;0];
else
    OC(:,b)=[0;0;0;0;0];
end
%δημιουργία πίνακα για το σημείο P
r=((5/8)^0.5)*cos(3*azimuth(i))*(cos(elevation(j))^3);
[PX(a,b),PY(a,b),PZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=((5/8)^0.5)*cos(3*azimuth(i+1))*(cos(elevation(j))^3);
[PX(a+1,b),PY(a+1,b),PZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=((5/8)^0.5)*cos(3*azimuth(i+1))*(cos(elevation(j+1))^3);
[PX(a+2,b),PY(a+2,b),PZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=((5/8)^0.5)*cos(3*azimuth(i))*(cos(elevation(j+1))^3);
[PX(a+3,b),PY(a+3,b),PZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=((5/8)^0.5)*cos(3*azimuth(i+1))*(cos(elevation(j))^3);
[PX(a+4,b),PY(a+4,b),PZ(a+4,b)]= ...

```

```

    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    PC(:,b)=[1;1;1;1;0];
else
    PC(:,b)=[0;0;0;0;0];
end
%δημιουργία πίνακα για το σημείο Q
r=((5/8)^0.5)*sin(3*azimuth(i))*(cos(elevation(j))^3);
[QX(a,b),QY(a,b),QZ(a,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
r=((5/8)^0.5)*sin(3*azimuth(i+1))*(cos(elevation(j))^3);
[QX(a+1,b),QY(a+1,b),QZ(a+1,b)]= ...
    sph2cart(azimuth(i+1),elevation(j),abs(r));
r=((5/8)^0.5)*sin(3*azimuth(i+1))*(cos(elevation(j+1))^3);
[QX(a+2,b),QY(a+2,b),QZ(a+2,b)]= ...
    sph2cart(azimuth(i+1),elevation(j+1),abs(r));
r=((5/8)^0.5)*sin(3*azimuth(i))*(cos(elevation(j+1))^3);
[QX(a+3,b),QY(a+3,b),QZ(a+3,b)]= ...
    sph2cart(azimuth(i),elevation(j+1),abs(r));
r=((5/8)^0.5)*sin(3*azimuth(i))*(cos(elevation(j))^3);
[QX(a+4,b),QY(a+4,b),QZ(a+4,b)]= ...
    sph2cart(azimuth(i),elevation(j),abs(r));
if r>=0
    QC(:,b)=[1;1;1;1;0];
else
    QC(:,b)=[0;0;0;0;0];
end
b=b+1;
end
end
%plot W
figure(1);
fill3(WX,WY,WZ,WC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot X
figure(2);
fill3(XX,XY,XZ,XC);
light;
lighting phong;
shading interp;
axis equal;
axis off;

```

```

view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot Y
figure(3);
fill3(YX,YY,YZ,YC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot Z
figure(4);
fill3(ZX,ZY,ZZ,ZC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot R
figure(5);
fill3(RX,RY,RZ,RC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot S
figure(6);
fill3(SX,SY,SZ,SC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;

```

```

axis([-1 1 -1 1 -1 1]);
%plot T
figure(7);
fill3(TX,TY,TZ,TC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot U
figure(8);
fill3(UX,UY,UZ,UC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot V
figure(9);
fill3(VX,VY,VZ,VC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot K
figure(10);
fill3(KX,KY,KZ,KC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot L
figure(11);

```

```

fill3(LX,LY,LZ,LC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot M
figure(12);
fill3(MX,MY,MZ,MC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot N
figure(13);
fill3(NX,NY,NZ,NC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot O

```

```

figure(14);
fill3(OX,OY,OZ,OC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot P
figure(15);
fill3(PX,PY,PZ,PC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);
%plot q
figure(16);
fill3(QX,QY,QZ,QC);
light;
lighting phong;
shading interp;
axis equal;
axis off;
view(-40,30);
material([0.5 0.8 1]);
camlight headlight;
axis([-1 1 -1 1 -1 1]);

```


Παράρτημα 6 Απόδειξη για γωνίες προσανατολισμού των καψών του Soundfield μικροφώνου.

Σκοπός είναι να γίνει αντιληπτό με ποιες γωνίες θα λάμβαναν το ηχητικό πεδίο οι κάψες των μικροφώνων, ώστε να είναι δυνατός ο υπολογισμός του τύπου μετατροπής του A-Format σε B-Format. Καλό θα ήταν λοιπόν να ήταν γνωστές οι γωνίες προσανατολισμού που έχουν οι κάψες των μικροφώνων. Αυτές βρίσκονται στις πλευρές ενός τετραέδρου. Για τον υπολογισμό αυτών των γωνιών θα πρέπει να είναι γνωστές οι κλίσεις των πλευρών του τετραέδρου. Για τον υπολογισμό της κλίσης κάθε πλευράς αρκεί να είναι γνωστά τρία σημεία πάνω σε αυτή την επιφάνεια. Για λόγους ευκολίας θα χρησιμοποιηθούν οι κορυφές κάθε πλευράς. Οι κορυφές ενός τετραέδρου που σχηματίζουν και τις τέσσερις πλευρές είναι τέσσερις. Για την δημιουργία του τετραέδρου αρχικά θα δημιουργήσουμε ένα ισόπλευρο τρίγωνο βρίσκοντας τις συντεταγμένες των τριών κορυφών και στη συνέχεια θα ψάξουμε να βρούμε την τέταρτη κορυφή. Για λόγους ευκολίας και λιγότερων πράξεων θεωρούμε ότι το ισόπλευρο τρίγωνο θα βρίσκεται στο οριζόντιο επίπεδο με $z=0$. Έστω ένα σημείο $\mathbf{a} (0,0,0)$ που είναι το κέντρο του ισόπλευρου τριγώνου και τις συντεταγμένες της πρώτης κορυφής $\mathbf{b} (1,0,0)$. Οι υπόλοιπες κορυφές θα πρέπει να βρίσκονται στις περιμέτρο ενός κύκλου με ακτίνα 1.

Η εξίσωση που μας δίνει όλα τα σημεία του κύκλου με κέντρο \mathbf{a} και ακτίνα 1 είναι:

$$x^2+y^2+z^2=1$$

Σχέση 49:

όπου x,y,z οι συντεταγμένες κάθε σημείου που βρίσκεται την περιμέτρο του κύκλου. Το z θα ισούται με το 0 καθώς είναι αρχική παραδοχή. Θεωρούμε τα υπόλοιπα σημεία του ισόπλευρου $\mathbf{c} (c_1,c_2,0)$, $\mathbf{d}(d_1,d_2,0)$. Αφού το τρίγωνο είναι ισόπλευρο θα πρέπει:

$$|bc|=|cd|=|db| \Leftrightarrow (c_1-1)^2+(c_2)^2=(d_1-c_1)^2+(d_2-c_2)^2=(1-d_1)^2+d_2^2$$

Σχέση 50:

Αντικαθιστώντας τις σχέσεις 49 στην σχέση 50 έχουμε:

$$(c_2^2+c_1^2)+1-2c_1=(d_2^2+d_1^2)+1-2d_1 \Rightarrow c_1=d_1$$

Σχέση 51:

Αντικαθιστώντας την 51 στη 50 έχουμε:

$$(c_1 - 1)^2 + (c_2)^2 = (1 - d_1)^2 + d_2^2 \Rightarrow c_2^2 = d_2^2$$

Σχέση 52:

Καθώς όμως $c_1 = d_1$ και $c_3 = d_3 = 0$ πρέπει c_2 να διαφέρει από το d_2 , πράγμα που σημαίνει ότι $c_2 = -d_2$. Αντικαθιστώντας την σχέση 51,52 στη σχέση 50 έχουμε:

$$\begin{aligned} (c_1 - 1)^2 + (c_2)^2 &= (d_1 - c_1)^2 + (d_2 - c_2)^2 \Rightarrow (c_1)^2 + 1 - 2c_1 + (c_2)^2 = (d_2)^2 + (c_2)^2 - 2d_2c_2 \Rightarrow \\ (c_1)^2 + 1 - 2c_1 &= (d_2)^2 - 2d_2c_2 \Rightarrow (c_1)^2 + 1 - 2c_1 = 1 - (c_1)^2 - 2d_2c_2 \Rightarrow 2(c_1)^2 - 2c_1 = -2d_2c_2 \Rightarrow \\ 2(c_1)^2 - 2c_1 &= 2c_2^2 \Rightarrow (c_1)^2 - c_1 = 1 - c_1^2 \Rightarrow c_1^2 - 0.5c_1 - 0.5 = 0 \Rightarrow c_1 = -0.5 \text{ ή } c_1 = 1 \end{aligned}$$

Σχέση 53:

Θέτοντας στην σχέση 49 $c_1 = 1$ προκύπτει ότι $c_2 = 0$ που είναι άτοπο καθώς τα σημεία c και d θα συνέπιπταν. Άρα προκύπτει ότι $c_1 = -0.5$. Αντικαθιστώντας τα c_1 και d_1 στη σχέση 49 βγάζουμε ότι $c_2 = \pm(0.75)^{0.5}$ και $d_2 = \pm(0.75)^{0.5}$. Θα δεχτούμε την θετική λύση για το c_2 και την αρνητική για το d_2 . Τελικά έχουμε $c_2 = (0.75)^{0.5}$ και $d_2 = -(0.75)^{0.5}$.

Το μήκος του ευθύγραμμου τμήματος bc είναι:

$$|bc| = (-0.5 - 5)^2 + ((0.75)^{0.5})^2 = 3$$

Σχέση 54:

Η τέταρτη κορυφή του τετράεδρου θα πρέπει να ισαπέχει από τις υπόλοιπες κορυφές κατά το μήκος $|bc|$. Άρα θα ισχύει:

$$|zb| = (1 - z_1)^2 + (-z_2)^2 + (-z_3)^2 = 3$$

Σχέση 55:

$$|zc| = (0.5 - z_1)^2 + (0.75^{0.5} - z_2)^2 + (-z_3)^2 = 3$$

Σχέση 56:

$$|zc| = (0.5 - z_1)^2 + (-0.75^{0.5} - z_2)^2 + (-z_3)^2 = 3$$

Σχέση 57:

Από σχέση 56-57 έχουμε:

$$4(0.75)^{0.5}z_2 = 0 \Rightarrow z_2 = 0$$

Σχέση 58:

Αντικαθιστώντας την 10 στην αφαίρεση των σχέσεων 49-51 έχουμε:

$$(1 - z_1)^2 - (0.5 - z_1)^2 - 0.75 = 0 \Rightarrow 1 - 0.25 - z_1 - 0.75 = 0 \Rightarrow z_1 = 0$$

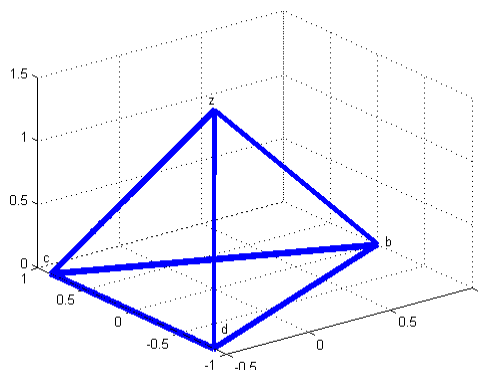
Σχέση 59:

Αντικαθιστώντας τις 58, 59 στη σχέση 55 έχουμε:

$$1+z_3^2=1.5^2+0.75 \Rightarrow z_3=\pm(2)^{0.5}$$

Σχέση 60:

θα δεχτώ την θετική τιμή άρα $z_3=(2)^{0.5}$. Στην εικόνα (42) αυτό το πολύεδρο (σχεδίαση με το matlab).



Εικόνα 42: Απεικόνιση του τετράεδρου που σχηματίστηκε από τις κορυφές b,c,d,z στο Matlab. Στις πλευρές αυτού τοποθετούνται οι κάψες του Soundfield μικροφώνου.

Για την εύρεση του προσανατολισμού που έχουν οι πλευρές θα πάρουμε το γινόμενο των ανυσμάτων της κάθε πλευράς.

Για την πλευρά bcd ισχύει:

$$cd \times cb = [(-0.75^{0.5}-0.75^{0.5})(0-0) - (0-0)(0-0.75^{0.5}), (0-0)*(1+0.5) - (-0.75^{0.5}-0.75^{0.5})(0+0), (-0.5+0.5)(0-0.75^{0.5})-(-0.75^{0.5}-0.75^{0.5})(1+0.5)]=[0,0,2.5980]$$

Σχέση 61:

Για την πλευρά zbc ομοίως ισχύει.

$$bz \times bc=[-1.2247,-2.1213,-0.8660]$$

Σχέση 62:

Για την πλευρά zcd ομοίως ισχύει:

$$cz \times cd=[2.4494,0,-0.8660]$$

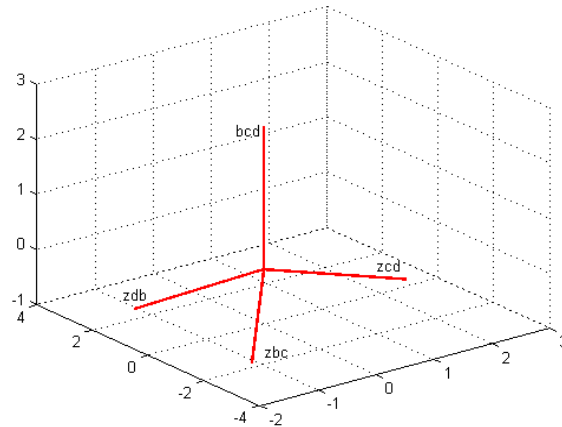
Σχέση 63:

Τέλος για την πλευρά zdb ομοίως ισχύει:

$$dz \times db=[-1.2247,2.1213,-0.8660]$$

Σχέση 64:

Αυτά τα διανύσματα απεικονίζονται στην εικόνα 43.



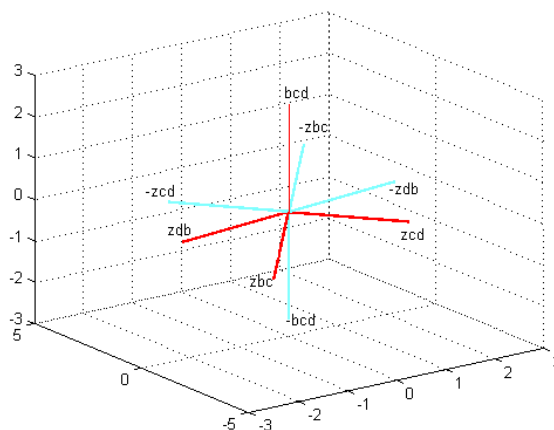
Εικόνα 43: Απεικόνιση των διανυσμάτων που είναι κάθετα στις πλευρές του τετράεδρου.

Το \cos της γωνίας που σχηματίζει η πλευρά bcd με την zbc βγαίνει αν πάρουμε το εσωτερικό τους γινόμενο και το διαιρέσουμε με το γινόμενο των μέτρων των δύο διανυσμάτων. Για να βρούμε την γωνία πρέπει να πάρουμε την ανάστροφη συνάρτηση του \cos του τελικού αποτελέσματος. Ο τύπος υπολογισμού λοιπόν είναι :

$$\theta = \cos^{-1}(bcd * zbc / |bcd| * |zbc|) = 109.47^\circ$$

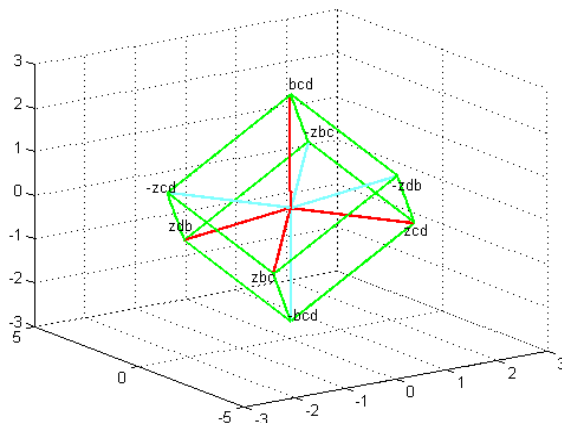
Σχέση 65:

Ομοίως υπολογίστηκε ότι όλα τα διανύσματα που είναι κάθετα στις πλευρές **έχουν μεταξύ τους την ίδια διαφορά στις γωνίες**. Επίσης έχουν και το ίδιο μήκος (βλέπε Σχέση 54). Αυτό παραπέμπει στην σχέση που έχουν οι τέσσερις διαγώνιοι ενός κύβου. Δημιουργώντας λοιπόν και τα αρνητικά αυτών των διανυσμάτων παίρνουμε στην πραγματικότητα ολόκληρα τα ευθύγραμμα τμήματα των διαγωνίων (βλέπε εικόνα 44).



Εικόνα 44: Απεικόνιση των διανυσμάτων των καθέτων στις πλευρές του τετράεδρου (με κόκκινο χρώμα), μαζί με τα αρνητικά τους (με θαλασσί).

Ενώνοντας το τέλος κάθε διανύσματος με το τέλος των υπολοίπων αρνητικών που δεν ανήκουν στην ίδια ευθεία δημιουργείται ένας κύβος (βλέπε εικόνα 45).



Εικόνα 45: Ο κύβος που δημιουργείται με την ένωση των θετικών με τα αρνητικά διανύσματα που δεν ανήκουν στην ίδια ευθεία.

Για να μπορέσει να ληφθεί κατάλληλα το ηχητικό πεδίο θα βόλευε αν όλα διανύσματα σχημάτιζαν ίδια ή αντίθετη κατακόρυφη και αζιμούθια γωνία. Αρχίζοντας λοιπόν περιστρέφουμε όλα τα διανύσματα ως προς τον άξονα y προσπαθώντας να βρούμε μια γωνία 'u' περιστροφής όπου τα διανύσματα bcd και zbc να σχημάτιζαν αντίθετη γωνία z (δηλαδή να έχουν αντίθετα z). Περιστρέφοντας λοιπόν τα διανύσματα ως προς 'u' ισχύει:

$$\begin{aligned} bcd' &= [bcd_1 * \cos(u) + bcd_3 * \sin(u), bcd_2, -bcd_1 * \sin(u) + bcd_3 * \cos(u)] \\ &\text{και} \\ zbc' &= [-zbc_1 * \cos(u) + zbc_3 * \sin(u), zbc_2, -zbc_1 * \sin(u) + zbc_3 * \cos(u)] \end{aligned}$$

Σχέση 66:

Θέλουμε όμως τα z των bcd' και zbc' να είναι αντίθετα άρα από την σχέση 66 έχουμε:

$$\begin{aligned} -bcd_1 * \sin(u) + bcd_3 * \cos(u) &= zbc_1 * \sin(u) - zbc_3 * \cos(u) \Rightarrow 2.5980 \cos(u) = -1.2247 \sin(u) + 0.8660 \cos(u) \Rightarrow \\ \Rightarrow \tan(u) &= -1.4142 \Rightarrow u = -54,7358 \text{ ή } u = 125.2642, \text{ αφού και οι δύο ισχύουν θα δεχτώ την τιμή} \\ \mathbf{u} &= \mathbf{-54,7358}. \end{aligned}$$

Σχέση 67:

Αντικαθιστώντας το 'u' στη σχέση 66 έχουμε ότι οι συντεταγμένες του **bcd'** είναι **[-2.1213, 0, 1.4999]**, του **zbc'**=[**0.0000, -2.1213, -1.5000**], **zcd'**=[**2.1213, 0, 1.5000**], **zdb'**=[**0.0000, 2.1213, -1.5000**]. Παρατηρούμε ότι όλα τα διανύσματα έχουν το ίδιο z. Ταυτόχρονα βλέπουμε ότι όλα τα διανύσματα θα έχουν μηδενικό x ή y. Στη συνέχεια θέλουμε να περιστρέψουμε όλα τα διανύσματα ως προς τον άξονα z ώστε να σχηματίζουν γωνία $\pm 45^\circ$. Έστω ότι θέλουμε να περιστρέψουμε το διάνυσμα bcd κατά μία γωνία 'w' ώστε να σχηματίζει γωνία 45 μοιρών με τον άξονα x, θα ισχύει λοιπόν:

$$\begin{aligned} Bcd'' &= [bcd'_1 * \cos(w) - bcd'_2 * \sin(w), bcd'_1 * \sin(w) + bcd'_2 * \cos(w), bcd'_3] = \\ &= [(|bcd'|)(2^{0.5})/2, bcd'_3, (|bcd'|)(2^{0.5})/2, bcd'_3] = \end{aligned}$$

Σχέση 68:

Για τα x της σχέσης 20 έχουμε:

$$bcd'_1 * \cos(w) - bcd'_2 * \sin(w) = |bcd'| * \cos(45) * \cos(\text{elevation})$$

Σχέση 69:

Για τα y της σχέσης 20 έχουμε:

$$bcd'_1 * \sin(w) + bcd'_2 * \cos(w) = |bcd'| * \sin(45) * \cos(\text{elevation})$$

Σχέση 70:

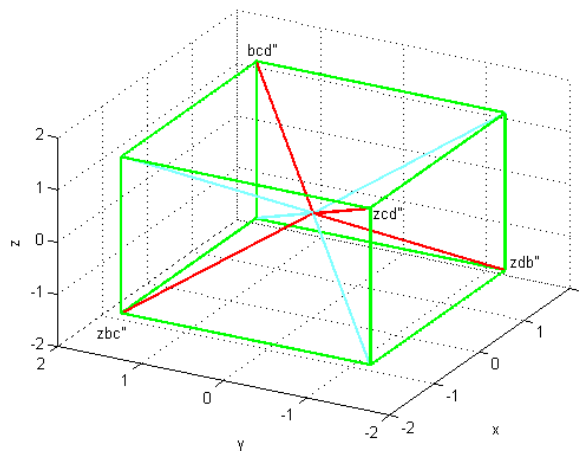
Αφαιρώντας τις σχέσεις 70-69 έχουμε:

$$(bcd'_1 + bcd'_2) * \sin(w) + (bcd'_2 - bcd'_1) * \cos(w) = 0 \Rightarrow \tan(w) = 1 \Rightarrow w = 45^\circ \text{ ή } w = 225^\circ$$

επειδή όμως οι τιμές του x βρίσκονται στην αρχή του τέταρτου τεταρτημόριου, για να σχηματίζει γωνία 45° θα πρέπει να δεχτώ την δεύτερη τιμή άρα **w=225°**

Σχέση 71:

Αντικαθιστώντας λοιπόν το 'w' στην εξίσωση 20 και περιστρέφοντας στη συνέχεια όλα τα διανύσματα κατά 'w' ως προς τον άξονα z, έχω **bcd'' = [1.5000, 1.5000, 1.4999]**, **zbc'' = [-1.5000, 1.5000, -1.5000]**, **zcd'' = [-1.5000, -1.5000, 1.5000]**, **zdb'' = [1.5000, -1.5000, -1.5000]**. Στην εικόνα (46) φαίνεται η κατεύθυνσεις των καψών του μικροφώνου και ο κύβος που σχηματίζεται.



Εικόνα 46: Απεικόνιση των διανυσμάτων κατεύθυνσης των καψών των τεσσάρων μικροφώνων με την κόκκινη γραμμή, τα αρνητικά τους με θαλασσιά και ο κύβος που σχηματίζεται. Στις κορυφές του κύβου που ενώνονται με τις κόκκινες γραμμές θα μπορούσαν να τοποθετηθούν οι κάψες των μικροφώνων με διευθύνσεις τις διαγωνίους (κόκκινες γραμμές).

Η αζιμούθια και η κατακόρυφη γωνία που θα έχει κάθε κάψα είναι **bcd(45.0000 , 35.2642)**, **zbc(135.0001 , -35.2636)**, **zcd(-135.0000 , 35.2646)**, **zdb(-45.0001 -35.2636)**. Κάνοντας την αντιστοίχιση με τις ονομασίες που έχουν δοθεί στις τέσσερις κάψες του Soundfield μικροφώνου έχουμε **FUL(45.0000 , 35.2642)**, **BDL(135.0001 , -35.2636)**, **BUR(-135.0000 , 35.2646)**, **FDR(-45.0001 -35.2636)**.

Παράρτημα 7 *Κώδικας Matlab για δημιουργία κβαντοποιητή με δυνατότητα επιλογής του τύπου κβαντοποίησης, dither , τεχνικών noise shape. Υποστηρίζει πολυκάναλα σήματα εισόδου.*

```

1. function [OutSignal]=Quantiser(InSignal,N,Smax,MidriserFlag,NSFlag,DitherFlag);
2.
3. % Συνάρτηση y=Quantiser(InSignal,N,Smax,MidriserFlag,NSFlag,DitherFlag);
4. % Αυτη η συνάρτηση κβαντοποιεί την ένταση του σηματος εισόδου σε 2^N
5. % διακριτές εντάσεις, όσο αναφορά την περίπτωση κβαντοποίησης με τεχνική midriser,
6. % ενώ με 2^N-1 εντάσεις στην περίπτωση χρήσης τεχνικής midtread. Το ελάχιστο βήμα
7. % κβαντισμού και στις δύο περιπτώσεις είναι min=2Smax/(2^N-1). Οι τιμές εξόδου
8. % του κβαντιστή κυμαίνονται στα [-Smax,Smax] στην περίπτωση χρήσης
9. % τεχνικής midriser, ενώ στα [-Smax+min/2,Smax-min/2]. Οι μεταβλητές αυτής
10. % της συνάρτησης αντιστοιχούν:
11. %
12. % InSignal: αντιστοιχεί στο σήμα που πρόκειται να κβαντιστεί.
13. % N: Αντιστοιχεί στο bit rate του κβαντιστή
14. % Smax: Αντιστοιχεί στα όρια τιμών που μπορεί να πάρει ο κβαντιστής.
15. % MidriserFlag: Αυτή η μεταβλητή μπορεί να πάρεις τιμές (0,1) όπου με την
16. % τιμή '0' γίνεται κβαντισμός με τεχνική midtread ενώ με την τιμή
17. % '1' γίνεται κβαντισμός με τεχνική midriser.
18. % NSFlag: Με την τιμή αυτής της μεταβλητής καθορίζεται αν θα γίνει χρήση
19. % noise shape κατά την χρήση του dither. Οι τιμες που μπορεί να πάρει είναι
20. % (0,2,3), όπου με το '0' δεν γίνεται χρήση noise shape, ενώ με τις τιμές
21. % '2' και '3' χρησιμοποιειται noise shape δεύτερης και τρίτης τάξης
22. % αντίστοιχα.
23. %DitherFlag Με την τιμή αυτής της μεταβλητής καθορίζεται αν θα γίνει χρήση dither
24. % κατά την διαδικασία της κβαντοποίησης, και σε ποιά κατηγορία θα ανήκει
25. % αυτό. Οι τιμές που μπορεί να πάρει είναι (0,1,2,3) όπου με την τιμή '0'
26. % δεν χρησιμοποιειείται το dither κατα την κβαντοποίησης (ουσιαστικά δεν
27. % χρησιμοποιειείται και τεχνική noise shape), με την τιμή '1' γίνεται
28. % χρήση RPDF dither, με τιμή '2' TPDF dither ενώ με την τιμή '3'
29. % γίνεται χρήση HighPass TPDF dither.
30.
31. %Ελεγχος για το όρισμα NSFlag.

```

```

32.  if (or((NSFlag==0),or((NSFlag==2),(NSFlag==3)))-1)
33.  error('Οι τιμές που μπορεί να πάρει το (NSFlag) είναι (0,2,3)')
34.  end
35.  % Αρχικές τιμές.
36.  OutSignal=[];
37.  dither=[];
38.  % Εύρεση γραμμών και στηλών του σήματος εισόδου.Ορίζω ως 'st' τις στήλες
39.  % και 'gr' τις γραμμές.
40.  [gr st]=size(InSignal);
41.
42.  % Υπολογισμός του αριθμού των σημείων που πρέπει να υπολογιστούν.
43.  npoints=gr;
44.
45.  % Υπολογισμός του βήματος κβαντισμού 'min'
46.  min=(2*Smax)/(2^N-1);
47.  % Εύρεση του τρόπου που θα γίνει ο κβαντισμός.
48.  if (MidriserFlag==0)
49.      % Σε αυτόν τον τομέα ελέγχεται το όρισμα DitherFlag αν έχει σωστή τιμή και
50.      % ανάλογα προσθέτει στο σήμα εισόδου το κατάλληλο dither που έχει
51.      % επιλεγεί (RPDF, TPDF, Highpass TPDF).
52.      if (DitherFlag==0)
53.          dither=zeros(size(InSignal));
54.          InSignal=InSignal+dither;
55.          elseif (or((DitherFlag==1),or((DitherFlag==2),(DitherFlag==3))))
56.              dither=MakeDither(DitherFlag,min,gr,st);
57.              InSignal=InSignal+dither;
58.      else error('Οι τιμές που μπορεί να πάρει το (DitherFlag) είναι (0,1,2,3)')
59.      end
60.      % Διαδικασία κβαντοποίησης χωρίς την χρήση noise shape. Η χρήση της
61.      % παρένθεσης (1-MidriserFlag) ή μόνο του MidriserFlag εξασφαλίζει αν η
62.      % κβαντοποίηση θα γίνει με την μέθοδο Midriser ή την μέθοδο Midtread.
63.      % Ανάλογα με την τιμή (0,1) που θα πάρει μηδενίζονται οι παράγοντες οι
64.      % οποίοι δεν ανήκουν σε αυτή την μέθοδο κβαντισμού.
65.      if(NSFlag==0)
66.          OutSignal=min*floor(InSignal/min+0.5);
67.      % Διαδικασία εύρεσης των τιμών εξόδου του κβαντιστή που ξεπερνάει τα όρια
68.      % του με αποτέλεσμα την δημιουργία ψαλιδισμού. Στο σήμα εξόδου υπάρχει το
69.      % min*(1-MidriserFlag)*0.5 για την περίπτωση που η κβαντοποίηση γίνεται με
70.      % μέθοδο midtreat, όπου η ελάχιστη και μέγιστη τιμή είναι αυξημένη και
71.      % μειωμένη κατά min/2 αντίστοιχα.
72.          i=find(OutSignal>=Smax);
73.          OutSignal(i)=Smax-min*0.5;
74.          i=find(OutSignal<=-Smax);
75.          OutSignal(i)=-Smax+min*0.5;

```

```

76.      % Διαδικασία κβαντοποίησης με την χρήση noise shape δεύτερης τάξης.
77.      elseif (NSFlag==2)
78.          for j=1:st,
79.              f1=0;
80.              f2=0;
81.              for i=1:1:gr,
82.                  % Οι συντελεστές έχουν παρθει απο το αντίστοιχο εργαλείο QuantiseAudio.m
83.                  % του Κ. Α.Φλώρου στο μάθημα του δευτέρου εξαμήνου του μεταπτυχιακού.
84.                  w=2*f1-f2;
85.                  OutSignal(i,j)=min*floor((InSignal(i,j)-w)/min+0.5);
86.                  % Διαδικασία εύρεσης των τιμών εξόδου του κβαντιστή που ξεπερνάει τα όριά
87.                  % του με αποτέλεσμα την δημιουργία ψαλιδισμού. Στο σημα εξόδου υπάρχει το
88.                  % min*(1-MidriserFlag)*0.5 για την περίπτωση που η κβαντοποίηση γίνεται με
89.                  % μέθοδο midtreat, όπου η ελάχιστη και μέγιστη τιμή είναι αυξημένη και
90.                  % μειωμένη κατά min/2 αντίστοιχα.
91.                  if (OutSignal(i,j)>=Smax);
92.                      OutSignal(i,j)=Smax-min*0.5;
93.                  end
94.                  if (OutSignal(i,j)<=-Smax);
95.                      OutSignal(i,j)=-Smax+min*0.5;
96.                  end
97.                  f=OutSignal(i,j)-InSignal(i,j)+w;
98.                  f2=f1;
99.                  f1=f;
100.             end
101.         end
102.         % Διαδικασία κβαντοποίησης με την χρήση noise shape τρίτης τάξης.
103.         elseif (NSFlag==3)
104.             for j=1:st,
105.                 f1=0;
106.                 f2=0;
107.                 f3=0;
108.                 for i=1:1:gr,
109.                     % Οι συντελεστές έχουν παρθει απο το αντίστοιχο εργαλείο QuantiseAudio.m
110.                     % του Κ. Α.Φλώρου στο μάθημα του δευτέρου εξαμήνου του μεταπτυχιακού.
111.                     w=3*f1-3*f2+f3;
112.                     OutSignal(i,j)=min*floor((InSignal(i,j)-w)/min+0.5);
113.                     % Διαδικασία εύρεσης των τιμών εξόδου του κβαντιστή που ξεπερνάει τα όριά
114.                     % του με αποτέλεσμα την δημιουργία ψαλιδισμού. Στο σημα εξόδου υπάρχει το
115.                     % min*(1-MidriserFlag)*0.5 για την περίπτωση που η κβαντοποίηση γίνεται με
116.                     % μέθοδο midtreat, όπου η ελάχιστη και μέγιστη τιμή είναι αυξημένη και
117.                     % μειωμένη κατά min/2 αντίστοιχα.
118.                     if (OutSignal(i,j)>=Smax);
119.                         OutSignal(i,j)=Smax-min*0.5;

```

```

120.                 end
121.                 if (OutSignal(i,j)<=-Smax)
122.                     OutSignal(i,j)=-Smax+min*0.5;
123.                 end
124.                 f=OutSignal(i,j)-InSignal(i,j)+w;
125.                 f3=f2;
126.                 f2=f1;
127.                 f1=f;
128.             end
129.         end
130.     end
131.
132.% Υλοποίηση κβαντιστη Midriser
133.elseif (MidriserFlag==1)
134.    % Σε αυτόν τον τομέα ελέγχεται το όρισμα DitherFlag αν έχει σωστή τιμή και
135.    % ανάλογα προσθέτει στο σήμα εισόδου το κατάλληλο dither που έχει
136.    % επιλεγεί (RPDF, TPDF, Highpass TPDF).
137.    if (DitherFlag==0)
138.        dither=zeros(size(InSignal));
139.        InSignal=InSignal+dither;
140.        elseif (or((DitherFlag==1),or((DitherFlag==2),(DitherFlag==3))))
141.            dither=MakeDither(DitherFlag,min,gr,st);
142.            InSignal=InSignal+dither;
143.    else error('Οι τιμές που μπορεί να πάρει το (DitherFlag) είναι (0,1,2,3)')
144.    end
145.    % Διαδικασία κβαντοποίησης χωρίς την χρήση noise shape. Η χρήση της
146.    % παρένθεσης (1-MidriserFlag) ή μόνο του MidriserFlag εξασφαλίζει αν η
147.    % κβαντοποίηση θα γίνει με την μέθοδο Midriser ή την μέθοδο Midtread.
148.    % Ανάλογα με την τιμή (0,1) που θα πάρει μηδενίζονται οι παράγοντες οι
149.    % οποίοι δεν ανήκουν σε αυτή την μέθοδο κβαντισμού.
150.    if(NSFlag==0)
151.        OutSignal=min*floor(InSignal/min)+min*0.5;
152.    % Διαδικασία εύρεσης των τιμών εξόδου του κβαντιστή που ξεπερνάει τα όριά
153.    % του με αποτέλεσμα την δημιουργία ψαλιδισμού. Στο σήμα εξόδου υπάρχει το
154.    % min*(1-MidriserFlag)*0.5 για την περίπτωση που η κβαντοποίηση γίνεται με
155.    % μέθοδο midtreat, όπου η ελάχιστη και μέγιστη τιμή είναι αυξημένη και
156.    % μειωμένη κατά min/2 αντίστοιχα.
157.        i=find(OutSignal>=Smax);
158.        OutSignal(i)=Smax;
159.        i=find(OutSignal<=-Smax);
160.        OutSignal(i)=-Smax;
161.    % Διαδικασία κβαντοποίησης με την χρήση noise shape δεύτερης τάξης.
162.    elseif (NSFlag==2)
163.        for j=1:st,

```

```

164.         f1=0;
165.         f2=0;
166.         for i=1:1:gr,
167.             % Οι συντελεστές έχουν παρθει απο το αντίστοιχο εργαλείο QuantiseAudio.m του
168.             % Κ. Α.Φλώρου στο μάθημα του δευτέρου εξαμήνου του μεταπτυχιακού.
169.             w=f1-0.5*f2;
170.             OutSignal(i,j)=min*floor((InSignal(i,j)-w)/min)+min*0.5;
171.             % Διαδικασία εύρεσης των τιμών εξόδου του κβαντιστή που ξεπερνάει τα όριά
172.             % του με αποτέλεσμα την δημιουργία ψαλιδισμού. Στο σημα εξόδου υπάρχει το
173.             % min*(1-MidriserFlag)*0.5 για την περίπτωση που η κβαντοποίηση γίνεται με
174.             % μέθοδο midtreat, όπου η ελάχιστη και μέγιστη τιμή είναι αυξημένη και
175.             % μειωμένη κατά min/2 αντίστοιχα.
176.             if (OutSignal>=Smax);
177.                 OutSignal(i)=Smax;
178.             end
179.             if (OutSignal<=-Smax);
180.                 OutSignal(i)=-Smax;
181.             end
182.             f=OutSignal(i)-InSignal(i)+w;
183.             f2=f1;
184.             f1=f;
185.         end
186.     end
187.     % Διαδικασία κβαντοποίησης με την χρήση noise shape τρίτης τάξης.
188. elseif (NSFlag==3)
189.     for j=1:st,
190.         f1=0;
191.         f2=0;
192.         f3=0;
193.         for i=1:1:gr,
194.             % Οι συντελεστές έχουν παρθει απο το αντίστοιχο εργαλείο QuantiseAudio.m
195.             % του Κ. Α.Φλώρου στο μάθημα του δευτέρου εξαμήνου του μεταπτυχιακού.
196.             w=f1-f2+(1/3)*f3;
197.             OutSignal(i,j)=min*floor((InSignal(i,j)-w)/min)+min*0.5;
198.             % Διαδικασία εύρεσης των τιμών εξόδου του κβαντιστή που ξεπερνάει τα όριά
199.             % του με αποτέλεσμα την δημιουργία ψαλιδισμού. Στο σημα εξόδου υπάρχει το
200.             % min*(1-MidriserFlag)*0.5 για την περίπτωση που η κβαντοποίηση γίνεται με
201.             % μέθοδο midtreat, όπου η ελάχιστη και μέγιστη τιμή είναι αυξημένη και
202.             % μειωμένη κατά min/2 αντίστοιχα.
203.             if (OutSignal>=Smax);
204.                 OutSignal(i,j)=Smax;
205.             end
206.             if (OutSignal<=-Smax)
207.                 OutSignal(i,j)=-Smax;

```

```
208.             end
209.             f=OutSignal(i,j)-InSignal(i,j)+w;
210.             f3=f2;
211.             f2=f1;
212.             f1=f;
213.         end
214.     end
215. end
216. else error('Οι τιμές που μπορεί να πάρει το (MidriserFlag) είναι (0,1)')
217.end
```

Παράρτημα 8 Κώδικας Matlab για δημιουργία dither με δυνατότητα επιλογής ανάμεσα σε RPDF, TPDF και Highpass TPDF.

```

1. function [dith]=MakeDither(dithtype,min,gr,st);
2.
3. % function [dither]=MakeDither(DitherType,min,gr,st)
4. %
5. % Αυτή η συνάρτηση δημιουργεί θόρυβο dither μήκους 'gr' και 'st' καναλιών.
6. % Με αυτή την συνάρτηση είναι δυνατό να δημιουργηθεί θόρυβος dither RPDF
7. % εύρους [-min/2, min/2], TPDF και HighPass TPDF εύρους [-min,min].
8. % Οι παράμετροι που μπορεί να πάρει η συνάρτηση είναι:
9. %
10. % Dithertype : Με αυτό το όρισμα καθορίζεται ο τρόπος που δημιουργίας του
11. %                dither. Οι τιμές που μπορεί να πάρει είναι [1,2,3] οι οποίες
12. %                αντιστοιχούν σε:
13. %                1, RPDF με εύρος τιμών [-min/2, min/2]
14. %                2, TPDF με εύρος τιμών [-min, min]
15. %                3, HighPass TPDF με εύρος τιμών [-min, min]
16. %
17. % min:          Με αυτό το όρισμα καθορίζεται το βήμα κβαντισμού.
18. % gr:           Με αυτό το όρισμα καθορίζεται αριθμός των δειγμάτων του
19. %                dither που πρόκειται να δημιουργηθούν (κάθε δείγμα
20. %                αντιστοιχεί σε χρονικό διάστημα που ισούτε με 1/sr)
21. % st:           Με αυτό το όρισμα καθορίζονται ο αριθμός των καναλιών που
22. %                πρόκειται να δημιουργηθεί το dither.
23.
24. % Αρχικές τιμές
25. a1=3453;
26. a2=2945;
27. m=65536;
28. c1=1;
29. c2=1;
30. dithran1=1531;
31. dithran2=18531;
32. dithtemp1=0;
33. dithtemp2=0;

```

```

34. dithtemp=0;
35. dith=[];
36.
37.
38. % ===== Γεννήτρια RPDF Dither =====
39.
40. if (dithtype==1)
41.     a=min/65536;
42.     b=a*32767;
43.     for j=1:st,
44.         for i=1:gr,
45.             dithran1=mod((dithran1*a1+c1),m);
46.             dith(i,j)=a*dithran1-b;
47.         end
48.     end
49. end
50. % =====
51.
52. % ===== Γεννήτρια TPDF Dither =====
53. if (dithtype==2)
54.     a=min/(65536*2);
55.     b=a*32767;
56.     for j=1:st,
57.         for i=1:gr,
58.             dithran1=mod((dithran1*a1+c1),m);
59.             dithran2=mod((dithran2*a2+c2),m);
60.             dithtemp1=a*dithran1-b;
61.             dithtemp2=a*dithran2-b;
62.             dith(i,j)=(dithtemp1+dithtemp2);
63.         end
64.     end
65. end
66. % =====
67.
68. % ===== Γεννήτρια High Pass TPDF Dither =====
69. if (dithtype==3)
70.     a=min/(65536);
71.     b=a*32767;
72.     for j=1:st,
73.         for i=1:gr,
74.             dithran1=mod((dithran1*a1+c1),m);
75.             dithtemp1=a*dithran1-b;
76.             dith(i,j)=(dithtemp1-dithtemp2)/2;
77.             dithtemp2=dith(i,j);

```



```
78.         end
79.     end
80. end
81. % =====
```


Παράρτημα 9 Κώδικας Matlab για κωδικοποίηση του σήματος A-Format Ambisonic σε B-Format.

```

1. function [OutSignal]=encodeabformat(FUL,BDL,FDR,BUR)
2.
3. %function [OutSignal]=encodeabformat(FUL,BDL,FDR,BUR)
4. %Αυτή η συνάρτηση κάνει την κωδικοποίηση των τεσσάρων σημάτων ενός A-Format
5. %Ambisonic (FUL, BDL, BUR, FDR) στα τέσσερα σήματα του B-format (W,Z,Y,Z).
6. %Αυτό που εξάγεται από αυτή την συνάρτηση είναι ένα τετρακάνναλο σήμα, όπου
7. %το πρώτο κανάλι (το κάθε κανάλι αντιστοιχεί στον ίδιο αριθμό στήλης του
8. %σήματος εξόδου) αντιστοιχεί στο σήμα W, το δεύτερο στο X, τρίτο στο Y
9. %και το τελευταίο στο Z. Θεωρήθηκε ότι αυτή η εφαρμογή πρόκειται να
10. %λειτουργήσει στα 16 bit, με δειγματοληψία 44100 και διάστημα αναφοράς
11. %των εντασεων στο |Smax|=1. Επίσης κατά την διαδικασία της κβαντοποίησης
12. %πρόκειται να χρησιμοποιηθεί κβαντιστής με dither τύπου TPDF και χρήση
13. %τεχνικής διαμόρφωσης θορύβου 3ης τάξης, ενώ θα γίνει κβαντοποίηση Midtreat.
14.
15. %Αν κάποιος ήθελε τις προηγούμενες συνθήκες να τις τροποποιήσει ή να τις
16. %ορίζει από την συνάρτηση, αρκεί να προσθέσει ορίσματα στην συνάρτηση
17. %εισοδου ή να αλλάξει τις αρχικές τιμές.
18.
19. %Αρχικές τιμές
20. OutSignal=[];
21. N=16;
22. MidriserFlag=0;
23. NSFlag=3;
24. Smax=1;
25. DitherFlag=2;
26. W=[];
27. X=[];
28. Y=[];
29. Z=[];
30. ALL=[];
31.
32. %Συνάρτηση μετατροπής του σήματος του A-Format Ambisonic σε B-Format.
33. W=0.3333*(FUL+BDL+BUR+FDR);
34. X=1.7320*(FUL-BDL+FDR-BUR);

```

```
35. Y=1.7320*(FUL+BDL-FDR-BUR);
36. Z=1.7321*(FUL-BDL-FDR+BUR);
37. %Δημιουργία ενός τετρακάναλου σήματος.
38. ALL=[W,X,Y,Z];
39. %Κβαντοποίηση όλων των σημάτων.
40. [OutSignal]=Quantiser(ALL,N,Smax,MidriserFlag,NSFlag,DitherFlag);
```

Παράρτημα 10 Κώδικας Matlab για σύνθεση του σήματος B-Format Ambisonic.

1. `function [OutSignal]=synthesisbformatkin(InSignal,Azimuth,Elevation)`
- 2.
3. `%function [OutSignal]=synthesisbformatkin(InSignal,Theta,Azimuth)`
4. %Αυτή η συνάρτηση συνθέτει τα τέσσερα σήματα του B-format (W,Z,Y,Z) βάση ενός
5. %μονοφωνικού σήματος και της οριζόντιας και κατακόρυφης γωνίας εντοπισμού του.
6. %Οι τιμές των εισόδων Elevation και Azimuth αντιστοιχούν στην κατακόρυφη
7. %φ και την αζιμούθια θ βάση των σφαιρικών συντεταγμένων. Το Elevation και το Azimuth
8. %είναι πίνακες με σύνταξη [τιμή(1) θέσητιμής(2) τιμή(2) θέσητιμής(3)...]. Η θέση
9. %της τελικής τιμής και τον δύο πινάκων πρέπει να είναι μεγαλύτερη του 200
10. %και θα καθορίσει
11. %και το μέγεθος (γραμμές) του σήματος εξόδου. Όπως φαίνεται από την σύνταξη οι ελάχιστες
12. %στήλες που πρόκειται να έχουν αυτοί οι πίνακες είναι 3. Στη περίπτωση που το
13. %σημα δεν είναι μονοφωνικό η συνάρτηση θα λάβει ως σήμα προς επεξεργασία το
14. %πρώτο κανάλι αυτού. Σημείωση: το σημα εισόδου πρέπει να έχει μήκος
15. %μεγαλύτερο των 200 δειγμάτων. Ταυτοχρονα η μέγιστη τιμή
16. %Οι τιμές που παίρνουν είναι για το Azimuth [0,360] ενώ για το Elevation [-90,90].
17. %Αυτό που εξάγεται από αυτή την συνάρτηση είναι ένα τετρακάναλο σήμα, όπου
18. %το πρώτο κανάλι (το κάθε κανάλι αντιστοιχεί στον ίδιο αριθμό στήλης του
19. %σήματος εξόδου) αντιστοιχεί στο σήμα W, το δεύτερο στο X, τρίτο στο Y
20. %και το τελευταίο στο Z. Θεωρήθηκε ότι αυτή η εφαρμογή πρόκειται να
21. %λειτουργήσει στα 16 bit, με δειγματοληψία 44100 και διάστημα αναφοράς
22. %των εντασεων στο |Smax|=1. Επίσεις κατά την διαδικασία της κβαντοποίησης
23. %πρόκειται να χρησιμοποιηθεί κβαντιστής με dither τύπου TPDF και χρήση
24. %τεχνικής διαμόρφωσης θορύβου 3ης τάξης, ενώ θα γίνει κβαντοποίηση Midtreat.
- 25.
26. %Αν κάποιος ήθελε τις προηγούμενες συνθήκες να τις τροποποιήσει ή να τις
27. %ορίζει από την συνάρτηση, αρκεί να προσθέσει ορίσματα στην συνάρτηση
28. %εισοδου ή να αλλάξει τις αρχικές τιμές.
- 29.
30. %Αρχικές τιμές
31. `OutSignal=[];`
32. `N=16;`
33. `MidriserFlag=0;`
34. `NSFlag=3;`

```

35. Smax=1;
36. DitherFlag=2;
37. W=[];
38. X=[];
39. Y=[];
40. Z=[];
41. ALL=[];
42. r=1;
43. w1=[];
44. elevtable=[];
45. w3=[];
46. azimtable=[];
47. w2=0;
48. w4=0;
49. %Ακύρωση των υπόλοιπων καναλιών του InSignal αν υπάρχουν, καθώς και των γραμμών
50. %Azimuth,Elevation.
51. InSignal=InSignal(:,1);
52. Azimuth=Azimuth(1,:);
53. Elevation=Elevation(1,:);
54. %Δημιουργία πίνακα fadein και fadeout 100 δειγμάτων.
55. fadeout=1:-1/99:0;
56. fadeout=fadeout';
57. fadein=1-fadeout;
58.
59. [a,b]=size(Azimuth);
60. [c,d]=size(Elevation);
61. [e,g]=size(InSignal);
62. clear a c g
63. %Έλεγχος για το αν το μήκος των πινάκων Azimuth και Elevation είναι μονός
64. %αριθμός, εάν είναι συνταγμένος σωστά και αν οι το μέγεθος των πινάκων
65. %που πρόκειται να δημιουργηθούν θα έχει το ίδιο μήκος.
66. if(or(mod(b,2)==0,mod(d,2)==0))
67.     error('λάθος σύνταξη στο Azimuth ή στο Elevation')
68. end
69. if(or(b<3,d<3))
70.     error('λάθος σύνταξη στο Azimuth ή στο Elevation')
71. end
72.
73. w5=0;
74. for i=1:2:b-2,
75.     if (w5>=Azimuth(i+1))
76.         error('λάθος σύνταξη στο Azimuth')
77.     else w5=(Azimuth(i+1));
78.     end

```

```

79. end
80.
81. w6=0;
82. for i=1:2:d-2,
83.     if (w6>=Elevation(i+1))
84.         error('λάθος σύνταξη στο Elevation')
85.     else w6=Elevation(i+1);
86.     end
87. end
88.
89. if (Azimuth(b-1)~=Elevation(d-1))
90.     error('δεν υπάρχει συμφωνία στο μέγεθος του σήματος που πρόκειται να δημιουργηθεί')
91. end
92.
93. if (or(e<=200,Azimuth(b-1)<=200))
94.     error('το μέγεθος του σήματος εισόδου καθώς και οι μέγιστη τιμή (θέσης τιμής) του Azimuth
    και Elevation πρέπει να είναι μεγαλύτερη του 200')
95. end
96. %Έλεγχος αν η γωνία 'φ' ξεπερνάνε τα επιτρεπτά όρια
97. for i=1:2:d,
98.     if (abs(Elevation(i))>90)
99.         error('Υπάρχει τιμή Elevation εκτός των επιτρεπτών ορίων ')
100.    end
101. end
102. clear w5 w6
103. %Δημιουργία πίνακα για τις τιμές του Azimuth.
104. for i=1:2:b-2,
105.     if (Azimuth(i)==Azimuth(i+2))
106.         w1=[];
107.         w1(1:(Azimuth(i+1)-w2-r+1):Azimuth(i));
108.     else
109. w1=Azimuth(i):(Azimuth(i+2)-Azimuth(i))/(Azimuth(i+1)-w2-r):Azimuth(i+2);
110.     end
111. azimtable=[azimtable,w1(2-1*r:Azimuth(i+1)-w2+1-1*r)];
112. w2=Azimuth(i+1);
113. r=0;
114. end
115. clear w2 w1
116. azimtable=azimtable';
117. %Δημιουργία πίνακα για τις τιμές του elevation.
118. r=1;
119. for i=1:2:d-2,
120.     if (Elevation(i)==Elevation(i+2))
121.         w1=[];

```

```

122.     wl(1:(Elevation(i+1)-w4-r+1))=Elevation(i);
123.     else
124. wl=Elevation(i):(Elevation(i+2)-Elevation(i))/(Elevation(i+1)-w4-r):Elevation(i+2);
125.     end
126. elevtable=[elevtable,wl(2-1*r:Elevation(i+1)-w4+1-1*r)];
127. w4=Elevation(i+1);
128. r=0;
129. end
130. clear w1 w4 r
131. elevtable=elevtable';
132.
133. %Δημιουργία συνολικού σήματος εισόδου.
134. if Azimuth(b-1)<=e
135.     InSignalTel=InSignal(1:Azimuth(b-1));
136.     InSignalTel(Azimuth(b-1)-99:Azimuth(b-1))=InSignal(Azimuth(b-1)-99:Azimuth(b-1)).*fadeout;
137.     clear InSignal
138. else x1=floor(Azimuth(b-1)/(e-100));
139.     x2=mod(Azimuth(b-1),(e-100));
140.     InSignalCross=InSignal(1:e-100);
141.     InSignalFirst=InSignalCross;
142.     InSignalCross(1:100)=InSignalCross(1:100).*fadein+InSignal(e-99:e).*fadeout;
143.     w8=[];
144.     InSignalTel=InSignalFirst;
145.     for i=1:x1,
146.         InSignalTel=[InSignalTel;w8];
147.         w8=InSignalCross;
148.     end
149.     clear w8
150.     sizeInSignal=x1*(e-100)+x2;
151.     InSignalTel=[InSignalTel;InSignalCross(1:x2)];
152.     InSignalTel(sizeInSignal-99:sizeInSignal)=InSignalTel(sizeInSignal-99:sizeInSignal).*fadeout;
153. end
154. clear InSignal InSignalCross InSignalFirst
155.
156. %Συνάρτηση σύνθεσης του σήματος B-Format Ambisonic.
157. W=(1/(2^0.5))*InSignalTel;
158. X=cosd(azimtable).*cosd(elevtable).*InSignalTel;
159. Y=sind(azimtable).*cosd(elevtable).*InSignalTel;
160. Z=sind(azimtable).*InSignalTel;
161. clear InSignalTel
162. %Δημιουργία ενός τετρακάναλου σήματος.
163. ALL=[W,X,Y,Z];

```


164. *clear W X Y Z*

165. *%Κβαντοποίηση όλων των σημάτων.*

166. *[OutSignal]=Quantiser(ALL,N,Smax,MidriserFlag,NSFlag,DitherFlag);*

Παράρτημα 11 Κώδικας Matlab για αποκωδικοποίηση ηχητικού πεδίου σε διάταξη Ambisonics με δυνατότητα κωδικοποίησης και για Binaural.

1. `function`
`[OutSignal,binaural]=decoderwithbinaural(InSignal,computebinauralflag,coefficientindex)`
- 2.
3. `%function`
`[OutSignal,binaural]=decoderwithbinaural(InSignal,computebinauralflag,coefficientindex)`
4. `%Αυτή η συνάρτηση υπολογίζει τα σηματα που θα τροφοδοτήσουν μια συστηχία ηχειων, βάση του σήματος`
5. `%B-Format και των κατάλληλων συντελεστών, ανάλογα με την διάταξη των`
6. `%ηχειων. Ταυτοχρονα έχει την δυνατότητα να εξάγει ένα στερεοφωνικό σήμα που`
7. `%θα είναι η κωδικοποίηση του τελικού ηχητικού πεδίου που δημιουργείται κατά την Ambisonic`
8. `%αναπαραγωγή, σε Binaural. Τα σήματα εξόδου είναι δύο, το OutSignal το`
9. `%οποίο αποτελείται από N κανάλια (στήλες) τα οποία αντιστοιχούν στα σήματα`
10. `%που πρέπει θα τροφοδοτηθεί η διάταξη των ηχειων για την ανασύνθεση του`
11. `%ηχητικού πεδίου, και ένα στερεοφωνικό σήμα (δύο στήλες) το οποίο είναι η`
12. `%κωδικοποίηση του τελικού ηχητικού πεδίου σε Ambisonic. Τα ορίσματα εισόδου`
13. `%είναι τρία, το 'InSignal'το οποίο αποτελείται από τα τέσσερα κανάλια`
14. `% (τέσσερις στήλες) του B-Format ambisonic,το 'computebinauralflag' που`
15. `%παιρνει τιμές (1,0) και καθορίζει αν θα γίνει υπολογισμός του σήματος`
16. `%binaural ή όχι αντίστοιχα, και τέλος το 'coefficientindex' ο οποίος είναι`
17. `%ένας πίνακας με δυο ομάδες συντελεστών των (W,X,Y,Z) (για υψηλές και`
18. `%χαμηλές συχνότητες)για το κάθε ηχείο και την αζιμουθια και κατακόρυφη γωνία αυτού.`
19. `%Ο πίνακας 'coefficientindex' αποτελείται από τόσες γραμμές όσα είναι και τα ηχεί αυτού`
20. `%, ενώ η σύνταξη κάθε γραμμής είναι`
21. `%(azimuth, elevation, WH,XH,YH,ZH,WL,XL,YL,ZL) όπου (WH,XH,YH,ZH) και (WL,XL,YL,ZL) οι`
22. `%συντελεστές των σημάτων (W,X,Y,Z) για την δημιουργια των σημάτων των ηχείων για`
23. `%τις υψηλές και χαμηλές συχνότητες αντίστοιχα.`
24. `%Οι τιμές που παίρνουν είναι για το azimuth [0,360] ενώ για το elevation [-90,90].`
25. `%θεωρήθηκε ότι αυτή η εφαρμογή πρόκειται να`
26. `%λειτουργήσει στα 16 bit, με δειγματοληψία 44100 και διάστημα αναφοράς`
27. `%των εντασεων στο |Smax|=1. Επίσεις κατά την διαδικασία της κβαντοποίησης`
28. `%πρόκειται να χρησιμοποιηθεί κβαντιστής με dither τύπου TPDF και χρήση`
29. `%τεχνικής διαμόρφωσης θορύβου 3ης τάξης, ενώ θα γίνει κβαντοποίηση Midtreat.`
- 30.
31. `%Αν κάποιος ήθελε τις προηγούμενες συνθήκες να τις τροποποιήσει ή να τις`

```

32. %ορίζει από την συνάρτηση, αρκεί να προσθέσει ορίσματα στην συνάρτηση
33. %εισοδου ή να αλλάξει τις αρχικές τιμές.
34.
35. %Αρχικές τιμές
36. OutSignalall=[];
37. N=16;
38. MidriserFlag=0;
39. NSFlag=3;
40. Smax=1;
41. DitherFlag=2;
42. ALL=[];
43. binaural=zeros(size(InSignal,1),2);
44.
45. % Μεγέθη ορισμάτων εισόδου
46. [a,b]=size(InSignal);
47. [e,g]=size(coefficentindex);
48.
49. %Έλεγχος για το κατα πόσον τα ορίσματα εισόδου είναι σωστά, ή έχουν κάποιο
50. %λάθος.
51.
52. %Έλεγχος για το αν το σήμα εισόδου έχει 4 κανάλια, ώστε να είναι ένα
53. %Ambisonic B-Format.
54. if (b~=4)
55.     error('Το σήμα εισόδου δεν αποτελείται από τέσσερα κανάλια (W,X,Y,Z)')
56. end
57.
58. %Έλεγχος αν το σημα 'coefficentindex' αποτελείται από δέκα στήλες
59. %(azimuth, elevation,WH,XH,YH,ZH,WL,XL,YL,ZL)
60. if (g~=10)
61.     error('Το coefficentindex αποτελείται από δέκα στήλες εκ των οποίων (azimuth,
        elevation,WH,XH,YH,ZH,WL,XL,YL,ZL)')
62. end
63.
64. %Έλεγχος στο σημα 'coefficentindex' αν το elevation παίρνει τιμές εκτός
65. %ορίων.
66. if (max(abs(coefficentindex(:,2))>90)
67.     error('Στο coefficentindex το elevation παίρνει τιμές στο διάστημα [-90,90]')
68. end
69.
70. %Δημιουργία φίλτρου lowpass (σχεδιασμένο με το fdatool του matlab)
71.
72. Fs = 44100; % Sampling Frequency
73.
74. Fpass = 400; % Passband Frequency

```

```

75. Fstop = 500;                % Stopband Frequency
76. Dpass = 0.028774368332;    % Passband Ripple
77. Dstop = 3.1622776602e-005; % Stopband Attenuation
78. dens = 20;                 % Density Factor
79.
80. % Calculate the order from the parameters using FIRPMORD.
81. [N, Fo, Ao, W] = firpmord([Fpass, Fstop]/(Fs/2), [1 0], [Dpass, Dstop]);
82.
83. % Calculate the coefficients using the FIRPM function.
84. lowpass = firpm(N, Fo, Ao, W, {dens});
85. lowpass=lowpass';
86. clear Fs Fstop Fpass Dstop Dpass dens N Fo Ao W
87.
88. %Δημιουργία φίλτρου highpass (σχεδιασμένο με το fdatool του matlab)
89.
90. % Equiripple Highpass filter designed using the FIRPM function.
91.
92. % All frequency values are in Hz.
93. Fs = 44100; % Sampling Frequency
94.
95. Fstop = 400;                % Stopband Frequency
96. Fpass = 500;                % Passband Frequency
97. Dstop = 3.1622776602e-005; % Stopband Attenuation
98. Dpass = 0.028774368332;    % Passband Ripple
99. dens = 20;                 % Density Factor
100.
101. % Calculate the order from the parameters using FIRPMORD.
102. [N, Fo, Ao, W] = firpmord([Fstop, Fpass]/(Fs/2), [0 1], [Dstop, Dpass]);
103.
104. % Calculate the coefficients using the FIRPM function.
105. highpass = firpm(N, Fo, Ao, W, {dens});
106. highpass=highpass';
107.
108. clear Fs Fstop Fpass Dstop Dpass dens N Fo Ao W
109.
110. %Φιλτράρισμα των σημάτων W,X,Y,Z για τις υψηλές και χαμηλές συχνότητες.
111. Whigh=conv(highpass,InSignal(:,1));
112. Whigh=Whigh((size(highpass(:,1),1)-1)/2+1:size(Whigh(:,1),1)-(size(highpass(:,1),1)-1)/2);
113. Xhigh=conv(highpass,InSignal(:,2));
114. Xhigh=Xhigh((size(highpass(:,1),1)-1)/2+1:size(Xhigh(:,1),1)-(size(highpass(:,1),1)-1)/2);
115. Yhigh=conv(highpass,InSignal(:,3));
116. Yhigh=Yhigh((size(highpass(:,1),1)-1)/2+1:size(Yhigh(:,1),1)-(size(highpass(:,1),1)-1)/2);
117. Zhigh=conv(highpass,InSignal(:,4));
118. Zhigh=Zhigh((size(highpass(:,1),1)-1)/2+1:size(Zhigh(:,1),1)-(size(highpass(:,1),1)-1)/2);

```

```

119.Wlow=conv(lowpass,InSignal(:,1));
120.Wlow=Wlow((size(lowpass(:,1),1)-1)/2+1:size(Wlow(:,1),1)-(size(lowpass(:,1),1)-1)/2);
121.Xlow=conv(lowpass,InSignal(:,2));
122.Xlow=Xlow((size(lowpass(:,1),1)-1)/2+1:size(Xlow(:,1),1)-(size(lowpass(:,1),1)-1)/2);
123.Ylow=conv(lowpass,InSignal(:,3));
124.Ylow=Ylow((size(lowpass(:,1),1)-1)/2+1:size(Ylow(:,1),1)-(size(lowpass(:,1),1)-1)/2);
125.Zlow=conv(lowpass,InSignal(:,4));
126.Zlow=Zlow((size(lowpass(:,1),1)-1)/2+1:size(Zlow(:,1),1)-(size(lowpass(:,1),1)-1)/2);
127.
128. %Δημιουργία ενός συνολικού σήματος
129.all_filter=[Whigh,Xhigh,Yhigh,Zhigh,Wlow,Xlow,Ylow,Zlow];
130. %Εύρεση μέγιστης απόλυτης τιμής από όλα τα σήματα (στήλες) του all_filter
131.norm=max(max(abs(all_filter)));
132. %Normalize, διαίρεση όλων των τιμών του πίνακα all_filter με την μέγιστη απόλυτη τιμή αυτού.
133.all_filter=all_filter*norm;
134.clear norm
135. %Δημιουργία συνολικού σήματος εξόδου για όλα τα ηχεία.
136.for i=1:e,
137.     S=(all_filter(:,1)*coefficientindex(i,3)+...
138.         all_filter(:,2)*coefficientindex(i,4)*cosd(coefficientindex(i,
139.         1))*cosd(coefficientindex(i,2))+...
140.         all_filter(:,3)*coefficientindex(i,5)*sind(coefficientindex(i,
141.         1))*cosd(coefficientindex(i,2))+...
142.         all_filter(:,4)*coefficientindex(i,6)*sind(coefficientindex(i,2))+...
143.         all_filter(:,5)*coefficientindex(i,7)+...
144.         all_filter(:,6)*coefficientindex(i,8)*cosd(coefficientindex(i,
145.         1))*cosd(coefficientindex(i,2))+...
146.         all_filter(:,7)*coefficientindex(i,9)*sind(coefficientindex(i,
147.         1))*cosd(coefficientindex(i,2))+...
148.         all_filter(:,8)*coefficientindex(i,10)*sind(coefficientindex(i,2)));
149.     OutSignalall=[OutSignalall S];
150.end
151. %Normalize σήματος εξόδου.
152.norm=max(max(abs(OutSignalall)));
153.OutSignalall=OutSignalall/norm;
154.
155. %Κβαντοποίηση όλων των σημάτων για τα ηχεία.
156.N=16;
157.[OutSignal]=Quantiser(OutSignalall,N,Smaz,MidriserFlag,NSFlag,DitherFlag);
158.clear OutSignalall
159.
160. %Binaural
161.TotL=zeros(size(OutSignal(:,1),1),1);
162.TotR=zeros(size(OutSignal(:,1),1),1);
163.if (computebinauralflag==1)

```

```

160.     i=find(abs(coefficentindex(:,1))>=360);
161.     coefficentindex(i)=mod(coefficentindex(i),360);
162.     clear i
163.     qw=find(coefficentindex(:,2)<=-45);
164.     if (numel(qw)~=0)
165.         error ('Για να γίνει κωδικοποίηση σε binaural τα ηχεία δεν πρέπει να έχουν γωνία
ανύψωσης μικρότερη των -45 μοιρών')
166.     end
167.         %Κβαντοποίηση γωνιών ηχείων σε αυτές που είναι δυνατές
168. for i=1:e,
169.     if (abs(coefficentindex(i,2))<25)
170.         st=round(coefficentindex(i,2)/10)*10;
171.         gr=round(coefficentindex(i,1)/5)*5;
172.     elseif (abs(coefficentindex(i,2))<35)
173.         st=round(coefficentindex(i,2)/10)*10;
174.         gr=round(coefficentindex(i,1)/6)*6;
175.     elseif (abs(coefficentindex(i,2))<45)
176.         st=round(coefficentindex(i,2)/10)*10;
177.         gr=round(coefficentindex(i,1)/6.5)*6.5;
178.     elseif (coefficentindex(i,2)<55)
179.         st=round(coefficentindex(i,2)/10)*10;
180.         gr=round(coefficentindex(i,1)/8)*8;
181.     elseif (coefficentindex(i,2)<65)
182.         st=round(coefficentindex(i,2)/10)*10;
183.         gr=round(coefficentindex(i,1)/10)*10;
184.     elseif (coefficentindex(i,2)<75)
185.         st=round(coefficentindex(i,2)/10)*10;
186.         gr=round(coefficentindex(i,1)/15)*15;
187.     elseif (coefficentindex(i,2)<85)
188.         st=round(coefficentindex(i,2)/10)*10;
189.         gr=round(coefficentindex(i,1)/30)*30;
190.     else
191.         st=round(coefficentindex(i,2)/10)*10;
192.         gr=0;
193.     end
194.     if gr==0
195.         %to sprintf γράφει την γραμμή με τις μεταβλητές ενώ το cellstr βάζει σε αυτή ''
196.         L=cellstr(sprintf('C:\\diplomatiki\\references\\full\\L%de000a.wav \n', st));
197.         elseif gr<100
198.         L=cellstr(sprintf('C:\\diplomatiki\\references\\full\\L%de0%da.wav \n', st,gr));
199.     else
200.         L=cellstr(sprintf('C:\\diplomatiki\\references\\full\\L%de%da.wav \n', st,gr));
201.     end
202.

```

```

203.
204.     if gr==0
205.         %to sprintf γράφει την γραμμή με τις μεταβλητές ενώ το cellstr
206.         %βάζει σε αυτή ''
207.         R=cellstr(sprintf('C:\\diplomatiki\\references\\full\\R%de000a.wav \n', st));
208.         elseif gr<100
209.             R=cellstr(sprintf('C:\\diplomatiki\\references\\full\\R%de0%da.wav \n', st,gr));
210.         else
211.             R=cellstr(sprintf('C:\\diplomatiki\\references\\full\\R%de%da.wav \n', st,gr));
212.
213.     end
214.     Left=wavread(char(L(1)));
215.     Right=wavread(char(R(1)));
216.     OutL=conv(Left,OutSignal(:,i));
217.     OutL=OutL(1:size(OutL,1)-size(Left,1)+1);
218.     TotL=TotL+OutL;
219.     OutR=conv(Right,OutSignal(:,i));
220.     OutR=OutR(1:size(OutR,1)-size(Right,1)+1);
221.     TotR=TotR+OutR;
222.
223.
224. end
225.
226.
227.end
228.
229. binaural=[TotL,TotR];
230. %Normalize σήματος εξόδου.
231. norm=max(max(abs(binaural)));
232. binaural=binaural/norm;
233. %κβαντοποίηση του σήματος
234. [binaural]=Quantiser(binaural,N,Smax,MidriserFlag,NSFlag,DitherFlag);
235.end

```


Παράρτημα 12 Κώδικας Matlab για τον σχηματισμό του πίνακα των συντελεστών, για την δημιουργία των σημάτων που πρέπει να τροφοδοτηθούν σε κάθε ηχείο, κατά την διαδικασία της αποκωδικοποίησης.

```

1. function [coefficientindextable,table_rv_re,dH,dL]=coefficientindex(todoflag,speakers)
2.
3. %function [coefficientindextable,table_rv_re,dH,dL]=coefficientindex(todoflag,speakers)
4. %Αυτή η συνάρτηση δημιουργεί τον πίνακα με τις γωνίες (αζιμούθια και ανύψωσης) που έχουν
5. %τα ηχεία και τους συντελεστές με τους οποίους πρέπει να προστεθούν οι σφαιρικές αρμονικές
6. %ώστε να προκύψει το σήμα που πρέπει να τροφοδοτηθεί το κάθε ηχείο. Ταυτόχρονα μπορούν να
7. %εξαχθούν και τα rv και re ώστε να μπορεί να γίνει ψυχοακουστική αποτίμηση του συστήματος.
8. %Μπορούν επίσης να εξαχθούν οι τιμές των dH και dL για οι οποίες είναι οι βέλτιστες στη
9. %περίπτωση που η τιμή todoflag=4.Η μορφή που θα έχει το σήμα εξόδου 'coefficientindextable'
    θα είναι :
10. %[azimuth, elevation, WH,XH,YH,ZH,WL,XL,YL,ZL], ενώ το table_rv_re είναι
11. %[azim rHv rHe VHzim EHazim rLv rLe VLazim ELazim]; όπου οι λέξεις τα
12. %ορίσματα που περιέχουν το azim αφορούν την αντιληπτή γωνία προέλευσης του
13. %ήχου, ενώ η πρώτη τιμή ορίζει την πραγματική γωνία.
14. %Τα σήματα εξόδου είναι δύο, τα (todoflag,speakers). Ο χρήστης με το
15. %πρωτο όρισμα 'todoflag' μπορεί να ορίσει με ποιά μεθοδολογία θα υπολογιστούν οι
16. %συντελεστες. Εκτός από την περίπτωση που ο χρήστης ορίσει todoflag=1, θα πρέπει
17. %να δώσει καταλληλο αριθμό ηχείων ώστε να υπολογιστούν οι συντελεστές (διάταξη ηχειων στο
    οριζόντιο
18. %επίπεδο). Ας δούμε όμως σε τι αντιστοιχούν οι τιμές του ορίσματος 'todoflag':
19. %'todoflag'          1, πίνακες σύμφωνα με τον Gerzon για 5 ηχεία με ηχεία
20. %                    τοποθετημένα στις 0,45,-45,130,-130 και 0 γωνία
21. %                    ανυψησης σε όλα.
22. %                    2, Για αίθουσες συναυλιών σύμφωνα με τον
23. %                    D.Malham (οριζόντιο επίπεδο)
24. %                    3, Για studio σύμφωνα με τον Gerzon (οριζόντιο επίπεδο)
25. %                    4, Farina et al, 2001 (γίνεται εύρεση από το σύστημα
26. %                    της καταλληλότερης τιμής για το d)
27.
28.
29. %καθορισμός αρχικής γωνίας και βήματος.
30. if (todoflag==1)

```

```

31. elseif (nargin==1)
32.     error('πρέπει να ορίσεις τον αριθμό των ηχείων που επιθυμείς και να είναι μεγαλύτερος
του 4')
33. elseif (speakers<4)
34.     error('πρέπει να ορίσεις τον αριθμό των ηχείων που επιθυμείς και να είναι μεγαλύτερος
του 4')
35. else
36.     bhma=(360/speakers);
37.     if (mod(speakers,2)==0)
38.         f0=bhma/2;
39.     else
40.         f0=0;
41.     end
42. end
43.
44.
45. %έλεγχος για τις τιμές του todoflag
46. if (or(todoflag==4,or(todoflag==1,or(todoflag==2,todoflag==3))))
47. else
48.     error('Οι τιμές που μπορεί να πάρει το όρισμα todoflag μπορεί να είναι μεταξύ των τιμών
(1,2,3,4)')
49. end
50.
51. %Αρχικές τιμές.
52. coefficientindextable=[];
53. table_rv_re=[];
54. PH=zeros(24,1);
55. VHx=zeros(24,1);
56. VHy=zeros(24,1);
57. EH=zeros(24,1);
58. EHx=zeros(24,1);
59. EHy=zeros(24,1);
60. PL=zeros(24,1);
61. VLx=zeros(24,1);
62. VLy=zeros(24,1);
63. EL=zeros(24,1);
64. ELx=zeros(24,1);
65. ELy=zeros(24,1);
66. rHv=[];
67. rHe=[];
68. rLv=[];
69. rLe=[];
70. azim=[];
71. dH=[];
72. dL=[];

```

```

73. bathtelH=ones(360,4);
74. bathtelL=ones(360,4);
75. coefficientindextabletest=[];
76. %Έλεγχος για το ποιό πίνακα θα φτιάξει
77. if todoflag==1
78.
79.     %Δημιουργία πίνακα συντελεστών
80.     coefficientindextable=[0 0 0.38324 0.37228 0 0 0.34190 0.23322 0 0;...
81.                             45 0 0.44022 0.3307 0.7650 0 0.26813 0.5401 0.7146 0;...
82.                             130 0 0.78238 0.8607 0.5532 0 0.56092 0.7756 0.5961 0;...
83.                             230 0 0.78238 0.8607 0.5532 0 0.56092 0.7756 0.5961 0;...
84.                             315 0 0.44022 0.3307 0.7650 0 0.26813 0.5401 0.7146 0];
85.     %Δημιουργία πίνακα rv και re (Μόνο για οριζόντιο επίπεδο).
86.     for i=1:5,
87.         for j=1:24,
88.             %Δημιουργία σημάτων B-Format
89.             %Δημιουργία των αθροισμάτων (συνολικό σήμα για την πίεση και την ένταση).
90.             PH(j,1)=PH(j,1)+coefficientindextable(i,
91. 3)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,4)*...
92.             cosd(coefficientindextable(i,1))+...
93.             sind((j-1)*15)*coefficientindextable(i,
94. 5)*sind(coefficientindextable(i,1));
95.             VHx(j,1)=VHx(j,1)+(coefficientindextable(i,
96. 3)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,4)*...
97.             cosd(coefficientindextable(i,1))+...
98.             sind((j-1)*15)*coefficientindextable(i,5)*sind(coefficientindextable(i,
99. 1))) *...
100.            cosd(coefficientindextable(i,1));
101.            VHy(j,1)=VHy(j,1)+(coefficientindextable(i,
102. 3)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,4)*...
103.            cosd(coefficientindextable(i,1))+...
104.            sind((j-1)*15)*coefficientindextable(i,5)*sind(coefficientindextable(i,
105. 1))) *...
106.            sind(coefficientindextable(i,1));
107.            VLx(j,1)=VLx(j,1)+(coefficientindextable(i,
108. 7)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,8)*...
109.            cosd(coefficientindextable(i,1))+...

```

```

110.          sind((j-1)*15)*coefficientindextable(i,9)*sind(coefficientindextable(i,
111.          1)))*...
112.          sind(coefficientindextable(i,1));
113.          EH(j,1)=EH(j,1)+(coefficientindextable(i,
114.          3)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,4)*...
115.          cosd(coefficientindextable(i,1))+...
116.          sind((j-1)*15)*coefficientindextable(i,5)*sind(coefficientindextable(i,
117.          1)))^2;
118.          EHX(j,1)=EHX(j,1)+((coefficientindextable(i,
119.          3)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,4)*...
120.          cosd(coefficientindextable(i,1))+...
121.          sind((j-1)*15)*coefficientindextable(i,5)*sind(coefficientindextable(i,
122.          1)))^2)*...
123.          cosd(coefficientindextable(i,1));
124.          EHY(j,1)=EHY(j,1)+((coefficientindextable(i,
125.          3)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,4)*...
126.          cosd(coefficientindextable(i,1))+...
127.          sind((j-1)*15)*coefficientindextable(i,5)*sind(coefficientindextable(i,
128.          1)))^2)*...
129.          sind(coefficientindextable(i,1));
130.          EL(j,1)=EL(j,1)+(coefficientindextable(i,
131.          7)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,8)*...
132.          cosd(coefficientindextable(i,1))+...
133.          sind((j-1)*15)*coefficientindextable(i,9)*sind(coefficientindextable(i,
134.          1)))^2;
135.          ELX(j,1)=ELX(j,1)+((coefficientindextable(i,
136.          7)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,8)*...
137.          cosd(coefficientindextable(i,1))+...
138.          sind((j-1)*15)*coefficientindextable(i,9)*sind(coefficientindextable(i,
139.          1)))^2)*...
140.          cosd(coefficientindextable(i,1));
141.          ELY(j,1)=ELY(j,1)+((coefficientindextable(i,
142.          7)*2^(-0.5)+cosd((j-1)*15)*coefficientindextable(i,8)*...
143.          cosd(coefficientindextable(i,1))+...
144.          sind((j-1)*15)*coefficientindextable(i,9)*sind(coefficientindextable(i,
145.          1)))^2)*...
146.          sind(coefficientindextable(i,1));
147.          azim(j,1)=(j-1)*15;
148.          end
149.          end
150.
151.elseif (or(todoflag==2,todoflag==3))
152.    %Αρχικές τιμές.
153.    if (todoflag==2)
154.      GH1=(8/(3*speakers))^0.5/2;
155.      GH2=(8/(3*speakers))^0.5/2;
156.      GL1=(8/(3*speakers))^0.5/2;
157.      GL2=(8/(3*speakers))^0.5/2;

```

```

146.     else
147.         GH1=(8/(4*speakers))^0.5/2;
148.         GH2=(8/(2*speakers))^0.5/2;
149.         GL1=(8/(6*speakers))^0.5/2;
150.         GL2=(16/(3*speakers))^0.5/2;
151.
152.     end
153.
154. %δημιουργία πίνακα συντελεστών
155.     for i=1:speakers,
156.         hxeio=[f0+(i-1)*bhma 0 GH1 GH2 GH2 0 GL1 GL2 GL2 0];
157.         coefficientindextable=[coefficientindextable;hxeio];
158.         for j=1:24,
159.             %Δημιουργία των αθροισμάτων (συνολικό σήμα για την πίεση και την ένταση) αφού
160.             πρώτα %δημιουργηθούν τα σήματα B-Format και γίνει το Decoding για κάθε ηχείο.
161.             PH(j,1)=PH(j,1)+hxeio(1,3)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,4)*cosd(hxeio(1,1))
162.             +...
163.             sind((j-1)*15)*hxeio(1,5)*sind(hxeio(1,1));
164.             VHx(j,1)=VHx(j,1)+
165.             (hxeio(1,3)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,4)*cosd(hxeio(1,1)))+...
166.             sind((j-1)*15)*hxeio(1,5)*sind(hxeio(1,1))*cosd(hxeio(1,1));
167.             VHy(j,1)=VHy(j,1)+
168.             (hxeio(1,3)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,4)*cosd(hxeio(1,1)))+...
169.             sind((j-1)*15)*hxeio(1,5)*sind(hxeio(1,1))*sind(hxeio(1,1));
170.             PL(j,1)=PL(j,1)+hxeio(1,7)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,8)*cosd(hxeio(1,1))
171.             +...
172.             sind((j-1)*15)*hxeio(1,9)*sind(hxeio(1,1));
173.             VLx(j,1)=VLx(j,1)+
174.             (hxeio(1,7)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,8)*cosd(hxeio(1,1)))+...
175.             sind((j-1)*15)*hxeio(1,9)*sind(hxeio(1,1))*cosd(hxeio(1,1));
176.             VLy(j,1)=VLy(j,1)+
177.             (hxeio(1,7)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,8)*cosd(hxeio(1,1)))+...
178.             sind((j-1)*15)*hxeio(1,9)*sind(hxeio(1,1))*sind(hxeio(1,1));
179.             EH(j,1)=EH(j,1)+(hxeio(1,3)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,4)*cosd(hxeio(1,1))
180.             +...
181.             sind((j-1)*15)*hxeio(1,5)*sind(hxeio(1,1)))^2;
182.             EHx(j,1)=EHx(j,1)+
183.             ((hxeio(1,3)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,4)*cosd(hxeio(1,1)))+...
184.             sind((j-1)*15)*hxeio(1,5)*sind(hxeio(1,1)))^2)*cosd(hxeio(1,1));
185.             EHy(j,1)=EHy(j,1)+
186.             ((hxeio(1,3)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,4)*cosd(hxeio(1,1)))+...
187.             sind((j-1)*15)*hxeio(1,5)*sind(hxeio(1,1)))^2)*sind(hxeio(1,1));
188.             EL(j,1)=EL(j,1)+(hxeio(1,7)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,8)*cosd(hxeio(1,1))
189.             +...
190.             sind((j-1)*15)*hxeio(1,9)*sind(hxeio(1,1)))^2;
191.             ELx(j,1)=ELx(j,1)+
192.             ((hxeio(1,7)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,8)*cosd(hxeio(1,1)))+...

```

```

182.             sind((j-1)*15)*hxeio(1,9)*sind(hxeio(1,1))^2*cosd(hxeio(1,1));
183.                                     ELy(j,1)=ELy(j,1)+
((hxeio(1,7)*2^(-0.5)+cosd((j-1)*15)*hxeio(1,8)*cosd(hxeio(1,1))+...
184.             sind((j-1)*15)*hxeio(1,9)*sind(hxeio(1,1))^2)*sind(hxeio(1,1));
185.             azim(j,1)=(j-1)*15;
186.             end
187.         end
188.
189.end
190.
191.if (todoflag~=4)
192.    %Τελικές τιμές
193.        VHx=VHx./PH;
194.        VHy=VHy./PH;
195.        VLx=VLx./PL;
196.        VLy=VLy./PL;
197.        EHx=EHx./EH;
198.        EHy=EHy./EH;
199.        ELx=ELx./EL;
200.        ELy=ELy./EL;
201.        rHv=((VHx).^2+(VHy).^2).^0.5;
202.        rHe=((EHx).^2+(EHy).^2).^0.5;
203.        rLv=((VLx).^2+(VLy).^2).^0.5;
204.        rLe=((ELx).^2+(ELy).^2).^0.5;
205.        a=find(VHx~=0);
206.        VHazim(a)=atand(VHy(a)./VHx(a));
207.        e=find(VHx(a)<0);
208.        VHazim(a(e))=VHazim(a(e))+180;
209.        b=find(VHx==0);
210.        c=find(VHy(b)<0);
211.        VHazim(b(c))=270;
212.        b(c)=[];
213.        VHazim(b)=90;
214.        VHazim=VHazim';
215.        a=find(VHazim<0);
216.        VHazim(a)=360+VHazim(a);
217.        a=find(EHx~=0);
218.        EHazim(a)=atand(EHy(a)./EHx(a));
219.        e=find(EHx(a)<0);
220.        EHazim(a(e))=EHazim(a(e))+180;
221.        b=find(EHx==0);
222.        c=find(EHy(b)<0);
223.        EHazim(b(c))=270;
224.        b(c)=[];

```

```

225.     EHazim(b)=90;
226.     EHazim=EHazim';
227.     a=find(EHazim<0);
228.     VHazim(a)=360+EHazim(a);
229.     a=find(VLx~=0);
230.     VLazim(a)=atand(VLy(a)./VLx(a));
231.     e=find(VLx(a)<0);
232.     VLazim(a(e))=VLazim(a(e))+180;
233.     b=find(VLx==0);
234.     c=find(VLy(b)<0);
235.     VLazim(b(c))=270;
236.     b(c)=[];
237.     VLazim(b)=90;
238.     VLazim=VLazim';
239.     a=find(VLazim<0);
240.     VLazim(a)=360+VLazim(a);
241.     a=find(ELx~=0);
242.     ELazim(a)=atand(ELy(a)./ELx(a));
243.     e=find(ELx(a)<0);
244.     ELazim(a(e))=ELazim(a(e))+180;
245.     b=find(ELx==0);
246.     c=find(ELy(b)<0);
247.     ELazim(b(c))=270;
248.     b(c)=[];
249.     ELazim(b)=90;
250.     ELazim=ELazim';
251.     a=find(ELazim<0);
252.     ELazim(a)=360+ELazim(a);
253.     table_rv_re=[azim rHv rHe VHazim EHazim rLv rLe VLazim ELazim];
254.     s=size(table_rv_re,1);
255.     table_rv_re(s+1,:)=table_rv_re(1,:);
256.     table_rv_re(s+1,1)=360;
257.     %Στρογγυλοποίηση
258.     table_rv_re=round(table_rv_re*10000)/10000;
259.end
260.
261.if (todoflag==4)
262.     table_rv_re=zeros(360,9);
263.     for d=1.00:(1.99-1.0)/99:1.99
264.         GH1=((2-d)/2);
265.         GH2=(d/2);
266.         GL1=((2-d)/2);
267.         GL2=(d/2);
268.         %αρχικές τιμές

```

```

269. PH=zeros(360,1);
270. VHx=zeros(360,1);
271. VHy=zeros(360,1);
272. EH=zeros(360,1);
273. EHx=zeros(360,1);
274. EHy=zeros(360,1);
275. PL=zeros(360,1);
276. VLx=zeros(360,1);
277. VLy=zeros(360,1);
278. EL=zeros(360,1);
279. ELx=zeros(360,1);
280. ELy=zeros(360,1);
281. coefficientindextabletest=[];
282. %δημιουργία πίνακα συντελεστών
283.     for i=1:speakers,
284.         hxeio=[f0+(i-1)*bhma 0 GH1 GH2 GH2 0 GL1 GL2 GL2 0];
285.         coefficientindextabletest=[coefficientindextabletest;hxeio];
286.         for j=1:360,
287.             %Δημιουργία των αθροισμάτων (συνολικό σήμα για την πίεση και την ένταση) αφού
                πρώτα
288.             %δημιουργηθούν τα σήματα B-Format και γίνει το Decoding για κάθε ηχείο.
289.             PH(j,1)=PH(j,1)+hxeio(1,3)*2^(-0.5)+cosd((j-1))*hxeio(1,4)*cosd(hxeio(1,1))+...
290.                 sind((j-1)*15)*hxeio(1,5)*sind(hxeio(1,1));
291.             VHx(j,1)=VHx(j,1)+(hxeio(1,3)*2^(-0.5)+cosd((j-1))*hxeio(1,4)*cosd(hxeio(1,1))
                +...
292.                 sind((j-1))*hxeio(1,5)*sind(hxeio(1,1)))*cosd(hxeio(1,1));
293.             VHy(j,1)=VHy(j,1)+(hxeio(1,3)*2^(-0.5)+cosd((j-1))*hxeio(1,4)*cosd(hxeio(1,1))
                +...
294.                 sind((j-1))*hxeio(1,5)*sind(hxeio(1,1)))*sind(hxeio(1,1));
295.             PL(j,1)=PL(j,1)+hxeio(1,7)*2^(-0.5)+cosd((j-1))*hxeio(1,8)*cosd(hxeio(1,1))+...
296.                 sind((j-1))*hxeio(1,9)*sind(hxeio(1,1));
297.             VLx(j,1)=VLx(j,1)+(hxeio(1,7)*2^(-0.5)+cosd((j-1))*hxeio(1,8)*cosd(hxeio(1,1))
                +...
298.                 sind((j-1))*hxeio(1,9)*sind(hxeio(1,1)))*cosd(hxeio(1,1));
299.             VLy(j,1)=VLy(j,1)+(hxeio(1,7)*2^(-0.5)+cosd((j-1))*hxeio(1,8)*cosd(hxeio(1,1))
                +...
300.                 sind((j-1))*hxeio(1,9)*sind(hxeio(1,1)))*sind(hxeio(1,1));
301.             EH(j,1)=EH(j,1)+(hxeio(1,3)*2^(-0.5)+cosd((j-1))*hxeio(1,4)*cosd(hxeio(1,1))+...
302.                 sind((j-1))*hxeio(1,5)*sind(hxeio(1,1)))^2;
303.             EHx(j,1)=EHx(j,1)+((hxeio(1,3)*2^(-0.5)+cosd((j-1))*hxeio(1,4)*cosd(hxeio(1,1))
                +...
304.                 sind((j-1))*hxeio(1,5)*sind(hxeio(1,1)))^2)*cosd(hxeio(1,1));
305.             EHy(j,1)=EHy(j,1)+((hxeio(1,3)*2^(-0.5)+cosd((j-1))*hxeio(1,4)*cosd(hxeio(1,1))
                +...
306.                 sind((j-1))*hxeio(1,5)*sind(hxeio(1,1)))^2)*sind(hxeio(1,1));
307.             EL(j,1)=EL(j,1)+(hxeio(1,7)*2^(-0.5)+cosd((j-1))*hxeio(1,8)*cosd(hxeio(1,1))+...

```



```

308.             sind((j-1))*hxeio(1,9)*sind(hxeio(1,1))^2;
309.             ELx(j,1)=ELx(j,1)+((hxeio(1,7)*2^(-0.5)+cosd((j-1))*hxeio(1,8)*cosd(hxeio(1,1))
+...
310.             sind((j-1))*hxeio(1,9)*sind(hxeio(1,1))^2)*cosd(hxeio(1,1));
311.             ELy(j,1)=ELy(j,1)+((hxeio(1,7)*2^(-0.5)+cosd((j-1))*hxeio(1,8)*cosd(hxeio(1,1))
+...
312.             sind((j-1))*hxeio(1,9)*sind(hxeio(1,1))^2)*sind(hxeio(1,1));
313.             azim(j,1)=(j-1);
314.             end
315.         end
316.         VHx=VHx./PH;
317.         VHy=VHy./PH;
318.         VLx=VLx./PL;
319.         VLy=VLy./PL;
320.         EHx=EHx./EH;
321.         EHy=EHy./EH;
322.         ELx=ELx./EL;
323.         ELy=ELy./EL;
324.         rHv=((VHx).^2+(VHy).^2).^0.5;
325.         rHe=((EHx).^2+(EHy).^2).^0.5;
326.         rLv=((VLx).^2+(VLy).^2).^0.5;
327.         rLe=((ELx).^2+(ELy).^2).^0.5;
328.         VHazim=zeros(size(VHx,1),1);
329.         a=find(VHx~=0);
330.         VHazim(a)=atand(VHy(a)./VHx(a));
331.         e=find(VHx(a)<0);
332.         VHazim(a(e))=VHazim(a(e))+180;
333.         b=find(VHx==0);
334.         c=find(VHy(b)<0);
335.         VHazim(b(c))=270;
336.         b(c)=[];
337.         VHazim(b)=90;
338.         a=find(VHazim<0);
339.         VHazim(a)=360+VHazim(a);
340.         a=find(EHx~=0);
341.         EHazim=zeros(size(VHx,1),1);
342.         EHazim(a)=atand(EHy(a)./EHx(a));
343.         e=find(EHx(a)<0);
344.         EHazim(a(e))=EHazim(a(e))+180;
345.         b=find(EHx==0);
346.         c=find(EHy(b)<0);
347.         EHazim(b(c))=270;
348.         b(c)=[];
349.         EHazim(b)=90;

```

```

350.     a=find(EHazim<0);
351.     EHazim(a)=360+EHazim(a);
352.     a=find(VLx~=0);
353.     VLazim=zeros(size(VHx,1),1);
354.     VLazim(a)=atand(VLy(a)./VLx(a));
355.     e=find(VLx<0);
356.     VLazim(a(e))=VLazim(a(e))+180;
357.     b=find(VLx==0);
358.     c=find(VLy(b)<0);
359.     VLazim(b(c))=270;
360.     b(c)=[];
361.     VLazim(b)=90;
362.     a=find(VLazim<0);
363.     VLazim(a)=360+VLazim(a);
364.     a=find(ELx~=0);
365.     ELazim=zeros(size(VHx,1),1);
366.     ELazim(a)=atand(ELy(a)./ELx(a));
367.     e=find(ELx<0);
368.     ELazim(a(e))=ELazim(a(e))+180;
369.     b=find(ELx==0);
370.     c=find(ELy(b)<0);
371.     ELazim(b(c))=270;
372.     b(c)=[];
373.     ELazim(b)=90;
374.     a=find(ELazim<0);
375.     ELazim(a)=360+ELazim(a);
376.     table_rv_re_g_test=[azim rHv rHe VHazim EHazim rLv rLe VLazim ELazim];
377.     %δημιουργία πίνακα ελέγχου για τις τιμές που προκύπτουν βάση του
378.     %επιλεγμένου d.
379.     test=table_rv_re_g_test;
380.     test(:,2:3)=abs(1-test(:,2:3))/360;
381.     test(:,4)=abs(test(:,1)-test(:,4))/360;
382.     test(:,5)=abs(test(:,1)-test(:,5))/360;
383.     test(:,6:7)=abs(1-test(:,6:7))/360;
384.     test(:,8)=abs(test(:,1)-test(:,8))/360;
385.     test(:,9)=abs(test(:,1)-test(:,9))/360;
386.
387.     bathmologiaH1=find(test(:,2)<=bathtelH(:,1));
388.     bathmologiaH2=find(test(:,3)<=bathtelH(:,2));
389.     bathmologiaH3=find(test(:,4)<=bathtelH(:,3));
390.     bathmologiaH4=find(test(:,5)<=bathtelH(:,4));
391.
392.     bathmologiaL1=find(test(:,6)<=bathtelL(:,1));
393.     bathmologiaL2=find(test(:,7)<=bathtelL(:,2));

```

```

394.     bathmologiaL3=find(test(:,8)<=bathtelL(:,3));
395.     bathmologiaL4=find(test(:,9)<=bathtelL(:,4));
396.     %Ελεγχος αν οι τιμές βάση του d είναι καλύτερες από την προηγούμενη
397.     %καλύτερη
398.     if (2*numel(bathmologiaH2)+numel(bathmologiaH3)+numel(bathmologiaH4)>=720)
399.         if (2*numel(bathmologiaH2)+numel(bathmologiaH3)+numel(bathmologiaH4)==720)
400.             if numel(bathmologiaH1)>330
401.                 bathtelH=test(:,2:5);
402.                 table_rv_re(:,1:5)=table_rv_re_g_test(:,1:5);
403.                 coefficientindextable(:,1:5)=coefficientindextabletest(:,1:5);
404.                 dH=d;
405.             end
406.         else
407.             bathtelH=test(:,2:5);
408.             table_rv_re(:,1:5)=table_rv_re_g_test(:,1:5);
409.             coefficientindextable(:,1:5)=coefficientindextabletest(:,1:5);
410.             dH=d;
411.         end
412.     end
413.     if (2*numel(bathmologiaL1)+numel(bathmologiaL3)+numel(bathmologiaL4)>=720)
414.         if (2*numel(bathmologiaL1)+numel(bathmologiaL3)+numel(bathmologiaL4)==720)
415.             if numel(bathmologiaL2)>330
416.                 bathtelL=bathmologiaL;
417.                 table_rv_re(:,6:9)=table_rv_re_g_test(:,6:9);
418.                 coefficientindextable(:,7:10)=coefficientindextabletest(:,7:10);
419.                 dL=d;
420.             end
421.         else
422.             bathtelL=test(:,6:9);
423.             table_rv_re(:,6:9)=table_rv_re_g_test(:,6:9);
424.             coefficientindextable(:,7:10)=coefficientindextabletest(:,7:10);
425.             dL=d;
426.         end
427.     end
428.
429.     end
430.     clear test
431. table_rv_re=table_rv_re(1:15:360,:);
432. s=size(table_rv_re,1);
433. table_rv_re(s+1,:)=table_rv_re(1,:);
434. table_rv_re(s+1,1)=360;
435. %Στρογγυλοποίηση
436.     table_rv_re=round(table_rv_re*10000)/10000;
437. end

```