

Τμήμα Μουσικής Τεχνολογίας & Ακουστικής
Παράρτημα Ρεθύμνου
Α.Τ.Ε.Ι Κρήτης

Τελική Μίξη (Mastering) Προηχογραφημένου
Ηχητικού Υλικού σε Στερεοφωνική Μορφή και Αξιολόγηση Αυτού με
Βάση Αντικειμενικά & Υποκειμενικά Κριτήρια, Αναφορικά με την Αρχική
του Μορφή.

Καρράς Κων/νος
2012

Τελική Μίξη (Mastering) Προηχογραφημένου
Ηχητικού Υλικού σε Στερεοφωνική Μορφή και Αξιολόγηση Αυτού με
Βάση Αντικειμενικά & Υποκειμενικά Κριτήρια, Αναφορικά με την Αρχική
του Μορφή.

Σπουδαστής: Καρράς Κων/νος, Α.Μ: 553
Υπ. Καθηγητής: Φιτσανάκης Μίνως.

Περιεχόμενα:

Εισαγωγή.....	ix
Κεφάλαιο 1 ^ο : Psychoacoustics & Audio Mastering.....	1
1.1: Εισαγωγή στην ψυχοακουστική.....	2
1.2.1: Ανατομία του αυτιού / Εξωτερικό αυτί / Πτερύγιο (pinna).....	2
1.2.2: Ανατομία του αυτιού / Εξωτερικό αυτί / Ακουστικό κανάλι.....	4
1.2.3: Ανατομία του αυτιού / Μέσο αυτί / Τύμπανο.....	4
1.2.4: Ανατομία του αυτιού / Μέσο αυτί / Σφύρα, Άκμονας, Αναβολέας.....	4
1.2.5: Ανατομία του αυτιού / Μέσο αυτί / Ευσταχιανή σάλπιγγα.....	5
1.2.6: Εσωτερικό αυτί.....	6
1.2.7: Εσωτερικό αυτί / Αιθουσαίο.....	6
1.2.8 Εσωτερικό αυτί / Κοχλίας.....	7
1.3: Απόκριση του ακουστικού συστήματος.....	8
1.3.1: Ευαισθησία της ακοής στη διάρκεια των ήχων & χρονική ολοκλήρωση...8	
1.3.2: Χρονική διάκριση (Temporal discrimination).....	10
1.3.3: Όρια ακουστότητας (Threshold of audibility).....	11
1.3.4: Ευαισθησία σε αλλαγές έντασης, συχνότητας & χρ. διάρκειας.....	12
1.3.5: Συχνотική διάκριση (Frequency discrimination).....	12
1.3.6: Αντίληψη / Διάκριση της έντασης.....	14
1.3.7: Τονικό ύψος.....	16
1.3.7.1: Complex pitch.....	18
1.3.8: Χροιά (Timbre).....	20
1.3.9: Φαινόμενα μη γραμμικότητας του ακουστικού συστήματος.....	23
1.4: Masking effect.....	24
1.5: Sound localization.....	27
1.5.1: Εντοπισμός στο οριζόντιο επίπεδο.....	27
1.5.2: Εντοπισμός στο κάθετο επίπεδο.....	33
1.5.3: Εντοπισμός της απόστασης.....	36
1.5.4: Φαινόμενο προβαδίσματος (Precedence effect).....	37
1.5.5: Φαινόμενο συνάθροισης (Cocktail party effect).....	38
1.6: Ψυχοακουστική και Audio mastering / Συμπεράσματα.....	39
Κεφάλαιο 2 ^ο : Editing & Audio Mastering.....	42
2.1: Editing, fades και άλλες εργασίες.....	43
2.2: Άλλες εργασίες στο mastering.....	43

Κεφάλαιο 3 ^ο : Equalization & Audio Mastering.....	46
3.1: Κεντρική συχνότητα (f_c).....	47
3.2: Εύρος ζώνης του φίλτρου (bandwidth, BW) και Q.....	48
3.3: Ενίσχυση / Εξασθένιση (Boost / Cut) μιας συχνοτικής περιοχής.....	48
3.4: Η τάξη (Order) του φίλτρου.....	48
3.5.1: Παραμετρικά (presence) equalizer.....	49
3.5.2: Notch equalizer.....	50
3.5.3: High & Low shelving φίλτρα.....	51
3.5.4: Υψηπερατά & βαθυπερατά φίλτρα (High & Low).....	52
3.6: Βασικά εισαγωγικά σημεία αναφορικά με τη χρήση των equalizer.....	53
3.7: Τι είναι η τονική ισορροπία του ηχητικού υλικού.....	55
3.8: Τα ξεχωριστά μουσικά είδη.....	57
3.9: Τεχνικές equalization, παραμετρικά equalizer.....	58
3.10: Τεχνικές equalization, shelving equalizer.....	60
3.11: Τα shelving equalizer τύπου Baxandall	61
3.12: High pass & Low pass φίλτρα.....	63
3.13: Εφαρμογές equalizer.....	65
3.13.1: Διαπιστώνοντας τα όρια της ηχογράφησης.....	66
3.13.2: Equalization στη θεμελιώδη ή στις αρμονικές.....	66
Κεφάλαιο 4 ^ο : Compression & Audio Mastering.....	69
4.1: Δυναμικό εύρος.....	70
4.1.1: Μικροδυναμικές και Μακροδυναμικές.....	70
4.1.2: Οι δυναμικές στη μουσική ιστορία.....	70
4.1.3: Μείωση του δυναμικού εύρους.....	71
4.1.4: Αύξηση του δυναμικού εύρους.....	71
4.1.5: Τύποι τροποποίησης του δυναμικού εύρους.....	72
4.1.6: Manual gain riding.....	73
4.1.7: Αλλαγή των internal levels ενός μουσικού κομματιού.....	74
4.2: Compressors & Limiters αντικειμενικά χαρακτηριστικά.....	77
4.2.1: Transfer curves / Compressors & Limiters.....	77
4.2.2: Χρόνοι Attack και Release.....	82
4.2.3: Limiting εναντίων compression στο audio mastering.....	84
4.2.4: Ο πιο διαυγής ψηφιακός limiter.....	85
4.2.5: Compressors εν ώρα εργασίας.....	85
4.2.6: Τυπικές τιμές του ratio και του threshold.....	86
4.2.7: Ορισμένες επιπλέον και χρήσιμες ρυθμίσεις των compressors.....	87
4.2.8: Compression και monitoring.....	88
4.2.9: Επεξεργασία της δυναμικής περιοχής με multiband compression.....	88

4.2.10: Equalization ή multiband compression.....	90
4.2.11: Compression, στερεοφωνική εικόνα & depth μιας ηχογράφησης.....	91
4.3: Upward.....	91
4.3.1: Upward Compression.....	92
Κεφάλαιο 5 ^ο : Noise reduction & Audio Mastering.....	96
5.: Noise reduction - Εισαγωγή.....	97
5.1: Τεχνικές μείωσης θορύβου / Απλό φιλτράρισμα (simple filtering).....	99
5.2: Narrow band Expansion.....	99
5.3: Σύνθετο φιλτράρισμα (complex filtering).....	100
5.4: Εξειδικευμένοι επεξεργαστές (specialised processors).....	101
5.5: Το φαί του ενός είναι το δηλητήριο του άλλου.....	101
Κεφάλαιο 6 ^ο : Aliasing & Audio Mastering.....	103
6.1: Aliasing.....	104
6.2: Σύνδεση του audio mastering με το aliasing.....	111
Κεφάλαιο 7 ^ο : Jitter & Audio Mastering.....	112
7.1: Τι είναι το jitter.....	113
Κεφάλαιο 8 ^ο : Dithering & Audio Mastering.....	116
8.1: Dithering - Εισαγωγή.....	117
8.2: Θεωρητικό υπόβαθρο.....	117
8.3: Το φάσμα που προκύπτει.....	119
8.4: Όλα άρχισαν με τους Fletcher & Munson.....	119
8.5: Περιπτώσεις χρησιμοποίησης του flat dither.....	124
8.6: Υπάρχει λόγος για τη μη χρήση οποιουδήποτε τύπου dither.....	125
8.7: Αναφορικά με τη συμπίεση δεδομένων.....	126
8.8: Κάποια πρώτα συμπεράσματα (πριν συνεχίσουμε).....	128
8.9: Περαιτέρω εμβάθυνση στο dither.....	129
8.10: Noise shaping.....	133
8.11: Dither_So little noise - So much effect.....	137
8.12: Μια επανάληψη κι ένα παράδειγμα στο Sonic Studio.....	141
8.12.1: Τύποι dither στο Sonic Studio HD.....	142
8.12.2: Εφαρμόζοντας το dither στο Sonic Studio.....	144
8.12.3: Παράμετροι του dither.....	145
Κεφάλαιο 9 ^ο : Audio Mastering - Από τη θεωρία στην πράξη.....	148
9.: Εισαγωγή.....	149
9.1: Σαβίνα Γιαννάτου.....	149

9.2: Μιλτιάδης Πασχαλίδης & Θανάσης Γκαϊφίλιας.....	154
9.3: Wish Upon a Star - Frank Lines.....	157
9.4: Μίκης Θεοδωράκης - 1.....	162
9.5: Μίκης Θεοδωράκης - 2.....	165
9.6: Kurt Weill - Ballad of Mack the Knife.....	168
9.7: Kurt Weill - Canon Song.....	172
9.8: Kurt Weill - Polly's Song.....	175
9.9: Sergei Rachmaninov - Come let us Worship.....	178
9.10: Jean Claude Risset - Sud La mer le matin.....	181
9.11: P. Kokoras - Response.....	185
Κεφάλαιο 10 ^ο : Συμπεράσματα.....	190
10.: Συμπεράσματα - Αντί επιλόγου.....	191
Προτάσεις για τη συνέχεια.....	195
Βιβλιογραφία.....	196
Παραρτήματα.....	197
Παράρτημα Α.....	198
Παράρτημα Β.....	199
Παράρτημα Γ.....	212
Παράρτημα Δ.....	218
Παράρτημα Ε.....	220
Παράρτημα ΣΤ.....	223
Παράρτημα Ζ.....	231

Εισαγωγή:

Η πτυχιακή αυτή εργασία καλείται να καταπιαστεί με την τελική επεξεργασία (audio mastering) προηχογραφημένου ηχητικού υλικού σε στερεοφωνική μορφή. Για τον σκοπό αυτό θα χρησιμοποιηθούν ηχητικά υλικά διαφορετικής μουσικής υφής, θέτοντας ως στόχο την επεξεργασία και την ανάλυση τους στο τεχνικό – αντικειμενικό, καθώς και στο υποκειμενικό αυτό, επίπεδο το οποίο αφορά έναν mastering ηχολήπτη καθώς κι έναν μηχανικό ήχου γενικότερα.

Επιπλέον στόχος είναι η συγκριτική παράθεση του ηχητικού υλικού της προγενέστερης έναντι της μεταγενέστερης του μορφής.

Τα ηχητικά υλικά διαφορετικής μουσικής υφής κρίνονται απαραίτητα για την εξαγωγή συμπερασμάτων καθώς η διαδικασία του Mastering δεν ευνοεί με όμοιο τρόπο το κάθε ένα διαφορετικό μουσικό είδος.

Η πτυχιακή αυτή εργασία στο επίπεδο της γενικότερης σύλληψης της αφορά:

- Τον τρόπο με τον οποίο επεξεργαστήκαμε το κάθε ηχητικό υλικό καθώς και την αιτιολόγηση του λόγου που μας οδήγησε στην επιλογή του τρόπου αυτού.
- Συγκριτική αξιολόγηση του αρχικού με το τελικό ηχητικό υλικό ανά μουσικό είδος με βάση αντικειμενικά τεχνικά χαρακτηριστικά και κριτήρια (συχνοτική ανάλυση, ανάλυση στο πεδίο της δυναμικής των εντάσεων κκ). Συγκριτική αξιολόγηση του αρχικού με το τελικό ηχητικό υλικό ανά μουσικό είδος με βάση υποκειμενικά κριτήρια (ψυχοακουστικά φαινόμενα τα οποία προσδιορίζουν τον τρόπο με τον οποίο αντιλαμβάνεται ο ανθρώπινος εγκέφαλος την ηχητική πληροφορία).

Πιο συγκεκριμένα η διάρθρωση της παρούσας πτυχιακής εργασίας έχει δομηθεί με τέτοιο τρόπο ώστε να είναι διακριτή η αναφορά τόσο στο ψυχοακουστικό πεδίο, όσο και στο τεχνικό αντικειμενικό πεδίο. Σε ότι αφορά τώρα το κομμάτι εκείνο της εργασίας στο οποίο αναλύονται τα μουσικά έργα, επιχειρείται η σύνδεση τόσο των τεχνικών και αισθητικών παραμέτρων που καθόρισαν τις πρακτικές μας επιλογές, όσο και των ψυχοακουστικών αυτών παραγόντων που επίσης καθόρισαν τις επιλογές μας, χωρίς η διαχωριστική γραμμή στο επίπεδο της πρακτικής να είναι πάντοτε ευδιάκριτη.

Στο πρώτο κεφάλαιο ασχολούμαστε με την ψυχοακουστική. Η ψυχοακουστική είναι ένας συνολικός όρος που περιλαμβάνει τη φυσική κατασκευή του αυτιού, τις διαδρομές κίνησης του ήχου, την αντίληψη του ήχου και τις σχέσεις μεταξύ τους και ασχολούμαστε με το συγκεκριμένο γνωστικό τομέα διότι ένα από τα ζητούμενα της πτυχιακής αυτής εργασίας είναι ο προσδιορισμός σε ένα πρώτο επίπεδο της σχέσης των ψυχοακουστικών φαινομένων που λαμβάνουν χώρα και επηρεάζουν τον τρόπο με

τον οποίο αντιλαμβανόμαστε τον ήχο, με τις διεργασίες που λαμβάνουν χώρα στο audio mastering.

Στο δεύτερο κεφάλαιο ασχολούμαστε με το editing, τα fades και άλλες παρόμοιου τύπου εργασίες. Το συγκεκριμένο πεδίο εργασίας αποτελεί, θα μπορούσαμε σχηματικά να πούμε, τη μορφή, το φαίνεσθαι, το εξωτερικό περίβλημα της ενασχόλησης ενός audio mastering ηχολήπτη. Τέτοιες εργασίες είναι τα fade-ins και fade-outs, η σειρά των κομματιών, οι παύσεις, και τα cross-fades. Επίσης η μείωση ανάλυσης και η μονοφωνική συμβατότητα του ηχητικού υλικού, οι λίστες PQ και ISRC και τέλος ο ποιοτικός έλεγχος του τελικά παραγόμενου ηχητικού υλικού.

Στο τρίτο κεφάλαιο καταπιανόμαστε με το equalization στο audio mastering. Μία από τις βασικές αρχές του mastering είναι πως οποιαδήποτε ενέργεια σε οποιοδήποτε από τα χαρακτηριστικά του ηχητικού υλικού, επηρεάζει ολόκληρο το μουσικό κομμάτι και αυτό δε μπορεί παρά να αφορά και τη χρήση των equalizer. Έτσι σε αυτό το κεφάλαιο εξετάζουμε τους βασικούς τύπους ισοσταθμιστών, τα χαρακτηριστικά τους, τον τρόπο λειτουργίας τους και τελικά τη χρήση τους αναφορικά με το mastering. Τέλος εξετάζουμε ειδικούς τύπους φίλτρων και προχωρημένες τεχνικές επεξεργασίας που έχουν προκύψει στο πέρασμα των χρόνων.

Στο τέταρτο κεφάλαιο καταπιανόμαστε με το compression, κοινώς τη δυναμική επεξεργασία, του ηχητικού υλικού στο mastering. Αρχικά γίνεται μια ευρεία αναφορά στη σημασία που έχει το δυναμικό εύρος για το ηχητικό συμβάν καθώς και μια διάκριση ανάμεσα στα μικροδυναμικά και μακροδυναμικά στοιχεία για την καλύτερη δυνατή κατανόηση. Στη συνέχεια αναφερόμαστε στην επεξεργασία (μείωση, αύξηση) του δυναμικού εύρους, στους «τύπους» τροποποίησης του δυναμικού εύρους, στις τεχνικές που χρησιμοποιούνται και στον τρόπο με τον οποίο αυτές λαμβάνουν χώρα. Αναφερόμαστε στους επεξεργαστές (compressors, limiters, expanders), στα χαρακτηριστικά τους και τον τρόπο χρήσης τους. Τέλος αναφερόμαστε στην εξειδικευμένη τεχνική του parallel compression και τις δυνατότητες ή περιορισμούς της τεχνικής αυτής.

Στο πέμπτο κεφάλαιο ασχολούμαστε με το noise reduction του οποίου η χρησιμότητα στα πλαίσια του audio mastering επουδενί δε μπορεί να αμφισβητηθεί. Έτσι λοιπόν σε αυτό το κεφάλαιο εξετάζουμε τους τύπους θορύβου που επηρεάζουν το ηχητικό υλικό, τους τρόπους αντιμετώπισης του θορύβου, τις τεχνικές μείωσής του και τον τρόπο με τον οποίο αυτές πραγματοποιούνται.

Στο έκτο κεφάλαιο καταγινόμαστε με το aliasing, μια ειδική και πολύ δύσκολη περίπτωση κατά τη διαδικασία επεξεργασίας του ηχητικού υλικού κατά τη μετατροπή του από αναλογικό σε ψηφιακό και το αντίθετο. Κάτι τέτοιο έχει σαν αποτέλεσμα την παραγωγή παρείσιακτων συχνοτήτων οι οποίες δεν ήταν υπαρκτές αρχικά. Έτσι λοιπόν ασχολούμαστε με τις προϋποθέσεις που πρέπει να πληρούνται ώστε να μην εμφανιστεί αυτό το φαινόμενο.

Στο έβδομο κεφάλαιο ασχολούμαστε με το jitter. Σαν jitter θεωρούμε το λάθος χρονισμό στην μεταφορά δεδομένων κατά την σύνδεση δυο η περισσότερων ψηφιακών συσκευών. Δεδομένης της πολυπλοκότητας και της δυσκολίας του φαινομένου δεν επιχειρούμε να προτείνουμε λύσεις, μα απλώς να αποτυπώσουμε μια όσο το δυνατόν πιο πλήρη εικόνα του και να βγάλουμε κάποια πρώτα κρίσιμα συμπεράσματα.

Στο όγδοο κεφάλαιο καταπιανόμαστε με το dithering και τον τρόπο με τον οποίο επενεργεί στο μουσικό υλικό. Μας απασχόλησαν στο προκείμενο πλήθος παραγόντων που συσχετίζονται με την ψυχοακουστική. Γίνεται μια σε βάθος ανάλυση αρκετών παραγόντων και προκύπτουν χρήσιμα συμπεράσματα. Εξετάζουμε, επίσης, τους διαφόρους τύπους dither και τον τρόπο με τον οποίο αυτοί χρησιμοποιούνται. Τέλος εξετάζονται διάφορα παραδείγματα, περιγράφονται εφαρμογές του.

Στο ένατο κεφάλαιο επιχειρούμε να αποτυπώσουμε τις εργασίες που έλαβαν χώρα, από μέρους μας, σε μία πλειάδα μουσικών κομματιών διαφορετικών ειδών. Στο κεφάλαιο αυτό παρατίθενται συγκριτικά πλήθος διαγραμμάτων τα οποία αποτυπώνουν τις διαφοροποιήσεις σε ότι αφορά την προγενέστερη έναντι της μεταγενέστερης μορφής του κάθε μουσικού κομματιού. Τα διαγράμματα αυτά αφορούν τις δυναμικές του σήματος, τη συχνотική του απόκριση, όπως επίσης δίνονται και πίνακες οι οποίοι αποτυπώνουν με ξεκάθαρο τρόπο τις δυναμικές επίσης διαφοροποιήσεις.

Στο δέκατο κεφάλαιο μας απασχόλησε η επεξεργασία των συμπερασμάτων που προκύπτουν από το ένατο κεφάλαιο έτσι ώστε να μπορέσουμε να πατήσουμε πάνω σε μια όσο το δυνατόν πιο στερεή βάση και να απαντήσουμε στο ερώτημα που αφορά στο με ποιο τρόπο υιοθετήθηκε το audio mastering από τη μουσική βιομηχανία στο σύνολο της. Επιπλέον στόχο αποτελούσε το γεγονός του να δούμε τι επιπτώσεις, θετικές ή αρνητικές, επέφερε τελικά η επιλογή αυτή στον τρόπο που ακούμε και αντιλαμβανόμαστε το μουσικό συμβαινόμενο.

Προς το τέλος της πτυχιακής αυτής εργασίας παρατίθενται κάποιες προτάσεις για ενδεχόμενη συνέχεια ή προέκταση επί αυτής της θεματικής. Αμέσως μετά παρατίθεται ο βασικός κορμός της βιβλιογραφίας που χρησιμοποιήσαμε.

Τέλος δίδονται μια σειρά από παραρτήματα που πραγματεύονται εξειδικευμένα ζητήματα που αν και απαραίτητα θα ήταν εξαιρετικά δύσκολο να αφομοιωθούν από τον κυρίως κορμό του κειμένου της εργασίας που κατατίθεται. Έτσι στο **Α Παράρτημα** δίνονται δύο πολύ βοηθητικοί πίνακες για τη μετατροπή της τιμής του bandwidth (δηλ. σε Hz), αλλά εμείς επιθυμούμε να γνωρίζουμε την τιμή του Q, καθώς επίσης και το αντίστροφο. Στο **Β Παράρτημα** ασχολούμαστε με τη αναστροφή φάσης που επιφέρει η χρήση των equalizer στο μουσικό σήμα, καθώς και με τα διάφορα επιφαινόμενα, όπως επίσης και με τα αίτια τους, που προκύπτουν απο αυτή τη συνθήκη και βλέπουμε πως προκύπτουν πολύ ενδιαφέροντα πράγματα. Στο **Γ Παράρτημα** δίδονται δύο κώδικες του matlab που βοηθούν πολύ στην κατανόηση

της λειτουργίας των limiters. Στο **΄Δ Παράρτημα** υπολογίζονται, μέσω της ακουστικής, διάφορα μεγέθη που αφορούν το δυναμικό εύρος. Αυτό μας βοηθά να κατανοήσουμε καλύτερα το πώς λαμβάνουν χώρα οι διαδικασίες κατά τη διάρκεια της επεξεργασίας του δυναμικού εύρους του ηχητικού υλικού. Στο **΄Ε Παράρτημα** δίδονται οι κώδικες που χρησιμοποιούνται στο 6^ο κεφάλαιο (σελ. 106) και στο **΄ΣΤ Παράρτημα** οι κώδικες που χρησιμοποιούνται για την καλύτερη κατανόηση και επεξήγηση του φαινομένου του aliasing (κεφ. 6^ο, σελ. 105). Κλείνοντας στο **΄Ζ Παράρτημα** παρατίθενται οι γραφικές απεικονίσεις που εξάγαμε διαφόρων τύπων dither. Κάτι τέτοιο είναι εξαιρετικά χρήσιμο γιατί μας βοηθά πάρα πολύ στην κατανόηση του τι πραγματικά συντελείται κατά τη διάρκεια αυτής της φάσης στην επεξεργασία του ηχητικού υλικού. Κατά συνέπεια έχει εξαιρετικά σημαντικό πρακτικό αντίκτυπο στην στιγμή κατά την οποία κάποιος εργάζεται γιατί μπορεί να κατανοήσει τους διαφορετικούς τύπους dither και να επιλέξει τον καταλληλότερο.

**Psychoacoustics
&
Audio Mastering**

1. Ψυχοακουστική:

1.1 Εισαγωγή:

Το ακουστικό μας σύστημα και συγκεκριμένα το ανθρώπινο αυτί είναι το πρώτο όργανο το οποίο είναι έτοιμο να λειτουργήσει στο έμβρυο. Κάνει την εμφάνιση του απο τις πρώτες μέρες της σύλληψης, ενώ λειτουργικά έχει ολοκληρωθεί απο τον τέταρτο μόλις μήνα της εγκυμοσύνης. Στον πέμπτο μήνα της εμβρυακής ζωής είναι ένα όργανο που γνωρίζει ήδη να αποθηκεύει την ηχητική πληροφορία. Το αυτί είναι ένα αισθητήριο το οποίο μπορεί να λειτουργήσει σαν ενισχυτής και σαν αναλυτής φάσματος, ενώ μπορεί να διαχειριστεί ηχητικές εντάσεις σε μία κλίμακα η οποία είναι της τάξης του 10^{13} . Μπορεί να απορρίψει τους ήχους σε μια συνεστίαση πολλών ανθρώπων και να εστιάσει σε μία μόνο φωνή (φαινόμενο cocktail party). Μπορεί να προσαρμοσθεί σε πολύ δυνατούς ήχους και να αυτοπροστατευθεί απο το να πάθει κακό. Όλα αυτά και άλλα, μπορεί να τα πράττει έχοντας ένα όγκο γύρω στα δύο κυβικά εκατοστά.

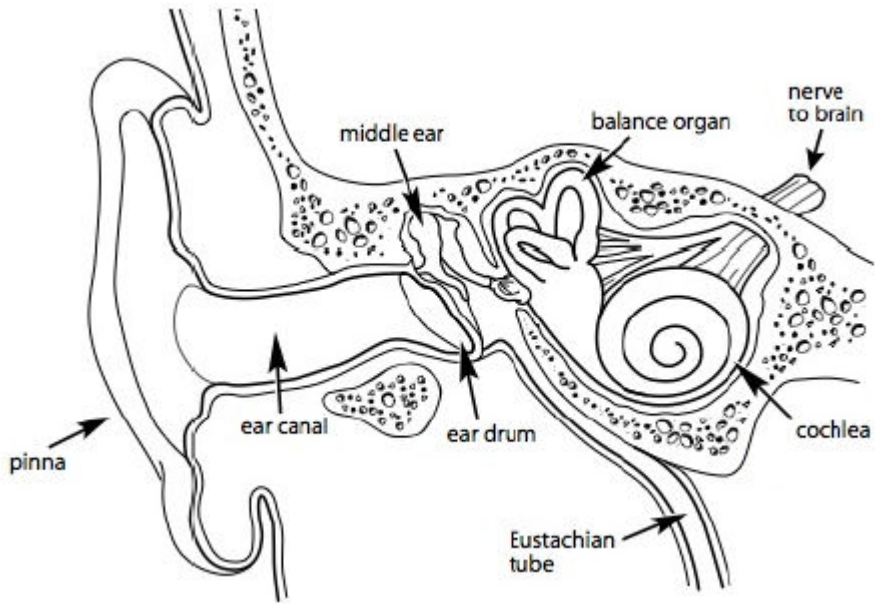
Η ψυχοακουστική είναι ένας, γενικά, πρόσφατος όρος που ασχολείται τόσο με την κατασκευή, όσο και με τη λειτουργία του αυτιού. Η μελέτη της κατασκευής του αυτιού είναι μια μελέτη ανατομίας. Η μελέτη της ανθρώπινης αντίληψης του ήχου βρίσκεται στο γενικότερο πλαίσιο της ψυχολογίας. Η ψυχοακουστική είναι ένας συνολικός όρος που περιλαμβάνει τη φυσική κατασκευή του αυτιού, τις διαδρομές κίνησης του ήχου, την αντίληψη του ήχου και τις σχέσεις μεταξύ τους.

Ορίζοντας το πεδίο της ψυχοακουστικής μπορούμε να πούμε ότι: ψυχοακουστική είναι ο τομέας ο οποίος εξετάζει τη σχέση μεταξύ των φυσικών ιδιοτήτων του ήχου (συχνότητα, ένταση και χρόνος) και της αίσθησης της οποίας αυτές προκαλούν .

1.2.1 Ανατομία του αυτιού / Εξωτερικό αυτί / Πτερύγιο (pinna):

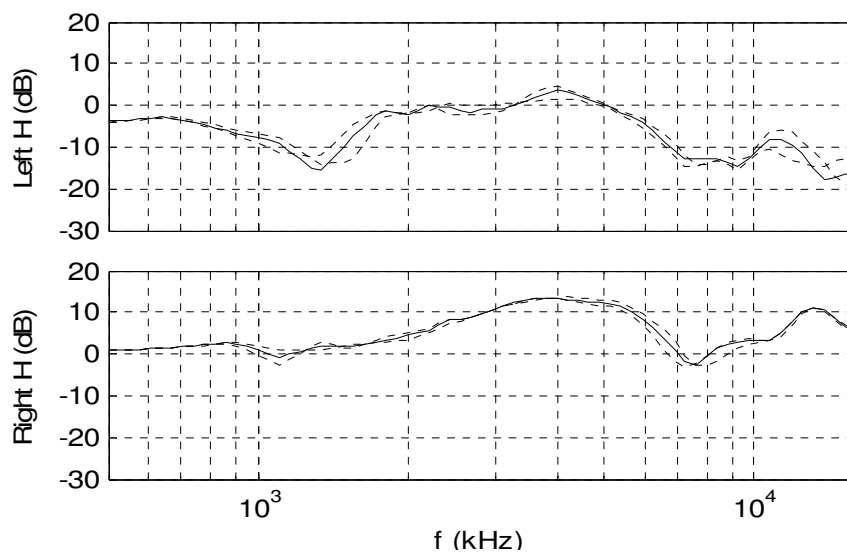
Τα τρία κύρια τμήματα του αυτιού που φαίνονται στο σχήμα 1.2.1.1 της επόμενης σελίδας είναι το εξωτερικό αυτί, το μέσο αυτί και το εσωτερικό αυτί. Το εξωτερικό αυτί αποτελείται απο το πτερύγιο και το ακουστικό κανάλι το οποίο τελειώνει στην ακουστική μεμβράνη ή τύμπανο όπου αρχίζει το μέσο αυτί.

Το πτερύγιο είναι μια «συσκευή» συλλογής ήχων. Αποτελείται κυρίως απο κόνδρο χωρίς κάποιους χρήσιμους μύες σε αυτό. Η μορφή του διαφέρει απο άνθρωπο σε άνθρωπο.



Σχ. 1.2.1.1: Σχηματική περιγραφή του εξωτερικού, του μέσου και του έσω αυτιού.

Η βασικότερη απο τις λειτουργίες που επιτελεί είναι να βοηθάει στον προσδιορισμό της θέσης μιας ηχητικής πηγής στο χώρο. Καθώς ο ήχος διαδίδεται απο την πηγή προς το αυτί, συναντά το ανθρώπινο σώμα, το κεφάλι και το πτερύγιο του αυτιού. Η αλληλεπίδραση αυτή έχει ως αποτέλεσμα τη διαφοροποίηση της ηχητικής πληροφορίας. Παρατηρούμε στο σχήμα 1.2.1.2 πως υπάρχει σημαντική διαφορά της ηχητικής πληροφορίας που φτάνει σε κάθε αυτί. Οι συναρτήσεις μεταφοράς που περιγράφουν τη διαφορά μεταξύ του ήχου της πηγής και του ήχου που περνάει απο το πτερύγιο, ονομάζονται Head Related Transfer Functions (HRTFs, θα αναφερθούμε εκτενέστερα σε επόμενη παράγραφο). Τέλος να αναφέρουμε ότι στο πτερύγιο οφείλονται, επίσης, οι διαφοροποιήσεις στις υψηλές συχνότητες.

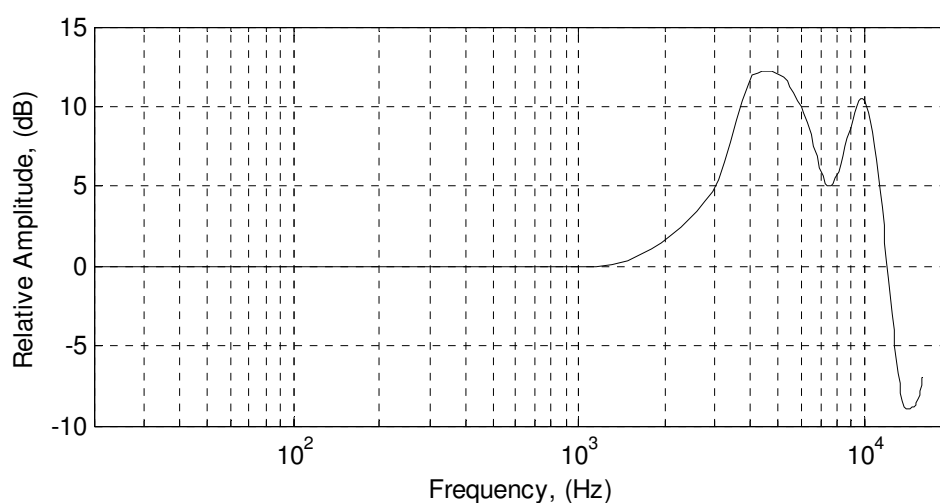


Σχ. 1.2.1.2: Τυπικό δείγμα Head Related Transfer Functions του αριστερού και του δεξιού αυτιού.

1.2.2 Ανατομία του αυτιού / Εξωτερικό αυτί / Ακουστικό κανάλι:

Το ακουστικό κανάλι έχει την ιδιότητα να αυξάνει την ηχηρότητα των ήχων που το διαπερνούν. Έχει μέση διάμετρο 0.7 cm και μήκος περίπου 2,5 cm. Το φαινόμενο του συντονισμού στο κανάλι του αυτιού αυξάνει την πίεση του ήχου στο τύμπανο για ορισμένες συχνότητες. Το μέγιστο βρίσκεται κοντά στη συχνότητα στην οποία ο σωλήνας των 2,5 cm έχει ένα τέταρτο του μήκους κύματος, δηλ. περίπου στα 3400 Hz. Η ενίσχυση που παρέχει κυμαίνεται στα 10 με 15 dB. Η απόκριση συχνότητας του ακουστικού καναλιού φαίνεται στο σχήμα 1.2.2.1 που ακολουθεί.

Μια άλλη λειτουργία του ακουστικού καναλιού και γενικότερα του εξωτερικού αυτιού, είναι αυτή της προστασίας του μέσου αυτιού από εξωτερικά σώματα και από τις αλλαγές στην υγρασία και τη θερμοκρασία.



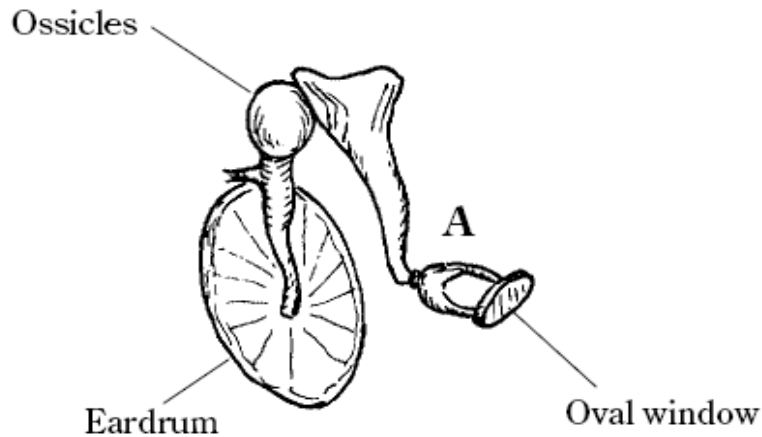
Σχ. 1.2.2.1: Συχνωτική απόκριση του ακουστικού καναλιού του αυτιού.

1.2.3 Ανατομία του αυτιού / Μέσο αυτί / Τύμπανο:

Η τυμπανική μεμβράνη είναι μια σχετικά διαφανής μεμβράνη με επιφάνεια από 55 έως 90 mm². Είναι κατασκευασμένη από λεπτά στρώματα ιστού του οποίου τα κεντρικά τμήματα είναι και τα πιο σημαντικά. Η μεμβράνη συγκρατείται στη θέση της από νευρικές ίνες. Το τύμπανο έχει την ικανότητα να αυξάνει την καμπυλότητα του για δυνατούς ήχους, κάτι που έχει ως συνέπεια να αυξάνει, ταυτόχρονα, και τη σκληρότητα του.

1.2.4 Ανατομία του αυτιού / Μέσο αυτί / Σφύρα, Άκμονας, Αναβολέας:

Στο μέσο αυτί υπάρχουν τρία (3) οστά τα οποία είναι ενωμένα μεταξύ τους με μύες. Τα οστά αυτά είναι η σφύρα, ο άκμονας και ο αναβολέας. Η σφύρα είναι ενωμένη με το τύμπανο ενώ ο αναβολέας είναι συνδεδεμένος με το ελλειψοειδές παράθυρο



Σχ. 1.2.4.1: Οστά στο μέσο αυτί. Πιο συγκεκριμένα φαίνονται το τύμπανο, η σφύρα, ο άκμονας, ο αναβολέας και το ελλειψοειδές παράθυρο.

όπως φαίνεται και στο σχήμα 1.2.4.1. Τα οστά αυτά είναι τα μικρότερα που υπάρχουν στο ανθρώπινο σώμα. Η λειτουργία τους αφορά τη μετάδοση της ενέργειας από την ταλάντωση του τυμπάνου, εν συνεχεία τη μεταφορά της στο ελλειψοειδές παράθυρο και από εκεί, τέλος, στον κοχλία.

1.2.5 Ανατομία του αυτιού / Μέσο αυτί / Ευσταχιανή σάλπιγγα:

Συνδετικό στοιχείο του άνω μέρους του λαιμού, το οποίο βρίσκεται πίσω από τη ρινική κοιλότητα, με το μέσο αυτί, αποτελεί η ευσταχιανή σάλπιγγα. Το τύμπανο λειτουργεί σαν σύστημα ακουστικής ανάρτησης ενάντια στην υποχωρητικότητα του αέρα που είναι παγιδευμένος στο μέσο αυτί. Η ευσταχιανή σάλπιγγα είναι κατάλληλα μικρή ώστε να μην καταστρέφει αυτή την υποχωρητικότητα.

Εξυπηρετεί, επίσης, μία δεύτερη λειτουργία εξισώνοντας τη στατική πίεση του αέρα στο μέσο αυτί με την εξωτερική ατμοσφαιρική πίεση, έτσι ώστε να λειτουργούν σωστά το τύμπανο και οι ευαίσθητες μεμβράνες του έσω αυτιού. Κάθε φορά που καταπίνουμε, η ευσταχιανή σάλπιγγα ανοίγει ισορροπώντας την πίεση στο μέσο αυτί. Αλλαγές στην πίεση του αέρα που συμβαίνουν, για παράδειγμα, όταν μπαίνουμε σε ένα ανελκυστήρα, όταν ανεβαίνουμε ένα βουνό, στα αεροπλάνα ή κάτω από το νερό θα είχαν ως αποτέλεσμα η τυμπανική μεμβράνη να μετακινείται για να προσαρμοστεί στην αλλαγή της πίεσης. Σε αυτές τις συνθήκες η τυμπανική μεμβράνη βρίσκεται σε έκταση, τεντωμένη, εξαιτίας της διαφοράς της πίεσης μεταξύ του μέσου και του εξωτερικού αυτιού. Αυτό έχει ως αποτέλεσμα, οι αλλαγές στην πίεση που προκαλούνται από το ηχητικό κύμα να μη μετατρέπονται σωστά σε μηχανική ενέργεια από τα οστά του μέσου αυτιού. Αυτό έχει ως συνέπεια τη μείωση (για ελάχιστο χρονικό διάστημα) της ακουστικής ικανότητας. Υπάρχουν, όμως, και περιπτώσεις, όπως για παράδειγμα σε καταδύσεις, όπου αλλάζει απότομα η εξωτερική πίεση να προκληθούν μόνιμες βλάβες

στο τύμπανο. Στην περίπτωση του αεροπλάνου, όπου αλλάζει απότομα ύψος, χρειάζεται κάποιος δεδομένος χρόνος ώστε η πίεση στο μέσο αυτί να ισορροπήσει με κατάποση.

Τέλος η ευσταχιακή σάλπιγγα έχει ακόμη μία σημαντική ιδιότητα: την απαγωγή των υγρών σε περιπτώσεις μόλυνσης του αυτιού.

1.2.6 Εσωτερικό αυτί:

Τοποθετημένο βαθιά στο πλάγιο μέρος του κρανίου, το εσωτερικό αυτί περιέχεται μέσα σε ένα μη συμπαγές οστό, το οποίο έχει πυραμοειδή μορφή και της οποίας η κορυφή είναι στραμμένη στο εσωτερικό και εμπρόσθιο τμήμα του κρανίου. Το εσωτερικό αυτί παρουσιάζεται ως μία οστέινη θήκη, η οποία καλείται λαβυρινθική κύστη ή οστέινος λαβύρινθος. Η δομή του κελύφους είναι πυκνή και περιέχει το μεμβρανώδη λαβύρινθο, ο οποίος αποτελείται από το αιθουσαίο και από τον κοχλία. Πρόκειται για ένα σύνολο ιδιαίτερα πολύπλοκο, λόγω της ανατομικής κατασκευής των διαφόρων στοιχείων από τα οποία απαρτίζεται. Η διάκριση δε των κοινών δεσμών που συνδέουν τα διάφορα μέρη του εσωτερικού αυτιού θεωρείται, και είναι, σχεδόν αδύνατη.



Σχ. 1.2.6.1: Απλοποιημένη απεικόνιση σε τομή του εσωτερικού αυτιού.

Οι δυνατότητες του εσωτερικού αυτιού καλύπτουν το εύρος από την αντίληψη των σωματικών κινήσεων μέχρι την ικανότητα ανίχνευσης και ανάλυσης των ελαχίστων ηχητικών πιέσεων. Ο μεμβρανώδης λαβύρινθος περιλαμβάνει δύο όργανα, το ένα ονομάζεται αιθουσαίο και το άλλο κοχλίας.

1.2.7 Εσωτερικό αυτί / Αιθουσαίο:

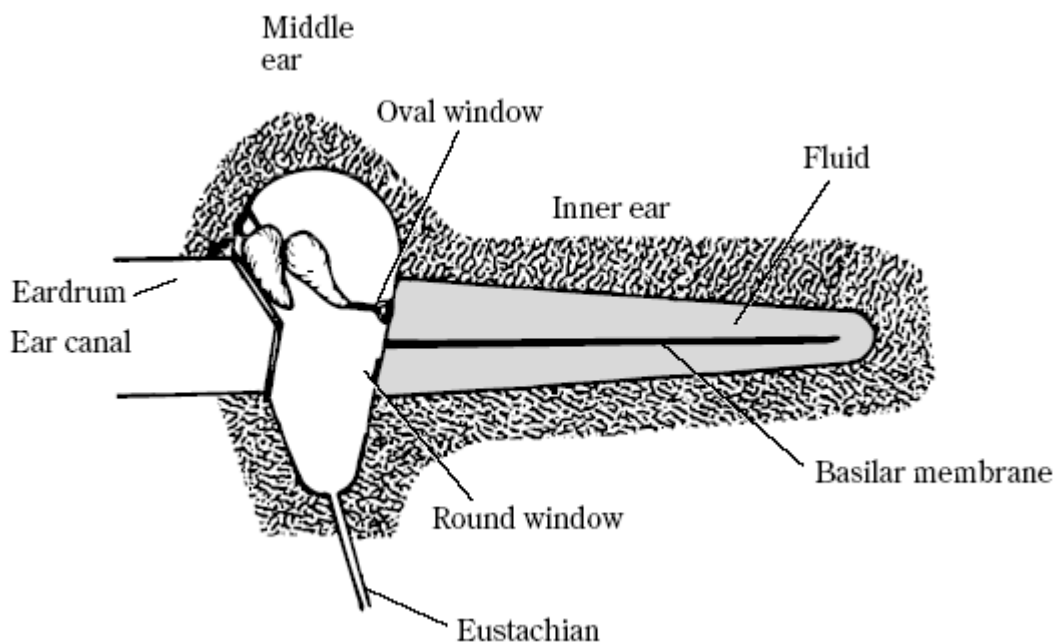
Το αιθουσαίο αποτελείται από το ελλειπτικό κυστίδιο, στο οποίο ακουμπούν οι ημικυκλικοί σωλήνες και το σφαιρικό κυστίδιο. Πιο συγκεκριμένα, περιλαμβάνει:

- i. Το ελλειπτικό κυστίδιο που οργανώνει, κυρίως, τις κινήσεις στο πεδίο του οριζοντίου άξονα και ιδιαίτερα στο επίπεδο του κεφαλιού.
- ii. Το σφαιρικό κυστίδιο που συμπεριφέρεται με ανάλογο τρόπο, αλλά επικεντρώνει περισσότερο στην καθετότητα και ιδιαίτερα στον τομέα της σπονδυλικής στήλης.
- iii. Τους ημικυκλικούς σωλήνες που τελειοποιούν τις πλάγιες γωνιώσεις ή τις γωνιώσεις στο εξωτερικό - οπίσθιο επίπεδο, καθώς επίσης και ενισχύουν την αντίληψη της στάσης του σώματος στις τρεις διαστάσεις του χώρου.

1.2.8 Εσωτερικό αυτί / Κοχλίας:

Το εσωτερικό αυτί και ιδιαίτερα ο κοχλίας είναι τα πιο σημαντικά μέρη της ανθρώπινης ακοής. Πολλά από τα όσα γνωρίζουμε για τη λειτουργία του κοχλίου προέρχονται από τις μελέτες του G. von Békésy.

Ο κοχλίας έχει τη μορφή σαλιγκαριού με σπείρα από δύομιση έως τρεις έλικες. Το μήκος του σε έκταση είναι περίπου 34 mm από τη βάση έως το ελικότρημα. Το εσωτερικό του είναι γεμάτο με υγρό και στη έκταση του χωρίζεται από τη βασική μεμβράνη. Ο κοχλίας επικοινωνεί με το μέσο αυτί, διαμέσου του οβάλ παραθύρου (oval window) στο οποίο είναι συνδεδεμένος ο αναβολέας. Ο κοχλίας τελειώνει στο στρογγυλό παράθυρο (round window), το οποίο δρα όπως ένας μηχανισμός εκτόνωσης της πίεσης. Στο σχήμα 1.2.8.1 βλέπουμε τον κοχλίο σε έκταση.



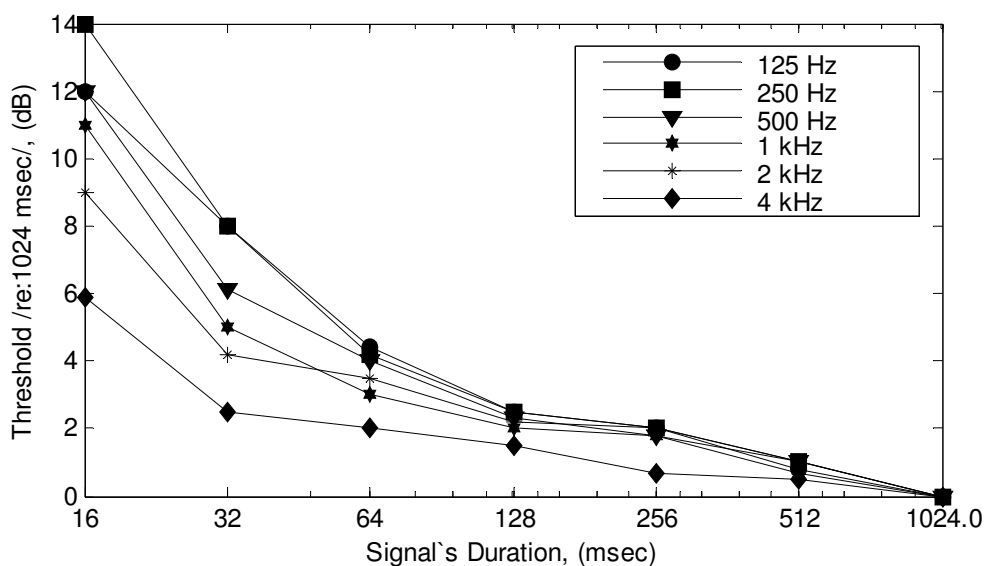
Σχ. 1.2.8.1: Στο σχήμα αυτό φαίνονται το μέσο αυτί, καθώς και ο κοχλίας σε έκταση.

1.3 Απόκριση του ακουστικού συστήματος:

Στις προηγούμενες παραγράφους ασχοληθήκαμε, περιληπτικά, με την ανατομία και τη φυσιολογία του ακουστικού συστήματος. Γνωρίζουμε, επίσης, ότι οι τρεις βασικές ιδιότητες του ήχου, που είναι η συχνότητα, η ένταση και η χρονική διάρκεια, μεταβαίνουν στον εγκέφαλο από τα νευρικά κύτταρα με τη μορφή ενός νευρικού κώδικα. Στις παραγράφους που ακολουθούν θα ασχοληθούμε με την ευαισθησία του ακουστικού μας συστήματος και με κάποια φαινόμενα μη γραμμικότητας που παρουσιάζονται σε αυτό.

1.3.1 Ευαισθησία της ακοής στη διάρκεια των ήχων & χρονική ολοκλήρωση:

Για τον προσδιορισμό της ευαισθησίας της ανθρώπινης ακοής, έχει μεγάλη σημασία να έχει προβλεφτεί η κατάλληλη χρονική διάρκεια των τόνων που θα χρησιμοποιηθούν για τον προσδιορισμό της. Το σχήμα 1.3.1.1 παρουσιάζει τα κατώφλια ακουστότητας για διάφορες συχνότητες συναρτήσει της χρονικής διάρκειας του σήματος. Οι τιμές αυτές προσδιορίζονται συσχετίζοντάς τες με τη χρονική διάρκεια που χρειάζεται για να ανιχνευθεί ένας τόνος διάρκειας 1024 msec αντίστοιχης συχνότητας.



Σχ. 1.3.1.1: Κατώφλιο ακουστότητας συναρτήσει της διάρκειας του σήματος για διαφορετικές συχνότητες.

Κάποιος θα μπορούσε να υποθέσει ότι όσο περισσότερο διαρκεί ένας ήχος, τόσο περισσότερο εύκολο είναι να τον ακούσουμε. Παρατηρούμε όμως ότι τα κατώφλια ακουστότητας, για διάρκεια μεγαλύτερη από 300 έως 500 msec, δεν αλλάζουν σημαντικά. Με άλλα λόγια σημαίνει ότι για χρονική διάρκεια μεγαλύτερη από

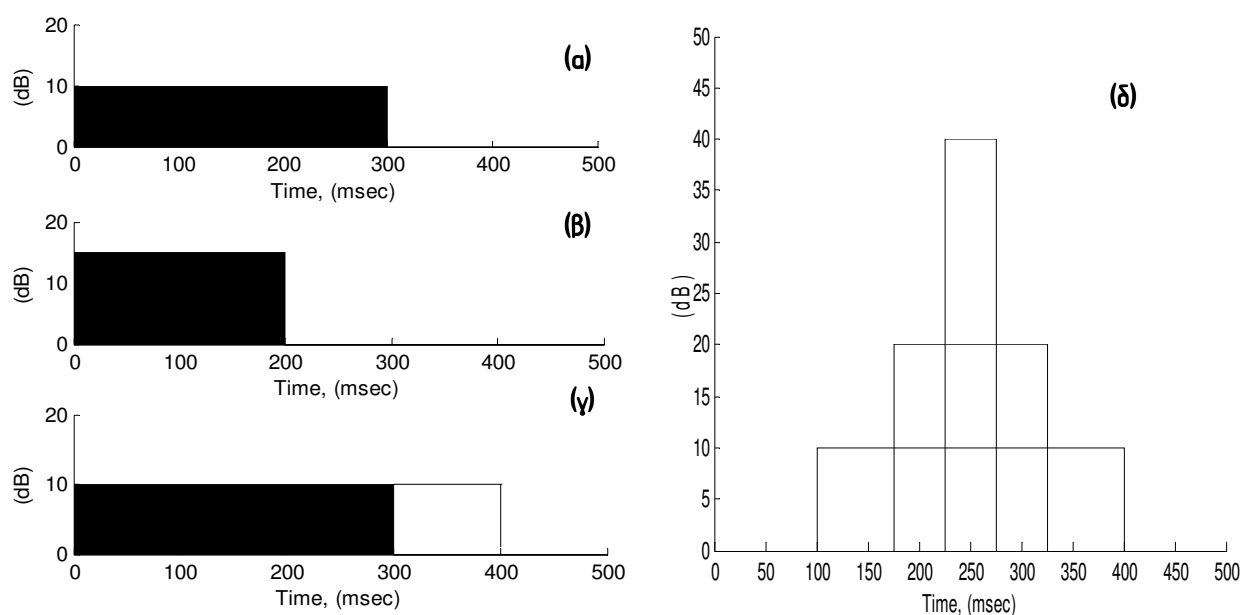
τις προαναφερθείσες τιμές, το σήμα δε γίνεται πιο εύκολο να ανιχνευθεί απο το ακουστικό μας σύστημα. Βλέπουμε, όμως, ότι για διάρκεια μικρότερη απο 300 msec πρέπει να αυξήσουμε τα dB προκειμένου ο ήχος να γίνει αντιληπτός. Συνεπώς, γίνεται κατανοητό, ότι για αύξηση της διάρκειας του σήματος παραπάνω απο 300 msec δεν προκύπτει αύξηση της ικανότητας αντίληψης του ήχου. Σαν το ακουστικό σύστημα να απαιτεί το χρονικό αυτό διάστημα για να έχει υψηλή απόδοση. Για σήματα κάτω των 300 msec χρειάζεται να αυξήσουμε την ένταση τους προκειμένου να γίνουν κατανοητά για σήματα πάνω απο τα 300 msec δε χρειάζεται να αυξήσουμε την ένταση τους προκειμένου να ανιχνεύσουμε καλύτερα κατώφλια ακουστότητας.

Αυτή η ιδιότητα του ακουστικού μας συστήματος, να ανιχνεύει ήχους ανάλογα με το ποσοστό της ενέργειας που περιέχουν, καλείται χρονική ολοκλήρωση (temporal integration).

Φαίνεται πως ένα σήμα πρέπει να έχει μια οριακή ενέργεια προκειμένου να ανιχνευθεί. Αυτό παρουσιάζεται στο σχήμα 1.3.1.2. Γνωρίζουμε ότι η ενέργεια ισούται με το γινόμενο της δύναμης επί το χρόνο, πράγμα που δίδεται απο την πολύ απλή σχέση:

$$E = P \cdot t \quad (1.3.1.1)$$

Στο παράδειγμα (α) του σχήματος 1.3.1.2 η σκιασμένη επιφάνεια αντιστοιχεί στο ποσό της ενέργειας που απαιτείται προκειμένου το σήμα να ανιχνευθεί. Στο παράδειγμα (β), επειδή έχει μειωθεί ο χρόνος του σήματος, έχει αυξηθεί το ποσοστό που απαιτείται. Στο παράδειγμα (γ) βλέπουμε ότι δεν απαιτείται παραπάνω δύναμη ή χρόνος προκειμένου να ανιχνευθεί το σήμα. Στο παράδειγμα (δ) παρουσιάζονται τρεις διαφορετικές περιπτώσεις σημάτων, όπου το καθένα απο αυτά έχει διαφορετική σχέση διάρκειας - ποσοστού έντασης, των οποίων, όμως, η αίσθηση δε διαφοροποιείται αναφορικά με τον τρόπο που τα αντιλαμβάνεται η ανθρώπινη ακοή.



Σχ. 1.3.1.2: Σχηματική αναπαράσταση της χρονικής ολοκλήρωσης (Temporal Integration)

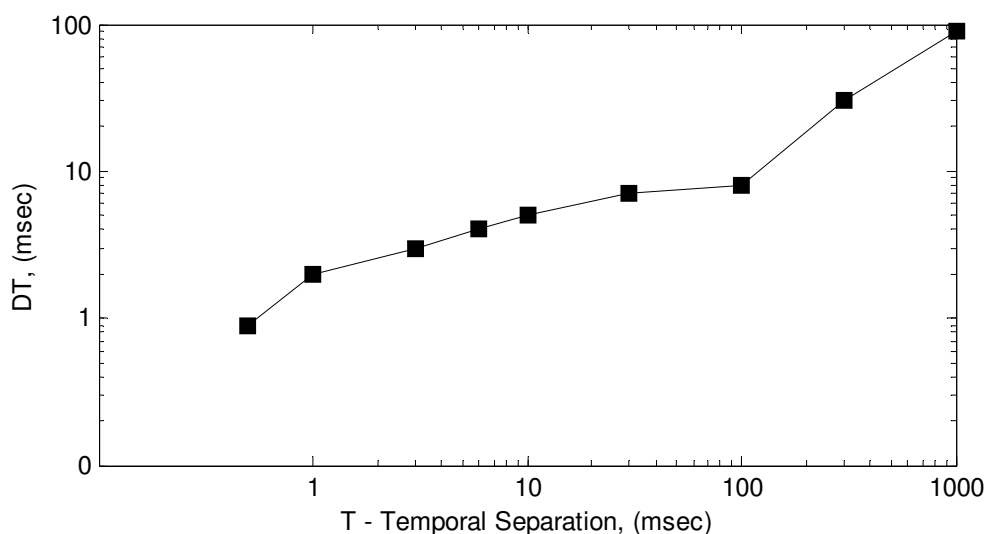
1.3.2 Χρονική διάκριση (Temporal Discrimination):

Αν ζητηθεί από ακροατές να διακρίνουν τη διαφορά μεταξύ ενός ημιτονοειδούς σήματος διάρκειας 50 msec και ενός 60 msec, τότε θα πρέπει να καθοριστούν περισσότερες μεταβλητές για τη διεξαγωγή αυτού του πειράματος. Όπως είδαμε και στο σχήμα 1.3.1.1, ήχοι διαφορετικής διάρκειας έχουν διαφορετικά κατώφλια ακουστότητας, οπότε τα σήματα των 50 και 60 msec θα έχουν διαφορές στη διακρισιμότητα τους. Επίσης, το φάσμα ενός ήκου εξαρτάται από τη διάρκεια του. Αυτοί οι δύο παράγοντες κάνουν τη χρονική διάκριση πιο δύσκολη.

Ορισμένοι ερευνητές έχουν προσπαθήσει να ξεπεράσουν αυτά τα προβλήματα χρησιμοποιώντας μία άλλου τύπου μέθοδο. Στους ακροατές παρουσιάζεται ένας ήχος διάρκειας 170 msec και στη συνέχεια, έπειτα από κενό της τάξης των 10 msec, ξαναπαρουσιάζεται ο ίδιος ήχος των 170 msec. Στη συνέχεια η μέτρηση επαναλαμβάνεται με διαφορά, όμως, 15 msec μεταξύ των δύο ήχων. Ο ακροατής καλείται να κρίνει εάν η χρονική διάρκεια ανάμεσα στους δύο ήχους είναι ίδια και για τις δύο μετρήσεις. Αν κριθεί σωστά, δηλ. ότι υφίσταται διαφορά, τότε μπορεί να υποθέσει κάποιος ότι το ακουστικό μας σύστημα μπορεί να διακρίνει αυτή τη διαφορά των 5 msec.

Το σχήμα 1.3.2.1 παρουσιάζει τα δεδομένα από ένα τέτοιο πείραμα. Ο σταθερός ήχος, διάρκειας 170 msec, αποτελείται από μία συχνότητα 1000 Hz στα 85 dB SPL. Συνεπώς, στο παράδειγμα που δώσαμε πριν, το T είναι ο χρόνος των 10 msec και το ΔT είναι η διαφορά 15 - 10 msec.

Σύμφωνα, λοιπόν, με τη μέτρηση αυτή, αλλά και σύμφωνα με άλλες μετρήσεις, όσο αυξάνεται ο χρόνος T , τόσο αυξάνεται και το απαιτούμενο ΔT . Το κλάσμα, όμως, του Weber δεν είναι σταθερό και αυτό γιατί για T ίσο με 1 msec το ΔT είναι ίσο με 2 msec, οπότε το κλάσμα είναι ίσο με 2. Όταν το T ισούται με 300 msec, το ΔT ισούται με 30 msec και άρα το κλάσμα ισούται με 0.1.

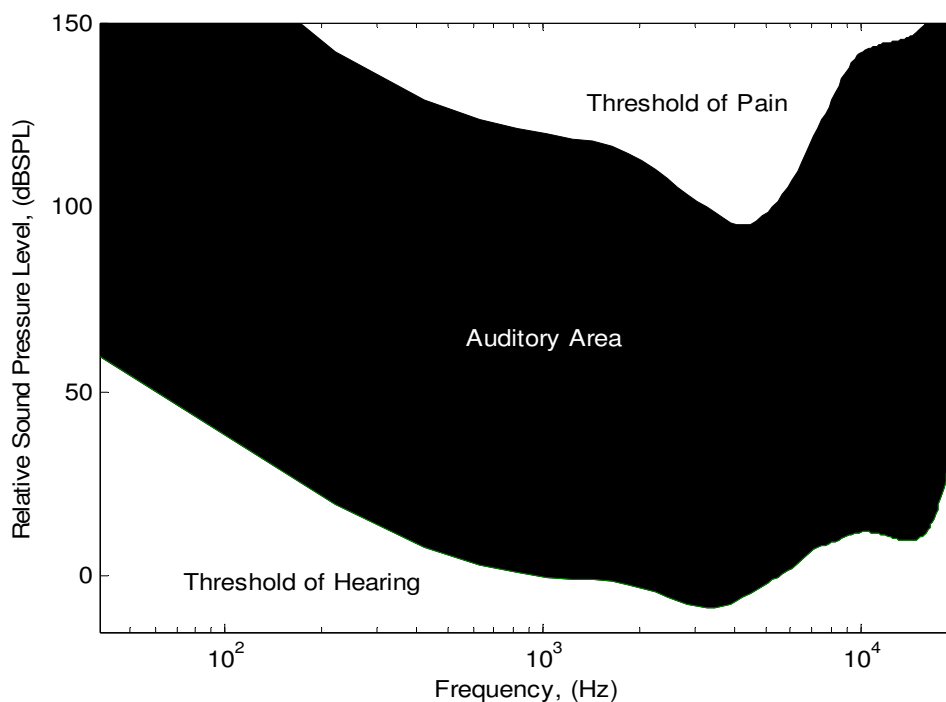


Σχ. 1.3.2.1: Σχηματική παρουσίαση πειραματικών δεδομένων για το Temporal Discrimination

1.3.3 Όρια ακουστότητας (Threshold of audibility):

Το πρώτο όριο ακουστότητας με το οποίο θα ασχοληθούμε, είναι το λεγόμενο κατώφλι ακουστότητας.

Το κατώφλι ακουστότητας ενός ήχου είναι η ελάχιστη αντιληπτή ένταση του ήχου αυτού, δίχως την παρουσία άλλων ήχων. Το διάγραμμα του σχήματος 1.3.3.1 που δίδεται παρακάτω παρουσιάζει μετρήσεις για τα όρια της ακουστότητας. Πιο συγκεκριμένα παρουσιάζει το κατώφλι ακουστότητας, καθώς και το λεγόμενο όριο του πόνου. Συνακόλουθα το εν λόγω διάγραμμα παρουσιάζει το εύρος και την έκταση της ανθρώπινης ακοής. Σύμφωνα δε με τα συμπεράσματα της προηγούμενης παραγράφου, η χρονική διάρκεια των μετρήσεων αυτών ήταν μεγαλύτερη των 400 msec. Οι μετρήσεις που παρουσιάζονται αποτελούν το κατά ISO πρότυπο 226:2003. Τα πρότυπα αυτά χρησιμοποιούνται όταν ακουστικές συσκευές κατασκευάζονται ή όταν ακουστικές μετρήσεις διεξάγονται.



Σχ. 1.3.3.1: Το κατώφλι ακουστότητας και το όριο του πόνου σε γραφική μορφή. Η σκιασμένη περιοχή υποδηλώνει το εύρος της ανθρώπινης ακοής.

Απο την κάτω καμπύλη του σχήματος (Threshold of hearing) παρατηρούμε μια αύξηση του κατωφλίου για συχνότητες γύρω από τα 1000 Hz και πάνω από τα 4000 Hz. Οι συναρτήσεις μεταφοράς (transfer functions) του εξωτερικού και του μέσου αυτιού θεωρείται ότι εξηγούν τη μορφή που έχει το κατώφλι ακουστότητας για τις συχνότητες αυτές. Ιδιαίτερα η συνάρτηση μεταφοράς του μέσου αυτιού ευθύνεται για την απώλεια ευαισθησίας στις χαμηλές συχνότητες. Το γεγονός ότι η κατασκευή του μέσου και του εξωτερικού αυτιού έχει μάζα και σαν αποτέλεσμα αδράνεια, τα κάνει να

λειτουργούν ως ένα είδος low - pass φίλτρου, με αποτέλεσμα να υπάρχει απώλεια της ευαισθησίας και στις υψηλές συχνότητες.

Στο κατώφλι της ακουστότητας 0 dB SPL αντιστοιχούν σε 20 μPa , τα οποία είναι μία πολύ χαμηλή τιμή πίεσης. Έχει υπολογιστεί ότι για αυτή την τιμή πίεσης και για συχνότητα 1000 Hz, η τυμπανική μεμβράνη, καθώς δονείται, μετακινείται κατά μία απόσταση περίπου ίση με τη διάμετρο ενός ατόμου του υδρογόνου.

Εκτός από το κατώφλι ακουστότητας, υπάρχει και το λεγόμενο όριο του πόνου (Threshold of pain). Για τη μέτρηση αυτού του μεγέθους ο ακροατής καλείται να πει πότε ο ήχος του προκαλεί πόνο.

Οι μεταξύ τους αναλογίες (μεταξύ των δύο αυτών ορίων δηλαδή) παραμένουν σχετικά σταθερές σαν συνάρτηση της συχνότητας. Επομένως, προκύπτει, πως η δυναμική περιοχή του ακουστικού μας συστήματος αλλάζει συναρτησίως της συχνότητας. Είναι περίπου στα 130 dB SPL για τη συχνότητα των 1000 Hz και γύρω στα 90 dB SPL για τη συχνότητα των 100 Hz.

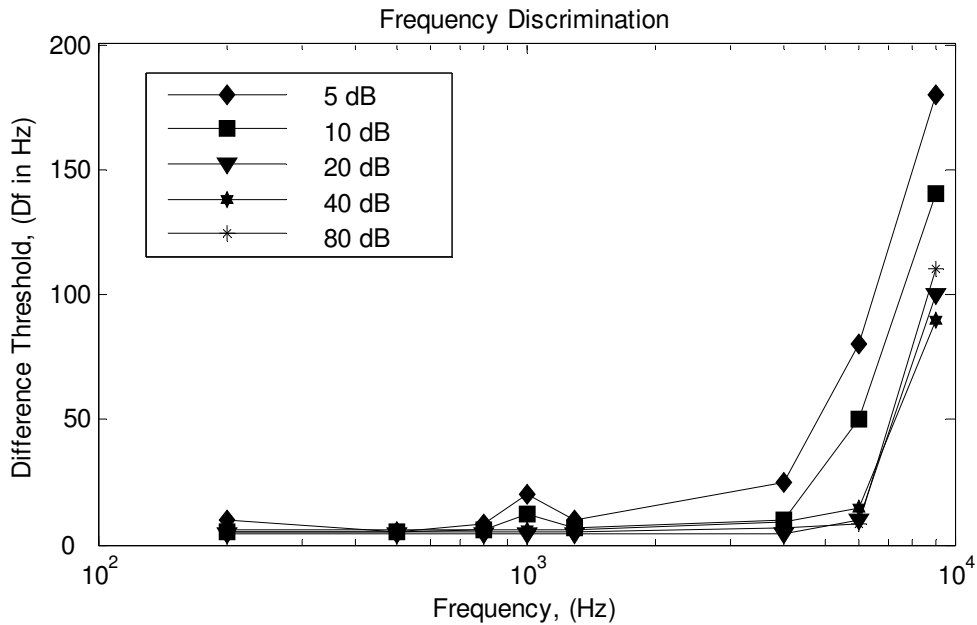
1.3.4 Ευαισθησία σε αλλαγές έντασης, συχνότητας & χρονικής διάρκειας:

Αν και τα όρια της ακουστότητας καθορίζουν την έκταση των συχνοτήτων και το εύρος των εντάσεων που μπορούμε να ακούσουμε, εντούτοις δεν περιγράφουν την ευαισθησία μας στις αλλαγές των μεγεθών αυτών.

Στις αρχές του 1800 ο Weber παρατήρησε ότι είναι εύκολο να διακρίνουμε τη διαφορά ανάμεσα σε βάρος του 1 κιλού και των 2 κιλών, αλλά όχι τόσο εύκολο τη διαφορά μεταξύ 100 και 101 κιλών, παρόλο που διαφορά παραμένει 1 κιλό. Βρέθηκε ότι η διαφορά ανάμεσα σε βάρη που μπορεί να ανιχνευθεί, είναι ανάλογη του μικρότερου από τα δύο βάρη. Αυτό σημαίνει πως αν η ελάχιστη αντιληπτή διαφορά (just noticeable difference - jnd) για ένα βάρος 1 κιλού είναι 0.1 κιλά, τότε για το βάρος των 100 κιλών είναι 10 κιλά (η αναλογία 0.1/1 είναι ίδια με την αναλογία 10/100). Ο Weber όρισε αυτή τη σχέση ως $\Delta S = \text{constant}$, όπου ΔS είναι η jnd και S είναι η μικρότερη από τις δύο ποσότητες προς διάκριση. Η ποσότητα $\Delta S/S$ ονομάζεται κλάσμα του Weber (Weber fraction). Το ερώτημα που προκύπτει τώρα, και μας αφορά, είναι εάν το κλάσμα του Weber ισχύει για τη συχνότητα και τη χρονική διάρκεια.

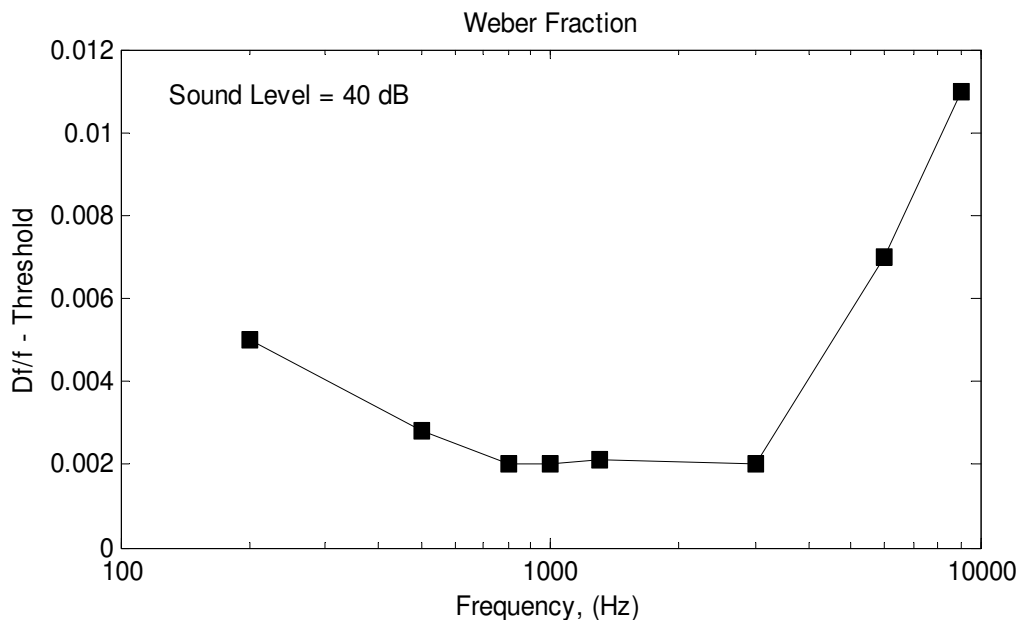
1.3.5 Συχνοτική διάκριση (Frequency Discrimination):

Το σχήμα 1.3.5.1 δείχνει τις τιμές του Δf (των διαφορών συχνότητας) οι οποίες απαιτούνται για να γίνει ακουστή μια διαφοροποίηση στη συχνότητα σε σχέση με μία συχνότητα f . Οι διαφορετικές καμπύλες στο σχήμα αντιπροσωπεύουν διαφορετικές εντάσεις. Καταδεικνύεται πως η τιμή του Δf αυξάνει για τιμές του f πάνω από τα 1000 Hz.



Σχ. 1.3.5.1: Διαφορές συχνότητας (Δf), για διάφορες τιμές έντασης, που προσδιορίζουν το Frequency Discrimination.

Στο σχήμα 1.3.5.2 που ακολουθεί φαίνεται ότι το κλάσμα του Weber παραμένει σταθερό στη μεσαία περιοχή συχνοτήτων. Στην περιοχή αυτή ισχύει ότι $\frac{\Delta f}{f} = 0.002$ ή 0.2 %. Αυτό σημαίνει ότι το Δf μπορεί να είναι ακόμη και 1 Hz σε κάποιες περιπτώσεις. Για παράδειγμα αν το f είναι 500 Hz και το $\frac{\Delta f}{500} = 0.002$, τότε $\Delta f = 500 \cdot 0.002 = 1 \text{ Hz}$.



Σχ. 1.3.5.2: Σχηματική αναπαράσταση του κλάσματος του Weber για συχνότητες $\Delta f/f$.

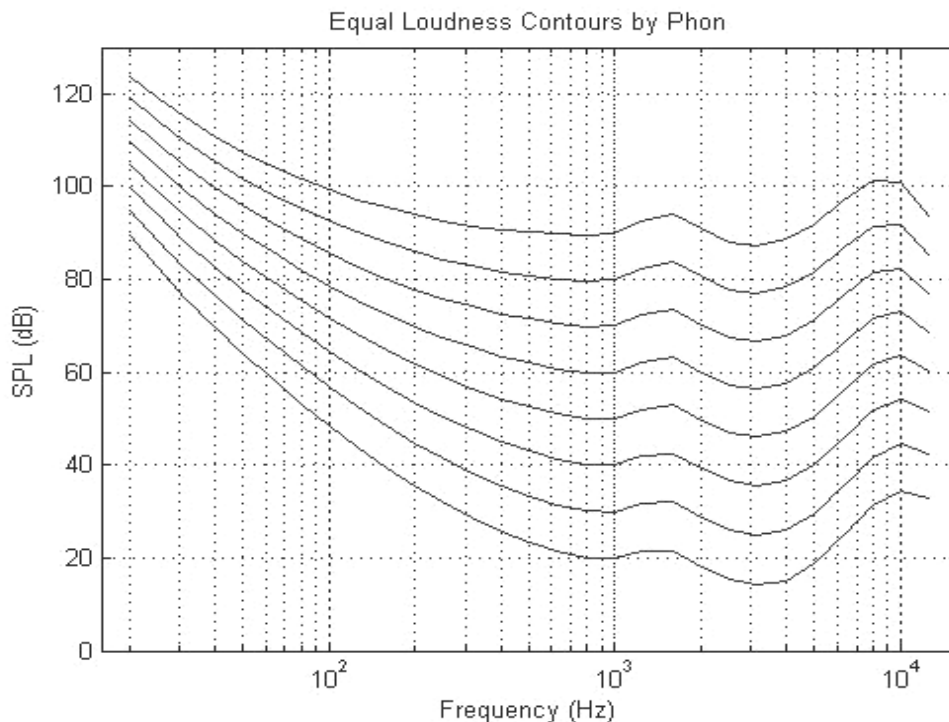
1.3.6 Αντίληψη / Διάκριση της έντασης:

Η αναγνώριση της έντασης ενός ήχου είναι ένα φαινόμενο αρκετά περίπλοκο. Ο όρος που χρησιμοποιείται για την περιγραφή της έντασης του υποκειμενικού αισθήματος είναι η Ακουστότητα (Loudness). Αυτή, λοιπόν, αντιστοιχεί στην ένταση (Intensity) του ήχου, αλλά όχι μόνο σε αυτή. Πιο συγκεκριμένα, η ακουστότητα ενός ήχου εξαρτάται από την ένταση και τη συχνότητα του, με τρόπο που δεν είναι σε όλες τις περιπτώσεις όμοιος, αλλά διαφέρει για κάθε διαφορετική τιμή του μεγέθους της έντασης. Κοντολογίς, το ανθρώπινο αισθητήριο της ακοής έχει μια κάποια καμπύλη απόκρισης η οποία δεν είναι επίπεδη. Επιπροσθέτως, όπως φαίνεται και στο σχήμα 1.3.6.1 στο οποίο παρουσιάζονται οι καμπύλες ίσης ακουστότητας κατά ISO 226:2003, η καμπύλη αυτή (η οποία είναι η ανεστραμμένη αυτών της ίσης ακουστότητας) είναι διαφορετική για διαφορετικές τιμές έντασης.

Όπως έχουμε ήδη δει, από την παράγραφο 1.2.2, το αυτί μπορεί να θεωρηθεί σαν ένας ανοικτός στη μία και κλειστός στην άλλη μεριά σωλήνας. Ο έξω ακουστικός πόρος έχει μήκος περίπου 2.5 cm, οπότε η θεμελιώδης συχνότητα συντονισμού του σωλήνα προκύπτει:

$$f_1 = \frac{c}{4L} = \frac{343}{0.1} = 3430 \text{ Hz} \quad (1.3.6.1)$$

Στη συχνότητα αυτή το αυτί παρουσιάζει τη μεγαλύτερη ευαισθησία. Όπως παρατηρούμε σε αυτή τη συχνότητα περίπου παρουσιάζουν ελάχιστο και οι καμπύλες ίσης ακουστότητας.



Σχ. 1.3.6.1: Οι καμπύλες ίσης ακουστότητας όπως αυτές προκύπτουν κατά το πρότυπο ISO 226:2003.

Οι καμπύλες αυτές στηρίζονται στις εργασίες που έγιναν από τους Fletcher και Steinberg αρχικά και Fletcher και Munson στη συνέχεια, στα εργαστήρια της Bell. Μεταγενέστερες μετρήσεις που έγιναν, κυρίως από ιάπωνες σε συνεργασία με ερευνητές από τις Ηνωμένες Πολιτείες, τη Δανία και τη Γερμανία είχαν αποτελέσματα που οδήγησαν στην αναθεώρηση του προτύπου (ISO 226:2003, βλ. σχ. 1.3.6.1).

Μονάδα μέτρησης της ακουστότητας είναι το phon. Η στάθμη ακουστότητας ενός ήχου σε phons ορίζεται σαν η στάθμη πίεσης ενός απλού ήχου συχνότητας 1 kHz που ηχεί το ίδιο δυνατά με αυτό τον ήχο για το μέσο ακροατή. Δηλαδή 1 phon είναι η στάθμη ακουστότητας ενός απλού ήχου με $f = 1 \text{ kHz}$ και $SPL = 1 \text{ dB}$. Εξ' αιτίας αυτού του ορισμού η κλίμακα των phons είναι λογαριθμική όπως αυτή της στάθμης ηχητικής πίεσης.

Τα phons δε μας δίνουν τη δυνατότητα να εκτιμήσουμε πόσο πιο δυνατός (ακουστός) είναι ένας ήχος από ένα άλλο (ένας ήχος με στάθμη ακουστότητας 60 phons δεν είναι δύο φορές πιο δυνατός από ένα με στάθμη ακουστότητας 30 phons). Τα αποτελέσματα μετρήσεων σε πολλούς ακροατές έδειξαν ότι η ακουστότητα ενός ήχου διπλασιάζεται όταν η στάθμη αναβαίνει κατά 10 phons. Για να κατατάξουμε τους ήχους ως προς την ακουστότητα χρησιμοποιούμε το sone. Η σχέση των δύο μονάδων είναι:

$$S = 2^{\frac{P-40}{10}} \quad (1.3.6.2)$$

Συνεπώς αν:

$$P_2 = P_1 + 10 \Rightarrow S_2 = 2^{\frac{(P_1+10)-40}{10}} = 2^{\frac{P_1-40}{10}} 2^{\frac{10}{10}} \Rightarrow S_2 = 2 \cdot S_1$$

Εν συνεχεία, λογαριθμίζοντας την (1.3.6.2):

$$\log S = \frac{P-40}{10} \log 2 \Rightarrow P-40 = \frac{10}{\log 2} \log S \Rightarrow$$

$$P = 40 + 33.3 \log S \quad (1.3.6.3)$$

Τα κύρια χαρακτηριστικά που προκύπτουν απ' όλα τα παραπάνω και κυρίως από τις καμπύλες ίσης ακουστότητας είναι πως το ανθρώπινο αυτί είναι πολύ ευαίσθητο στη μεσαία περιοχή συχνοτήτων και κυρίως γύρω από τα 4000 Hz. Η ευαισθησία μειώνεται ως ένα βαθμό στις υψηλές συχνότητες, αλλά η κυρίως μείωση της ευαισθησίας του αυτιού αφορά τις χαμηλές συχνότητες. Το χαρακτηριστικό αυτό είναι έντονο στις χαμηλές εντάσεις (λίγα phons), ενώ στις υψηλότερες σταδιακά γίνεται λιγότερο έντονο. Δηλαδή όσο προχωράμε προς τις υψηλότερες εντάσεις η καμπύλη του αυτιού τείνει να γίνεται περισσότερο επίπεδη.

1.3.7 Τονικό ύψος:

Απο την απλή φυσική σκοπιά, ύψος ενός τόνου είναι η συχνότητα του. Στους σύνθετους ήχους το ύψος χαρακτηρίζεται απο τις συχνότητες των αρμονικών, αλλά και απο την ηχητική πίεση και την κυματομορφή τους. Για το αυτό, το ύψος είναι το χαρακτηριστικό της ακουστικής αίσθησης με βάση το οποίο οι ήχοι κατατάσσονται και συγκρίνονται. Το ύψος επηρεάζεται σημαντικά απο την ηχητική ένταση στις πολύ χαμηλές συχνότητες και στις πολύ υψηλές συχνότητες. Όταν η ηχητική πίεση αυξάνεται, το ύψος ενός τόνου χαμηλής συχνότητας ελαττώνεται, ενώ αντίθετα, αυτό ενός τόνου υψηλής συχνότητας αυξάνει.

Η ακουστική μονάδα που χρησιμοποιούμε για να περιγράψουμε το ύψος ενός ήχου λέγεται mel. Ένας απλός ήχος, συχνότητας 1 kHz και στάθμης ακουστότητας 40 rhaps ορίζεται να έχει ύψος 1000 mels. Έτσι ένας ήχος που ακούγεται να έχει διπλάσιο ύψος απο αυτόν, λέμε ότι έχει ύψος 2000 mels.

Δύο ήχοι, τώρα, με πολύ κοντινές συχνότητες ($\omega_1 \approx \omega_2$) θα ακουστούν να έχουν το ίδιο ύψος. Όμως το πλάτος, άρα και η ακουστότητα, του συνισταμένου ήχου ξέρουμε ότι μεταβάλλεται με συχνότητα ίση με τη διαφορά των συνιστωσών συχνοτήτων. Το ολικό κύμα που προκύπτει περιγράφεται απο τη σχέση:

$$p = p_1 + p_2 = p_0 \cos \omega_1 t + p_0 \cos \omega_2 t \Rightarrow p = 2p_0 \cos \frac{\omega_1 - \omega_2}{2} t \cdot \cos \frac{\omega_1 + \omega_2}{2} t \quad (1.3.7.1)^1$$

Δηλαδή θα είναι ένα κύμα με συχνότητα $\frac{\omega_1 + \omega_2}{2}$ και με χρονικά μεταβαλλόμενο πλάτος που δίδεται απο τη σχέση:

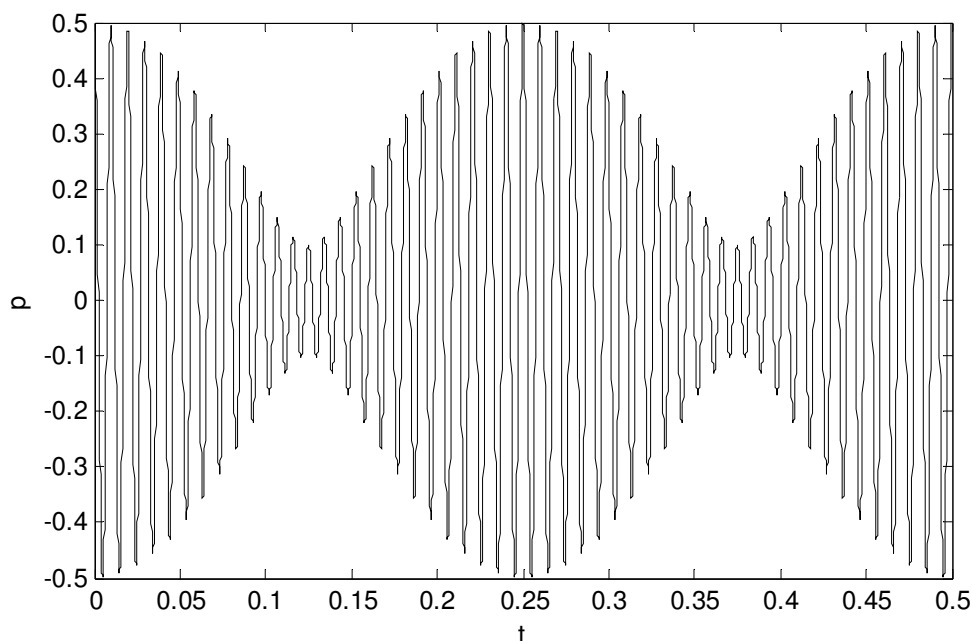
$$p'_0 = 2p_0 \cos \frac{\Delta\omega t}{2} \quad (1.3.7.2)$$

Τα ηχητικά αυτά κύματα με περιοδική μεταβολή του πλάτους τους, εμφανίζονται όταν έχουμε ταυτόχρονη διάδοση δύο ηχητικών κυμάτων με κοντινές συχνότητες, και λέγονται διακροτήματα. Δύο διαδοχικά μέγιστα συμβαίνουν όταν:

$$\frac{\Delta\omega t}{2} = 2\eta\pi \Rightarrow t_\eta = \frac{4\eta\pi}{\Delta\omega} \quad (1.3.7.3)$$

¹ Η σχέση (1.3.6.1) προκύπτει με τη χρήση της τριγωνομετρικής ταυτότητας:

$$\cos a + \cos b = 2 \cos \frac{a+b}{2} \cos \frac{a-b}{2}$$



Σχ. 1.3.7.1: Διακρότημα που προκύπτει από δύο καθαρούς τόνους συχνοτήτων 100 Hz & 104 Hz.

Σε ένα πλήρη κύκλο μεταβολής του πλάτους υπάρχουν δύο διακροτήματα. Συνεπώς, η περίοδος των διακροτημάτων δίδεται από τη σχέση:

$$2T_b = t_{\eta+1} - t_{\eta} = \frac{4\pi}{\Delta\omega} \Rightarrow T_b = \frac{2\pi}{\Delta\omega} \quad (1.3.7.4)$$

και η συχνότητα:

$$f_b = \frac{\Delta\omega}{2\pi} = |f_1 - f_2| \quad (1.3.7.5)$$

Διαγραμματικά το διακρότημα δίδεται στο σχ. 1.3.7.1 από το οποίο γίνεται σαφής η χρονική μεταβολή του πλάτους του σήματος.

Προχωρώντας παραπέρα να ορίσουμε, καταρχήν, τι εννοούμε όταν λέμε τονικό ύψος. Το τονικό ύψος (pitch), είναι η υποκειμενική απόκριση του αυτιού στη συχνότητα.

Όμως, όπως ο ήχος δεν περιέχει κάποια άμεσα μετρήσιμη μεταβλητή για την ακουστότητα, έτσι δεν περιέχει και κάποια άμεσα μετρήσιμη μεταβλητή για το τονικό ύψος. Η ακουστότητα και το τονικό ύψος είναι ιδιότητες της αντίληψης που παρέχουν πληροφορίες για το περιεχόμενο το οποίο εκπέμπουν οι οποιεσδήποτε ηχητικές πηγές. Διαφορετικά φυσικά χαρακτηριστικά του ήχου προκαλούν την αντίληψη της ακουστότητας και διαφορετικά αυτή του τονικού ύψους. Και αυτά τα χαρακτηριστικά είναι αποτέλεσμα της επεξεργασίας που κάνει το ανθρώπινο ακουστικό σύστημα στις φυσικές μεταβλητές του ήχου όπως η συχνότητα, η ένταση και ο χρόνος.

Δύο βασικές θεωρίες για την αντίληψη του τονικού ύψους έχουν αναπτυχθεί. Η πρώτη είναι ευρύτερα γνωστή ως χωρική θεωρία ή θεωρία της συχνότητας· η δεύτερη

είναι γνωστή ως η θεωρία της περιοδικότητας ή ως χρονική θεωρία. Σύμφωνα με τη χωρική θεωρία, ο κοχλίας του συστήματος ακοής μετατρέπει μία δόνηση ως προς το χρόνο σε ένα ερέθισμα κατά μήκος της βασικής μεμβράνης που με τη σειρά του προκαλεί μία χωρικά εξαρτώμενη διέγερση των νευρώνων. Η χωρική θεωρία εξηγεί μερικά από τα φαινόμενα της ακουστικής αντίληψης, αλλά αποτυγχάνει να εξηγήσει κάποια άλλα.

Σύμφωνα, τώρα, με τη θεωρία της περιοδικότητας του τονικού ύψους, το αυτί εκτελεί μία χρονική ανάλυση του ηχητικού κύματος. Δηλαδή η χρονική κατανομή των παλμών που μεταφέρονται με το ακουστικό νεύρο, έχει κωδικοποιηθεί (μέσα σε αυτή) τη χρονική δομή του ηχητικού κύματος.

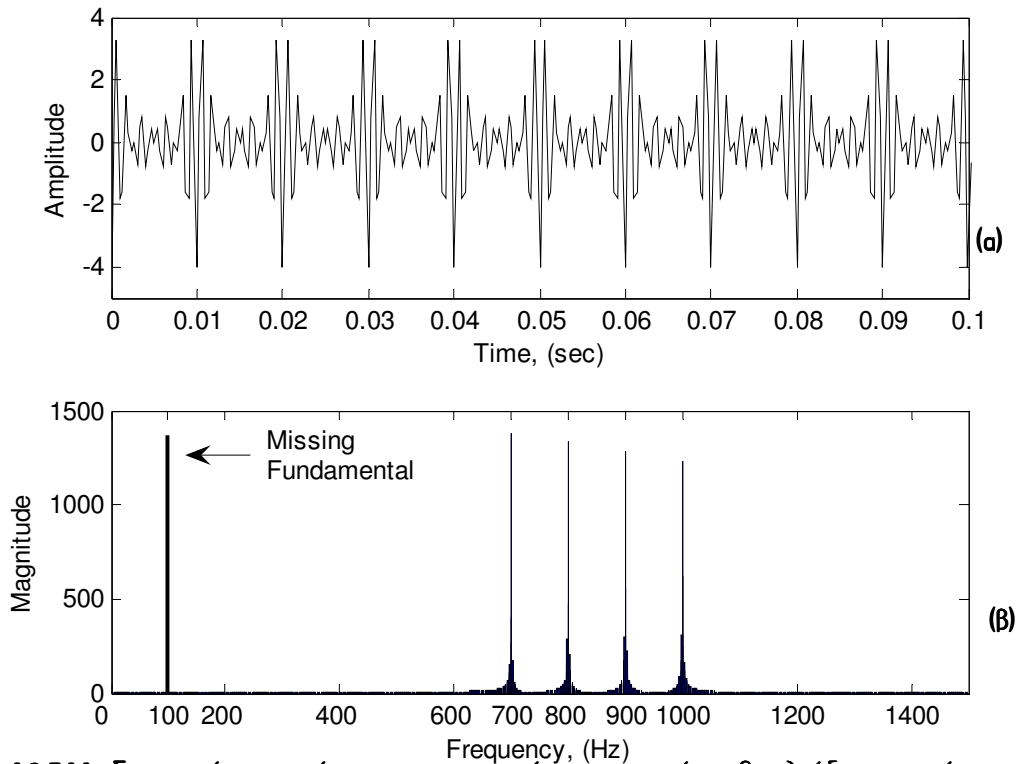
Για την αντίληψη του τονικού ύψους, ο Von Békésy βρήκε ότι ένας ήχος με συχνότητα κάτω από τα 1000 Hz πρέπει η διάρκεια του να είναι ίση με 3 έως 9 περιόδους προκειμένου να γίνει αντιληπτό το τονικό του ύψος. Πάνω από τα 1000 Hz αυτή η διάρκεια είναι 10 msec ανεξαρτήτως της συχνότητας του ήχου.

Πειράματα πάνω στο τονικό ύψος, συνήθως διεξάγονται με μια διαδικασία όπου ο ακροατής καλείται να απαντήσει εάν το τονικό ύψος δύο ήχων είναι το ίδιο. Οι τόνοι τους οποίους καλείται να συγκρίνει το πειραματικό υποκείμενο συνήθως είναι απλοί ημιτονοειδής τόνοι. Όπως έχουμε ήδη δει, για τη μέτρηση του τονικού ύψους ως υποκειμενικό αίσθημα, χρησιμοποιείται το mel, σε αντίθεση με τη συχνότητα που είναι ένα αντικειμενικό μέγεθος και για την οποία χρησιμοποιείται το Hz. Η κλίμακα του mel έχει κατασκευαστεί με τέτοιο τρόπο ώστε διπλασιασμός του αριθμού των mel συνεπάγεται και το διπλασιασμό του υποκειμενικού τονικού ύψους. Τέλος, η κλίμακα από 0 έως 2400 mel, καλύπτει όλο το ακουστικό φάσμα από 16 έως 16000 Hz.

1.3.7.1 Complex Pitch:

Γνωρίζουμε πως σίγουρα υπάρχει μία ισχυρή συσχέτιση μεταξύ του τονικού ύψους ενός ήχου και των συχνοτήτων από τις οποίες αυτός αποτελείται. Αυτό σημαίνει ότι εάν ένας σύνθετος ήχος έχει ένα φάσμα συχνοτήτων που μπορούν να αναλυθούν από το ακουστικό μας σύστημα, τότε αυτή η φασματική κατανομή, όπως για παράδειγμα η συχνότητα με τη μεγαλύτερη ένταση, είναι συσχετισμένα με το τονικό ύψος του ήχου.

Όμως, σε ορισμένες περιπτώσεις, οι ακροατές μπορούν να αντιληφθούν φασματική συνιστώσα του ήχου για την οποία δεν υπάρχει ενέργεια στη συγκεκριμένη συχνότητα του φάσματος για την οποία καθορίστηκε το τονικό ύψος του ήχου. Για παράδειγμα, υπάρχουν άνθρωποι (και δεν είναι λίγοι), που αναφέρουν τονικό ύψος 100 Hz για ήχο ο οποίος αποτελείται από το άθροισμα συχνοτήτων 700, 800, 900 και 1000 Hz. Ενώ δηλαδή δεν υπάρχει καθόλου ενέργεια στα 100 Hz οι ακροατές κρίνουν



Σχ. 1.3.7.1.1: Σχηματική αναπαράσταση του φαινομένου της χαμένης θεμελιώδους συχνότητας. (α) Η κυματομορφή του ήχου, ο οποίος αποτελείται από τα 700, 800, 900 & 1000 Hz. (β) Το φάσμα του συγκεκριμένου ήχου. Στα 100 Hz σημειώνεται η χαμένη θεμελιώδης, η οποία, στην πραγματικότητα δεν υπάρχει.

ότι το τονικό ύψος του ήχου καθορίζεται από τη συχνότητα των 100 Hz. Από τα παραπάνω προκύπτει ότι: η περίπτωση, στην οποία, ένα συγκεκριμένο τονικό ύψος μπορεί να συσχετισθεί με έναν σύνθετο ήχο και μάλιστα με τη θεμελιώδη συχνότητα του, ακόμη και στην περίπτωση όπου η συχνότητα αυτή δεν υπάρχει στο φάσμα του, ονομάζεται ύψος χαμένου θεμελιωτή (case of the missing fundamental ή αλλιώς missing fundamental pitch).

Το φάσμα και η κυματομορφή του σύνθετου τόνου του παραδείγματος μας δίδονται στο σχήμα 1.3.7.1.1. Παρατηρούμε ότι υπάρχει ένα διάστημα της τάξης των 10 msec μεταξύ των κορυφών στην κυματομορφή του ήχου. Αφού, λοιπόν, η συχνότητα που έχει περίοδο 10 msec είναι αυτή των 100 Hz τότε είναι πιθανό το ανθρώπινο ακουστικό σύστημα να αντιλαμβάνεται τη συχνότητα των 100 Hz εξαιτίας της περιοδικότητας αυτής των 10 msec στην κυματομορφή. Στην πραγματικότητα πολλοί ήχοι οι οποίοι έχουν μια περιοδικότητα στην κυματομορφή τους, θα έχουν κι ένα τονικό ύψος αντίστοιχο της περιόδου αυτής, το οποίο ονομάζεται: τονικό ύψος εκ της περιοδικότητας (periodicity pitch). Είναι, επίσης, πιθανό το ύψος της χαμένης θεμελιώδους να οφείλεται σε κάποιου είδους φασματική ανάλυση που κάνει το ανθρώπινο αυτί.

Ακόμη, όμως, κι αν η τιμή της περιοδικότητας μπορεί να ευθύνεται για το τονικό ύψος της χαμένης θεμελιώδους, υπάρχουν περιπτώσεις για τις οποίες αυτό δεν ισχύει. Για παράδειγμα, ένας ήχος ο οποίος δεν έχει θεμελιώδη, αλλά έχει τις συχνότητες 400, 600, 800, 1000, 1200 και 1400 Hz μπορεί να θεωρηθεί απο ακροατές ότι έχει τα τονικά ύψη των 200, 400 ή ακόμη και των 600 Hz.

Ένα άλλο χαρακτηριστικό του ακουστικού μας συστήματος, είναι να αναγνωρίζει περισσότερα απο ένα τονικά ύψη. Αν π.χ. έχουμε έναν ήχο ο οποίος αποτελείται απο συχνότητες των 250, 500, 825, 1000, 1250 και 1500 Hz, και δεδομένου όπως παρατηρούμε ότι όλες οι συχνότητες είναι πολλαπλάσια των 250 Hz εκτός απο τη συχνότητα των 825 Hz, τότε έχει διαπιστωθεί ότι εάν μια συχνότητα διαφέρει πάνω απο 8 % απο κάποια άλλη αρμονική του ήχου προκαλείται στους ακροατές η αίσθηση ότι στον ήχο αυτό υπάρχουν δύο τονικά ύψη.

Αν και το τονικό ύψος εξαρτάται απο τη συχνότητα, η σχέση που συνδέει τα δύο μεγέθη δεν είναι γραμμική. Δηλαδή ίσες μεταβολές στη συχνότητα δεν συνεπάγονται, απαραίτητα, ίσες μεταβολές στην αίσθηση του τονικού ύψους. Επιπροσθέτως η στάθμη του ήχου, καθώς και η κυματομορφή του, παίζουν σημαντικό ρόλο στον καθορισμό του τονικού του ύψους. Αν η στάθμη αυξηθεί, χωρίς να αλλάξει η συχνότητα του ήχου, η αίσθηση του ύψους χαμηλώνει. Το φαινόμενο αυτό είναι ιδιαίτερα αισθητό για ήχους των οποίων η συχνότητα είναι μικρότερη των 300 Hz. Για συχνότητες μεταξύ 500 και 3000 Hz, η αίσθηση του τονικού ύψους είναι σχεδόν ανεξάρτητη της στάθμης και εξαρτάται μόνο απο τη συχνότητα. Για ήχους με συχνότητα μεγαλύτερη των 4000 Hz, η αίσθηση του τονικού ύψους φαίνεται να αυξάνει, για την ίδια συχνότητα, όσο αυξάνει η στάθμη του ήχου.

1.3.8 Χροιά (Timbre):

Τα τρία κυριότερα χαρακτηριστικά ενός ήχου είναι η τονικότητα, η ακουστότητα (με τα οποία ασχοληθήκαμε στις αμέσως προηγούμενες παραγράφους) και η χροιά (timbre). Σύμφωνα με το American National Standards Institute, η χροιά μπορεί να οριστεί ως εξής: *«Η χροιά αποτελεί εκείνο το γνώρισμα της ακουστικής εντύπωσης που μας προκαλεί ένας ήχος, εξαιτίας της οποίας, ένας ακροατής μπορεί να κρίνει αν δύο ήχοι οι οποίοι έχουν παρουσιαστεί με τον ίδιο τρόπο και έχουν την ίδια ακουστότητα και τονικότητα είναι όμοιοι ή διαφορετικοί»* (ANSI 1960).

Η χροιά ενός ήχου εξαρτάται απο διάφορους παράγοντες. Ο πιο σημαντικός είναι η συχνοτική του σύνθεση, δηλαδή οι συχνότητες τις οποίες εμπεριέχει και απο τις οποίες αποτελείται (όπως οι αρμονικές). Η συχνοτική σύνθεση, συνεπώς επηρεάζει τη χροιά του ήχου. Ένας ήχος μπορεί να αποτελείται απο συχνότητες οι οποίες μπορεί να είναι αρμονικές της θεμελιώδους ή και όχι. Σημαντική είναι και η ενέργεια που εμπεριέχεται στην κάθε συχνότητα ενός ήχου για την αντίληψη της χροιάς. Η χροιά ενός ήχου επηρεάζεται απο τη σχέση μεταξύ της ενέργειας που υπάρχει στη θεμελιώδη και

της ενέργειας που υφίσταται στις αρμονικές· όπως επίσης και με τη σχέση της ενέργειας που περιέχουν μεταξύ τους οι αρμονικές.

Για την αντίληψη της χροιάς ενός οργάνου σημασία έχει, επίσης, ο τρόπος με τον οποίο μεταβάλλεται η σχέση μεταξύ της ενέργειας της θεμελιώδους και των αρμονικών σε διάφορες τονικότητες. Διαφορετικές νότες από ένα μουσικό όργανο δεν έχουν αναγκαστικά το ίδιο σχετικό φάσμα. Μια χαμηλή νότα πιάνου περιέχει μικρό ποσοστό ενέργειας στη θεμελιώδη συχνότητα, ενώ αντίθετα παρέχει το μεγαλύτερο ποσοστό της ενέργειας στις υψηλότερες αρμονικές. Μια υψηλή, τονικά, νότα πιάνου, έχει μια ισχυρή θεμελιώδη και ασθενέστερες αρμονικές. Αν, τώρα, ένας τόνος από ένα όργανο αναλυθεί φασματικά και το φάσμα χρησιμοποιηθεί για τη δημιουργία και περαιτέρω τόνων, τότε οι τόνοι που θα προκύψουν δε θα μοιάζουν ότι ακούγονται από το ίδιο μουσικό όργανο.

Γνωρίζουμε πως το εύρος από την 1^η έως την 6^η περίπου αρμονική ενός μουσικού οργάνου, αναλύεται ξεχωριστά για την κάθε μία από αυτές τις αρμονικές από το ακουστικό μας σύστημα. Αυτές είναι επίσης οι αρμονικές που παίζουν ένα ρόλο ιδιαίτερα σημαντικό για την αντίληψη της χροιάς του μουσικού οργάνου. Αρμονικές από την 4^η έως την 7^η και πάνω οι οποίες δε διαχωρίζονται μεταξύ τους από το ανθρώπινο αυτί, έχουν μια διαφορετική σημασία στην αντίληψη της χροιάς. Δηλαδή οι αρμονικές αυτές (4^η – 7^η και πάνω) δεν αναλύονται ξεχωριστά με αποτέλεσμα να μη γίνονται αντιληπτές ως ξεχωριστές αρμονικές· γίνονται αντιληπτές ως ομάδα αρμονικών.

Όργανα τα οποία έχουν αρμονικές πάνω από την 5^η ή 7^η κατά τη διάρκεια της steady state² φάσης τους, είναι η τρομπέτα και το σαξόφωνο. Η χροιά για αυτές τις περιπτώσεις μπορεί να χαρακτηριστεί ως «φωτεινή». Όργανα τα οποία δεν έχουν μεγάλο ποσοστό ενέργειας, κατά τη steady state φάση τους, για τις αρμονικές αυτές είναι το κλαρίνο, το όμποε και το φλάουτο. Σε σύγκριση με το σαξόφωνο και την τρομπέτα, η χροιά των οργάνων αυτών θα μπορούσε να χαρακτηριστεί ως λιγότερο φωτεινή.

Ο Helmholtz (1877, *On the sensation of tone*) κατέληξε σε τέσσερις κανόνες οι οποίοι επηρεάζουν την αντίληψη της χροιάς ενός ήχου βάσει της συχνοτικής του σύνθεσης. Οι κανόνες αυτοί εξακολουθούν να είναι σήμερα τόσο σημαντικοί, όσο ήταν και την εποχή που εκφράστηκαν. Αυτοί είναι:

- i) Απλοί τόνοι (simple tones), χωρίς αρμονικές έχουν έναν απαλό, ευχάριστο ήχο, χωρίς τραχύτητα, μα ανεπαρκή σε ισχύ και χροιά όχι ιδιαίτερα ενδιαφέρουσα.
- ii) Τόνοι με μουσικότητα (musical tones), έχουν αρκετή ενέργεια στις αρμονικές τους μέχρι, περίπου, και την 6^η. Σε σύγκριση με τους

² Τα στάδια της κυματομορφής του ήχου είναι τρία. Το onset, ο ήχος τη χρονική περίοδο από τη σιωπή μέχρι την αρχή της νότας. Το steady state, η κύρια περίοδος διάρκειας του ήχου. Και τέλος το offset, η χρονική περίοδος που μεσολαβεί από το τέλος της νότας ως ότου επανέλθουμε στη σιωπή.

απλούς τόνους είναι πιο πλούσιοι ηχητικά, ενώ εάν δεν έχουν ενέργεια στις αρμονικές πάνω από την έκτη, μπορούν να έχουν ήχο αρκετά απαλό και ευχάριστο.

- iii) Εάν μόνο οι αρμονικές με περιττό αριθμό είναι παρούσες στο φάσμα, η ποιότητα της χροιάς μπορεί να χαρακτηριστεί είτε ως κενή, είτε ως άδεια (hollow). Όταν υπάρχει μεγάλος αριθμός τέτοιων αρμονικών, η ποιότητα μπορεί να χαρακτηριστεί ένρινη (nasal). Όταν η θεμελιώδης συχνότητα υπερισχύει, η χροιά είναι πλούσια· μα όταν η θεμελιώδης συχνότητα δεν υπερισχύει ικανοποιητικά σε σχέση με τις υψηλότερες αρμονικές, τότε η χροιά χαρακτηρίζεται ως φτωχή.
- iv) Όταν οι αρμονικές πάνω από την 6^η ή την 7^η έχουν αρκετή ενέργεια, τότε η χροιά μπορεί να είναι διαπεραστική ή και σκληρή (cutting or rough).

Στον πίνακα που ακολουθεί, βλέπουμε παραδείγματα από διάφορα ακουστικά, μουσικά όργανα για τα οποία η ακουστική συμπεριφορά τους αντιστοιχεί σε κάποιον από τους κανόνες του Helmholtz.

Helmholtz rule	Human hearing modeling spectrogram	Example timbre descriptions	Example acoustic instruments
1	fo dominates	Pure Soft Simple Pleasant Dull at low pitch Free from roughness	Tuning fork Wide stopped organ flues Baroque flute
2	Harmonics dominate	Sweet & soft Rich Splendid Dark Dull Less shrill Bland	French horn, tuba Modern flute Recorder Open organ flues Soft sung sounds
3	Odd harmonics dominate	Hollow Nasal	Clarinet Narrow stopped organ flues
4	Striations dominate	Cutting Rough Bright Brilliant Shrill Brash	Oboe, bassoon Trumpet, trombone Loud sung sounds Bowed instruments Harmonium Organ needs

Πιν. 1.3.8.1: Οι κανόνες του Helmholtz και παραδείγματα για διάφορα μουσικά - ακουστικά όργανα.

1.3.9 Φαινόμενα μη γραμμικότητας του ακουστικού συστήματος:

Γενικά ένα γραμμικό σύστημα, είναι αυτό που στην έξοδο του δεν προσθέτει νέες συχνότητες σε σχέση με την είσοδο του (αν και μπορεί να επηρεάζει το πλάτος και τη φάση της εισόδου του), ενώ ένα μη γραμμικό σύστημα, είναι αυτό που προσθέτει νέες συχνότητες στην έξοδο του. Για να ξεχωρίσουμε την επίδραση του μη γραμμικού συστήματος στην έξοδο του, θα σχηματοποιήσουμε δύο κατηγορίες. Η πρώτη είναι όταν ένας ήχος αποτελούμενος από μία μόνο συχνότητα f_1 εισέρχεται στο σύστημα και το αποτέλεσμα της εξόδου του συστήματος αποτελείται από το άθροισμα δύο ημιτονοειδών συχνοτήτων f_1 και f_2 . Όταν στο μη γραμμικό σύστημα εισέλθει ένας ήχος συχνότητας f_1 , τότε η έξοδος του συστήματος μπορεί να περιέχει και αρμονικές της συχνότητας εισόδου, δηλ. f_1, f_2, f_3 . Όταν, τώρα, στο μη γραμμικό σύστημα εισέλθουν συχνότητες f_1 και f_2 , τότε μπορεί να έχουμε πάλι αρμονικές όπως στην προηγούμενη περίπτωση, αλλά θα έχουμε επίσης, και συχνότητες που αποτελούν συνδυασμούς των συχνοτήτων f_1 και f_2 . Αυτές οι συχνότητες ονομάζονται combination tones και μπορεί να είναι συχνότητες οι οποίες προστίθενται (summation tones) ή συχνότητες διαφοράς (difference tones). Τυπικές συχνότητες πρόσθεσης είναι οι $f_1+f_2, 2f_1+f_2, f_1+2f_2$. Τυπικές συχνότητες διαφοράς είναι οι $f_1-f_2, 2f_1-f_2, 2f_2-f_1$.

Σε ότι αφορά το ακουστικό μας σύστημα, αυτό έχει μη γραμμικά χαρακτηριστικά τα οποία παράγουν αρμονικές (aural harmonics) και τόνους συνδυασμού (combination tones). Αν για παράδειγμα υποθέσουμε ότι η απόκριση του αυτιού στην ηχητική πίεση είναι:

$$r = ap + bp^2$$

τότε όταν δύο συχνότητες ακουστούν μαζί:

$$p = p_1 \cos(\omega_1 t) + p_2 \cos(\omega_2 t)$$

η απόκριση του αυτιού θα είναι:

$$r = a(p_1 \cos \omega_1 t + p_2 \cos \omega_2 t) + b(p_1^2 \cos^2 \omega_1 t + p_2^2 \cos^2 \omega_2 t + 2p_1 p_2 \cos \omega_1 t \cos \omega_2 t)$$

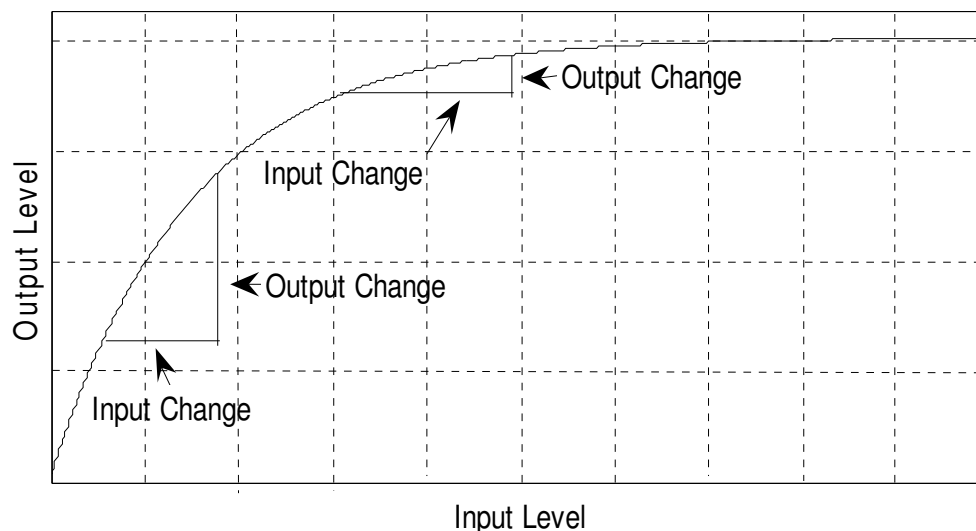
χρησιμοποιώντας τις: $\cos^2 a = \frac{1}{2}(\cos 2a + 1)$ και $\cos a \cdot \cos b = \frac{1}{2}[\cos(a + b) + \cos(a - b)]$

παίρνουμε τελικά:

$$r = a(p_1 \cos \omega_1 t + p_2 \cos \omega_2 t) + b \left[\frac{p_1^2 + p_2^2}{2} + \frac{p_1}{2} \cos 2\omega_1 t + \frac{p_2}{2} \cos 2\omega_2 t + p_1 p_2 \cos(\omega_1 + \omega_2)t + p_1 p_2 \cos(\omega_1 - \omega_2)t \right]$$

που όπως βλέπουμε, τελικά προκύπτουν 6 όροι με διαφορετική συχνότητα ο καθένας.

Τέλος η μη γραμμικότητα του ακουστικού μας συστήματος έχει τη μορφή του σχήματος 1.3.9.1 και καλείται compressive nonlinear.



Σχ. 1.3.9.1: Γραφική αναπαράσταση της μη γραμμικότητας του ανθρώπινου αισθητηρίου της ακοής.

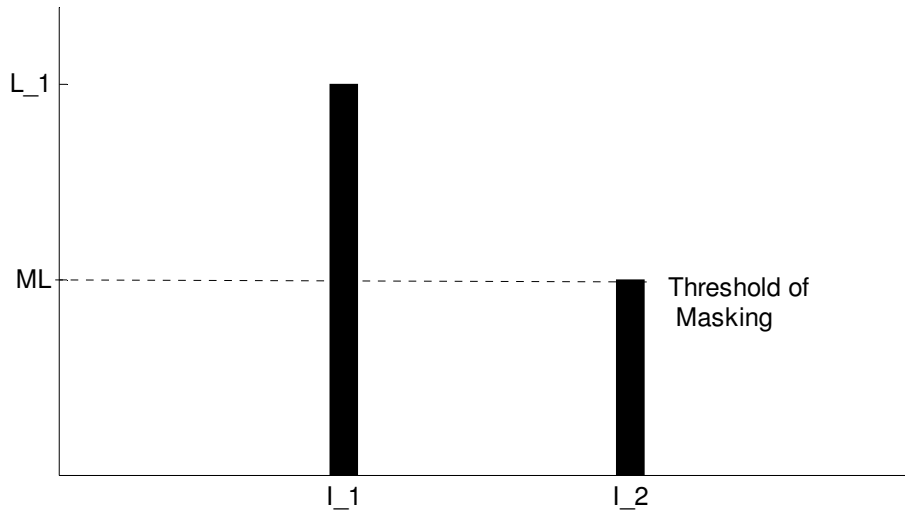
1.4 Masking effect:

Οι ήχοι στη ζωή μας ακούγονται τις περισσότερες φορές μαζί ή σε κοντινά μεταξύ τους χρονικά διαστήματα. Μέσω αυτού που αποκαλείται masking effect, μας δίδεται η δυνατότητα να εξετάσουμε το πώς η παρουσία ενός ήχου επηρεάζει την αντίληψη ενός άλλου και γενικότερα με την αλληλεπίδραση των ήχων. Η αλλαγή στο κατώφλι ακουστότητας ενός ήχου, παρουσία ενός άλλου, αποτελεί το πεδίο εξέτασης και εξαγωγής συμπερασμάτων σχετικά με το masking effect.

Γενικά, είναι εμπειρικά γνωστό ότι μη επιθυμητοί ήχοι, μέσω της έντασης τους, μπορούν να παρεμβάλλονται και να καθιστούν μη ακουστούς άλλους, χρήσιμους ήχους, τους οποίους τελικά δυσκολευόμαστε ή αδυνατούμε να αντιληφθούμε. Κοντολογίς, η ύπαρξη δύο ήχων με εντάσεις I_1 και I_2 αντίστοιχα και με $I_1 > I_2$, οδηγεί ενδεχόμενα σε πλήρη απόκρυψη του ήχου με I_2 , ο οποίος μόνος του είναι απολύτως άνετα ακουστός.

Επίσης, είναι λογικά αναμενόμενο, η εντονότερη παρουσία του masking να υφίσταται όταν συχνοτικά οι ήχοι με εντάσεις I_1 και I_2 ταυτίζονται. Θεωρούμε, λοιπόν, ότι οι ήχοι με τις παραπάνω εντάσεις είναι ίδιας συχνότητας και φάσης.

Αυτό που προκύπτει είναι ότι ενώ η I_1 υφίσταται αθροίζοντας τη I_2 . Το level, συνεπώς αυξάνει και η απαιτούμενη διαφορά για την αντίληψη, από το ανθρώπινο αυτί, αυτής της αύξησης είναι περίπου 0.2 - 0.4 dB, τιμή σταθερή πρακτικά για όλη τη μεσαία περιοχική συχνότητων. Αποτελεί δηλαδή το just noticeable difference για το



Σχ. 1.4.1: Σχηματική αναπαράσταση για το παράδειγμα δύο ήχων εντάσεων I_1 και I_2 .

sound level]. Συνεπώς, απο τα παραπάνω προκύπτει:

$$jnd_{s,l} = 10 \log \frac{I_1 + I_2}{I_1} \dots \approx 0.2 - 0.4 \text{ dB} \quad (1.4.1)$$

Όπως φαίνεται και στο σχήμα 1.4.1, το $jnd_{s,l}$, ως η ελάχιστη αντιληπτή διαφορά, έχει την ιδιότητα να καθορίζει συνάμα και μια οριακή τιμή του I_2 , κάτω απο την οποία ο συγκεκριμένος ήχος δεν θα ακούγεται, θα είναι δηλαδή masked. Μπορούμε να ονομάσουμε αυτή την τιμή Threshold of masking, αυτή η τιμή παράλληλα ισούται και με το masking level (ML). Οπότε, απο τη (1.4.1):

$$\frac{I_1 + I_2}{I_1} = 10^{\frac{jnd_{s,l}}{10}} \Rightarrow \frac{I_2}{I_1} = 10^{\frac{jnd_{s,l}}{10}} - 1 \Rightarrow I_2 = I_1 \left(10^{\frac{jnd_{s,l}}{10}} - 1 \right)$$

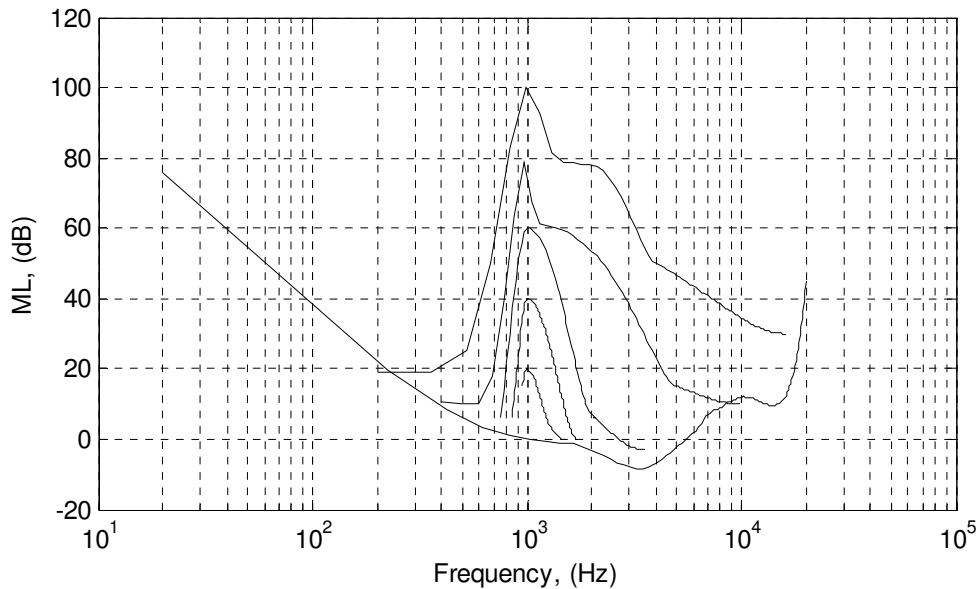
συνεπώς, $ML \equiv L_2$, δηλαδή:

$$ML = L_1 + 10 \log \left(10^{\frac{jnd_{s,l}}{10}} - 1 \right) \quad (1.4.2)$$

Μπορούμε, τώρα, να βρούμε που κυμαίνεται η τιμή του ML:

$$jnd_{s,l} = 0.2 \rightarrow ML = L_1 - 13.2 \text{ dB}$$

$$0.4 \rightarrow ML = L_1 - 10 \text{ dB}$$



Σχ. 1.4.2: Το masking level (ML) συναρτήσει της συχνότητας, εξαιτίας ενός masking tone με συχνότητα 1 kHz. Παρατηρούμε το ML για διάφορες τιμές έντασης.

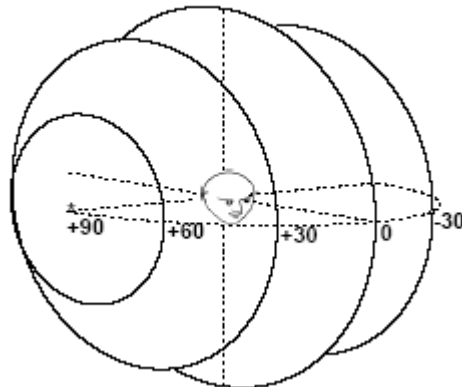
Μια πιο πλήρη εικόνα του φαινομένου αποκτούμε επεκτεινόμενοι στην περίπτωση όπου ο masked tone έντασης I_2 είναι τυχαίας συχνότητας f και δεν ταυτίζεται αναγκαστικά με τον masking tone f_0 έντασης I_1 . Το σχήμα 1.4.2 παρουσιάζει ακριβώς αυτό.

Παρατηρούμε, ότι για διάφορες τιμές έντασης, όταν η I_2 ταυτίζεται με τον masking tone $f_0 = 1 \text{ kHz}$, το ML ταυτίζεται, πρακτικά, με αυτό που εξάγει η σχέση (1.4.2) για κάθε τιμή έντασης. Διαφοροποιούμενη, όμως, η f βλέπουμε ότι το ML ελαττώνεται στην περιοχή γύρω από την f_0 , για να καταλήξει σε μηδενική τιμή αρκετά μακριά από αυτή. Αξίζει, τέλος, να προσέξουμε ότι για μικρές εντάσεις το ML ελαττώνεται σχεδόν συμμετρικά γύρω από την f_0 , ενώ αντίθετα σε υψηλότερες εντάσεις το φαινόμενο του masking effect απλώνεται περισσότερο στις υψηλές συχνότητες σε σχέση με τις χαμηλές. Αυτό οφείλεται στη μη γραμμικότητα του ακουστικού μας συστήματος και μάλιστα έχει να κάνει με την περιγραφή που δίδεται στο σχήμα 1.3.9.1 και αφορά την απόκριση του αυτιού συναρτήσει των εντάσεων που γίνεται αποδέκτης καθώς έχει την τάση να παράγει αρμονικές και τόνους συνδυασμού.

1.5 Sound Localization:

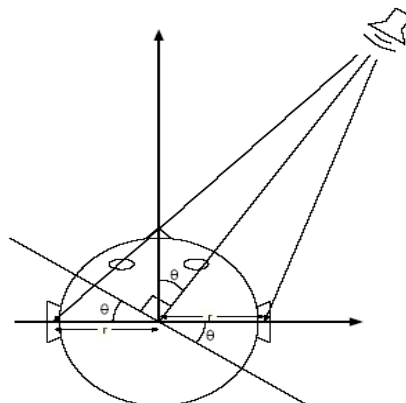
Το ανθρώπινο ακουστικό σύστημα έχει την ικανότητα να εντοπίζει τη θέση από την οποία εκπέμπεται ένας ήχος. Η πηγή ενός ήχου μπορεί να εντοπιστεί στις τρεις χωρικές διαστάσεις. Στο οριζόντιο επίπεδο, στο κάθετο επίπεδο και στην απόσταση από τον ακροατή.

1.5.1 Εντοπισμός στο οριζόντιο επίπεδο:

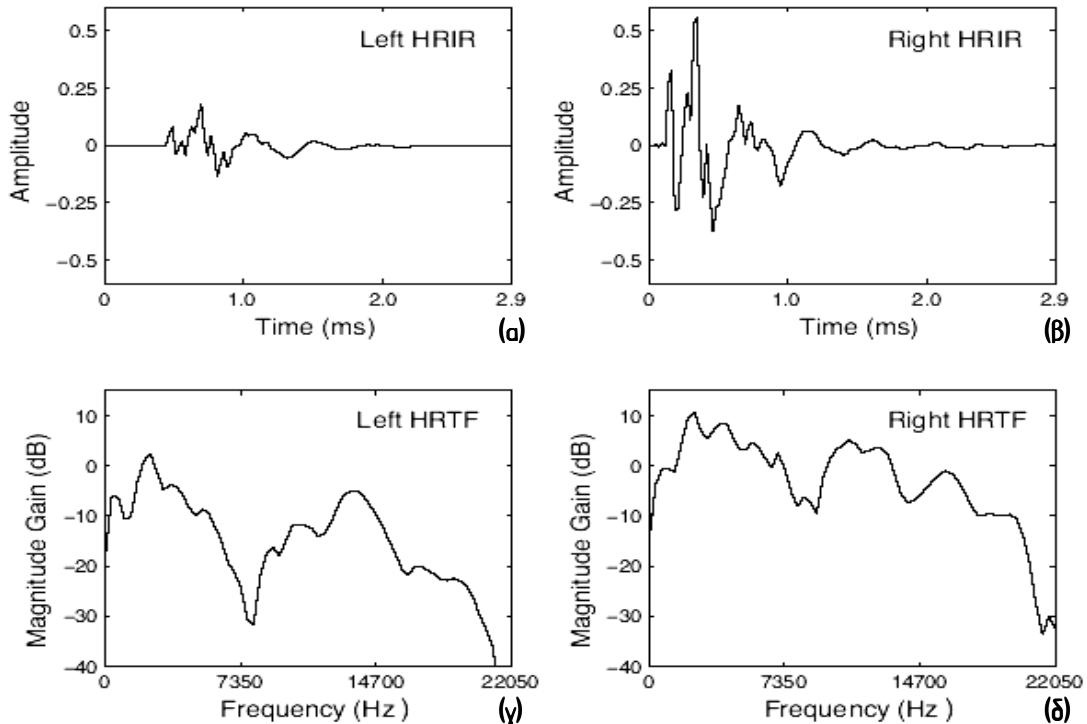


Σχ. 1.5.1.1: Παρουσίαση του οριζοντίου επιπέδου σε σύστημα σφαιρικών πολικών συντεταγμένων.

Για να κατανοήσουμε τον εντοπισμό του ήχου στο οριζόντιο επίπεδο, ας υποθέσουμε ότι κάποιος κάθεται σε ένα ανοικτό περιβάλλον, ακούγοντας μια ηχητική πηγή, δίχως να μετακινεί το κεφάλι του και η οποία πηγή είναι τοποθετημένη μπροστά και δεξιά του (σχ. 1.5.1.2). Το σχήμα 1.5.1.3 που δίδεται στην επόμενη σελίδα (σελ. 27) μας δίνει πληροφορίες σε ότι αφορά το χρόνο άφιξης του ηχητικού σήματος σε κάθε αυτί, πληροφορίες σε ότι αφορά την ένταση του ηχητικού σήματος για το κάθε αυτί, αλλά και τις Head Related Transfer Functions (HRTFs) που προκύπτουν για το κάθε αυτί από το συγκεκριμένο ηχητικό γεγονός.



Σχ. 1.5.1.2: Χωροθέτηση πηγής ακροατή για την εξαγωγή συμπερασμάτων αναφορικά με την αντίληψη του ήχου στο οριζόντιο επίπεδο.



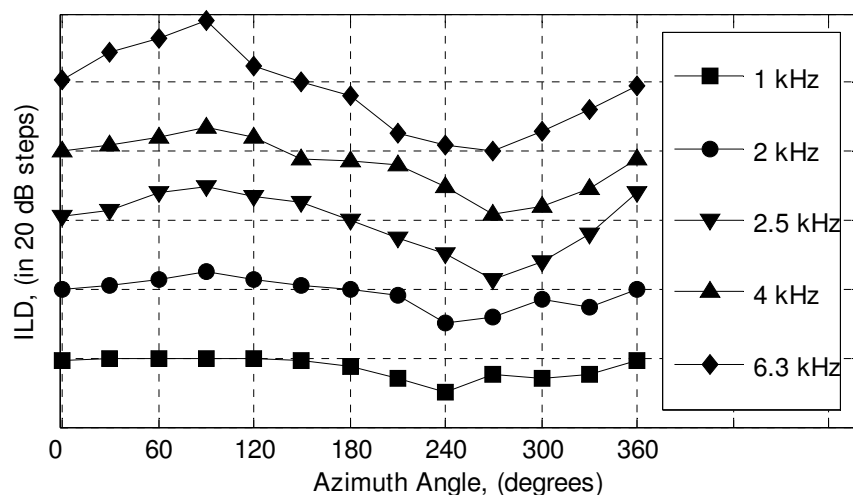
Σχ. 1.5.1.3: Συναρτήσεις μεταφοράς της κρουστικής απόκρισης (Head Related Impulse Responses, HRIRs) (α) & (β), καθώς και οι συναρτήσεις μεταφοράς της συχνотικής απόκρισης (Head Related Transfer Functions, HRTFs) (γ) & (δ) οι οποίες ηχογραφήθηκαν στο ανδρείκελο KEMAR για πηγή τοποθετημένη μπροστά και στα δεξιά του KEMAR, με γωνία αζιμούθιου τις 40° και γωνία ανύψωσης τις 0° . Παρατηρούμε πως το ερέθισμα γίνεται περισσότερο αισθητό στο δεξί αυτί, καθώς επίσης ότι φτάνει νωρίτερα στο δεξί απ' ότι στο αριστερό αυτί. (Τα στοιχεία αντλήθηκαν από τη βάση δεδομένων του CIPIC, University of California).

Παρατηρούμε ότι ο ήχος χρειάζεται να ταξιδέψει μικρότερη απόσταση για να φτάσει στο δεξί, σε σχέση με το αριστερό αυτί. Συνεπώς θα φτάσει στο δεξί αυτί νωρίτερα. Αυτό έχει ως αποτέλεσμα να έχουμε μια ενδοακουστική χρονική διαφορά, (Interaural Time Difference, ITD), στην άφιξη του ήχου. Η ταχύτητα του ήχου στον αέρα είναι σχετικά σταθερή και ανεξάρτητη της συχνότητας. Επομένως, η ενδοακουστική χρονική διαφορά είναι θεωρητικά ίδια για όλες τις συχνότητες, ενώ όμως, η ενδοακουστική διαφορά φάσης (Interaural Phase Difference, IPD) θα εξαρτάται από τη συχνότητα του ήχου. Αυτό σημαίνει ότι αν ένας ήχος με συχνότητα 1000 Hz, ο οποίος έχει περίοδο 1 msec, φτάσει στο αριστερό αυτί 0.5 msec αργότερα από τη στιγμή που θα φτάσει στο δεξί αυτί, τότε είναι μισή περίοδο (ή αλλιώς 180°) εκτός φάσης με τον ήχο που φτάνει στο δεξί αυτί. Εάν, τώρα, ένας ήχος συχνότητας 500 Hz που έχει περίοδο 2 msec, φτάσει στο αριστερό αυτί 0.5 msec αργότερα απ' ότι στο δεξί, θα υπάρχει μόνο ένα τέταρτο της περιόδου (ή 90°) διαφορά φάσης μεταξύ των αυτιών. Συνεπώς οι δύο διαφορετικοί τόνοι, οι οποίοι έχουν και οι δύο την ίδια ενδοακουστική χρονική διαφορά, θα έχουν διαφορετική ενδοακουστική φασική διαφορά.

Υπάρχει, επίσης, και μία ενδοακουστική διαφορά έντασης (Interaural Level Difference, ILD). Δύο είναι οι λόγοι για αυτή τη διαφορά στην ένταση. Ο πρώτος είναι

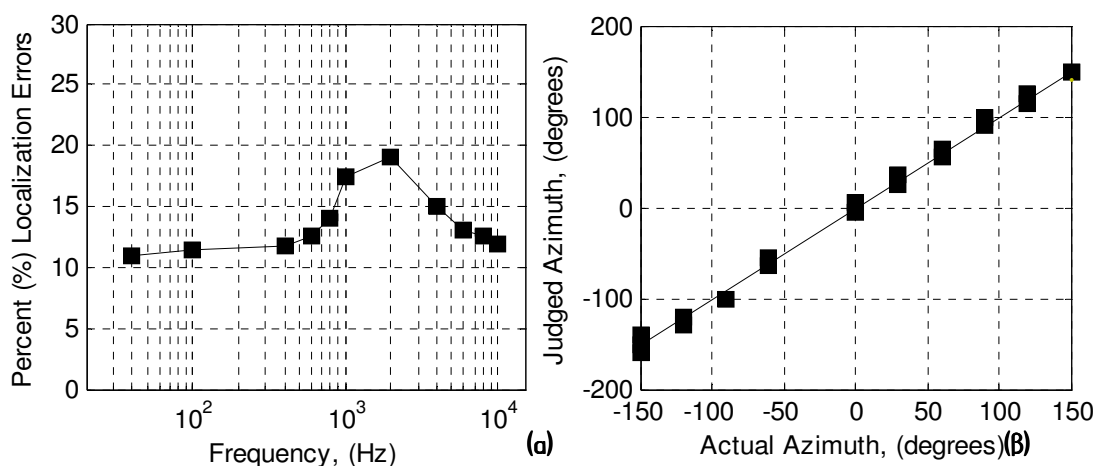
επειδή, ο ήχος για το παράδειγμα που αναφέραμε, φτάνει στο αριστερό αυτί αφού έχει φτάσει πρώτα στο δεξί, έχοντας διανύσει, παράλληλα, μεγαλύτερη απόσταση με αποτέλεσμα (λόγω του νόμου του αντιστρόφου τετραγώνου) να είναι μικρότερης έντασης. Η διαφορά των εντάσεων, όμως, εξαιτίας αυτού του λόγου είναι εξαιρετικά μικρή. Ο δεύτερος λόγος που θα αναφερθεί είναι και ο σημαντικότερος. Το κεφάλι, όπως και οποιοδήποτε άλλο αντικείμενο, μπορεί να προκαλέσει μια ηχητική σκίαση (sound shadow) που έχει ως αποτέλεσμα μειωμένη ένταση στο αυτί που βρίσκεται αντίθετα της πηγής, θέτοντας ως απαραίτητη, όμως, προϋπόθεση για να συμβεί αυτό ότι το μέγεθος του κεφαλιού είναι συγκρίσιμο με το μήκος κύματος του ήχου. Επειδή το μήκος κύματος είναι αντιστρόφως ανάλογο της συχνότητας, η ενδοακουστική διαφορά έντασης, η οποία προκαλείται από την ηχητική σκίαση του κεφαλιού, εξαρτάται από τη συχνότητα. Όσο υψηλότερη είναι η συχνότητα, τόσο μικρότερο είναι το μήκος κύματος και, συνεπώς, τόσο μεγαλύτερη η ηχητική σκιά που προκαλείται. Επομένως, το συμπέρασμα που προκύπτει, είναι ότι μεγάλες ενδοακουστικές διαφορές έντασης υφίστανται σε υψηλές συχνότητες και συνεπώς μπορούν να χρησιμοποιηθούν για τον εντοπισμό της ηχητικής πηγής.

Το σχήμα 1.5.1.4 που ακολουθεί, παρουσιάζει τις ενδοακουστικές διαφορές έντασης για διαφορετικές γωνίες και συχνότητες. Όπως παρατηρούμε η ενδοακουστική διαφορά έντασης διαφέρει σημαντικά και ιδιαίτερα για τις υψηλές συχνότητες. Παρατηρούμε, επίσης, ότι όσο αυξάνει η συχνότητα, τόσο αυξάνει και ILD. Το ακουστικό σύστημα δηλαδή, μπορεί να κρίνει για τον εντοπισμό μιας ηχητικής πηγής στον οριζόντιο άξονα (αν για παράδειγμα έρχεται από μία συγκεκριμένη γωνία), λαμβάνοντας υπόψη το δεδομένο ότι ο ήχος έρχεται πρώτα στο δεξί αυτί και ότι η ένταση είναι μεγαλύτερη.



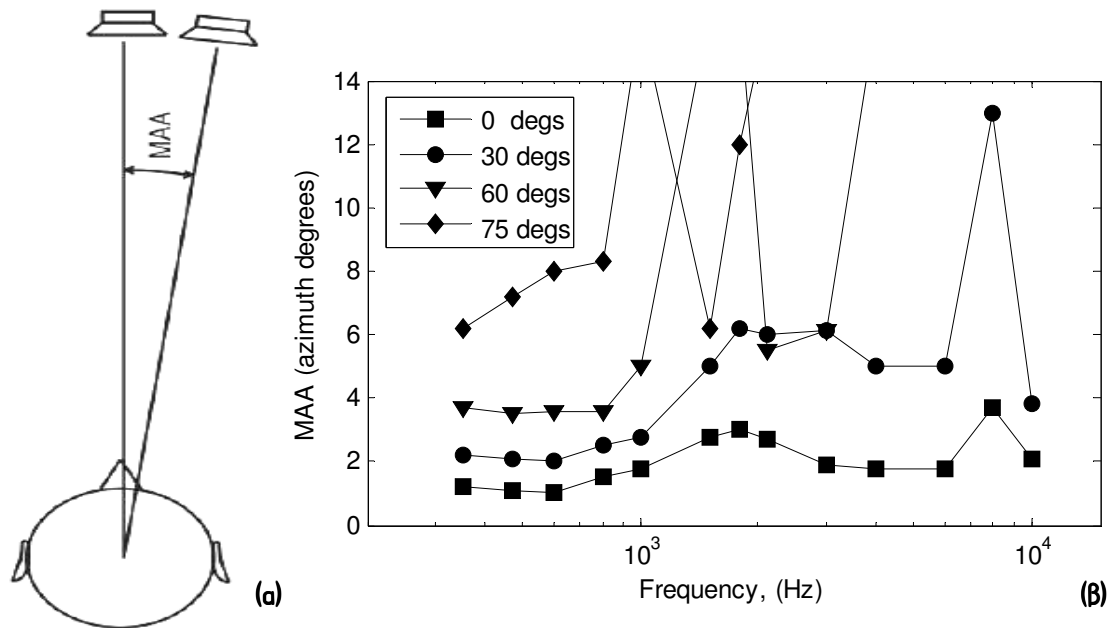
Σχ. 1.5.1.4: Ενδοακουστικές διαφορές έντασης (ILD) για διάφορες συχνότητες και για διάφορες γωνίες.

Τα διαγράμματα (α) και (β) του σχήματος 1.5.1.5 που ακολουθεί, δείχνουν πόσο καλά ένας ακροατής μπορεί να εντοπίσει τις ηχητικές πηγές. Το διάγραμμα (α) δείχνει τα σφάλματα στον εντοπισμό ημιτονοειδών ήχων για διαφορετικές συχνότητες. Τα στοιχεία αυτά μας δείχνουν ότι οι ακροατές κάνουν περισσότερα λάθη όταν καλούνται να εντοπίσουν ήχους συχνοτήτων που κυμαίνονται στη μεσαία μερισική, δηλαδή γύρω στα 2000 Hz, σε σχέση με ήχους υψηλότερης ή χαμηλότερης συχνότητας. Τα στοιχεία, τώρα, που παρουσιάζονται στο διάγραμμα (β), δείχνουν πόσο καλά ένας ακροατής εντοπίζει ήχους με ευρύ φάσμα. Όπως φαίνεται, από τα αποτελέσματα, το ανθρώπινο ακουστικό σύστημα είναι πολύ καλό στο να εντοπίζει broadband noise στο οριζόντιο επίπεδο.



Σχ. 1.5.1.5: (α) Διάγραμμα του ποσοστού σφαλμάτων εντοπισμού των ήχων. (β) Διάγραμμα εντοπισμού broadband noise στο οριζόντιο επίπεδο.

Εν συνεχεία, τα διαγράμματα 1.5.1.6(a) και 1.5.1.6(β) της επόμενης σελίδας παρουσιάζουν τα δεδομένα από ένα πείραμα στο οποίο οι ακροατές, με δεμένα τα μάτια, καλούνταν να διακρίνουν μεταξύ των θέσεων δύο μικρών ηχείων βάσης, όπου το καθένα ήταν σε απόσταση περίπου 100 cm από τον κάθε ακροατή. Ο πιο μικρός διαχωρισμός σε ότι αφορά τη γωνία μεταξύ των ηχείων, την οποία μπορεί να αντιληφθεί ο ακροατής, καλείται ελάχιστη ακουστή γωνία (Minimum Audible Angle, MAA). Συνεπώς, όπως είναι εύλογο, η ελάχιστη ακουστή γωνία σε μοίρες μετρήθηκε συναρτήσει της συχνότητας. Οι διάφορες καμπύλες όπως φαίνονται στο διάγραμμα 1.5.1.6(β) παρουσιάζουν διάφορες γωνίες για τις οποίες έγιναν οι μετρήσεις. Για παράδειγμα, σε ότι αφορά την καμπύλη με γωνία 0°, σημαίνει ότι τα ηχεία ήταν τοποθετημένα ακριβώς μπροστά από τον ακροατή. Συνακόλουθα η καμπύλη για γωνία 75° σημαίνει ότι τα ηχεία ήταν τοποθετημένα σε γωνία 75° προς το ένα αυτί. Παρατηρούμε ότι ο ακροατής χρειάζεται όλο και μεγαλύτερο διαχωρισμό αναφορικά με τη γωνία μεταξύ των ηχείων προκειμένου να ξεχωρίσει την οποιαδήποτε διαφορά στην τοποθέτησή τους όσο αυτά μετακινούνταν από τη θέση ακριβώς μπροστά του



Σχ. 1.5.1.6: (α) Σχηματική αναπαράσταση της μεθόδου μέτρησης για την ελάχιστη ακουστική γωνία (MAA). (β) Μετρήσεις για διάφορες γωνίες και συχνότητες της ελάχιστης ακουστικής γωνίας (τα στοιχεία αντλήθηκαν από τη βάση δεδομένων του CIPIC, University of California).

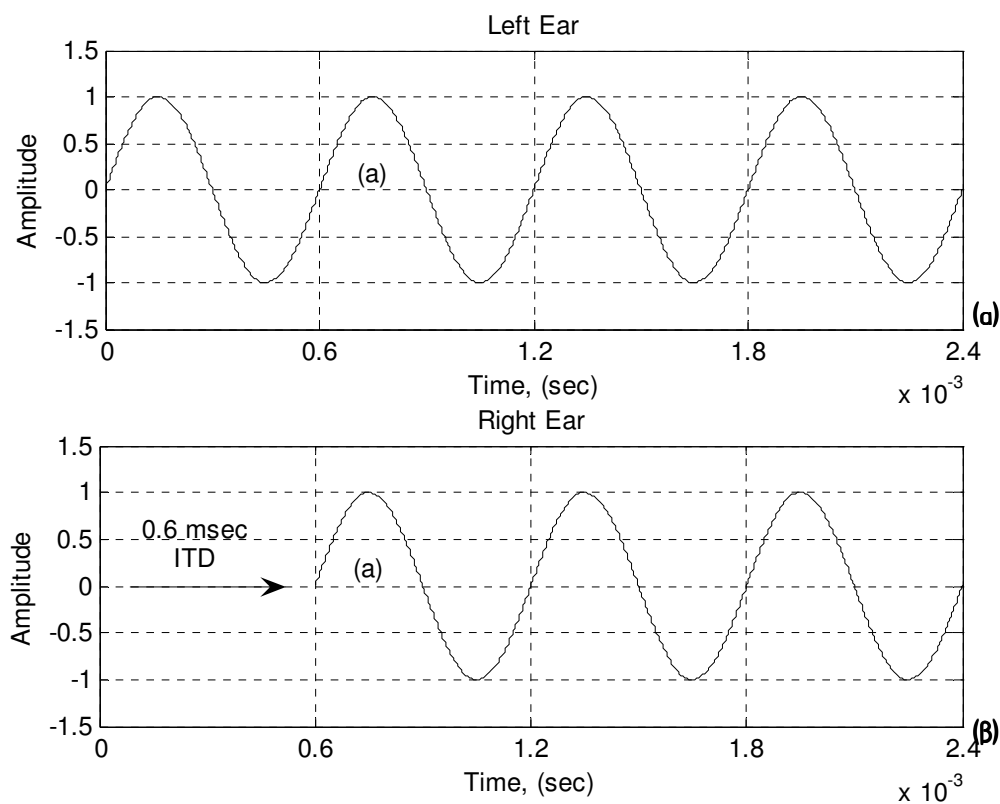
προς το ένα από τα δύο αυτιά (ενώ, προφανώς, το κεφάλι του ακροατή παρέμενε σταθερό). Με άλλα λόγια, η αλλαγή στη θέση μπορεί πιο εύκολα να διευκρινιστεί όταν είναι μπροστά από τον ακροατή, απ' ό,τι στην περίπτωση που βρίσκεται στο πλάι. Τα αποτελέσματα από την ελάχιστη ακουστική γωνία μας δείχνουν, όπως και το διάγραμμα (α) του σχήματος 1.5.1.5, ότι οι ακροατές κάνουν περισσότερα λάθη όταν καλούνται να εντοπίσουν ηχητικές πηγές των οποίων το συχνοτικό περιεχόμενο τοποθετείται στη μεσαία περιοχή του φάσματος.

Οι Stevens και Newman, καθώς και νωρίτερα από αυτούς ο Rayleigh, πίστευαν ότι στη μεσαία συχνοτική περιοχή υπάρχουν οι συχνότητες για τις οποίες οι ενδοακουστικές χρονικές διαφορές, αλλά και οι ενδοακουστικές διαφορές έντασης, είναι σχετικά μικρές ώστε να χρησιμοποιηθούν για τον εντοπισμό της ηχητικής πηγής. Το συμπέρασμα που προέκυψε από τις μελέτες των παραπάνω είναι ότι υπάρχουν δύο στοιχεία στα οποία οφείλεται, κατά κύριο λόγο, ο εντοπισμός της θέσης μιας ηχητικής πηγής. Το πρώτο στοιχείο αναφέρεται στην ενδοακουστική χρονική διαφορά (ITD), στην οποία οφείλεται η λήψη πληροφοριών από τον ακροατή για τον εντοπισμό της διεύθυνσης ήχων χαμηλής συχνότητας. Και το δεύτερο αναφέρεται στην ενδοακουστική διαφορά έντασης (ILD), στην οποία οφείλεται η παροχή πληροφοριών στον ακροατή σχετικά με τη διεύθυνση ήχων υψηλών συχνοτήτων. Η θεωρία αυτή των δύο στοιχείων είναι ευρύτερα γνωστή ως διπλή θεωρία εντοπισμού (duplex theory of localization).

Η ηχητική σκιά που προκαλείται από το κεφάλι μας δείχνει γιατί η ενδοακουστική διαφορά στην ένταση (ILD) είναι μικρή για τις χαμηλές συχνότητες, και συνεπώς γιατί οι ακροατές δεν μπορούν να τη χρησιμοποιήσουν για τον προσδιορισμό της διεύθυνσης εκπομπής των χαμηλών αυτών συχνοτήτων. Η ILD μειώνεται ανάλογα όσο μειώνεται και η συχνότητα (αύξηση του μήκους κύματος), οπότε σε χαμηλές συχνότητες η ενδοακουστική διαφορά έντασης μπορεί να μην είναι αρκετά υψηλή ώστε να είναι ανιχνεύσιμη. Έτσι εξηγείται και το κενό στο γράφημα 1.5.1.6(β) από τα 20 - 300 Hz.

Μία εξήγηση, τώρα, του γιατί η ενδοακουστική χρονική διαφορά (ITD) μπορεί να παρέχει εντοπισμό μόνο σε χαμηλές συχνότητες παρουσιάζεται στο σχήμα 1.5.1.7. Ας υποθέσουμε ότι ένας ήχος συχνότητας 1666 Hz καταφθάνει πρώτος στο αριστερό αυτί. Επομένως, ο ημιτονοειδής ήχος που θα φτάσει και στα δύο αυτιά θα έχει τη μορφή του διαγράμματος 1.5.1.7(α) για το αριστερό αυτί και του διαγράμματος 1.5.1.7(β) για το δεξί αυτί. Οι διαφορές που προκύπτουν, μπορούν να χρησιμοποιηθούν από το ακουστικό σύστημα για να καθορίσουν ότι η ηχητική πηγή είναι πιο κοντά στο αριστερό αυτί.

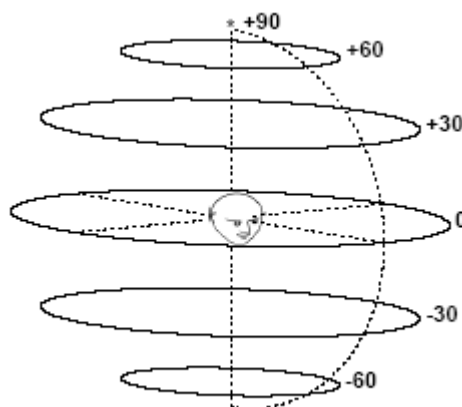
Γνωρίζουμε ότι η περίοδος του ήχου των 1666 Hz είναι σχεδόν ίση με 0.6 msec και υποθέτουμε ότι ο μέγιστος χρόνος (Δt), που είναι απαραίτητος για έναν ήχο, ώστε να διανύσει την απόσταση από το ένα αυτί στο άλλο είναι επίσης 0.6 msec. Έτσι ένας ήχος με συχνότητα 1666 Hz θα παρουσιάζοταν, στα δύο αυτιά, όπως στο σχήμα 1.5.1.7.



Σχ 1.5.1.7: Παρουσίαση της ενδοακουστικής χρονικής διαφοράς για απλό τόνο συχνότητας 1666 Hz και για διαφορά δρόμου (Δt) ίση με 0.6 msec (α) για το αριστερό αυτί και (β) για το δεξί αυτί.

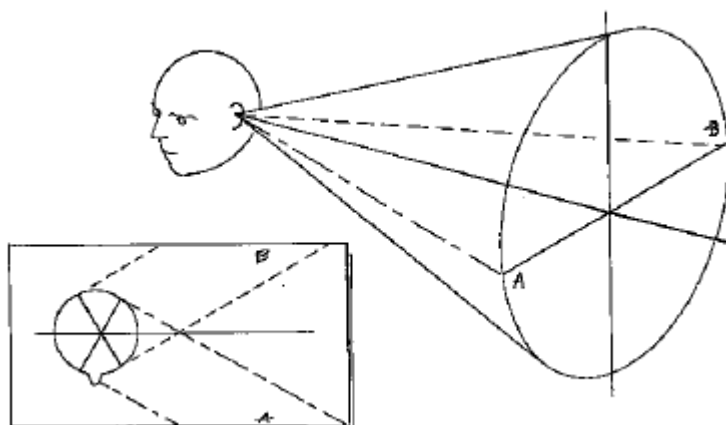
Παρατηρούμε ότι αν και ο ημιτονοειδής ήχος προσεγγίζει πρώτα το αριστερό αυτί και η πρώτη περίοδος παρουσιάζεται, ενώ δεν ισχύει το ίδιο για το δεξί αυτί του ακροατή, οι κυματομορφές είναι όμοιες από ένα σημείο και μετά (δηλ. από τη δεύτερη περίοδο και μετά, που σημειώνεται με το λατινικό α μέσα σε παρένθεση, (a)). Επομένως, εκτός από την πρώτη περίοδο, δεν υπάρχει διαφορά μεταξύ των δύο σημάτων, οπότε ο ακροατής θα μπορούσε να εντοπίσει τον ήχο σε μία θέση ακριβώς μπροστά του. Το συμπέρασμα αυτό, όμως, θα ήταν λάθος, εφόσον είδαμε ότι ο ήχος παρουσιάστηκε πιο κοντά στο αριστερό αυτί. Αυτού του είδους η σύγκυση δε θα υφίσταται για συχνότητες μικρότερες των 1666 Hz, μα θα ισχύει για συχνότητες μεγαλύτερες από τη συχνότητα αυτή, υποθέτοντας πως ο χρόνος που χρειάζεται ο ήχος για να διανύσει την απόσταση από το ένα αυτί στο άλλο είναι 0.6 msec. Επομένως, η ενδοακουστική χρονική διαφορά (ITD) θα προκαλεί διαφορετικά συμπεράσματα αναφορικά με τον εντοπισμό του ήχου, όταν η συχνότητα είναι πολύ μεγάλη (μικρό μήκος κύματος) σε σχέση με τις διαστάσεις του πλάτους του κεφαλιού.

1.5.2 Εντοπισμός στο κάθετο επίπεδο:



Σχ. 1.5.2.1: Παρουσίαση του κάθετου επιπέδου σε σύστημα σφαιρικών πολικών συντεταγμένων.

Στην προηγούμενη παράγραφο είδαμε ότι η ITD και η ILD είναι μεγάλης σημασίας για τον εντοπισμό μιας πηγής στο οριζόντιο επίπεδο. Αν το κεφάλι ενός ακροατή παρέμενε σταθερό, τότε υπάρχει ένας διαφορετικός αριθμός θέσεων που θα μπορούσαν να έχουν τις ίδιες τιμές διαφορών έντασης και χρόνου (ITD & ILD). Για παράδειγμα ένας ήχος, υπο συνθήκες, που βρίσκεται μπροστά από τον ακροατή θα είχε τις ίδιες ενδοακουστικές διαφορές με έναν ήχο που βρίσκεται στο πίσω μέρος του ακροατή, τις ίδιες ενδοακουστικές διαφορές με έναν ήχο που βρίσκεται ακριβώς από πάνω και τις ίδιες με έναν ήχο που βρίσκεται ακριβώς από κάτω του ακροατή



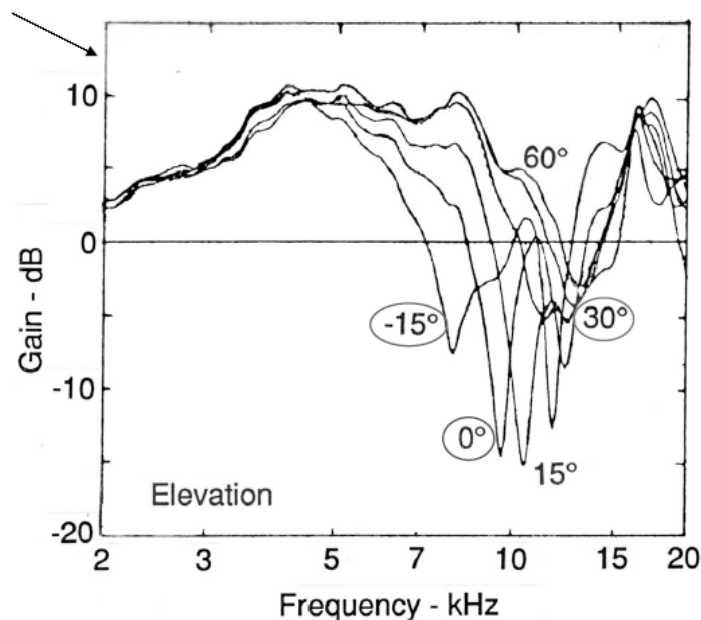
Σχ. 1.5.2.2: Σχηματική αναπαράσταση του φαινομένου του κώνου της σύγχυσης (cone of confusion).

Το επίπεδο αυτό, σε σχήμα υπερβολής, είναι μια περιοχή σύγχυσης (cone of confusion) αναφορικά με τον εντοπισμό της ηχητικής πηγής. Για κάθε ηχητική πηγή υπάρχει μία περιοχή σύγχυσης όπου οι ενδοακουστικές διαφορές είναι ίδιες. Οι περιοχές σύγχυσης λαμβάνουν χώρα όταν το κεφάλι παραμένει σταθερό. Μικρές κινήσεις του κεφαλιού, δίνουν την ικανότητα σε έναν ακροατή να εντοπίσει την ηχητική πηγή. Ακόμα, όμως, και στις περιπτώσεις που δε μετακινούμε το κεφάλι μας είμαστε σε θέση να εντοπίσουμε ήχους σε περιοχές σύγχυσης. Δηλαδή μπορούμε να εντοπίσουμε τη διεύθυνση του ήχου (μπρός - πίσω, πάνω ή κάτω). Ισχύει, πάντως ότι όταν κάνουμε λάθη σχετικά με τον εντοπισμό μιας ηχητικής πηγής, αυτά συνήθως οφείλονται στον κώνο της σύγχυσης.

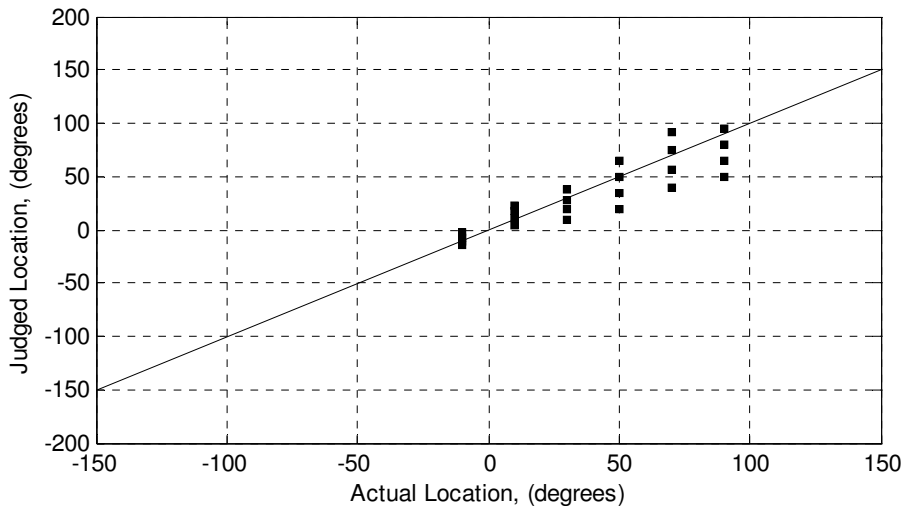
Οι ηχητικές πηγές, όταν οι ήχοι βρίσκονται εντός των περιοχών σύγχυσης, μπορούν να εντοπιστούν. Για να συμβεί αυτό, όμως, το ακουστικό μας σύστημα πρέπει να αξιοποιήσει κάποια επιπλέον δεδομένα. Τα δεδομένα αυτά τα λαμβάνει από τις λεγόμενες συναρτήσεις μεταφοράς σχετιζόμενες με το κεφάλι (Head Related Transfer Functions, HRTFs). Το πτερύγιο, κυρίως, του αυτιού προκαλούν μια μικρή ηχητική σκιά. Η επίδραση της είναι πιο σημαντική για τις υψηλές συχνότητες και αυτό γιατί τα μήκη κύματος αυτών των συχνοτήτων είναι κοντά στις διαστάσεις του πτερυγίου. Στην περίπτωση όπου ο ήχος είναι σύνθετος, όπως για παράδειγμα ο θόρυβος, διαφορετικές συχνότητες του ήχου θα ενισχυθούν και θα ελαττωθούν και κάποιες άλλες θα καθυστερήσουν χρονικά. Αυτή η καθυστέρηση θα οδηγήσει σε διαφορές φάσεων και ως αποτέλεσμα προκύπτει ένα ευρύ φάσμα φάσεων για τις HRTFs. Το ποσοστό της ενίσχυσης ή της ελάττωσης καθώς και της καθυστέρησης θα εξαρτάται από τη διεύθυνση από την οποία προέρχεται ο ήχος. Για παράδειγμα η επίδραση του πτερυγίου είναι μεγαλύτερη για ήχους που έρχονται από το πίσω μέρος του κεφαλιού σε σχέση με ήχους που έρχονται από μπροστά. Κατά συνέπεια οι HRTFs

για έναν σύνθετο ήχο θα διαφέρουν ανάλογα με τη σχέση που έχει η πηγή του ήχου με το ανθρώπινο σώμα. Επειδή, τώρα, οι μεγαλύτερες αλλαγές στις HRTFs συμβαίνουν για υψηλές συχνότητες, εξαιτίας της αλληλεπίδρασης μεταξύ του μήκους κύματος και του μεγέθους του εμποδίου, δεν είναι περίεργο που οι πιο σημαντικές ενδείξεις για τον προσδιορισμό της διεύθυνσης του ήχου στο κάθετο επίπεδο συμβαίνουν για τις υψηλές συχνότητες. Το σχήμα 1.5.2.3 παρουσιάζει τις HRTFs για διάφορες γωνίες στον κάθετο άξονα. Επειδή η ηχητική πηγή βρίσκεται στο κάθετο επίπεδο (δηλ. ανάμεσα στα μάτια) η ένταση που φτάνει σε κάθε αυτί είναι περίπου η ίδια. Παρατηρούμε ότι γύρω από την περιοχή των 10 kHz, εμφανίζονται έντονα βυθίσματα (notches) που όμως μετακινούνται με την αύξηση ή τη μείωση αντίστοιχα της γωνίας. Συνεπώς, προκύπτει, πως οι HRTFs χρησιμοποιούνται από την ανθρώπινη ακοή για τον κάθετο εντοπισμό μιας ηχητικής πηγής. Οι HRTFs χρησιμοποιούνται, επίσης, και για το διαχωρισμό του αν η πηγή είναι πίσω ή μπροστά από τον ακροατή. Το σχήμα 1.5.2.4 δείχνει την ικανότητα ενός ακροατή να αντιληφθεί την κάθετη τοποθέτηση μιας ηχητικής πηγής με ευρύ φάσμα. Το γεγονός του ότι η διακύμανση των δεδομένων είναι μεγάλη, μας δηλώνει ότι το ακουστικό μας σύστημα δεν είναι τόσο ικανό στο να διακρίνει την τοποθέτηση ενός ήχου στο κάθετο επίπεδο, όσο είναι αναφορικά με την αναγνώριση της τοποθέτησης ενός ήχου στο οριζόντιο επίπεδο.

Τέλος, μπορούμε να πούμε συμπερασματικά, ότι ο εντοπισμός της διεύθυνσης ενός ήχου στον κάθετο άξονα είναι το αποτέλεσμα της αλληλεπίδρασης του ήχου με τα πτερύγια των αυτιών, με το κεφάλι και με το σώμα.



Σχ. 1.5.2.3: Παρουσίαση των HRTFs για τον κάθετο άξονα και για διάφορες γωνίες.



Σχ. 1.5.2.4: Εντοπισμός της διεύθυνσης εκπομπής στον κάθετο άξονα για έναν broadband noise.

1.5.3 Εντοπισμός της απόστασης:

Για τον εντοπισμό μιας ηχητικής πηγής σε σχέση με την απόσταση της, πολύ λιγότερα είναι γνωστά συγκριτικά με τις προηγούμενες περιπτώσεις. Μια πιθανή ένδειξη για την απόσταση είναι η ένταση. Γενικά οι πιο δυνατοί ήχοι βρίσκονται πιο κοντά σε σχέση με τους πιο μακρινούς. Μπορεί, όμως, οι δυνατοί ήχοι να είναι μακριά και οι ήχοι χαμηλής συχνότητας να είναι κοντά. Κατά συνέπεια, η ένταση του ήχου μπορεί να αποτελεί ένδειξη για την απόσταση του, μόνον εάν ο ακροατής έχει πληροφορίες για την αναμενόμενη ένταση μιας ηχητικής πηγής. Για παράδειγμα, εάν κάποιος μιλάει, τότε γνωρίζουμε απο την εμπειρία μας στον καθημερινό κόσμο πώς αυτό πρέπει να ακούγεται και κατά συνέπεια αυτό μπορεί να δώσει πληροφορίες σε κάποιον για την απόσταση της ηχητικής πηγής.

Οι πρώτες ανακλάσεις απο μια ηχητική πηγή, παρέχουν ενδείξεις για τον προσδιορισμό της απόστασης της πηγής. Η αναλογία του άμεσου ήχου που φθάνει στον ακροατή σε σχέση με τις πρώτες ανακλάσεις μειώνεται όσο πιο μακριά βρίσκεται η ηχητική πηγή. Κοντά στον ακροατή η ένταση του ανακλώμενου ήχου είναι μικρότερη σε σχέση με τον άμεσο ήχο που φθάνει σε αυτόν. Γενικά όμως, μπορούμε να πούμε, ότι οι ακροατές έχουν πολύ μικρότερη ικανότητα στο να διακρίνουν την απόσταση μιας ηχητικής πηγής, απο την ικανότητα τους να προσδιορίζουν τη θέση της στο οριζόντιο ή στο κάθετο επίπεδο.

1.5.4 Φαινόμενο προβαδίσματος (Precedence effect):

Απο την καθημερινή μας εμπειρία γνωρίζουμε, ότι ακόμα και σε δωμάτια που υπάρχουν πολλές ανακλάσεις απο τους τοίχους, το πάτωμα και το ταβάνι, έχουμε την ικανότητα να εντοπίζουμε μια ηχητική πηγή. Οι αντηχήσεις προκαλούν ένα πολύπλοκο ηχητικό σήμα το οποίο προέρχεται απο πολλές κατευθύνσεις. Πως το ακουστικό μας σύστημα τότε, συλλέγει τα κατάλληλα στοιχεία που θα το οδηγήσουν στον επιτυχή προσδιορισμό της θέσης της ηχητικής πηγής;

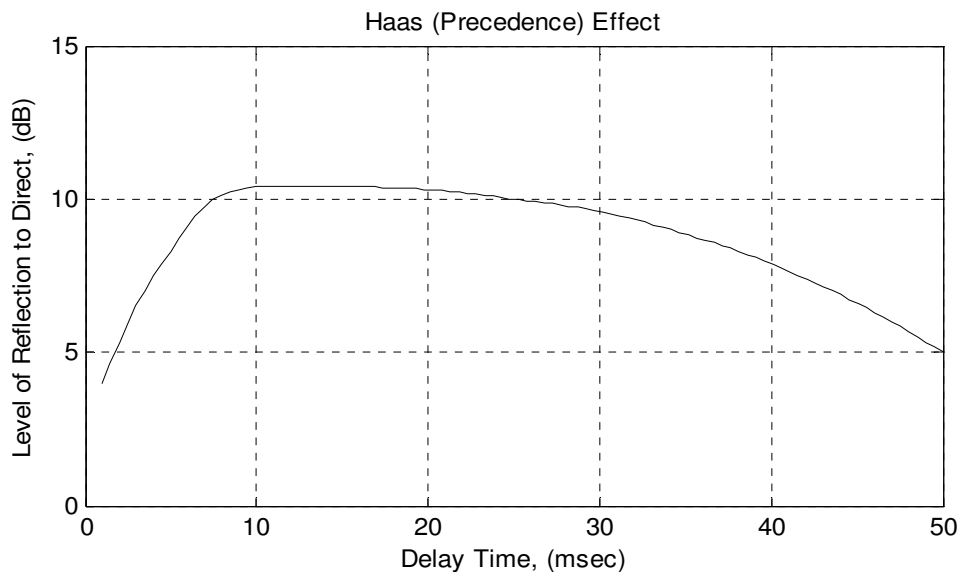
Ένας μεγάλος αριθμός πειραμάτων έχει δείξει ότι το πρώτο κύμα που φτάνει στα αυτιά του ακροατή, κυριαρχεί αναφορικά με τον εντοπισμό της θέσης της ηχητικής πηγής (νόμος του πρώτου μετώπου κύματος). Το ακουστικό μας σύστημα φαίνεται ότι αξιοποιεί αυτόν το νόμο του πρώτου μετώπου κύματος.

Ο ακουστικός μας μηχανισμός ολοκληρώνει τις εντάσεις του ήκου σε μικρά χρονικά διαστήματα. Σαν παράδειγμα θα μπορούσαμε να αναφέρουμε ότι στην περίπτωση που βρισκόμαστε στο εσωτερικό μιας αίθουσας, το αυτί και ο εγκέφαλος έχουν την αξιοσημείωτη ικανότητα να συλλέγουν όλες τις ανακλάσεις που καταφθάνουν εντός του χρονικού πεδίου των 50 msec, περίπου, μετά την πρόσληψη του άμεσου (direct) ήκου και να τις προσθέτουν (δηλ. ολοκληρώνουν) έτσι ώστε να δημιουργείται η εντύπωση ότι όλοι αυτοί οι ήχοι έρχονται απο τη διεύθυνση της αρχικής πηγής, αν και στην πραγματικότητα υφίστανται ανακλάσεις απο άλλες διευθύνσεις. Η ηχητική ενέργεια που ολοκληρώνεται μέσα σε αυτή την περίοδο δίνει και μία επιπρόσθετη αίσθηση ηχηρότητας.

Δε θα πρέπει να μας εκπλήσσει το γεγονός πως το ανθρώπινο αυτί συγχωνεύει όλους τους ήχους που καταφθάνουν σε αυτό εντός ενός ορισμένου χρονικού πλαισίου. Η ακουστική αυτή συγχώνευση λειτουργεί καλύτερα κατά τη διάρκεια των πρώτων 20 – 30 msec. Πέρα απο τα 50 msec και μέχρι τα 80 msec επικρατούν οι διακριτές αντηχήσεις³.

Ο Helmut Haas μελέτησε σε βάθος το φαινόμενο και κατέληξε στο συμπέρασμα πως αν ένας ήχος και η επανάληψη του είναι της ίδιας στάθμης και φθάσουν στον ακροατή με μικρή χρονική διαφορά (π.χ. 20 msec), τότε ο ήχος που θα φτάσει πρώτος καθορίζει τη φαινόμενη θέση της πηγής που ο ακροατής αντιλαμβάνεται ως μοναδική. Ως ορισμό για το φαινόμενο μπορούμε να αναφέρουμε το εξής: «*Φαινόμενο Haas ή προβαδίσματος (Precedence effect), είναι το φαινόμενο εκείνο κατά το οποίο το ανθρώπινο ακουστικό σύστημα δε μπορεί να διαχωρίσει τη θέση ενός ήκου απο αυτή της αντήχησής του, εάν η μεταξύ τους χρονική διαφορά είναι μικρότερη της τάξεως των 30 msec*». [15]

³ Εκτενέστερη αναφορά επί του θέματος της πρόσληψης του αντηχητικού πεδίου και των καθυστερημένων ανακλάσεων βλέπε στο παράρτημα Β όπου εξετάζεται το εν λόγω φαινόμενο συναρτήσει των ψυχοακουστικών ιδιοτήτων που προκύπτουν απο το delay που εισάγουν τα ΕQ.



Σχ. 1.5.4.1: Η καμπύλη του φαινομένου του προβαδίσματος (Precedence ή αλλιώς Haas effect).

Ο Haas, για τα πειράματά του, τοποθέτησε άτομα σε απόσταση τριών μέτρων από δύο μεγάφωνα που ήταν τοποθετημένα έτσι ώστε έτεμναν μία γωνία 45° , ενώ η γραμμή συμμετρίας των ακροατών διχοτομούσε τη γωνία αυτή. Οι συνθήκες ήταν περίπου ανηχοϊκές. Ζητήθηκε, από τους ακροατές, να ρυθμίσουν την ένταση έως ότου ο ήχος από το απευθείας μεγάφωνο εξισωνόταν με τον ήχο από το καθυστερημένο μεγάφωνο. Στη συνέχεια ο Haas προχώρησε στη μελέτη της μεταβολής της καθυστέρησης. Όπως φαίνεται στο σχήμα 1.5.4.1 ο Haas βρήκε ότι στην περιοχή καθυστερήσεων από 5 έως 35 msec η ένταση του ήχου από το καθυστερημένο μεγάφωνο θα πρέπει να αυξηθεί περισσότερο από 10 dB σε σχέση με τον απευθείας πριν ακουστεί σε διακριτή επανάληψη (ηχώ). Μέσα σε ένα δωμάτιο, η ενέργεια από ανάκλαση που φτάνει στο ανθρώπινο αυτί σε χρόνο εντός των 35 msec, ολοκληρώνεται με τον απευθείας ήχο και γίνεται αντιληπτή ως τμήμα του απευθείας ήχου και όχι ως διακριτή επανάληψη.

1.5.5 Φαινόμενο συνάθροισης (Cocktail party effect):

Η ικανότητα του ακουστικού μας συστήματος να ακούει και να καταλαβαίνει τι λέει ένας ομιλητής κατά τη διάρκεια της ομιλίας πολλών, ονομάζεται cocktail party effect. Αυτή η ικανότητα του αυτιού μας, πηγάζει από το γεγονός ότι το ακουστικό μας σύστημα μπορεί να αλλάζει το χρόνο που κάνει το σήμα να φτάσει από τα αυτιά μας στον εγκέφαλο, με αποτέλεσμα για το σήμα που θέλουμε να ακούσουμε να αθροίζεται η ένταση και από τα δύο αυτιά. Με αυτό τον τρόπο το σήμα γίνεται περισσότερο ευδιάκριτο.

Ως παράδειγμα μπορούμε να φανταστούμε τις φωνές δύο ομιλητών οι οποίοι διαβάζουν ένα διαφορετικό ο καθένας κείμενο και τους οποίους τους έχουμε ηχογραφήσει. Αν ακροαστούμε τους δύο αυτούς ομιλητές μέσω δύο μεγαφώνων έχοντας τοποθετήσει και τους δύο στο κέντρο της στερεοφωνικής εικόνας, τότε ίσως, στην καλύτερη περίπτωση, να μπορέσουμε να αντιληφθούμε κάποιες μεμονωμένες λέξεις, χωρίς ωστόσο να μπορέσουμε να καταλάβουμε τα όσα αναφέρει έστω ο ένας εκ των δύο ομιλητών. Αν, εν συνεχεία, ακροαστούμε τους δύο ομιλητές και πάλι διαμέσου δύο μεγαφώνων μόνο που αυτή τη φορά έχοντας τους τοποθετήσει στα άκρα της στερεοφωνικής εικόνας, δηλ. τέρμα δεξιά τον ένα και τέρμα αριστερά τον άλλο. Σε αυτή την περίπτωση μπορούμε να συγκεντρωθούμε και να κατανοήσουμε τα όσα λέει ο ένας από τους δύο ομιλητές εύκολα.

1.6 Ψυχοακουστική και audio mastering:

Ένα από τα ζητούμενα της πτυχιακής αυτής εργασίας είναι ο προσδιορισμός σε ένα πρώτο επίπεδο της σχέσης των ψυχοακουστικών φαινομένων που λαμβάνουν χώρα και επηρεάζουν τον τρόπο με τον οποίο αντιλαμβανόμαστε τον ήχο, με τις διεργασίες που λαμβάνουν χώρα στο audio mastering. Σε αυτή την παράγραφο θα επιδιώξουμε τη συσχέτιση αυτή. Στα επόμενα κεφάλαια δε, που αναφέρονται οι διάφορες στο mastering τεχνικές, καθώς και στα συμπεράσματα που προκύπτουν από τις ακροάσεις των διαφόρων υποκειμένων στο ηχητικό υλικό, θα εμβαθύνουμε, όπου κρίνεται αναγκαίο αυτό, στη σχέση αυτή.

Στην παράγραφο 1.3.1 λοιπόν, αναφερόμαστε στην ευαισθησία της ακοής αναφορικά με τη διάρκεια των ήχων. Προκύπτει πως όσο μεγαλύτερη διάρκεια έχει ένας ήχος (με όριο τα 300 - 500 msec) τόσο πιο εύκολα τον αντιλαμβανόμαστε. Συνακόλουθα ήχοι με μικρότερης διάρκειας, για να γίνουν το ίδιο ακουστοί, θα πρέπει να αυξήσουμε τη στάθμη τους σε dB. Έτσι λοιπόν, στο μουσικό σήμα που συναντάμε και κρουστικούς ήχους (π.χ. ταμπούρο, μπότα) και οι οποίοι έχουν μικρή χρονική διάρκεια, καλούμαστε στο mastering να τους ενισχύσουμε (αν αυτό χρειάζεται) προκειμένου να γίνουν ευκολότερα αντιληπτοί. Ταυτόχρονα ενισχύουμε με αυτόν τον τρόπο και την αίσθηση του ρυθμικού μέρους του μουσικού κομματιού. Τα παραπάνω μπορούν να γίνουν με τεχνικές όπως είναι το multiband compression, το multiband expansion καθώς και με το limiting.

Στην παράγραφο 1.3.5 τώρα, εξετάζεται η ικανότητα του ακουστικού μας συστήματος στο να διακρίνει τις διάφορες συχνотικές περιοχές μεταξύ τους. Αυτό έχει μεγάλη σημασία για τη χρήση των equalizer όπως πολύ εύκολα μπορεί να γίνει αντιληπτό.

Εν συνεχεία, στην παράγραφο 1.3.6 καταπιανόμαστε με την αντιληπτικότητα του ακουστικού μας συστήματος αναφορικά με την ένταση, καθώς και με τη διακριτική του ικανότητα σχετικά με αυτή. Είναι επομένως, πολύ σημαντικό να γνωρίζουμε με πιο

τρόπο ανταποκρίνεται το αυτί μας στις διάφορες εντάσεις, όπως επίσης και στις διακυμάνσεις τους κατά τη διάρκεια ενός μουσικού κομματιού μιας και πλήθος τεχνικών αφορά ακριβώς την ένταση του ηχητικού υλικού. Τέτοιες τεχνικές είναι το limiting, το upward compression, το upward expansion, κ.ο.κ. Επιπροσθέτως, το ζήτημα της έντασης ενός συγκεκριμένου ήχου αφορά και το τονικό του ύψος, καθώς γνωρίζουμε ότι ήχοι ίδιας συχνότητας, διαφορετικής όμως έντασης μπορούν να γίνουν αντιληπτοί με διαφορετικό τονικό ύψος από τον άνθρωπο.

Στην παράγραφο 1.3.7.1 ασχολούμαστε με το λεγόμενο complex pitch, δηλαδή ασχολούμαστε με το τονικό ύψος σύνθετων τόνων οι οποίοι εμπεριέχουν μία θεμελιώδη συχνότητα και ένα πλήθος από αρμονικές. Στην παράγραφο αυτή εξετάσαμε, επίσης, και το φαινόμενο της χαμένης θεμελιώδους. Από αυτά που προκύπτουν από την παράγραφο 1.3.7.1 εφοδιαζόμαστε με περισσότερες πληροφορίες αναφορικά με τις τεχνικές enhancement στο mastering. Πιο συγκεκριμένα διαπιστώνουμε πως μπορούμε να ενισχύσουμε το βαθύ χαμηλό μιας μίξης ενισχύοντας τις αρμονικές μιας συγκεκριμένης συχνότητας περιοχής (υπάρχουν και ειδικευμένα plug - in για αυτό το σκοπό). Παραπέρα μπορούμε σε θεωρητικό επίπεδο να δημιουργήσουμε στον ακροατή την ψευδαίσθηση ότι υφίσταται ενέργεια σε συχνότητες περιοχές για τις οποίες στην πραγματικότητα δε συμβαίνει κάτι τέτοιο. Τέλος όπως γίνεται εύκολα αντιληπτό, τα παραπάνω είναι πολύ χρήσιμα και για τη χρήση των equalizer.

Στην παράγραφο 1.3.8 τώρα, ασχοληθήκαμε με τη χοιρά. Από αυτή την παράγραφο προκύπτει πως η χοιρά ενός ήχου εξαρτάται από τις αρμονικές τις οποίες εμπεριέχει, την ενέργεια η οποία εμπεριέχεται σε κάθε αρμονική, τη σχέση θεμελιώδους - αρμονικών και τέλος για την αντίληψη της χοιράς ρόλο παίζουν και τα χαρακτηριστικά της κυματομορφής του ήχου. Στην παράγραφο αυτή είδαμε, επίσης, και τους τέσσερις κανόνες του Helmholtz που αφορούν τη χοιρά. Από τα παραπάνω προκύπτει πως τεχνικές όπως το limiting και το compression που επηρεάζουν τα χαρακτηριστικά της κυματομορφής του ήχου, θα επηρεάζουν κατά συνέπεια και την αντίληψη της χοιράς του ίδιου του ήχου. Επίσης από τους τέσσερις κανόνες του Helmholtz προκύπτουν χρήσιμα συμπεράσματα για τη χρήση των equalizer. Για παράδειγμα από το 2^ο κανόνα προκύπτει πως ένας ήχος ηχεί ευχάριστα στην περίπτωση που το μεγαλύτερο ποσοστό της ενέργειας του βρίσκεται μέχρι και την 6^η αρμονική. Έτσι εξηγείται και το γιατί όταν το παρακάνουμε με τη χρήση των EQ στις υψηλές συχνότητες, ο ήχος αποκτά μια σκληρή και τραχιά αίσθηση (πράγμα που επιβεβαιώνει τον 4^ο κανόνα του Helmholtz).

Εν συνεχεία, στην παράγραφο 1.3.9 ασχοληθήκαμε με τη μη γραμμικότητα του ανθρώπινου συστήματος της ακοής. Βλέπουμε ότι το ανθρώπινο αυτί παράγει συχνότητες (ως ένα μη γραμμικό σύστημα) που δεν υπάρχουν στο φάσμα του ήχου. Είδαμε ότι οι τόνοι συνδυασμού μπορούν να είναι είτε πρόσθεσης, είτε διαφοράς. Είδαμε, επίσης, ότι δε συμπεριφέρεται με τον ίδιο τρόπο για διαφορετικές εντάσεις. Από αυτά συνεπάγεται πως οι επεξεργασίες στο mastering μπορεί σε ορισμένες

περιπτώσεις να μη στηρίζονται σε ορθολογικά - αντικειμενικά κριτήρια, αλλά σε αισθητικά κριτήρια. Σε κάθε περίπτωση, όμως, τον καλύτερο μας σύμβουλο αποτελούν τα αυτιά μας.

Συνακόλουθα στην παράγραφο 1.4 εξετάζουμε το φαινόμενο του masking. Είδαμε με πιο τρόπο ήχοι ίδιας ή παραπλήσιας συχνότητας μπορούν να επικαλύψουν άλλους ήχους. Είδαμε επίσης, για ποιες συχνοτικές περιοχές το εν λόγω φαινόμενο τείνει να αποκτά τα μέγιστα του και πως αυτό εξαρτάται από την ένταση. Από τα όσα έχουμε αναφέρει εύκολα προκύπτει ότι το φαινόμενο αυτό μπορεί να εμφανιστεί, με αρνητικές συνέπειες, σε πλήθος τεχνικών στο audio mastering. Τέτοιες τεχνικές είναι το equalization, το multiband compression, το upward compression και άλλες. Συνεπώς θα πρέπει να είμαστε προσεκτικοί, συναρτήσει του ότι επεξεργασόμαστε ολοκληρωμένες μίξεις και όχι μεμονωμένα μουσικά όργανα.

Στις παραγράφους, τώρα, 1.5, 1.5.1 και 1.5.2 ασχολούμαστε με το sound localization τόσο στο οριζόντιο όσο και στο κάθετο επίπεδο. Ασχοληθήκαμε, επίσης, με τις HRTFs, με τις ενδοακουστικές διαφορές έντασης, χρόνου και φάσης (ILD, ITD & IPD) και με την ελάχιστη αντιληπτή γωνία (MAA) μεταξύ άλλων. Κοντολογίς και σε ότι αφορά τη σχέση όλων των παραπάνω με το audio mastering, καταδεικνύεται η απαίτηση για ένα όσο το πιο δυνατό καλύτερα ισορροπημένο τονικά ηχητικό υλικό και με την όσο το δυνατόν ευρύτερη στερεοφωνική εικόνα, χωρίς θολό κέντρο και με την όσο το δυνατόν καλύτερη χωροθέτηση των μουσικών οργάνων. Αυτά απαιτούνται για την όσο το δυνατόν καλύτερη αντίληψη του μουσικού συμβάντος από τον ακροατή. Συνεπώς γίνεται κατανοητό, πως τα παραπάνω, αφορούν ολόκληρη την αλυσίδα της μουσικής παραγωγής, από την ηχογράφιση μέχρι το mastering: και συνεπώς αφορούν και το σύνολο των επεξεργασιών που λαμβάνουν χώρα κατά τη διάρκεια του mastering. Πιο συγκεκριμένα, όμως, και για την καλύτερη κατανόηση, έστω μιας πτυχής του φαινομένου, δίδεται στο παράρτημα Β της παρούσης εργασίας, ένα παράδειγμα διαπλάτυνσης της στερεοφωνικής εικόνας με τη χρήση equalizer, το οποίο και θα εξετάσουμε παρακάτω.

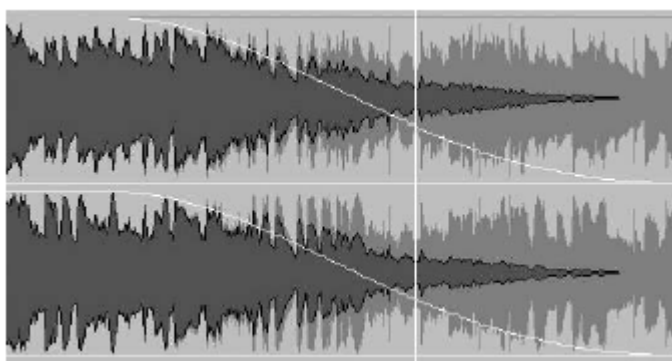
Κλείνοντας τη συγκεκριμένη παράγραφο, στην οποία προσπαθήσαμε να συσχετίσουμε ορισμένα φαινόμενα της ψυχοακουστικής με τις διεργασίες στο audio mastering, καλό θα ήταν να αναφερθεί πως η επιρροή των ψυχοακουστικών φαινομένων αφορά οποιαδήποτε επιστήμη ασχολείται με τον ήχο. Την ακουστική, την ηχοληψία, το audio mastering κ.ο.κ. Ειδικά, πάντως, για το mastering είναι κομβικής σημασίας διότι αφενός το πεδίο εργασίας αφορά ολοκληρωμένες στερεοφωνικές και πολυκάναλες μίξεις και αφετέρου διότι όπως έχουμε ήδη αναφέρει μας δίνει τη δυνατότητα να αναπτύξουμε τεχνικές για την περαιτέρω καλύτερευση ορισμένων συνιστωσών του ηχητικού υλικού, δίχως αυτό να λειτουργεί αναιρετικά σε σχέση με όλες τις άλλες.

**Editing
&
Audio Mastering**

2.1 Editing, fades και άλλες εργασίες:

Μετά τη συλλογή του μουσικού - ηχητικού υλικού και την πιθανή μεταφορά του στα επιθυμητά μέσα, η διαδικασία του mastering μπορεί να χωριστεί σε ορισμένες βασικές κατηγορίες. Η πρώτη από αυτές είναι το μοντάζ, οι οποίες διορθώσεις χρειάζεται να γίνουν, καθώς και άλλες τεχνικές λειτουργίες.

Τα ψηφιακά συστήματα επεξεργασίας ήχου, συχνά αποκαλούμενα στη μουσική τεχνολογία με τον όρο Digital Audio Workstations (DAWs), μας επιτρέπουν πια να έχουμε τον έλεγχο στην ηχογραφημένη μουσική όσο πότε άλλοτε. Η δυνατότητα της οπτικής επαφής με το ηχητικό κύμα είναι σε πολλές περιπτώσεις μείζονος σημασίας για την ακριβή επεξεργασία του ηχητικού υλικού. Στο mastering μπορεί να χρειαστεί μοντάζ σε διάφορες περιπτώσεις, όπως όταν για παράδειγμα όταν ένα κομμάτι πρέπει να μεγαλώσει ή να μικρύνει σε διάρκεια (επανάληψη ή αφαίρεση αποσπασμάτων). Υπάρχει ακόμη και η περίπτωση να χρειαστεί ο συνδυασμός αποσπασμάτων από διάφορες μίξεις ενός μουσικού κομματιού με σκοπό τη δημιουργία μιας καλύτερης τελικής μίξης. Μια από τις σημαντικότερες επεξεργασίες είναι αυτή των fade-ins και fade-outs όπου κι όταν χρειάζονται (κι όπως θα δούμε παρακάτω δεν χρειάζονται μόνο στην αρχή και στο τέλος των μουσικών κομματιών). Μπορεί να φαίνεται σαν μια απλή διαδικασία, όμως απαιτείται μεγάλη προσοχή ώστε τα fades να είναι απόλυτα ομαλά να ξεκινούν και να καταλήγουν από και σε πλήρη σιωπή. Τα σύγχρονα DAWs έχουν ειδικές ρυθμίσεις για τη συγκεκριμένη διαδικασία (βλ. σχ. 2.1)



Σχ. 2.1: Τυπικό fade-out σε ένα DWA σύστημα.

2.2 Άλλες εργασίες στο mastering:

Εκτός όμως από τα όσα αναφέρθηκαν στην προηγούμενη παράγραφο και τα όσα θα αναφερθούν στα επόμενα κεφάλαια, το mastering περιλαμβάνει και άλλες εργασίες πιο αυτονόητες καθώς και τεχνικές που δεν μπορούν να μπουν σε κάποια από αυτές τις κατηγορίες. Οι εργασίες στις οποίες αναφερόμαστε είναι:

Σειρά κομματιών, παύσεις, cross-fades: Το να μπουν τα κομμάτια στην επιλεγμένη σειρά για την ολοκλήρωση του άλμπουμ είναι κάτι που αποφασίζεται κατά τη διαδικασία του mastering. Μπορεί να φαίνεται μια απλή επιλογή, όμως πρέπει να έχει κανείς υπόψη του πως η σειρά των κομματιών σε ένα δίσκο είναι πάρα πολύ σημαντική δουλειά με επίπτωση στην αποδοχή του άλμπουμ, καθώς πολλές φορές μπορεί να παίξει ρόλο στην επιτυχία ή όχι μίας παραγωγής. Πρέπει λοιπόν αυτό να αποφασιστεί πολύ προσεκτικά μεταξύ των παραγωγών, καλλιτεχνών και του τεχνικού mastering. Η λεπτομέρεια στο mastering είναι τόσο σημαντική που δίνεται μεγάλη προσοχή ακόμα και για τη διάρκεια της παύσης (αν υπάρχει) μεταξύ των κομματιών. Για τη καλή ροή του άλμπουμ, η παύση δε θα πρέπει να είναι ούτε πολύ μικρή ούτε πολύ μεγάλη και θα πρέπει να δίνει την αίσθηση σωστού τέμπο και συγχρονισμού μεταξύ του τέλους του προηγούμενου και της αρχής του επόμενου κομματιού. Είναι επίσης πολύ πιθανό κάποια κομμάτια σε ένα άλμπουμ να συνδέονται ομαλά με το λεγόμενο cross-fade, τεχνική που επίσης εκτελείται στο mastering και απαιτεί μεγάλη λεπτομέρεια και προσοχή.

Μείωση ανάλυσης και μονοφωνική συμβατότητα: Η επεξεργασία του ήχου είναι τώρα πια κυρίως ψηφιακή και καταλήγει σε ψηφιακά μέσα. Έτσι λοιπόν οι περισσότεροι τεχνικοί mastering επιλέγουν να δουλεύουν σε υψηλή ανάλυση ψηφιακού ήχου (π.χ. 32bit, 96kHz) για όσο το δυνατόν λιγότερες απώλειες κατά την επεξεργασία του ήχου. Καθώς όμως τα ψηφιακά μέσα όπως το *CD* υποστηρίζουν χαμηλότερες αναλύσεις (*CD*: 16bit 44.1kHz, *DVD-Audio* : 24bit 96kHz) χρειάζεται να γίνει η μείωση ψηφιακής ανάλυσης - *bit rate* και η μείωση της δειγματοληψίας (down-sampling) του ψηφιακού ήχου. Όσο αφορά την μείωση του *bit rate* χρησιμοποιείται πάντα η τεχνική dithering για την μικρότερη δυνατή απώλεια του ηχητικού σήματος κατά την μετατροπή του (βλ. κεφάλαιο__). Στην μείωση της δειγματοληψίας χρησιμοποιούνται ειδικές μέθοδοι ή/και μηχανήματα, με πολύπλοκους αλγόριθμους για να επιτευχθεί καλή μετατροπή. Στο mastering είναι επίσης πολύ σημαντικό να εξασφαλιστεί η σωστή αναπαραγωγή και συμβατότητα του ηχητικού υλικού σε μονοφωνικά συστήματα, (τηλεοράσεις, ραδιόφωνο κ.α.), γνωστό ως mono compatibility. Δεν θα πρέπει να ξεχνάμε ότι όσο καλά κι αν ακούγεται ένα ηχητικό σήμα στη στερεοφωνία, είναι πιθανό μονοφωνικά να δημιουργούνται ακυρώσεις φάσης και προβλήματα διαχωρισμού των ήχων/οργάνων. Επομένως πάντα γίνεται ακρόαση του υλικού σε μονοφωνία και είναι πολύ σημαντικό να ελεγχθεί και να φιλτραριστεί για τυχόν ακυρώσεις φάσης του ήχου σε περιπτώσεις μονοφωνίας.

Λίστα PQ και ISRC: Όταν προετοιμαστεί ένα άλμπουμ για μαζική εκτύπωση θα πρέπει να συνοδεύεται από τις απαραίτητες τεχνικές πληροφορίες αναγκαίες για την διανομή του. Το *PQ log* είναι μια λίστα που περιέχει όλες τις λεπτομέρειες της διάρκειας και

παύσεων των κομματιών του άλμπουμ - εμπεριέχεται στο δίσκο και εκτυπώνεται ξεχωριστά για τον έλεγχό της. Ο *ISRC* (International Standard Recording Code), είναι ένας κωδικός μοναδικός για κάθε μουσικό κομμάτι αποτελούμενος από δώδεκα συνολικά ψηφία και γράμματα και εμπεριέχεται στη λίστα *PQ*.

Ποιοτικός έλεγχος: Όταν στο mastering ένα project έχει ολοκληρωθεί θα πρέπει να περάσει από τον ποιοτικό έλεγχο (quality control - CQ). Η διαδικασία αυτή περιλαμβάνει έλεγχο ακρόασης αλλά και αντικειμενικά τεχνικά τεστ. Στα μεγάλα mastering studios το *QC* γίνεται από ξεχωριστό τμήμα, ο τεχνικός του οποίου θα πρέπει να έχει μεγάλη εμπειρία, μουσικά αυτιά καθώς και τεχνικές γνώσεις. Όταν ο χώρος ακρόασης δεν είναι ιδανικός, η χρήση ακουστικών κατά τον ποιοτικό έλεγχο είναι πιθανή. Ακόμα περισσότερο όταν πρόκειται για πολυκάναλο ήχο, προβλήματα στα πίσω ηχεία θα εντοπισθούν ευκολότερα όταν γίνει απομονωμένη ακρόαση των καναλιών αυτών από ακουστικά. Τα ψηφιακά μέσα όπως τα *DAT*, *DDP Exabyte*, *PCM-1630*, *DLT*, *CD-R* και *DVD-R*, εμφανίζουν συχνά ψηφιακά σφάλματα γι' αυτό και όλα περιέχουν μηχανισμούς διόρθωσης σφαλμάτων. Τα σφάλματα που δεν διορθώνονται ενδέχεται να προκαλέσουν *clicks* και άλλους θορύβους στον ήχο. Καθώς όμως οι μηχανισμοί διόρθωσης λαθών λειτουργούν δεν μπορούμε να ξέρουμε σε τι κατάσταση είναι το μέσο. Έτσι υπάρχουν συστήματα, όπως το Clover System - *CD-R media evaluator*, που εφαρμόζουν ειδικά τεστ για την ανίχνευση και την παραγωγή λεπτομερούς λίστας των ψηφιακών σφαλμάτων που παρουσιάζει το μέσο. Εάν το μέσο που θα σταλεί στο εργαστήριο μαζικής παραγωγής περιέχει ψηφιακά λάθη πάνω από κάποιο όριο, τότε δεν θα γίνει δεκτό.

**Equalaization
&
Audio Mastering**

3. Βασικά εισαγωγικά σημεία αναφορικά με τη λειτουργία των equalizer:

3.1 Κεντρική συχνότητα (f_c):

Η κεντρική συχνότητα είναι το σημείο αναφοράς του φίλτρου. Υπάρχουν δύο διαφορετικές ερμηνείες της κεντρικής συχνότητας κάθε μία από τις οποίες εξαρτάται από τον τύπο του φίλτρου που χρησιμοποιείται.

Στα φίλτρα που έχουν παραμέτρους ρύθμισης του bandwidth (BW) ή του Q, η κεντρική συχνότητα ορίζεται ως το σημείο αυτό, στο πεδίο των συχνοτήτων, που τέμνει στη μέση την περιοχή που επηρεάζεται από το φίλτρο. Τις περισσότερες φορές η κεντρική συχνότητα είναι αυτή που επηρεάζεται περισσότερο από όλες τις άλλες σε αυτού του είδους τα φίλτρα.

Οι τύποι των φίλτρων, τώρα, που επηρεάζουν ολόκληρη τη συχνοτική περιοχή πριν ή μετά από μία καθορισμένη κεντρική συχνότητα (δηλ. τη συχνότητα αποκοπής του φίλτρου) ορίζουν τη συχνότητα αυτή, ως το σημείο αυτό στο οποίο έχουμε 3 dB αύξηση ή μείωση του πλάτους της συγκεκριμένης συχνοτικής περιοχής. Π.χ. ένα High Shelf φίλτρο το οποίο είναι σχεδιασμένο να ενισχύει το σήμα κατά 6 dB έχοντας ως συχνότητα αποκοπής τα 10 kHz, θα επιφέρει στη συχνότητα αυτή των 10 kHz μία αύξηση του πλάτους της τάξης των 3 dB.

Η κεντρική συχνότητα ορίζεται μαθηματικά με δύο τρόπους, αναλόγως με το αν χρησιμοποιούμε, για την περιγραφή του φίλτρου, ζώνες σταθερού εύρους (γραμμικός ο άξονας των συχνοτήτων) ή ζώνες σταθερού ποσοστιαίου εύρους (λογαριθμικός ο άξονας των συχνοτήτων).

Στην πρώτη περίπτωση η κεντρική συχνότητα ορίζεται από την αριθμητική μέση τιμή των δύο ακραίων συχνοτήτων f_1 & f_2 , δηλ.:

$$f_c = \frac{f_1 + f_2}{2} \quad (3.1.1)$$

Στη δεύτερη περίπτωση, η οποία είναι και η συνηθέστερη διότι ανταποκρίνεται στον τρόπο που αντιλαμβάνεται το ανθρώπινο αυτί τον ήχο, ορίζεται ως ο γεωμετρικός μέσος των δύο ακραίων συχνοτήτων f_1 & f_2 , δηλ.:

$$f_c = \sqrt{f_1 \cdot f_2} \quad (3.1.2)$$

3.2 Εύρος ζώνης του φίλτρου (bandwidth, BW) και Q:

Το bandwidth και το Q είναι δύο διαφορετικοί τρόποι για να περιγράψουμε το ίδιο πράγμα, το ζωνικό εύρος του φίλτρου. Το εύρος του φίλτρου ορίζεται στα σημεία όπου έχουμε πτώση κατά 3 dB εκατέρωθεν της κεντρικής συχνότητας. Το bandwidth αναπαριστά αυτό το εύρος σε Hz. Ένα bandwidth με τιμή 1000 σημαίνει ότι το εύρος ζώνης του φίλτρου είναι 1000 Hz μεταξύ των δύο ακραίων συχνοτήτων f_1 & f_2 όπου σημειώνεται η πτώση κατά 3 dB.

Το Q εκφράζει το εύρος ζώνης του φίλτρου αναφορικά με τον τρόπο που ακούμε. Όταν π.χ. το $Q=2$, τότε πάντα το εύρος του φίλτρου ισούται με μισή οκτάβα ανεξάρτητα από την τιμή της κεντρικής συχνότητας· όταν το $Q=1$, τότε το φίλτρο έχει εύρος μίας οκτάβας· όταν το $Q=4$, τότε το φίλτρο έχει εύρος το ένα τέταρτο της οκτάβας· τέλος, όταν το $Q=0.5$ τότε το φίλτρο έχει εύρος δύο οκτάβες. Με μαθηματικούς όρους το Q ισούται με την κεντρική συχνότητα διαιρεμένη από το bandwidth, δηλ.:

$$Q = \frac{f_c}{BW} \quad (3.2.1)$$

Κατά συνέπεια το bandwidth μπορεί να εκφραστεί ως εξής:

$$BW = \frac{f_c}{Q} \quad (3.2.2)$$

Όπως, επίσης, μπορεί να εκφραστεί και σύμφωνα με τον επίσημο ορισμό του:

$$BW = f_2 - f_1 \quad (3.2.3)$$

3.3 Ενίσχυση/εξασθένιση (Boost/Cut) μιας συχνοτικής περιοχής:

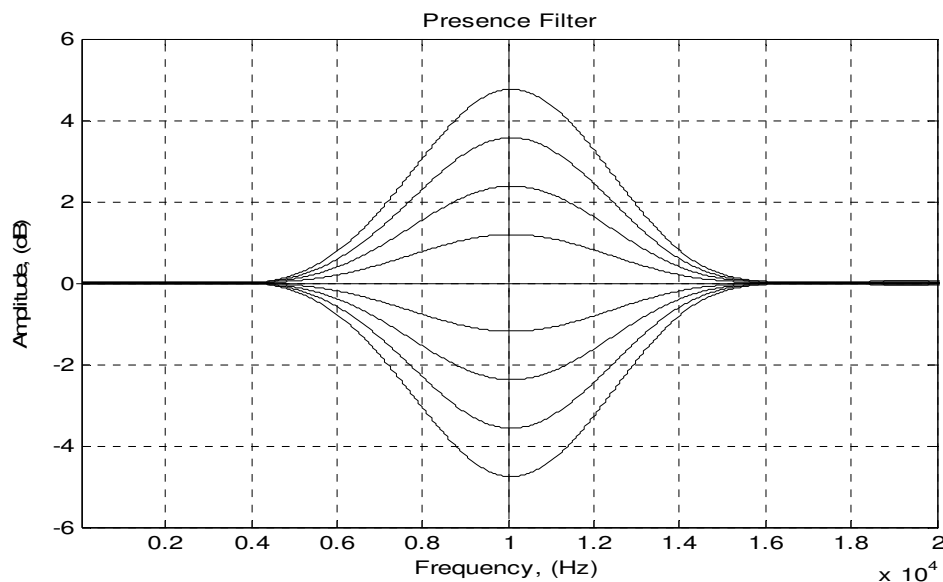
Το ποσοστό ενίσχυσης ή εξασθένισης εκφράζει τη μέγιστη επίδραση του φίλτρου στη συχνοτική περιοχή του ηχητικού υλικού όπου εφαρμόζεται. Το ποσοστό ενίσχυσης ή εξασθένισης αναπαρίσταται με θετικές ή αρνητικές τιμές των decibels (dB).

3.4 Η τάξη (Order) του φίλτρου:

Η τάξη του φίλτρου ορίζει την μεταβολή της καμπύλης του φίλτρου για τη συχνοτική περιοχή εφαρμογής του. Ένα 1^{ης} τάξης φίλτρο, εκφράζεται με τη μεταβολή της καμπύλης στο πεδίο των συχνοτήτων κατά 6 dB/oct. Κάθε αύξηση της τιμής της τάξης του φίλτρου προσθέτει ακόμα 6 dB/oct. Έτσι ένα φίλτρο 3^{ης} τάξης θα έχει μία καμπύλη η οποία θα εκφράζεται με την τιμή 18 dB/oct κ.ο.κ.

3.5.1 Παραμετρικά (presence) equalizer:

Το παραμετρικό equalizer ενισχύει ή εξασθενεί μια συγκεκριμένη περιοχή συχνοτήτων του audio φάσματος.



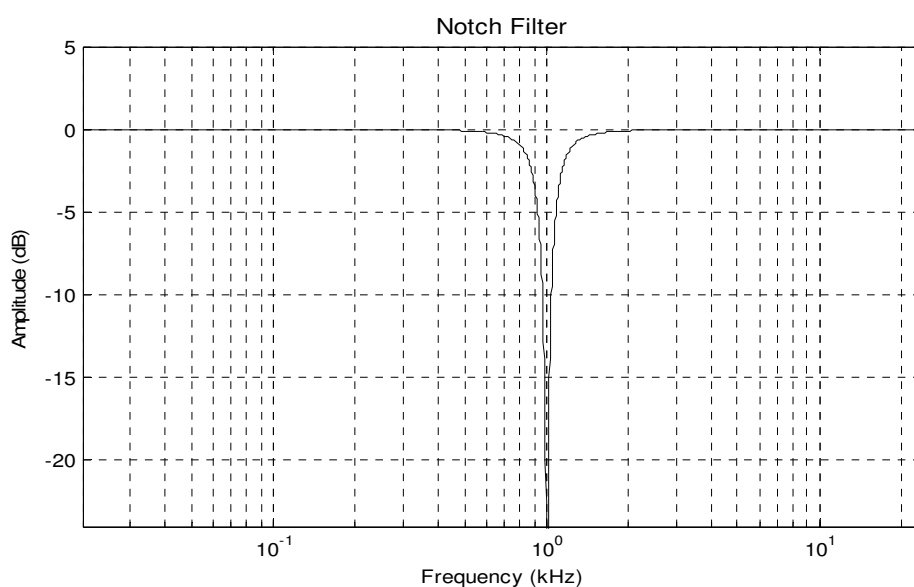
Σχ. 3.5.1.1: Ένα παραμετρικό equalizer με κεντρική συχνότητα (f_c) τα 10 kHz, με φαρδύ εύρος ζώνης και για διάφορες τιμές ενίσχυσης & εξασθένησης.

Υπάρχουν τρεις παράμετροι οι οποίες ερμηνεύουν τη λειτουργία του παραμετρικού equalizer.

- Η κεντρική συχνότητα (f_c): Είναι η συχνότητα, όπως έχουμε δει, που εκφράζει είτε την αριθμητική μέση τιμή των δύο ακραίων συχνοτήτων f_1 & f_2 , είτε το γεωμετρικό μέσο των αυτών τιμών. Ένα παραμετρικό equalizer μπορεί να τοποθετηθεί οπουδήποτε στο φάσμα και, συνεπώς, η κεντρική του συχνότητα μπορεί να είναι οποιαδήποτε από τα 20 Hz έως τα 20 kHz.
- Bandwidth (BW) & Q: Το bandwidth του φίλτρου ορίζεται στα σημεία όπου έχουμε πτώση κατά 3 dB εκατέρωθεν της κεντρικής συχνότητας. Το bandwidth εκφράζει το εύρος σε Hz. Το Q εκφράζει και αυτό το εύρος του φίλτρου αλλά με απόλυτες τιμές.
- Ενίσχυση ή εξασθένηση (Boost or Cut): Η ενίσχυση που προκύπτει από το φίλτρο, αφορά το ποσοστό του κέρδους στην τιμή του πλάτους της κεντρικής συχνότητας εκφραζόμενο σε dB. Η εξασθένηση που προκύπτει από το φίλτρο, αφορά ακριβώς το αντίθετο.

3.5.2 Notch equalizer:

Το notch EQ είναι μία ειδική περίπτωση του παραμετρικού. Του οποίου το κέρδος επί της κεντρικής συχνότητας τίθεται στο μείον άπειρο (θεωρητικά. Πρακτικά μία τιμή της τάξης των -24 dB περίπου, είναι συνήθως υπεραρκετή), δηλ. ουσιαστικά εξαλείφει τη συχνότητα αυτή στο υπό επεξεργασία ηχητικό σήμα.



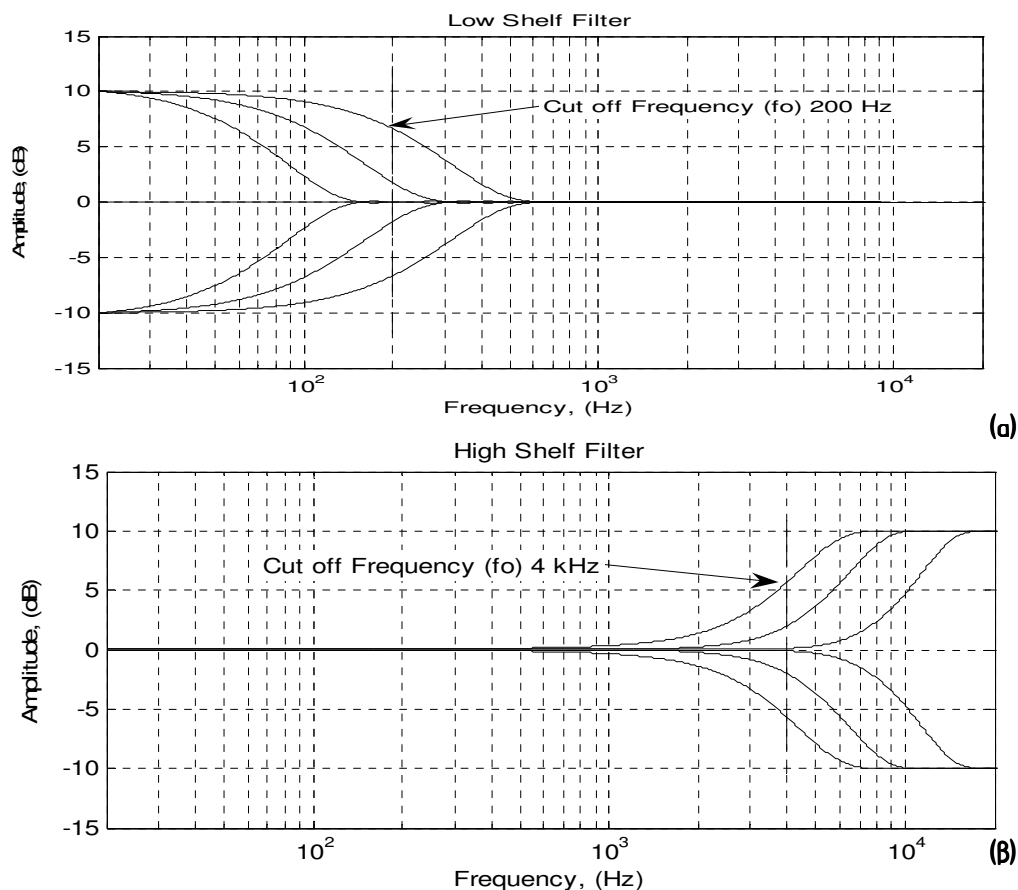
Σχ. 3.5.2.1: Ένα notch equalizer με κεντρική συχνότητα το 1 kHz και τιμή εξασθένησης στην συχνότητα αυτή τα -24 dB.

Χρειάζονται μόνο δύο παράμετροι για να εκφράσουμε ένα notch φίλτρο: η κεντρική συχνότητα και το bandwidth. Το gain του φίλτρου στην κεντρική συχνότητα είναι δεδομένο (θεωρητικά) στο μείον άπειρο, έτσι ώστε η κεντρική συχνότητα να εξαλείφεται.

- Κεντρική συχνότητα (f_c): Η κεντρική συχνότητα στο notch EQ ορίζεται με όμοιο τρόπο όπως και στα παραμετρικά. Το πεδίο εφαρμογής του notch φίλτρου είναι επίσης όμοιο με αυτό του παραμετρικού, δηλ. από τα 20 Hz έως το ήμισυ της συχνότητας Nyquist.
- Ενίσχυση ή εξασθένηση (Boost or Cut): Όμοια με το παραμετρικό equalizer.

3.5.3 High & Low - Shelving φίλτρα:

Τα shelving φίλτρα εφαρμόζουν ένα δεδομένο ποσοστό ενίσχυσης ή εξασθένισης σε όλο το συχνοτικό εύρος πριν ή μετά απο τη συχνότητα αποκοπής του φίλτρου.



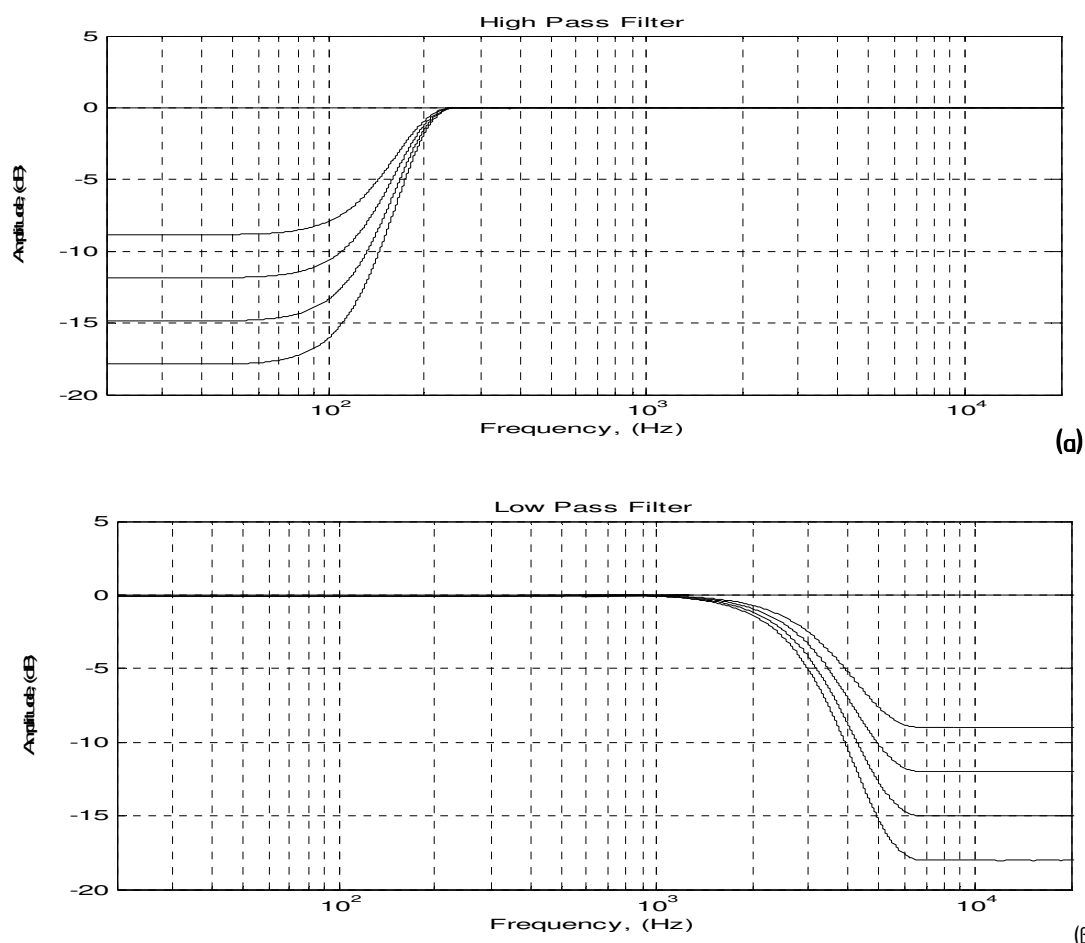
Σχ. 3.5.3.1: (α) Low Shelf φίλτρα για διαφορετικές συχνότητες αποκοπής και gain ± 10 dB. (β) High Shelf φίλτρα για διαφορετικές συχνότητες αποκοπής και gain ± 10 dB.

Τα shelving φίλτρα έχουν δύο μεταβλητές παραμέτρους:

- Τη συχνότητα αποκοπής (f_0): Η συχνότητα αποκοπής στα shelving ορίζεται ως το σημείο εκείνο όπου το σήμα ενισχύεται ή εξασθενεί 3 dB περισσότερο ή λιγότερο. Αν η τιμή ενίσχυσης ή εξασθένισης είναι μικρότερη από 3 dB τότε η f_0 ορίζεται στο σημείο όπου έχουμε το $\frac{1}{2}$ της τιμής αυτής.
- Ενίσχυση ή εξασθένιση (Boost or Cut): Το ποσοστό ενίσχυσης ή εξασθένισης στα shelving φίλτρα εφαρμόζεται είτε πριν (Low Shelf), είτε μετά (High Shelf) απο τη συχνότητα αποκοπής και για ολόκληρο το εύρος του φάσματος που απομένει.

3.5.4 Υψηπερατά & βαθυπερατά φίλτρα (High & Low):

Τα High και τα Low pass φίλτρα επιτρέπουν είτε μόνο στην χαμηλή, είτε μόνο στην υψηλή, αντίστοιχα, περιοχή να περάσει.



Σχ. 3.5.4.1: (α) Ένα υψηπερατό (High Pass) φίλτρο. (β) Ένα βαθυπερατό (Low Pass) φίλτρο.

Τα High & Low Pass φίλτρα ορίζονται και αυτά από δύο παραμέτρους:

- Την συχνότητα αποκοπής (f_0): Ορίζεται όμοια με τον τρόπο που ορίζεται για τα shelving,
- Την τάξη (n), ή αλλιώς η κλίση της καμπύλης, του φίλτρου: Την τάξη του φίλτρου εκφράζει ένας ακέραιος αριθμός με μία τιμή από 1 έως 4. Κατ' ουσία εκφράζει το ρυθμό πτώσης της καμπύλης του φίλτρου. Για $n=1$ έχουμε 6 dB/oct. Κάθε επιπλέον τάξη προσθέτει άλλα 6 dB, δηλ. για $n=2$ έχουμε 12 dB/oct, κ.ο.κ.

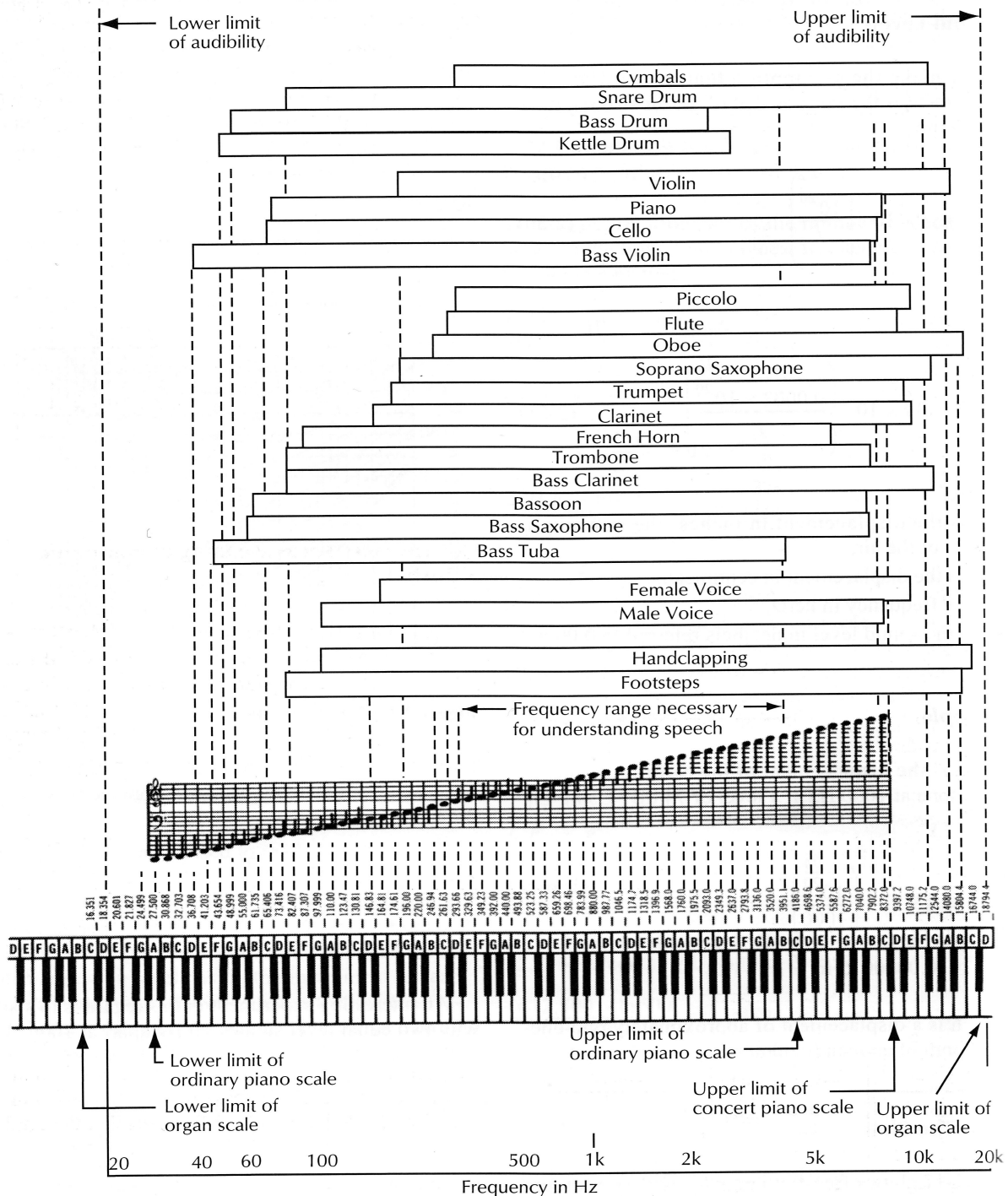
3.6 Βασικά εισαγωγικά σημεία αναφορικά με τη χρήση των equalizer:

Το mastering μπορεί να θεωρηθεί και ως η τέχνη του συμβιβασμού. Είναι το να γνωρίζει κάποιος τι μπορεί να είναι βοηθητικό για το ηχητικό υλικό και κατά συνέπεια να προβαίνει στις κινήσεις αυτές που θα το αναδεικνύουν. Μία από τις βασικές αρχές του mastering είναι πως οποιαδήποτε ενέργεια σε οποιοδήποτε από τα χαρακτηριστικά του ηχητικού υλικού, επηρεάζει ολόκληρο το μουσικό κομμάτι. Η παραπάνω αξιακή συνθήκη σημαίνει πως δε μπορούμε, στο mastering, να αντιγράψουμε τεχνικές που χρησιμοποιούνται σε άλλα συγγενικά, γνωστικά αντικείμενα (όπως π.χ. η ηχοληψία) και να τις εφαρμόσουμε. Η τεχνική του Equalization είναι μια ξεκάθαρα σαφής περίπτωση της οποίας η εφαρμογή είναι εξαιρετικά διαφορετική από εκείνη της ηχοληψίας. Για παράδειγμα, κατά τη διάρκεια του mastering, προσπαθώντας να ρυθμίσουμε το βαθύ χαμηλό μιας στερεοφωνικής μίξης, επηρεάζουμε παράλληλα και το πώς αντιλαμβανόμαστε την υψηλή περιοχή του φάσματος. Ομοίως, στην περίπτωση που το snare drum ακούγεται μουντό, αλλά τα φωνητικά ηχούν άριστα, τότε εννέα στις δέκα φορές τα φωνητικά θα υποφέρουν αν προσπαθήσουμε να διορθώσουμε τον ήχο του snare εφαρμόζοντας equalization.⁴ Τέτοιου είδους προβλήματα απαντώνται ακόμα και μεταξύ μουσικών οργάνων που βρίσκονται στην ίδια συχνотική περιοχή, για παράδειγμα αν εφαρμόσουμε EQ στη μπότα, τότε το μπάσο θα πάθει κάτι περισσότερο από το να επηρεαστεί απλώς, είτε προς το καλύτερο, είτε προς το χειρότερο. Με προσεκτικό, όμως, equalization σε μια στενή μπάντα συχνοτήτων, είναι πολύ πιθανό να περάσουμε «κάτω» από τη συχνотική περιοχή του μπάσου, στα 60 Hz περίπου, και να διορθώσουμε τη θεμέλιο της μπότας. Αν όμως, όπως πολύ συχνά συμβαίνει, το πρόβλημα δεν βρίσκεται στη θεμέλιο συχνότητα, αλλά στο αρμονικό περιεχόμενο (καταχρηστικά εδώ, μιας και τα κρουστά δεν έχουν αρμονικό φάσμα) τότε, προφανώς, θα επηρεάσουμε και το μπάσο καθώς επίσης και το κάτω άκρο των φωνητικών, του πιάνου ή της κιθάρας που τυχόν υπάρχουν στο μουσικό κομμάτι. Πολλές φορές δεν μπορούμε να πούμε πως ένα πρόβλημα μπορεί να επιλυθεί έως ότου προσπαθήσουμε για την επίλυση του. Σε κάθε περίπτωση δεν μπορούμε να υποσχεθούμε θαύματα. Όπως ένας ηχοληψτής δε μπορεί να πει πως θα βγάλει μια πολύ καλή μίξη από μία κακή ηχογράφιση, έτσι και ο τεχνικός mastering δε μπορεί να πει πως θα παράγει ένα καλό master από μία κακή μίξη.

Το πρώτο πράγμα που θα πρέπει να αποφασίσει κανείς όταν εφαρμόζει equalization είναι ποια περιοχή του συχνотικού φάσματος αντιστοιχεί σε ποιο τονικό

⁴ Έχουν επινοηθεί τεχνικές (πέρα από το απλό equalization) που μας επιτρέπουν να απομονώσουμε το ένα όργανο από το άλλο. Στο παράδειγμα μας είναι πιθανό να βελτιώσουμε την αίσθηση και τη διαύγεια του snare, καθώς και άλλων κρουστών, δίκως να επηρεάσουμε την τονικότητα των φωνητικών χρησιμοποιώντας έναν upward expander με το κατάλληλο attack και release. Αυτό συμβαίνει διότι έχουμε τη δυνατότητα να επιλέξουμε επακριβώς τη συχνотική περιοχή στην οποία θα εφαρμόσουμε το δυναμικό επεξεργαστή. Αυτού του τύπου οι τεχνικές (τις οποίες θα εξετάσουμε σε επόμενο κεφάλαιο) μπορούν να οριστούν και ως dynamic equalization.

στοιχείο του κάθε ήχου. Κάτι που μπορεί να βοηθήσει λίγο είναι να ξέρουμε ποιες συχνότητες αντιστοιχούν στις θεμελιώδεις συχνότητες της κάθε μουσικής νότας. Το σχήμα 2.6.1 δείχνει ακριβώς αυτό, σε συνδυασμό με το εύρος διαφόρων κοινών μουσικών οργάνων.



Σχ. 3.6.1: Το συχνοτικό εύρος διαφόρων μουσικών οργάνων, έχοντας ως σημείο αναφοράς το εκκλησιαστικό όργανο, το πιάνο και το εύρος της ανθρώπινης ακοής.

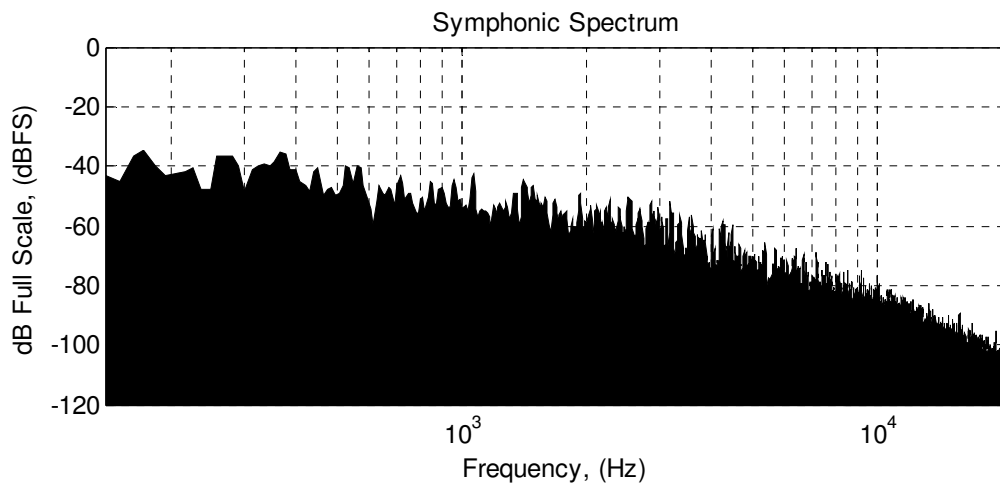
3.7 Τί είναι η τονική ισορροπία του ηχητικού υλικού:

Πιθανότατα ο κύριος λόγος για τον οποίο τα περισσότερα απο τα albums, τα οποία εξέρχονται απο τα studio ηχογραφήσεων και έρχονται στο mastering studio, είναι για να επικυρώσουν και να διασφαλίσουν μια ακριβή τονική ισορροπία του ηχητικού υλικού το οποίο περικλείουν. Το αποτέλεσμα που προκύπτει απο τα περισσότερα mastering studios είναι εξαιρετικά συνεπές στις προϋποθέσεις που τίθενται απο τα πριν και αυτό, εν πολλοίς, οφείλεται στο απόλυτα ακριβές monitoring που η πλειονότητα των mastering studios διαθέτει. Ενόσω είναι πιθανό, πολλές φορές, να βοηθήσουμε στην ανάδειξη ξεχωριστών - μεμονωμένων μουσικών οργάνων, το σημαντικότερο είναι, στην περίπτωση μας, να προσδώσουμε στο τελικό ηχητικό αποτέλεσμα μια συνολική τονική ισορροπία στο σύνολο του μουσικού κομματιού και κατά συνέπεια στο σύνολο του φάσματος που αυτό εμπεριέχει.

Ας ορίσουμε όμως το τι εννοούμε όταν αναφερόμαστε στον όρο «τονική ισορροπία». Ο ορισμός δίδεται με βάση τα τονικά χαρακτηριστικά που κυριαρχούν σε μια συμφωνική ορχήστρα και αυτό γιατί το ανθρώπινο αυτί αντιλαμβάνεται το φάσμα που προκύπτει απο το συγκεκριμένο σύνολο ως πλήρες και μη ενοχλητικό. Αν τροφοδοτήσουμε έναν φασματικό αναλυτή με ένα ηχογραφημένο συμφωνικό κομμάτι θα διαπιστώσουμε ένα σταδιακό roll off στις υψηλές συχνότητες (βλ. Σχ. 3.7.1). Έτσι θα πρέπει να ηχούν και τα περισσότερα είδη της λεγόμενης popular μουσικής.

Τα πάντα ξεκινούν απο τη μεσαία περιοχή του φάσματος. Εάν η μεσαία περιοχή σε μία π.χ. rock ηχογράφηση είναι υποτονική, τότε είναι σα να αφήνουμε τη βιόλα ή τα ξύλινα πνευστά έξω απο το συμφωνικό σύνολο. Οι θεμελιώδεις συχνότητες των φωνητικών, του πιάνου καθώς και των υπόλοιπων οργάνων πρέπει να είναι οι σωστές (απο την ηχογράφηση ακόμα ει δυνατόν), αλλιώς οτιδήποτε άλλο δε θα μπορέσει να σώσει την κατάσταση. Η προσφορά του τεχνικού mastering έγκειται στο να διασφαλίσει την καλή τονική ισορροπία του ηχητικού υλικού στο ακουστό συχνοτικό εύρος, πράγμα που σημαίνει πως δεν υπάρχουν περιοχές οι οποίες είναι περισσότερο ή λιγότερο προεξέχουσες ή το αντίθετο (εξαρτάται βεβαίως και απο το μουσικό είδος). Διασφαλίζει, επίσης, πως ο ήχος είναι ευχάριστος, αναλυτικός και με διαύγεια. Παρόλα αυτά, ορισμένα μουσικά είδη απαιτούν τα κύμβαλα να είναι π.χ. απλωμένα στη στερεοφωνική εικόνα και προς τα πίσω· άλλα, πάλι, απαιτούν την προβολή ορισμένων χαρακτηριστικών τους. Με το κατάλληλο monitoring και την απαιτούμενη εμπειρία μπορεί κανείς να γνωρίζει το πώς θα επιφέρει τα επιθυμητά αποτελέσματα.

Ενώ εμείς αναζητούμε, για τη δική μας διευκόλυνση, ένα απόλυτο standard σε ότι αφορά το τονικό ύψος του μουσικού κομματιού (και το έχουμε βρεί στη συμφωνική ορχήστρα), υπάρχουν πάμπολλες περιπτώσεις, στις οποίες μια ηχογράφηση έχει ηθελμένα αποκτήσει απο τον ηχολήπτη ένα ιδιαίτερο «χρώμα». Για παράδειγμα ένας οξύτερος, φωτεινότερος ήχος θα προκαλέσει το αυτί μας και αυτό θα ανταποκριθεί



Σχ. 3.7.1: Το συχνотικό φάσμα μιας συμφωνικής ορχήστρας. Παρατηρούμε το roll off στην υψηλή περιοχή του φάσματος.

και θα μάθει να τον αποδέχεται αμελώντας το γεγονός του ότι αποκλίνει από το πώς θα έπρεπε κανονικά να ηχεί ο συγκεκριμένος ήχος. Τέτοιου είδους φαινόμενα έχουμε συναντήσει σε αρκετά μουσικά είδη, τα οποία με αυτό τον τρόπο αποκτούν τα ιδιαίτερα χαρακτηριστικά τους. Εάν, όμως, στο συγκεκριμένο παράδειγμα εφαρμόσουμε EQ στις μεσαίες συχνότητες και αυξήσουμε το πλάτος τους, τότε σε σύγκριση με το αρχικό θα φανεί μουντό ίσως. Παρόλα αυτά ο μηχανικός mastering θα πρέπει να μπορεί να αναγνωρίζει και τις ιδιαιτερότητες του μουσικού κομματιού, αλλά και που αυτές, ενδεχομένως, συναντούν τα όρια τους, και αυτό για να μπορεί το ηχητικό υλικό να αναπαραχθεί κατάλληλα στην πλειονότητα των συστημάτων αναπαραγωγής. Το τελευταίο είναι ιδιαίτερα σημαντικό για τεχνικούς κι αισθητικούς λόγους. Χαρακτηριστικό παράδειγμα αποτελούν τα συριστικά σύμφωνα των φωνητικών. Θα πρέπει να μπορούμε να αναγνωρίσουμε πότε αυτά είναι αποδεκτά ως τέτοια και πότε θα πρέπει να ελεγχθούν. Ένα μουσικό κομμάτι με συριστικά, όταν αναπαράγεται από ένα ραδιοφωνικό σταθμό μπορεί να προκαλέσει ολέθρια αποτελέσματα λόγω των limiters που χρησιμοποιούνται για τον έλεγχο των φαινομένων προέμφασης στις υψηλές συχνότητες στη μπάνα των FM. Συνεπώς τα συριστικά μπορούν να προκαλέσουν την ενεργοποίηση του limiter με αποτέλεσμα τη μείωση της ευκρίνειας του μουσικού κομματιού λόγω της υπερβολικής συμπίεσης των υψηλών συχνοτήτων. Σε ακραίες περιπτώσεις μπορούν να χαθούν ακόμα και συλλαβές ή ολόκληρες λέξεις των στοίχων του μουσικού κομματιού (αυτό εξαρτάται από την τιμή της ρύθμισης του release στον limiter). Επιπροσθέτως, οι πέραν του δέοντος «φωτεινές» ηχογραφήσεις μπορεί να ακουστούν μουντές στο ράδιο. Κατανοούμε, συνεπώς, ότι η «λαμπρότητα» καθιστά ευάλωτο το μουσικό υλικό όταν αυτό καλείται να αναπαραχθεί μέσω ενός ραδιοφωνικού σταθμού.

3.8 Τα ξεχωριστά μουσικά είδη:

Ας κρατήσουμε το τονικό φάσμα που προκύπτει απο μια συμφωνική ορχήστρα ως σημείο αναφοράς για μουσικά είδη όπως το rock, την pop, τη jazz καθώς και για την παραδοσιακή μουσική (world & folk music). Τις περισσότερες φορές, αυτού του είδους η προσέγγιση αποδίδει τα μέγιστα κατά τη διάρκεια του equalization στο mastering. Υπάρχουν όμως και ορισμένες περιπτώσεις που η παραπάνω προσέγγιση δε λειτουργεί. Ορισμένα μουσικά είδη, σκόπιμα, απαιτούν ένα αρκετά διαφορετικό τονικό «ισοζύγιο» και για αυτά τα είδη η συμφωνική προσέγγιση δεν είναι η κατάλληλη. Π.χ. για ορισμένα είδη τα έντονα χαμηλά (ή για ορισμένα άλλα το αντίθετο) είναι απλώς ότι πρέπει. Μπορούμε να σκεφτούμε ένα reggae τραγούδι με το συμφωνικό φάσμα, αλλά με πολύ εντονότερα μπάσα. Όπως, επίσης, μπορούμε να σκεφτούμε ένα punk τραγούδι με σκληρό, τραχύ και δυνατό ήχο· τα punk φωνητικά μπορεί να είναι τραχιά και υψηλότερα ενώ, παράλληλα, το μουσικό μέρος να είναι ιδιαίτερα συμπαγές με έμφαση προς το κάτω μέρος του φάσματος. Προφανώς, μία τέτοια προσέγγιση στο equalization για ένα ολόκληρο album μπορεί να το καταστήσει κουραστικό στον ακροατή· μπορεί, όμως, να είναι ενδιαφέρον και μουσικά χρήσιμο όταν αυτό αποτελεί μέρος της αισθητικής προσέγγισης του μουσικού.

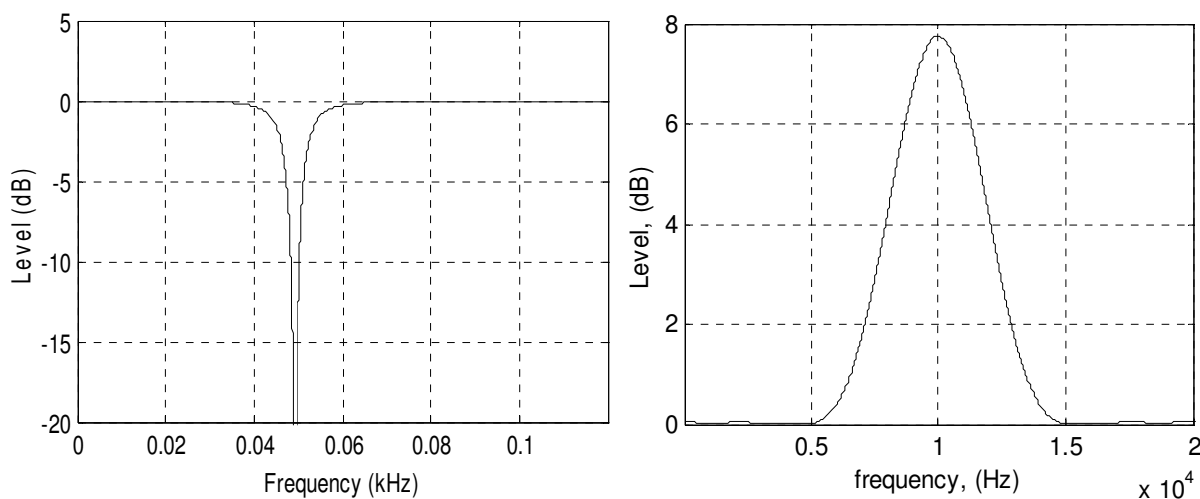
Το equalization (καθώς και άλλου τύπου επεξεργασίες) δεν επηρεάζει απλώς την τονικότητα ενός μουσικού κομματιού, επηρεάζει συνολικά το πώς αντιλαμβανόμαστε το μουσικό υλικό. Δε θα πρέπει ο τεχνικός mastering να επηρεάζει την ισορροπία των μουσικών οργάνων μεταξύ τους, την οποία ο ηχολήπτης έχει καθορίσει στο στάδιο της μίξης, παρά μόνο όταν αυτό κρίνεται απαραίτητο. Για παράδειγμα, αυξάνοντας τις χαμηλές για να πάρουμε έναν πιο ζεστό ήχο επηρεάζουμε αναπόφευκτα το level του μπάσου σε σχέση με το level των φωνητικών. Σε ορισμένες περιπτώσεις (και ανάλογα με το ποσοστό ενίσχυσης των χαμηλών) μπορεί όντως να αποκτήσουμε έναν πιο ζεστό ήχο, σε κάποιες άλλες, όμως, μπορεί το αποτέλεσμα να είναι ένα μουντό και χωρίς νεύρο μουσικό κομμάτι.

Συνεπώς, κατανοούμε πως μπορούμε να έχουμε μια αρχική προσέγγιση για το πώς οφείλει να είναι ένα τονικά ισορροπημένο μουσικό σύνολο, όμως, παρόλα αυτά, δεν υπάρχουν μαγικές συνταγές· τα πάντα εξαρτώνται απο το συχνοτικό περιεχόμενο του μουσικού κομματιού, απο το είδος του, καθώς και απο την προσέγγιση των μουσικών. Βεβαίως υπάρχουν και όρια τα οποία ο mastering ηχολήπτης θα πρέπει να καθορίσει με ακρίβεια.

3.9 Τεχνικές Equalization, παραμετρικά equalizer:

Υπάρχουν δύο βασικοί τύποι equalizer: τα παραμετρικά και τα shelving, τα οποία έχουν πάρει τα ονόματά τους από τη χαρακτηριστική μορφή της καμπύλης τους. Το παραμετρικό EQ χρησιμοποιείται εξαιρετικά από τους ηχοηλίπτες κυρίως κατά τη διάρκεια της μίξης. Το παραμετρικό equalizer επινοήθηκε από τον George Massenburg το 1967⁵, το παραμετρικό EQ μας παρέχει τρεις επιλογές: την κεντρική συχνότητα του φίλτρου, το εύρος του και το ποσοστό που θα ενισχύσει ή θα αποκόψει (boost or cut) για τη συγκεκριμένη περιοχή. Οι ηχοηλίπτες χρησιμοποιούν παραμετρικά EQ για κάθε μουσικό όργανο - κανάλι της κονσόλας ξεχωριστά, είτε ενισχύοντας μια συγκεκριμένη περιοχή για να αναδείξουν τη διαύγεια του μουσικού οργάνου, είτε μειώνοντας το ποσοστό της ενέργειας στη συγκεκριμένη περιοχή συχνοτήτων για να εξαλείψουν ή να μειώσουν τυχόν προβλήματα σε αυτή. Το παραμετρικό equalizer είναι επίσης το πιο δημοφιλές και για το mastering και αυτό συμβαίνει διότι μας δίνει τη δυνατότητα να επεμβαίνουμε στο ηχητικό υλικό με χειρουργική ακρίβεια (π.χ. με τα υπερβολικά resonant που μπορεί να συναντήσει κανείς σε ένα μπάσο). Ένα απλό (μη παραμετρικό) EQ έχει προκαθορισμένη συχνότητα αποκοπής και εύρος ζώνης και η μόνη επιλογή που μας παρέχει είναι η ρύθμιση του gain για κάθε μπάντα συχνοτήτων.

Το Q των equalizer ορίζεται μαθηματικά ως το αποτέλεσμα της κεντρικής συχνότητας του φίλτρου διαιρεμένη από το εύρος ζώνης του φίλτρου -bandwidth- (βλ. σχέση 3.2.1). Το εύρος ζώνης, με τη σειρά του, ορίζεται στα σημεία της καμπύλης του φίλτρου στα οποία μετράμε 3 dB πάνω (ή κάτω) από το μέγιστο (ή το ελάχιστο) σημείο της καμπύλης του φίλτρου (βλ. σχέση 3.2.3). Το παρακάτω σχήμα (Σχ. 3.9.1) μας δείχνει δύο τέτοια παραμετρικά equalizer.



Σχ. 3.9.1: Στο σχήμα απεικονίζονται δύο παραμετρικά EQ όπου στο 1^ο από αυτά αποκόπτεται κατά 20 dB τη συχνότητα των 50 Hz με ένα πολύ στενό εύρος ζώνης της τάξης των 4 Hz, δηλαδή 0.11 οκτάβες και Q-12.5. Το 2^ο ενισχύει τη συχνότητα των 10 kHz κατά 8 dB με ένα εύρος ζώνης της τάξης 4000 Hz, δηλαδή 0.58 οκτάβες και Q-2.5.

⁵ Το 1967 ο νεαρός (μόλις 19 ετών ήταν τότε) George Massenburg ξεκίνησε μια έρευνα για τη δημιουργία ενός ηλεκτρονικού κυκλώματος το οποίο θα είχε τη δυνατότητα να ρυθμίζει ανεξάρτητα το gain, το bandwidth και την κεντρική συχνότητα του equalizer. Στο κύκλωμα που κατασκεύασε έδωσε το όνομα παραμετρικό equalizer.

Όπως βλέπουμε στο σχήμα το πρώτο equalizer (αριστερά) αποκόπει την συχνότητα των 50 Hz κατά 20 db με ένα πολύ στενό $Q=12.5$ που ισοδυναμεί με 0.11 οκτάβες και συχνотικά μεταφράζεται σε ένα εύρος 4 Hz. Το δεύτερο equalizer (δεξιά) ενισχύει την συχνότητα των 10 kHz κατά 8 dB με ένα πιο φαρδύ αυτή τη φορά $Q=2.5$ που ισοδυναμεί με 0.58 οκτάβες και συχνотικά ισούται με ένα εύρος της τάξης των 4000 Hz⁶ το οποίο σημειώνεται στο σχήμα με τη λευκή οριζόντια διακεκομμένη γραμμή.

Η επιλογή για το πόσο στενό ή το πόσο φαρδύ θα είναι το Q εξαρτάται από τη συνθήκη στην οποία καλείται να ανταπεξέλθει. Η εφαρμογή equalizer με μικρή τιμή του Q (μεγάλο bandwidth) επιφέρει έναν περισσότερο φυσικό ήχο⁷ και για αυτό το λόγο χρησιμοποιείται συχνά με τιμές Q που κυμαίνονται από 0.6 έως 0.7 περίπου. Η χρήση equalizer με υψηλές τιμές του Q (μεγαλύτερες του 2, στενό bandwidth) γίνεται στις περιπτώσεις αυτές στις οποίες χρειάζεται ιδιαίτερη ακρίβεια, όπως όταν παρουσιάζονται φαινόμενα τύπου resonance ή όταν παρουσιάζεται κάποιος θόρυβος σε μια στενή συχνотική περιοχή. Τα σύγχρονα EQ μας παρέχουν τη δυνατότητα να εργαστούμε ακόμη και σε μία μόνο νότα ενός μουσικού οργάνου (δίνοντάς μας τη δυνατότητα επιλογής πολύ στενού Q). Επίσης οι υψηλότερες τιμές του Q χρησιμοποιούνται στο mastering όταν θέλουμε να δώσουμε έμφαση σε κάποιο μουσικό όργανο με τη μικρότερη δυνατή επιρροή στις συχνотικές περιοχές των υπολοίπων μουσικών οργάνων. Για παράδειγμα εάν θελήσουμε να ενισχύσουμε την χαμηλή περιοχή μιας αδύναμης μίξης, η οποία είναι πολύ πιθανό να έχει ένα πολύ υποτονικό bass guitar, τότε ενισχύοντας την περιοχή γύρω στους 80 κύκλους μπορεί να ενισχύσουμε το επίπεδο χαμηλών της μίξης, αλλά, προφανώς, θα επηρεάσουμε και τα φωνητικά: σε αυτή την περίπτωση στενεύουμε το εύρος του φίλτρου αυξάνοντας το Q μέχρις ότου το EQ πάψει να επηρεάζει τα φωνητικά. Μία κλασική τεχνική μέσα στο χρόνο, που χρησιμοποιείται για την αντιμετώπιση προβλημάτων συντονισμού των μουσικών οργάνων, είναι να ξεκινά κανείς την εφαρμογή του equalizer ενισχύοντας γενναιόδωρα μια οποιαδήποτε συχνотική περιοχή χρησιμοποιώντας μια μικρή τιμή του Q (φαρδύ bandwidth). Στη συνέχεια κάνουμε ένα sweep στο συχνотικό φάσμα έως ότου εντοπίσουμε το resonance. Αφού το εντοπίσουμε στενεύουμε το Q και αποκόπουμε το level της συγκεκριμένης συχνотότητας όσο χρειάζεται ώστε να πάψει να υφίσταται ο

⁶ Πολλά equalizers περιγράφουν το συχνотικό εύρος με την τιμή του bandwidth, ενώ άλλα το περιγράφουν με την τιμή του Q . Για το λόγο αυτό, στο παράρτημα Α, παρουσιάζονται δύο χρήσιμοι πίνακες για τη μετατροπή από Q σε bandwidth και το αντίστροφο.

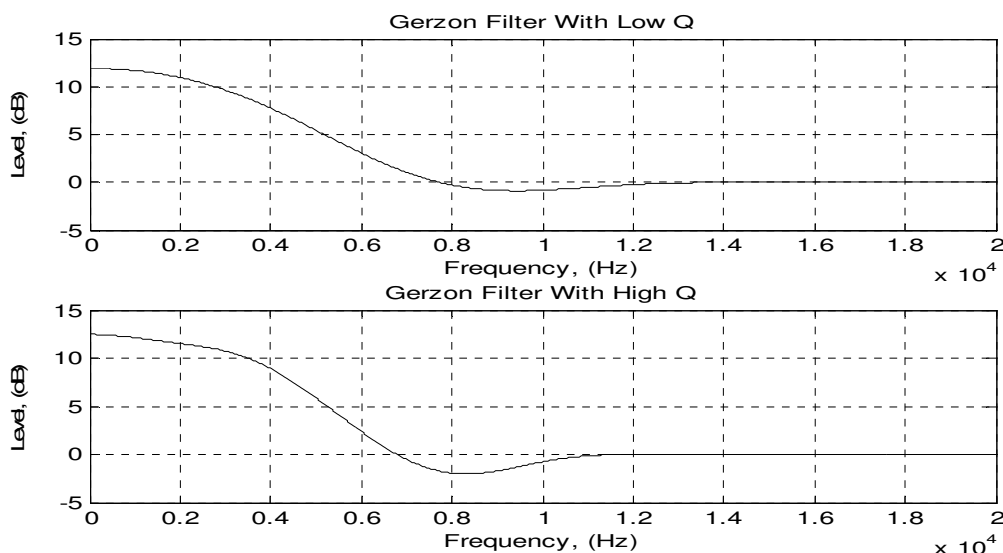
⁷ Η εξήγηση του λόγου για τον οποίο συμβαίνει αυτό είναι ιδιαίτερος απλή. Το Q δεν ορίζει τίποτε άλλο παρά μόνο την τάξη του φίλτρου. Όσο μεγαλύτερη η τάξη του φίλτρου τόσο πιο στενό το εύρος του. Όμως όσο μεγαλύτερη είναι η τάξη του φίλτρου τόσο περισσότερο αλλοιώνεται το σήμα λόγω της αναστροφής φάσης που επιφέρει το EQ στο υπό επεξεργασία σήμα, συνεπώς γίνεται περισσότερο αισθητή η επεξεργασία που έχει τελεστεί και η φυσικότητα του ήχου πάει περίπατο. Στο παράρτημα Β παρουσιάζονται ορισμένα παραδείγματα EQ διαφορετικής τάξης καθώς και η αναστροφή φάσης που επιφέρουν.

συγκεκριμένος συντονισμός. Μία άλλη επίσης συνηθισμένη τεχνική στο mastering με τη χρήση παραμετρικών equalizer είναι η λεγόμενη frequency feathering. Αυτό σημαίνει ότι αντί να εφαρμόζει κανείς μεγάλη αυξομείωση στο EQ σε μία συχνότητα, προσθέτει μικρές αυξομειώσεις σε παραπλήσιες συχνότητες. Για παράδειγμα αντί να προσθέσουμε +3 dB στα 100 Hz, θα προσθέσουμε +1.5 dB στα 100 Hz και +0,5 dB στα 80 και 120 Hz. Αυτό έχει σαν αποτέλεσμα πιο ήπιες και ομαλές αλλαγές στον ήχο.

Υπάρχει, επίσης, πλειάδα παραδειγμάτων που δεν θα είχε νόημα να αναφέρουμε. Το σίγουρο είναι πως, το παραμετρικό equalizer, αποτελεί ένα από τα κυριότερα εργαλεία στα χέρια (και τα αυτιά) του μηχανικού mastering. Απαιτεί εμπειρία καθώς και να μπορεί να αναγνωρίσει κανείς ως που μπορεί να ωθήσει τη συγκεκριμένη πρακτική σύμφωνα με τις ανάγκες του ηχητικού υλικού.

3.10 Τεχνικές Equalization, shelving equalizer:

Τα shelving equalizer επηρεάζουν το level, είτε όλης της συχνοτικής περιοχής των χαμηλών συχνοτήτων, είτε, αντίστοιχα, όλης της περιοχής που καταλαμβάνουν οι υψηλές συχνότητες, θέτοντας ως όριο μία αυστηρά καθορισμένη συχνότητα (συχνότητα αποκοπής του φίλτρου). Για παράδειγμα ένα υπερπερατό shelving φίλτρο στο 1.5 kHz θα επηρεάσει όλες τις συχνότητες από το 1.5 kHz και πάνω. Στο mastering τα shelving EQ παίζουν έναν ιδιαίτερα σημαντικό ρόλο καθώς καλούμαστε να εργαστούμε με ολοκληρωμένες στερεοφωνικές ή πολυκάναλες μίξεις. Μία ενδιαφέρουσα παραλλαγή της κλασικής καμπύλης του φίλτρου μπορεί να βρει κανείς στο plug - in της Waves, Renaissance EQ, καθώς και στη συσκευή Manley Massive Passive Stereo Tube EQ, πολύ χρήσιμα, ομολογουμένως, εργαλεία και τα δύο στο mastering. Η παραλλαγή αυτή φαίνεται στο σχήμα (Σχ. 2.10.1) που παρουσιάζεται στην επόμενη σελίδα. Αποτελεί πρόταση του ψυχοακουστικού ερευνητή Michael Gerzon και θεωρείται μια άκρως «εύηχη» καμπύλη για ένα shelving equalizer. Την πρόταση του αυτή ονόμασε resonant shelf. Αυτή μπορεί να περιγραφεί ως ο συνδυασμός ενός χαμηλοπερατού shelving φίλτρου που ενισχύει το κάτω φάσμα και ενός παραμετρικού equalizer που μειώνει το level με κεντρική συχνότητα, τη συχνότητα αποκοπής του shelving (ο ρόλος των φίλτρων μπορεί να αντιστραφεί για να μειώσουμε το level της συγκεκριμένης συχνοτικής περιοχής). Στην ουσία δηλαδή πρόκειται για ένα shelving EQ το οποίο μας παρέχει, πέρα από τη ρύθμιση της συχνότητας αποκοπής του φίλτρου και τη ρύθμιση της ενίσχυσης ή της ελάττωσης του level του, μία ακόμη ρύθμιση: αυτή του Q του φίλτρου. Και στη συγκεκριμένη περίπτωση το Q ορίζει την τάξη του φίλτρου όπως στα παραμετρικά equalizer. Η τάξη του φίλτρου εδώ, όπως και σε όλα τα shelving EQ, ορίζεται από το ρυθμό πτώσης της καμπύλης

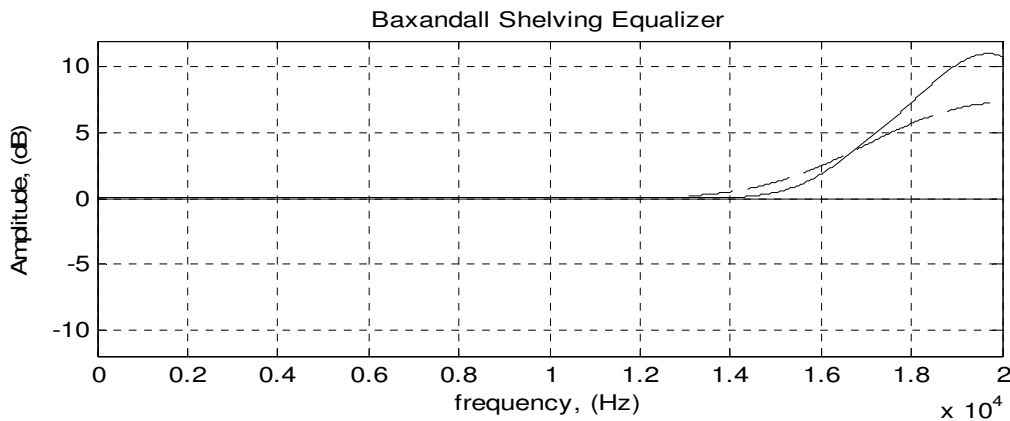


Σχ. 3.10.1: (α) Το λεγόμενο resonant shelf του Gerzon με μία μικρή τιμή του Q. (β) Το ίδιο φίλτρο, αυτή τη φορά, με μία υψηλότερη τιμή του Q. Το βύθισμα που παρουσιάζεται στην καμπύλη είναι το χαρακτηριστικό γνώρισμα του φίλτρου απο το οποίο προέκυψε και το όνομα του (resonant).

του φίλτρου (dB/oct.) που με τη σειρά της ορίζεται απο το Q όπως φαίνεται και στο παραπάνω σχήμα. Στο διάγραμμα (α) παρουσιάζεται ένα χαμηλοπερατό resonant shelving με ένα σχετικά χαμηλό $Q=2.1$ που ενισχύει την περιοχή απο τα 600 Hz και κάτω κατά 12 dB (μέγιστη τιμή). Η τιμή των 12 dB στις περισσότερες περιπτώσεις είναι μία υπερβολική τιμή, όμως στο παραπάνω σχήμα χρησιμοποιήθηκε για λόγους παρουσίασης. Παρόλα αυτά με μία τέτοιου είδους καμπύλη θα μπορούσε να αποφευχθεί η επιρροή των φωνητικών απο την ενίσχυση του κάτω άκρου με αποτέλεσμα να ακούγεται ο τραγουδιστής/τρια λες κι έχει κρύωμα και είναι με βουλομένη μύτη (προσπαθήστε να μιλήσετε κλείνοντας με το ένα σας χέρι τα ρουθούνια σας και θα καταλάβετε τι είναι αυτό που αποφεύγουμε με αυτού του είδους το φίλτρο). Το διάγραμμα (β) παρουσιάζει το ίδιο ακριβώς φίλτρο, μόνο που αυτή τη φορά η τιμή του Q είναι υψηλότερη ($Q=2.83$).

3.11 Τα shelving equalizer τύπου Baxandall και η χρήση τους στο audio mastering:

Γνωρίζουμε πως η συχνотική περιοχή που προσδίδει «λαμπρότητα» στο μουσικό υλικό είναι η περιοχή μεταξύ των 15 kHz και των 20 kHz, δηλαδή η υψηλότερη συχνотική περιοχή που μπορούμε να ακούσουμε, γνωστή και ως air band λόγω του «αέρα» που προσδίδει στον ήχο. Ένα ακριβές και αναλυτικό monitoring μπορεί να υποδηλώσει με σαφήνεια το πότε αυτή η περιοχή χρειάζεται να τονιστεί ή όχι. Το να ενισχύει κανείς την περιοχή αέρα αντενδεικνύεται στις περιπτώσεις αυτές όπου κάτι τέτοιο έχει ως αποτέλεσμα ο ήχος να αποκτά τραχύτητα ή στην περίπτωση όπου π.χ. τα κύμβαλα έρχονται πιο μπροστά και αλλοιώνεται η αίσθηση της κατά μήκος κωροθέτησης των οργάνων (depth) της μίξης.



Σχ. 3.11.1: Με τη συνεχή καμπύλη παρουσιάζεται ένα Baxandall shelving equalizer με μια περισσότερο κάθετη καμπύλη αποκοπής, σε αντίθεση με την διακεκομμένη όπου παρουσιάζει το ίδιο φίλτρο με πιο «ευγενικά» όμως χαρακτηριστικά. (Το διάγραμμα δίδει τον οριζόντιο άξονα των συχνοτήτων σε γραμμική ανάλυση για την καλύτερη παρουσίαση της υψηλής περιοχής).

Για τους παραπάνω λόγους έχει δημιουργηθεί μία τρίτη και ιδιαίτερα σημαντική καμπύλη η οποία είναι εξαιρετικά χρήσιμη στο mastering. Πρόκειται για την καμπύλη Baxandall (βλ. σχ. 3.11.1) που πήρε το όνομα της από τον Peter Baxandall ο οποίος πρώτος την παρουσίασε. Το Baxandall είναι ένα shelving EQ και όπως όλα τα shelving έτσι κι αυτό μπορεί να εφαρμοστεί στα άκρα του συχνοτικού φάσματος και να ενισχύσει ή να αποκόψει το πλάτος των συγκεκριμένων συχνοτικών περιοχών. Μπορούμε να προσομοιάσουμε το Baxandall φίλτρο με ένα παραμετρικό EQ (του οποίου το $Q=1$) όπου το τοποθετούμε στο άνω (ή στο κάτω) άκρο του φάσματος με την κεντρική του συχνότητα στα 20 kHz (ή στα 20 Hz). Όπως γίνεται εύκολα κατανοητό το μισό της καμπύλης που καλύπτει την περιοχή μετά τα 20 kHz (ή πριν τα 20 Hz) θα αγνοηθεί (σαν μια καμπάνα που την κόβουμε στη μέση) και το αποτέλεσμα που προκύπτει είναι μια βαθμιαία αύξηση του πλάτους των υψηλών που ξεκινά από τα 12 kHz και φτάνει στη μέγιστη τιμή ακριβώς στα 20 kHz (το αντίστοιχο για τις χαμηλές συχνότητες παράδειγμα έχει, όπως είναι φυσικό, αντίθετη φορά. Δηλαδή η μέγιστη τιμή συμβαίνει για την συχνότητα των 20 Hz και η ενίσχυση του πλάτους ελαττώνεται βαθμιαία μέχρι τα 125 Hz). Αυτή η μορφή της καμπύλης συχνά ανταποκρίνεται καλύτερα στις ανάγκες του αυτιού και η ενίσχυση της υψηλής (ή της χαμηλής) περιοχής με ένα Baxandall φίλτρο συχνά επιφέρει καλύτερα αποτελέσματα από οποιοδήποτε άλλο shelving equalizer.

Όμως θα πρέπει να είμαστε ιδιαίτερα προσεκτικοί όταν ενισχύουμε την περιοχή «αέρα» και αυτό γιατί με αυτόν τον τρόπο είναι δυνατό να δημιουργήσουμε sparklies⁸. Επίσης η ενίσχυση των υψηλών συχνοτήτων μπορεί με το πρώτο άκουσμα να τις

⁸ Τα λεγόμενα sparklies είναι μια ακόμη ορολογία στην audio αργκό. Θέλει να εκφράσει την παραμόρφωση που συχνά συναντάται στις υψηλές συχνότητες του φάσματος. Τα sparklies ηχούν όπως οι σπινθήρες και έχουν αποτελέσει την αιτία καταστροφής πολλών tweeter.

καθιστά ιδιαίτερα θελκτικές, σε μακροχρόνια όμως ακρόαση μπορεί να καταστήσουν το ηχητικό υλικό κουραστικό.

Ένας ακόμη λόγος για τον οποίο θα πρέπει να είμαστε ιδιαίτερα προσεκτικοί όταν ενισχύουμε τις υψηλές, είναι το γεγονός πως το ανθρώπινο αυτί μπορεί (για λόγους που αφορούν τη ψυχοακουστική κι έχουμε ήδη εξετάσει) να αντιληφθεί την ενίσχυση της περιοχής από τα 15 - 20 kHz ως την εκλέπτυνση της χαμηλομεσαίας συχνοτικής περιοχής, πράγμα που σημαίνει ότι με αυτόν τον τρόπο μεταβάλλεται δραστικά η ισορροπία της μίξης. Οι υψηλές αναδεικνύονται, μαζί τους όμως μεταβάλλεται και ο ήχος από τα κύμβαλα, ο ήχος του ταμπούρου ή ακόμα και τα φωνητικά ενός τενόρου ή μιας σοπράνο τραγουδίστριας. Αυτό συμβαίνει διότι μπορεί η θεμελιώδης συχνότητα ενός οργάνου να είναι χαμηλότερα στο φάσμα, οι αρμονικές του όμως, που του προσδίδουν και το ιδιαίτερο χρώμα του, μπορεί να εκτίνονται έως και την υψηλή περιοχή του φάσματος. Συνεπώς θα πρέπει, σε συμφωνία με τη βασική αξιωματική αρχή του mastering πως οποιαδήποτε ενέργεια σε μία παράμετρο του ηχητικού υλικού επηρεάζει ολόκληρο το ηχητικό υλικό, να δίνουμε ιδιαίτερη προσοχή στο equalization.

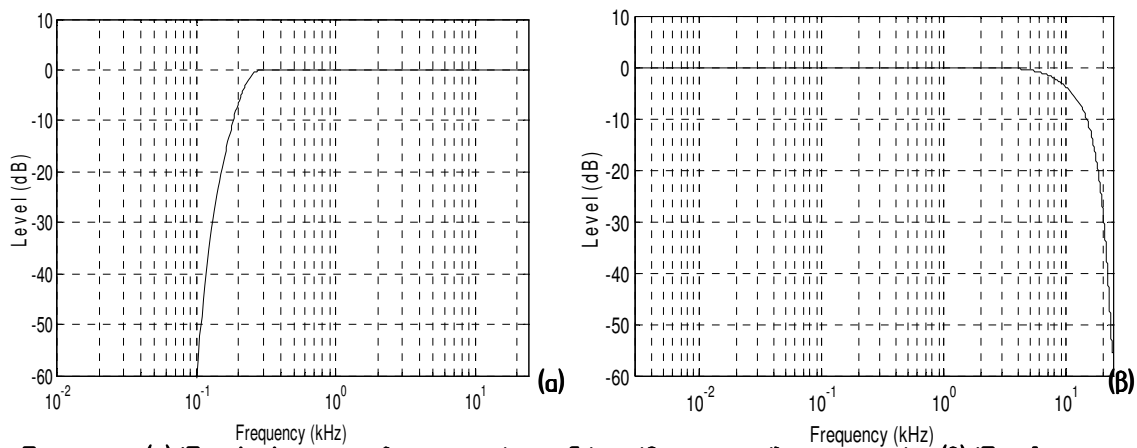
3.12 High - Pass & Low - Pass φίλτρα:

Στην επόμενη σελίδα και στο σχ. 3.12.1 παρουσιάζονται δύο φίλτρα, ένα low pass και ένα high pass. Το high pass επιτρέπει σε όλες τις συχνότητες από τα 300 Hz και πάνω να περάσουν ανεπηρέαστες και ουσιαστικά αποκόπτει όλες τις υπόλοιπες από αυτό το όριο και κάτω. Αυτό συμβαίνει διότι όπως παρατηρούμε έχει μία εξαιρετικά απότομη καμπύλη αποκοπής. Η συχνότητα αποκοπής του φίλτρου είναι τα 290 Hz (-3 dB) και η μείωση του πλάτους, λόγω του συγκεκριμένου φίλτρου, στα 100 Hz (σημ. ο οριζόντιος άξονας των συχνοτήτων είναι βαθμονομημένος σε kHz, συνεπώς ισχύει: $(10^{-1} \cdot 10^3) Hz = 100 Hz$) είναι της τάξης των -60 dB, διαπιστώνουμε, συνεπώς, ότι πρόκειται για ένα ιδιαίτερα δραστικό φίλτρο. Το low pass συμπεριφέρεται λίγο πιο «ευγενικά». Η συχνότητα αποκοπής (-3 dB) ορίζεται στα 10 kHz και μέχρι τη συχνότητα των 20 kHz έχει μειώσει το πλάτος του σήματος κατά -30 dB. Τα φίλτρα αυτά σχεδιάστηκαν στο Matlab, είναι δηλαδή ψηφιακά φίλτρα (FIR)⁹ και για αυτό κατορθώνουν τις συγκεκριμένες τιμές, τις οποίες, ομολογουμένως,

⁹ Ένα ψηφιακό φίλτρο N τάξης, το οποίο εφαρμόζεται σε μια ακολουθία x δειγματοληπτημένων δεδομένων, δημιουργεί μια νέα ακολουθία y που σχετίζεται με τη x σύμφωνα με την εξίσωση:

$$a_1 y_n + a_2 y_{n-1} + \dots + a_{N+1} y_{n-N} = b_1 x_n + b_2 x_{n-1} + \dots + b_{N+1} x_{n-N}$$

Η οποία εξίσωση μπορεί να γραφτεί (είναι η μορφή με την οποία συναντάται συχνότερα) και ως:



Σχ. 3.12.1: (α) Ένα high pass φίλτρο με μία σχεδόν κάθετη καμπύλη αποκοπής. (β) Ένα low pass φίλτρο με πιο «ευγενικά» χαρακτηριστικά.

δεν συναντάμε στα αναλογικά φίλτρα. Παρόλα αυτά, τέτοιου είδους φίλτρα με αυτές ή παρόμοιες τιμές μπορεί να συναντήσει κανείς σε επαγγελματικά mastering audio workstations όπως το Sonic Studio HD και το S.A.D.i.E.

Στο mastering τα high και low pass φίλτρα χρησιμοποιούνται για να επιλύσουν προβλήματα θορύβων τα οποία ίσως υφίσταται το ηχητικό υλικό, άλλα όπως θα δούμε παρακάτω μπορούν να προκαλέσουν με τη σειρά τους άλλα είδη προβλήματος. Αυτού του τύπου τα φίλτρα είναι γενικά δύσκολο να εφαρμοστούν στο mastering, διότι δεν μπορούν να χρησιμοποιηθούν με ακρίβεια στην όποια επιθυμητή συχνотική περιοχή, δεδομένης της φύσης της λειτουργίας τους: δεδομένου δηλαδή ότι επηρεάζουν ολόκληρο το φάσμα πριν ή μετά από μία καθορισμένη συχνότητα. Τα high pass φίλτρα χρησιμοποιούνται για την απαλοιφή θορύβων όπως τα rumbles και thumps.¹⁰ Τα low pass φίλτρα, τώρα, χρησιμοποιούνται ορισμένες φορές για την αντιμετώπιση του hiss¹¹, εφόσον όμως το ανθρώπινο αυτί είναι περισσότερο ευαίσθητο, για αυτό τον τύπο θορύβου, γύρω από την περιοχή των 3 kHz, πολλές φορές είναι καλύτερο να χρησιμοποιούμε ένα παραμετρικό equalizer με στενό εύρος, το οποίο μας παρέχει καλύτερα αποτελέσματα.

$$y_n = \frac{b_1 x_n + b_2 x_{n-1} + \dots + b_{N+1} x_{n-N} - a_1 y_n + a_2 y_{n-1} + \dots + a_{N+1} y_{n-N}}{a_1}$$

Αν οι όροι του φίλτρου: a_2, \dots, a_N είναι όλοι ίσοι με το μηδέν τότε το φίλτρο ονομάζεται FIR (finite impulse response). Σε κάθε άλλη περίπτωση ονομάζεται IIR (infinite impulse response).

¹⁰ - Rumble: είναι ένας χαμηλόσυχνος τύπος θορύβου. Μπορεί να προκύψει από τα μηχανικά μέρη συστημάτων όπως τα πικάπ και τα αναλογικά κασετόφωνα: το τονικό του ύψος δεν είναι σταθερό, έχουμε δηλαδή διαμόρφωση του ηχητικού σήματος.

- Thumps: τα thumps είναι μη μουσικοί χαμηλόσυχνοί ήχοι. Μπορούν να προκύψουν από μεγάλα σε διάρκεια click ή ακόμα κι από τυχόν ασυνέχειες στη ροή του σήματος κατά τη διάρκεια της εγγραφής του στο μέσο αναπαραγωγής.

¹¹ - Hiss: θόρυβος, υψηλής συχνότητας. Μπορεί να ξεκινά από την περιοχή των 2 kHz και να φτάνει μέχρι τα 20 kHz.

3.13 Εφαρμογές equalizer:

Τις περισσότερες φορές το να κάνει κανείς τις ίδιες ρυθμίσεις στο EQ και για τα δύο κανάλια (Left & Right) επιφέρει τα καλύτερα αποτελέσματα μιας και δε διαταράσσεται με αυτό τον τρόπο η στερεοφωνική εικόνα και η σχετική φάση μεταξύ των δύο καναλιών. Όμως, μερικές φορές, είναι απαραίτητο να επεξεργαστούμε με την τεχνική του EQ μόνο το ένα από τα δύο κανάλια. Για παράδειγμα, με ένα υπερτονισμένο hi-hat στα δεξιά, τα φωνητικά στο κέντρο όπως ακριβώς πρέπει να είναι και ένα αξιοπρεπέστατο crash στα αριστερά, η ενδεδειγμένη λύση για την τονική ισορρόπηση του ηχητικού υλικού είναι να εφαρμόσουμε equalizer στην υψηλή περιοχή του δεξιού καναλιού.

Ορισμένες φορές κύρια μουσικά όργανα της μίξης χρειάζονται βοήθεια από τον mastering ηχολόγητη. Η καλύτερη προσέγγιση, στο mastering, για τη διόρθωση τονικών σφαλμάτων στα διάφορα μουσικά όργανα είναι να ξεκινάμε αρχικά δίδοντας ευγενικές τιμές στο EQ. Δηλαδή μικρές τιμές ενίσχυσης ή εξασθένησης, μικρές τιμές του Q, κοκ. Δραστικότερη (αλλά και πάλι μέσα σε κάποια όρια) χρήση των equalizer μπορεί να γίνει μόνο στις περιπτώσεις αυτές που η παραπάνω προσέγγιση δε λειτουργεί.

Για παράδειγμα, εάν το σόλο ενός πιάνου σε ένα jazz κομμάτι θεωρούμε πως είναι αδύναμο, τότε προσπαθούμε να διορθώσουμε το γεγονός αυτό κάνοντας μία ή περισσότερες από τις παρακάτω κινήσεις:

- Εφαρμόζουμε EQ μόνο στη διάρκεια του κομματιού την οποία καταλαμβάνει το σόλο του πιάνου.
- Εφαρμόζουμε EQ μόνο στο κανάλι το οποίο περιέχει την περισσότερη ηχητική πληροφορία (χωροθέτηση του μουσικού οργάνου, L/R), στην περίπτωση που κάτι τέτοιο επιφέρει καλύτερα αποτελέσματα.
- Εφαρμόζουμε EQ μόνο στη συχνοτική αυτή περιοχή που μας απασχολεί για το συγκεκριμένο μουσικό όργανο. Είτε στις θεμελιώδεις συχνότητες, είτε στις αρμονικές, είτε και στις δύο αυτές συνιστώσες.

Μόνο στην περίπτωση που οι παραπάνω κινήσεις δε μας ικανοποιούν με το αποτέλεσμα που προσφέρουν καταφεύγουμε στην τελευταία και έσχατη λύση που μας απομένει. Αυτή είναι να αυξήσουμε το level ολόκληρης της περιοχής που μας ενδιαφέρει. Αυτό μπορεί να γίνει με δύο τρόπους, ο πρώτος αφορά το πεδίο των συχνοτήτων και την εφαρμογή equalizer στη συχνοτική περιοχή που μας αφορά. Η διαφορά του με τα προηγούμενα έγκειται στο γεγονός ότι εφαρμόζεται και στα δύο κανάλια αλλά και για ολόκληρη τη συχνοτική περιοχή. Ο δεύτερος αφορά το πεδίο του χρόνου και είναι ο λιγότερο ενδεδειγμένος τρόπος. Εφαρμόζοντας normalizing για τη χρονική διάρκεια του κομματιού που μας απασχολεί και ενισχύοντας τη χρονική αυτή περιοχή. Το μειονέκτημα και των δύο αυτών τελευταίων περιπτώσεων που

αναφέρθηκαν είναι ότι πολύ εύκολα μπορούν να γίνουν αντιληπτά ακόμη και απο ένα μέσο αυτί διότι επηρεάζουν καθοριστικά το τονικό ισοζύγιο του μουσικού κομματιού. Ως εκ τούτου χρησιμοποιούνται μόνο στις περιπτώσεις αυτές όπου τα μειονεκτήματα που προσδίδουν υποσκιάζονται απο τα πλεονεκτήματα τους.

3.13.1 Διαπιστώνοντας τα όρια της ηχογράφησης:

Υπάρχει, σε πολλούς, η εσφαλμένη αντίληψη πως στο mastering μπορούν να διορθωθούν σχεδόν τα πάντα. Είναι αλήθεια ότι μπορούν να διορθωθούν αρκετά πράγματα, όμως όχι τόσα, όσα θα περίμενε κανείς και σίγουρα όχι τα πάντα.

Λίγα πράγματα μπορούμε να κάνουμε, για παράδειγμα, σε ένα μουσικό κομμάτι του οποίου ορισμένες συχνотικές περιοχές είναι τονισμένες υπέρ του δέοντος και κάποιες άλλες όχι. Αν π.χ. ελαττώσουμε τη χαμηλή περιοχή (δημιουργώντας ένα roll off σε αυτή) για να δημιουργήσουμε ένα πιο σφικτό μπάσο, είναι βέβαιο ότι θα χάσουμε την ατάκα της μπότας. Ή αν δώσουμε μία επιπλέον λάμψη στην υψηλή περιοχή των φωνητικών, είναι επίσης βέβαιο ότι με αυτό τον τρόπο θα καταστήσουμε κουραστικό το ταμπούρο ή άλλα όργανα. Σε αυτές τις περιπτώσεις και υπό την προϋπόθεση πως δύναται να συμβεί κάτι τέτοιο, είναι καλύτερο να ζητάμε απο το studio ηχογράφησης και τον ηχολήπτη το επαναμιξάρισμα του ηχητικού υλικού. Αν κάτι τέτοιο δεν είναι δυνατό, τότε εφαρμόζουμε ορισμένες ιδιαίτερα εξειδικευμένες τεχνικές, όπως είναι το M/S equalization και τη χρήση multiband δυναμικών επεξεργασιών (θα αναφερθούμε σε αυτά σε επόμενο κεφάλαιο), για να φέρουμε στην επιφάνεια μουσικά όργανα τα όποια έχουν θαφτεί στη μίξη ή για να «κρύψουμε» κάποια άλλα. Αυτού του είδους οι τεχνικές μπορούν να επιφέρουν εξαιρετικά αποτελέσματα, ορισμένες φορές μη διακρινόμενα ακόμα και απο ένα νέο βελτιωμένο remix· αυτό όμως είναι συνήθως η εξαίρεση και όχι ο κανόνας.

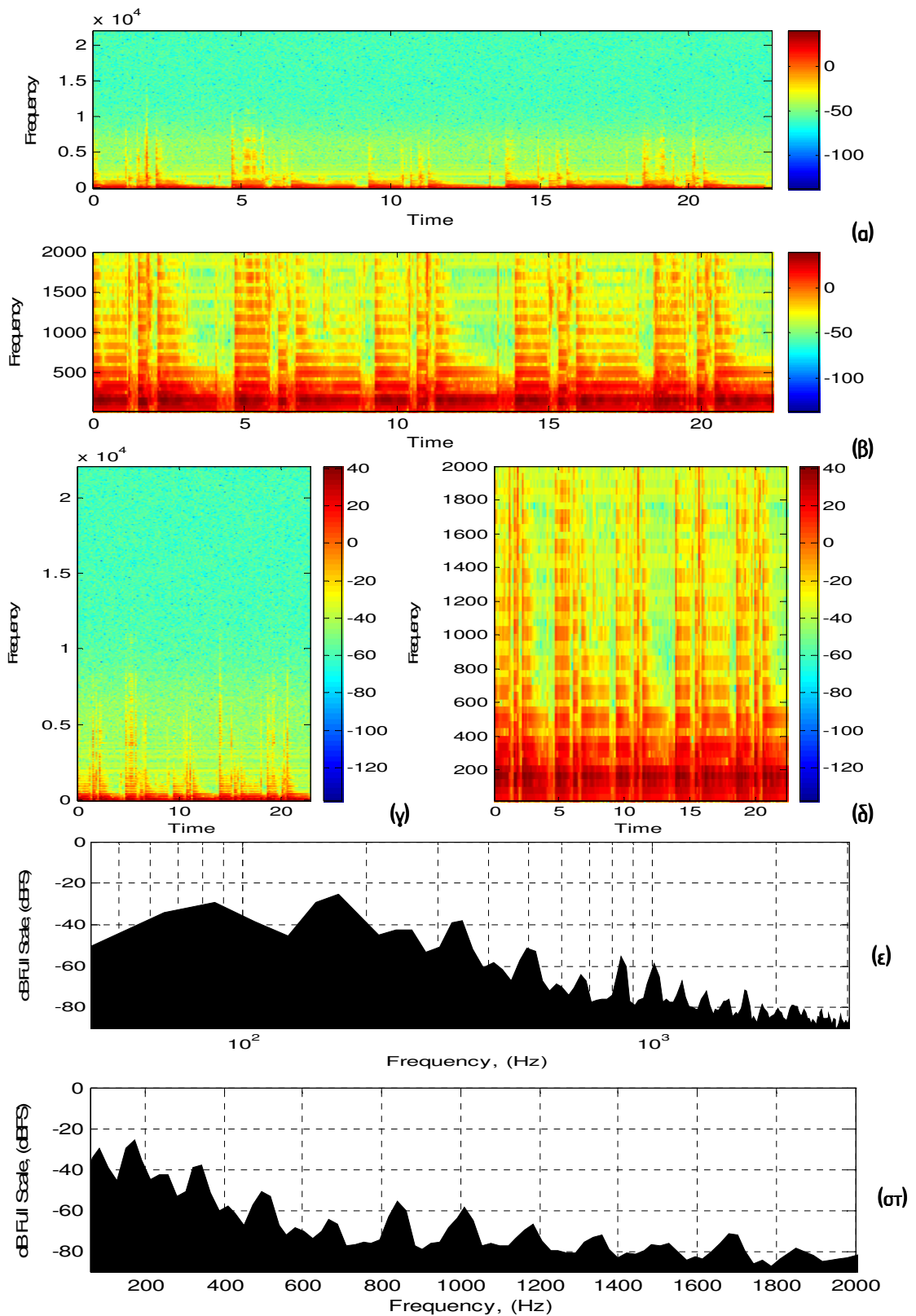
Στο mastering ισχύει γενικότερα πως όσο καλύτερη είναι μια μίξη, ευθέως ανάλογα θα είναι καλύτερο και το master που θα προκύψει απο αυτή. Πράγμα που σημαίνει πως μία πολύ καλή μίξη μπορεί να μη χρειασθεί καθόλου mastering!

3.13.2 Equalization στη θεμελιώδη ή στις αρμονικές;

Η πολύ υψηλή συχνотική περιοχή εμπεριέχει περισσότερο τις αρμονικές των μουσικών οργάνων. Με έκπληξη μπορεί κάποιος να διαπιστώσει πως η θεμελιώδης συχνότητα ενός crash μπορεί να είναι στο 1.5 kHz ή και ακόμη χαμηλότερα στο φάσμα. Όταν εφαρμόζουμε την επεξεργασία του equalization σε ένα μπάσο μουσικό όργανο είναι εύκολο να ξεγελαστούμε και να μπερδέψουμε τη θεμελιώδη συχνότητα με τη δεύτερη αρμονική. Τα φασματογραφήματα καθώς και τα διαγράμματα συχνότητας πλάτους που δίδονται στη σελ. 25 παρουσιάζουν όλα τους το ίδιο πράγμα· την ηχογράφηση ενός ηλεκτρικού μπάσου για 22 δευτερόλεπτα. Στα φασματογραφήματα ο

οριζόντιος άξονας είναι ο άξονας του χρόνου και ο κάθετος είναι των συχνοτήτων. Τα υψηλά πλάτη δίδονται με το κόκκινο χρώμα και όσο αυτά μειώνονται με πορτοκαλί, κίτρινο, πράσινο και τέλος μπλε. Τα φασματογραφήματα καλό είναι να διαβάζονται όπως μια παρτιτούρα. Ο οριζόντιος άξονας είναι ο άξονας του χρόνου και κινείται απο τα αριστερά προς τα δεξιά. Το τονικό ύψος του «φθόγγου» κινείται απο τα κάτω προς τα πάνω με αύξουσα σειρά.

Παρατηρούμε ότι η θεμελιώδης συχνότητα του μπάσου είναι περίπου στα 87 Hz, δηλ. στη νότα F2, η δεύτερη αρμονική, η οποία σημειωτέον έχει μεγαλύτερο πλάτος απο τη θεμελιώδη, είναι περίπου στα 174 Hz, η τρίτη αρμονική είναι περίπου στα 260 Hz με πλάτος μικρότερο της δεύτερης, τέλος η τέταρτη αρμονική είναι περίπου στα 350 Hz και έχει πλάτος μεγαλύτερο απο αυτό της τρίτης και μικρότερο απο αυτό της δεύτερης. Και πάει λέγοντας. Προκύπτει επομένως το ερώτημα του που να εφαρμόσουμε EQ, στη θεμελιώδη ή στις αρμονικές; Είναι εύκολο να ξεγελαστεί κανείς απο τις οκταβικές αναλογίες και τα μεταξύ τους πλάτη, συνεπώς η απόφαση που θα λάβουμε προκύπτει απο το αυτί και απο το πώς το συγκεκριμένο μουσικό όργανο ηχεί καλύτερα: ορισμένες φορές εφαρμόζουμε EQ στη θεμέλιο, ορισμένες στις αρμονικές τις και ορισμένες άλλες και στα δύο. Ένας τρόπος για την εφαρμογή του equalization σε αυτές τις περιπτώσεις είναι να μετακινούμε την κεντρική συχνότητα του EQ απο τη θεμέλιο στις αρμονικές και να λαμβάνουμε έπειτα τις αποφάσεις μας. Όμως, ο τρόπος αυτός, στο mastering, δεν είναι και ο πλέον ενδεδειγμένος διότι η εμπλοκή μας στο πεδίο των συχνοτήτων δεν εξαντλείτε σε ένα και μόνο μουσικό όργανο, αλλά σε μία ολόκληρη μίξη που, προφανώς, εμπεριέχει ένα πλήθος μουσικών οργάνων με διαφορετικές συχνοτικές περιοχές που ορισμένες φορές, όμως, συμπίπτουν μεταξύ τους. Συνακόλουθα, λοιπόν, η επιλογή ενίσχυσης των αρμονικών του ηλεκτρικού μπάσου, που όπως βλέπουμε στα παρακάτω φασματογραφήματα και διαγράμματα εκτείνονται έως και τα 2 kHz, επηρεάζουν κάλλιστα τα φωνητικά, τα τύμπανα στο σύνολο τους, τις κιθάρες, το πιάνο, κ.ο.κ. Σε αυτές τις περιπτώσεις επιλέγουμε τη συχνοτική αυτή περιοχή που επηρεάζει λιγότερο όλα τα υπόλοιπα μουσικά όργανα, παρά τη συχνοτική περιοχή που θα λειτουργούσε ιδανικά για το συγκεκριμένο μουσικό όργανο, αλλά θα επηρέαζε και όλα τα άλλα.



Σχ. 3.13.2.1: (α) Φασματογράφημα του ηλ. μπάσου για 22 sec. (β) Το φασματογράφημα του ίδιου ήκου επικεντρωμένο, όμως στην περιοχή απο τα 40 Hz έως τα 2 kHz. (γ) Όμοια με το (α) με μεγεθυμένο τον κάθετο (των συχνοτήτων) άξονα και σμικρμένο τον οριζόντιο (του χρόνου) άξονα. (δ) Όμοια με το (β) με μεγεθυμένο τον κάθετο (των συχνοτήτων) άξονα και σμικρμένο τον οριζόντιο (του χρόνου) άξονα. (ε) Διάγραμμα του πλάτους ανά συχνότητα του ίδιου ήκου σε λογαριθμική ανάλυση (οριζόντιου άξονα). (στ) Όμοια με το (ε), αυτή τη φορά σε γραμμική ανάλυση (οριζόντιου άξονα).

Compression & Audio Mastering

4.1 Δυναμικό εύρος:

Το δυναμικό εύρος ή αλλιώς η δυναμική διακύμανση (Dynamic range) ορίζεται ως η σχέση μεταξύ των δυνατώτερων και των σιγανότερων μερών του μουσικού κομματιού. Για αυτό δεν θα έπρεπε να συγχέεται με την ένταση ή με μία απόλυτη τιμή. Ο όρος δυναμική διακύμανση αφορά μόνο διαφορές. Στην popular μουσική αυτό τυπικά είναι μεταξύ 6 και 10 db, αλλά για κάποια είδη μουσικής μπορεί να είναι μικρό ως 1 db ή μεγάλο ως 15 db (πολύ σπάνια). Στην τυπική pop μουσική, απαλά μέρη 8 ως 15 db κάτω του υψηλότερου level είναι αποτελεσματικά μόνο για σύντομες περιόδους, αλλά στην κλασσική, την jazz και πολλά άλλα μουσικά είδη τα απαλά μέρη μπορούν να διαρκούν αρκετά λεπτά.

4.1.1 Μικροδυναμικές και Μακροδυναμικές:

Η τέχνη της διαχείρισης των δυναμικών μπορεί να χωριστεί στις Μικροδυναμικές και στις Μακροδυναμικές. Αποκαλούμε την ρυθμική έκφραση, την ακεραιότητα ή την ζωντάνια της μουσικής με τον όρο μικροδυναμικές. Αποκαλούμε μακροδυναμικές τις διαφορές έντασης μεταξύ των διαφόρων μερών ενός τραγουδιού ή ενός κύκλου τραγουδιών. Συνήθως οι δυναμικοί επεξεργαστές (όπως compressors, expanders) είναι καλύτεροι για μικροδυναμικές επεξεργασίες και η χειροκίνητη αλλαγή του level (manual gain riding) είναι καλύτερη για μακροδυναμικού τύπου επεξεργασίες. Οι μικροδυναμικές και μακροδυναμικές επεξεργασίες αυτού του τύπου συνήθως εφαρμόζονται παράλληλα και πολλές καλές συνθέσεις ενσωματώνουν τόσο μικροδυναμικές αλλαγές (πχ χτυπήματα κρουστών ή στιγμιαίες αλλαγές) όσο και μακροδυναμικές (π.χ. crescendos - decrescendos). Αν φανταστείτε ένα μουσικό άλμπουμ σαν ένα πλήρες γεύμα τότε η πρόοδος από τη σούπα στα ορεκτικά και απο εκεί στο κυρίως πιάτο και στο γλυκό, είναι η μακροδυναμική. Ο αντίκτυπος των μπαχαρικών σε κάθε μπουκιά είναι η μικροδυναμική.

4.1.2 Οι Δυναμικές στη Μουσική Ιστορία:

Οι αλλαγές στις δυναμικές ενός μουσικού κομματιού, γίνονται πολύ σημαντικές στη δυτική μουσική κάποια στιγμή μεταξύ των Μεσαιωνικών Γρηγοριανών ασμάτων και της κλασσικής περιόδου, όταν ο συνθέτης F.J.Haydn¹² μας ξαφνιάζει με το πρώτο ίσως δείγμα ταυτόχρονων μικροδυναμικών και μακροδυναμικών αλλαγών. Από την αρχαία εποχή πολλά «μη δυτικά» στυλ, όπως Αφρική, Αφρό-Καραϊβική, Ανατολή, Ινδικό, Μπαλί και άλλα ανατολίτικα είδη μουσικής έδωσαν έμφαση στο ρυθμό (μικροδυναμική κυρίως ως κρουστά) όσο και στη μελωδία, και στον 20^ο αιώνα της

¹² Surprise Symphony. No 94 in G, 1791.

ενσωμάτωσης ο δυνατός ρυθμός των κρουστών έγινε εξαιρετικά σημαντικός και για τις μορφές της δυτικής μουσικής.

4.1.3 Μείωση του Δυναμικού Εύρους:

Οι δυναμικές ενός τραγουδιού ή ενός κύκλου τραγουδιών είναι σημαντικές για μουσικούς και συνθέτες. Ως τεχνικοί (engineers) η αναφορά μας ως προς την ποιότητα του ήχου θα έπρεπε να είναι η ποιότητα ήχου μιας καλοστημένης ζωντανής παρουσίασης - συναυλίας. Θα έπρεπε να είμαστε σε θέση να πούμε εξ ακοής αν μια ηχογράφηση θα ωφεληθεί ή θα ζημιωθεί από την τροποποίηση των δυναμικών της. Πολλές ηχογραφήσεις έχουν ήδη υποστεί διάφορα στάδια επεξεργασίας του δυναμικού εύρους τους και μια χωρίς διακρίσεις ή περαιτέρω μείωση της δυναμικής μπορεί εύκολα να παρασύρει προς τα κάτω την καθαρότητα και την ποιότητα τους. Όμως συνήθως, τα μέσα καταγραφής και το προσχεδιασμένο ακουστικό περιβάλλον ακρόασης απλά δεν μπορούν να συμβαδίσουν με το πλήρες δυναμική εύρος της πραγματικής ζωής. Έτσι ο τεχνικός του mastering καλείται συχνά να ενισχύσει το level στα απαλά μέρη και/ ή να μειώσει τα δυνατά μέρη, μια μορφή χειροκίνητης συμπίεσης (manual compression). Μπορεί να μειώσουμε το δυναμικό εύρος όταν η αρχική διακύμανση είναι πολύ μεγάλη για το μέσο οικιακό περιβάλλον ή να βοηθήσουμε να γίνει η μίξη του ήχου πιο συναρπαστική, παχύτερη και πιο συμπαγής· καθώς και να αναδείξουμε εσωτερικές λεπτομέρειες ή να ισορροπήσουμε τις αλλαγές δυναμικής μέσα σε ένα τραγούδι αν ακούγονται υπερβολικές.

Η εμπειρία μας λέει τότε κάποιο μέρος ενός μουσικού κομματιού έχει πολύ χαμηλό level. Το μουσικό περιεχόμενο του μέρους αυτού καθώς και η χρονική στιγμή μέσα που παρουσιάζεται στο κομμάτι καθορίζει, επίσης, το αν πρέπει να αυξηθεί το level του ή όχι. Για παράδειγμα, μια απαλή εισαγωγή αμέσως μετά από ένα δυνατό τραγούδι μπορεί να πρέπει να αυξηθεί, αλλά ένα παρόμοιο απαλό μέρος στη μέση ενός κομματιού μπορεί να είναι μια χαρά. Αυτό συμβαίνει επειδή τα αυτιά αυτό-ρυθμίζουν την ευαισθησία τους μετά από κάποιο χρονικό διάστημα και μπορεί να μην είναι προετοιμασμένα για ένα άμεσο σιγανό level μετά από ένα πολύ δυνατό ή το αντίθετο. Γι αυτό και οι ενδείξεις των μετρητικών οργάνων είναι αδιάφορες σε αυτές τις περιπτώσεις.

4.1.4 Αύξηση του Δυναμικού Εύρους:

Μπορεί, επίσης, ένα τραγούδι να ακούγεται πιο συναρπαστικό αυξάνοντας την ένταση ενός peak, καθώς ένα μεγάλο μέρος του αντίκτυπου ενός τραγουδιού προέρχεται από τις δυναμικές μεταβολές του. Η σημασία έγκειται στο να αναγνωρίζουμε τότε η επεξεργασία του ηχητικού υλικού γίνεται ελάττωμα - το

ενδιαφέρον για την μουσική μπορεί να εμπλουτιστεί από την ποικιλία, αλλά υπερβολική ποικιλία είναι εξίσου κακή με την υπερβολική ομοιομορφία. Για να παρθούν αυτές οι αποφάσεις απαιτείται μουσικό γούστο, εμπειρία και ένα καλό σύστημα monitor. Η αύξηση του δυναμικού εύρους είναι γνωστή ως expansion. Μια άλλη αφορμή για την εφαρμογή του expansion είναι η αποκατάσταση ή η προσπάθεια αποκατάστασης της αίσθησης των δυναμικών που έχει χαθεί εξαιτίας πολλαπλών γενεών συμπίεσης η κορεσμού της ταινίας, στην περίπτωση αυτή ενισχύουμε τις μεταβολές του δυναμικού εύρους της ηχογράφησης.

4.1.5 Τύποι Τροποποίησης του Δυναμικού Εύρους:

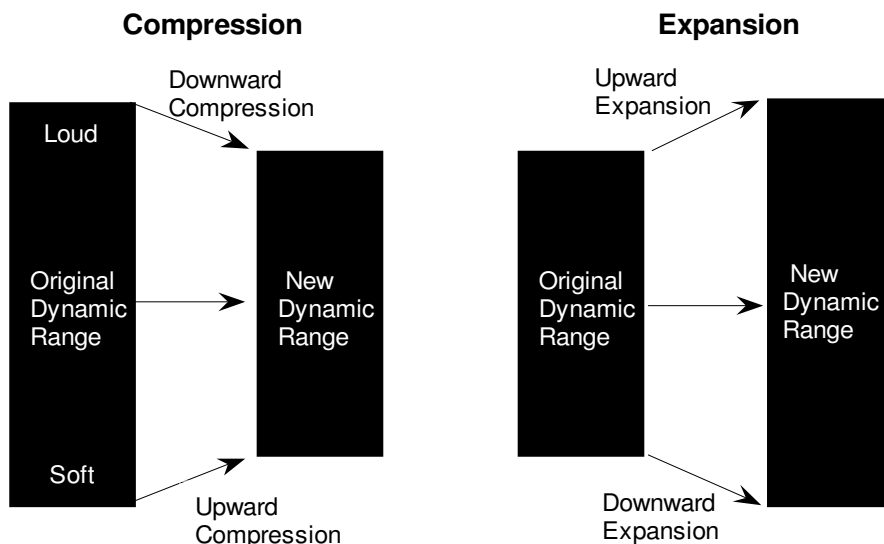
Χρησιμοποιούμε πάντα τον όρο compression για να περιγράψουμε τη μείωση της δυναμικού εύρους και expansion για να περιγράψουμε την αύξηση του. Υπάρχουν δύο τύποι επεξεργασίας τόσο για το compression, όσο και για το expansion, το ανοδικό compression, το καθοδικό compression, το ανοδικό expansion και το καθοδικό expansion, όπως φαίνεται και στο σχήμα 4.5.1 της επόμενης σελίδα.

i. Το downward compression είναι η πιο δημοφιλής μορφή τροποποίησης δυναμικών, λαμβάνει μέρη του ηχητικού υλικού με υψηλό level και τα συμπιέζει. Το limiting είναι μια ειδική περίπτωση συμπίεσης του ηχητικού υλικού με πολύ υψηλό ratio (εκτενέστερη παρουσίαση σε επόμενη παράγραφο). Για ευκρίνεια σε αυτή την εργασία θα χρησιμοποιούμε τον σύντομο όρο compressor για τον downward compressor εκτός αν χρειάζεται να τον διακρίνουμε από τον upward compressor.

ii. Το upward compression παίρνει τμήματα χαμηλού level και ενισχύει την έντασή τους. Ορισμένα παραδείγματα τέτοιου είδους επεξεργασίας είναι το κωδικοποιημένο μέρος ενός Dolby digital ή άλλου συστήματος αποθρομβοποίησης και το A.G.C¹³ που χρησιμοποιούν οι ραδιοφωνικοί σταθμοί για να κάνουν τα ιδιαίτερα χαμηλά ηχητικά μέρη του τηλεοπτικού προγράμματος δυνατότερα.

iii. Το upward expansion λαμβάνει μέρη του ηχητικού υλικού με υψηλό level και τα ενισχύει ακόμα περισσότερο. Οι upward expanders είναι πολύ σπάνιοι και πολύτιμοι, γιατί σε ικανά χέρια μπορούν να εμπλουτίσουν τις δυναμικές, να αποκαταστήσουν την μουσική διαύγεια ή να αποκαταστήσουν χαμένες δυναμικές. Ορισμένα παραδείγματα τέτοιων επεξεργασιών είναι το Dolby SR, το DBX Quantum Processor, τους διάφορους brand dynamics processors της Waves, και τον Weiss DS1 - MK2 όταν χρησιμοποιείται με ratio μικρότερο από 1:1.

¹³ Το A.G.C αποτελεί το ακρωνύμιο των λέξεων Automatic Gain Control.



Σχ. 4.1.5.1: Σχηματική αναπαράσταση του είδους των επεξεργασιών στο compression και το expansion.

iv. Το downward expansion είναι ο συνηθέστερος τύπος χρήσης των expanders, λαμβάνει μέρη του ηχητικού υλικού με χαμηλό level και τα αποδυναμώνει περισσότερο. Οι περισσότεροι downward expanders χρησιμοποιούνται για να μειώσουν τον θόρυβο που τυχόν υφίσταται στο ηχητικό υλικό, το hiss ή το leakage που έχει ηχογραφηθεί. Εν συνεχεία ένας expander με ratio $1:\infty$ είναι μία ειδική περίπτωση εφαρμογής του downward expansion με πολύ υψηλό λόγο (Noise Gate). Ορισμένες συσκευές που εφαρμόζουν το downward expansion είναι οι Kerex (Keyable Program Expander) και Drawmer, το Dolby και παρόμοια συστήματα αποθορυβοποίησης. Για τη μεγαλύτερη δυνατή σαφήνεια στην εργασία αυτή θα χρησιμοποιούμε τον σύντομο όρο expander για το downward expansion εκτός αν χρειάζεται να τον διακρίνουμε από τον upward expansion τύπο.

4.1.6 Manual Gain Riding:

Εισαγωγές με πολύ χαμηλό level, επίλογοι μουσικών κομματιών ή ακόμα και σημεία στο μέσο της διάρκειας του μουσικού κομματιού, απαιτούν ιδιαίτερη προσοχή. Αν το υψηλότερο level στο τραγούδι ακούγεται «όπως πρέπει» μετά την επεξεργασία αλλά η εισαγωγή ακούγεται πολύ soft, είναι καλύτερα απλά να ανεβάσουμε λίγο το level της εισαγωγής, κάνοντας χρήση των, πολύ βοηθητικών σε αυτές τις περιπτώσεις, τεχνικών τύπου editing για να φέρουμε το level του συγκεκριμένου χρονικού σημείου του ηχητικού κομματιού έτσι όπως πρέπει να είναι χρησιμοποιώντας μια ή περισσότερες από αυτές τις παρακάτω προσεγγίσεις:

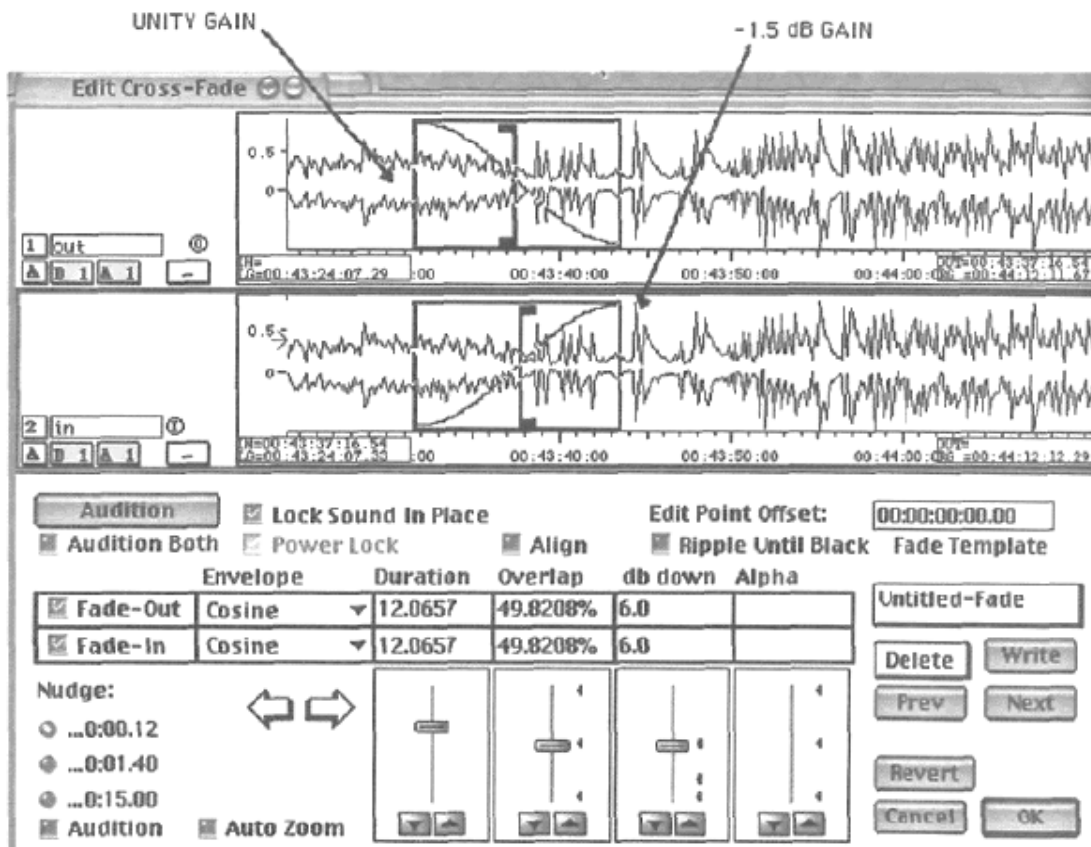
- Μερικές φορές ένα μακρύ βαθμιαίο decrescendo, με τη χρήση των fades, είναι η λύση, που μπορεί να γίνεται στο τέλος της εισαγωγής ή αργά κατά τους πρώτους στίχους του κυρίως σώματος του ηχητικού υλικού.

- Μερικές φορές μια σειρά επεμβάσεων $\frac{1}{4}$ ή $\frac{1}{2}$ dB και πάλι με fades, μειώνοντας το level του ήχου βήμα - βήμα, με αυτό το λόγο, σε κρίσιμα σημεία του ηχητικού κομματιού, επιφέρει θαυμάσια αποτελέσματα. Αυτό είναι χρήσιμο όταν δε θέλουμε να προσέξει ο ακροατής τη μείωση του gain.
- Ορισμένες φορές η αλλαγή του level στο μεταβατικό σημείο μεταξύ π.χ. του ιδιαίτερα υψηλού level μιας εισαγωγής, συγκριτικά με το level του κυρίως σώματος του ηχητικού υλικού, δημιουργεί μουσικό ενδιαφέρον, καθώς επίσης είναι και η μικρότερη δυνατή παρείσφρηση του τεχνικού αναφορικά με τις δυναμικές του μουσικού κομματιού.

Η αντίστροφη προσέγγιση, να δημιουργήσουμε σκόπιμα μια απαλότερη εισαγωγή ώστε το σώμα του τραγουδιού να φαίνεται δυνατότερο και να έχει αντίκτυπο κατά την εισαγωγή στο κύριο μέρος του μουσικού κομματιού μπορεί επίσης να λειτουργήσει ικανοποιητικά.

4.1.7 Αλλαγή των Internal Levels Ενός Μουσικού Κομματιού:

Ορισμένα μέρη του ηχητικού υλικού με πολύ χαμηλό level, προφανώς, θα πρέπει να ενισχυθούν. Αλλά αν οι μουσικοί προσπαθούν να παίξουν κάτι ευαίσθητα, το να σπρώξουμε το fader πολύ μακριά μπορεί να χαλάσει το συναίσθημα που προσπαθούν να αποδώσουν. Η τέχνη είναι να ξέρουμε πόσο να αυξήσουμε το level δίχως να χαθεί η αίσθηση της απαλότητας, καθώς και το να προσδιορίσουμε την ιδανική ταχύτητα με την οποία θα προστίθεται ένας συγκεκριμένος αριθμός απο dB συν το χρόνο (για τη συγκεκριμένη ηχητική περιοχή που μας ενδιαφέρει) χωρίς να γίνουμε αντιληπτοί. Σε ένα DAW οι φυσικές κινήσεις του fader μιας κονσόλας για παράδειγμα αντικαθίστανται από εντολές, crossfades ή τον σχεδιασμό μιας γραφικής παράστασης στο πεδίο του χρόνου συναρτήσει της κυματομορφής. Η πραγματική μαγεία του τεχνικού επεξεργασίας είναι να είναι τόσο αόρατος ώστε κανείς να μην μπορεί να αντιληφθεί ότι έχει συντελεστεί η οποιαδήποτε επεξεργασία. Αν γίνει αντιληπτό ότι ο ήχος έχει υποστεί την οποιαδήποτε επεξεργασία, αυτό σημαίνει πως δεν έχουμε κάνει καλά τη δουλειά μας. Μια τεχνική μείωσης του δυναμικού εύρους με τον λιγότερο επισφαλής και πιο επιβληθτικός τρόπος είναι η παρακάτω. Σε ένα DAW είναι πραγματικά εύκολο να περιγράψουμε την τεχνική αυτή γιατί μπορούμε και έχουμε οπτική επαφή με τη μορφή του κύματος (waveform). Ας υποθέσουμε, λοιπόν, ότι πρέπει να μειώσουμε το level σε ένα δυνατό τμήμα του ηχητικού υλικού. Το καλύτερο σημείο για να πράξουμε κάτι τέτοιο είναι στο τέλος του προηγούμενου low level μέρους πριν αρχίσει το μέρος αυτό του οποίου το level θέλουμε να μειώσουμε.

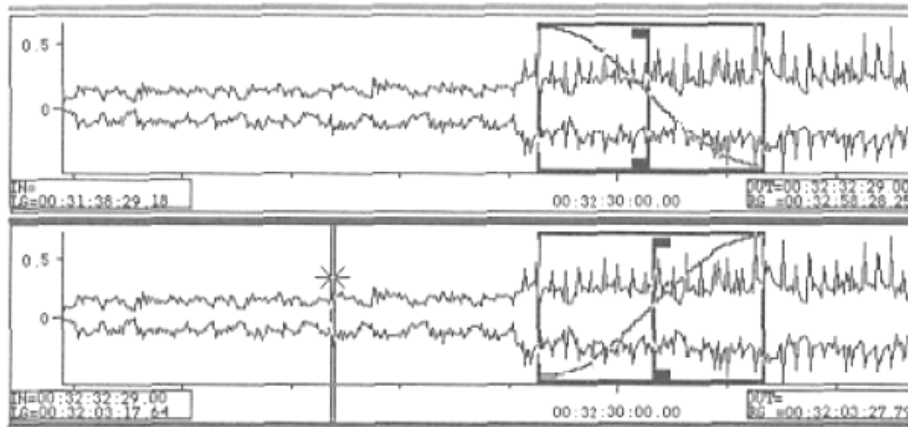


Σχ. 4.1.7.1: Η μοντέρνα εκδοχή του fade - riding. Παρατηρούμε ότι η μείωση του πλάτους της κυματομορφής εκτελείται από εκεί όπου παρουσιάζεται βύθισμα στην κυματομορφή, αμέσως πριν το σημείο εκείνο του οποίου το πλάτος θέλουμε να μειώσουμε.

Ψάχνουμε για ένα φυσικό deep της κυματομορφής πριν την αρχή του crescendo, και εφαρμόζουμε τη μείωση του level μέχρι να φτάσουμε στα επιθυμητά αποτελέσματα. Με αυτό τον τρόπο το δυνατό μέρος του ηχητικού υλικού δε θα χάσει την συγκριτική του επίδραση γιατί το αυτί κρίνει τα δυνατά μέρη ενός μουσικού κομματιού συναρτήσει πάντοτε των low level σημείων του.

Το σχήμα 3.1.7.1 που δίδεται αμέσως πιο πάνω προέρχεται από ένα σταθμό Sonic Solutions και εικονογραφεί την τεχνική αυτή. Η αλλαγή του level επιτυγχάνεται μέσω ενός crossfade (για το ίδιο κανάλι, αλλά επί δύο).

Μερικά τραγούδια αρχίζουν με μια πολύ απαλή εισαγωγή και αυτή μπορεί να πρέπει να αυξηθεί. Άλλα τραγούδια αρχίζουν απαλά και προχωρούν σε μεγάλη κλιμάκωση. Είναι καλό να αρχίζουμε την επεξεργασία του ηχητικού υλικού πηγαίνοντας κατευθείαν στο σημείο εκείνο όπου το level ξεκινάει να κλιμακώνεται. Αφού πετύχουμε έναν καλό ήχο με την κατάλληλη επεξεργασία επιστρέφουμε στην αρχή του μουσικού κομματιού και αν υπάρχει «χώρος» υπάρχει η πιθανότητα να μειώσουμε περαιτέρω το level της εισαγωγής, πράγμα που θα ενισχύσει το ακόλουθο κυρίως σώμα με την αντίθεση που δημιουργείται με αυτό τον τρόπο. Αυτό επίσης



Σχ. 4.1.7.2: Η εισαγωγή ενός μουσικού κομματιού με χαμηλό level. Εξασθενούμε ακόμη περισσότερο το level της συγκεκριμένης εισαγωγής και ενισχύουμε το level του κυρίως σώματος του ηχητικού υλικού.

μειώνει τον πειρασμό να ενισχύσουμε με άλλου τύπου επεξεργασίες το level του δυνατού μέρους τόσο ώστε να συνθλιβεται απο την υπερβολική συμπίεση που θα χρειαστεί για να συμβεί αυτό. Για παράδειγμα στο σχήμα 3.1.7.2 που δίδεται μειώσαμε το level της εισαγωγής ενός τραγουδιού και αργά εισαγάγαμε ένα crescendo (διάρκειας 20 δευτερολέπτων) που ενισχύει τη φυσική δομή του τραγουδιού καθώς οδεύει προς το πρώτο χορωδιακό του μέρος. Το πάνω μέρος μειώνει το gain κατά -1dB, το κάτω είναι στη unity gain (0dB), που επιτυγχάνεται στο τέλος του crossfade.

Όπως γίνεται εύκολα αντιληπτό, στα παραπάνω παραδείγματα δε χρησιμοποιήθηκαν ούτε compressors, ούτε expanders. Οι παραπάνω επεξεργασίες έλαβαν χώρα με τη χρήση fades. Η χρήση των fades στο audio mastering για την βελτίωση των μακροδυναμικών ενός μουσικού κομματιού συναντάται πολύ συχνά με μόνη προϋπόθεση, όμως, την ύπαρξη ενός επαγγελματικού ψηφιακού σταθμού εργασίας (DAW).

Τέλος ένα άλλο τρικ, είναι η αύξηση του διαστήματος πριν την εισαγωγή του τραγουδιού, που αυξάνει την δυναμική του επίδραση προεκτείνοντας την ένταση που προκαλεί η σιωπή. Είναι καλό να δίνουμε στο αυτί την ευκαιρία να συνηθίσει την ηχητική σιωπή και μετά να το κατακλύζουμε με πληροφορία! Οι καλύτεροι μουσικοί ξέρουν πώς να χρησιμοποιούν το διάστημα στη μουσική τους και θεωρούν τις ηχητικές σιωπές εξίσου σημαντικές με τις νότες.

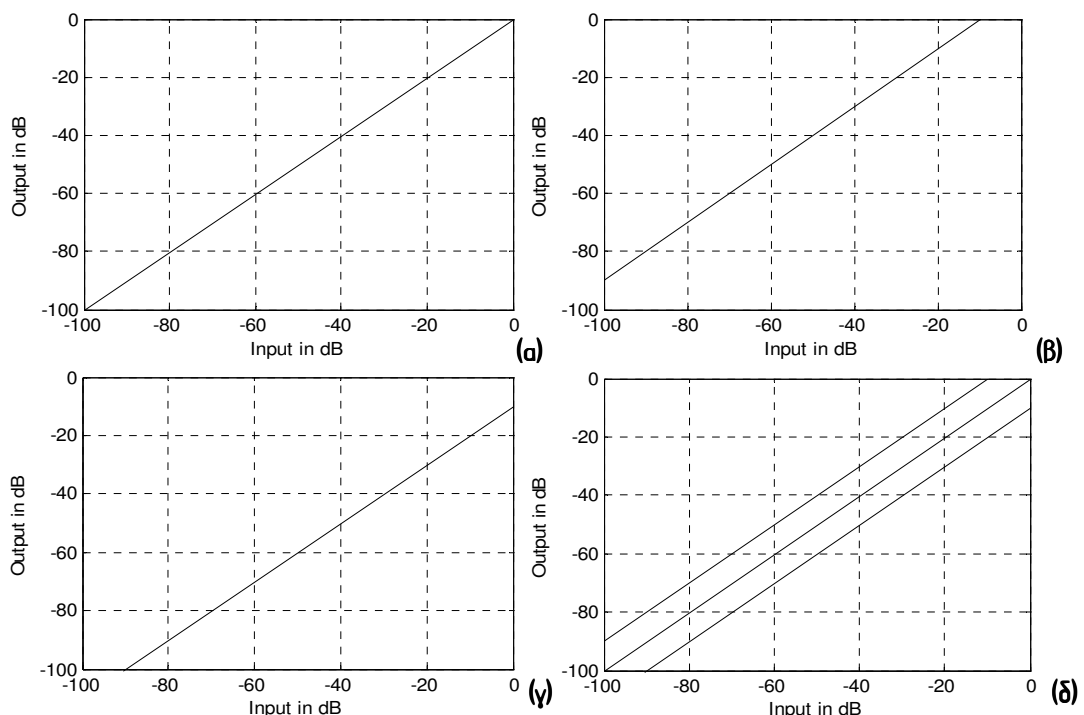
Οι μακροδυναμικοί χειρισμοί είναι ένα, παραμελημένο μερικές φορές, αλλά εξαιρετο εργαλείο στο οπλοστάσιο του τεχνικού mastering. Στις επόμενες παραγράφους θα προχωρήσουμε στη χρήση compressors, expanders και limiters για τον μικροδυναμικό χειρισμό του ηχητικού υλικού.

4.2 Compressors & Limiters Αντικειμενικά Χαρακτηριστικά:

Οι παράγραφοι που ακολουθούν αφορούν τους μικρόδυναμικούς χειρισμούς, που επιτυγχάνονται βασικά με την χρήση καθορισμένων δυναμικών επεξεργαστών (dynamics processors). Εδώ, λοιπόν, θα δούμε πως λειτουργούν οι downwards επεξεργαστές. Πριν περάσουμε όμως στο πεδίο λειτουργίας συσκευών όπως είναι οι compressors και οι expanders, πρέπει να μελετήσουμε τα αντικειμενικά χαρακτηριστικά των συσκευών αυτών.

4.2.1 Transfer Curves / Compressors & Limiters:

Ας ξεκινήσουμε με τα μετρίσιμα χαρακτηριστικά των επεξεργαστών καθοδικής συμπίεσης (downward compression), εν συντομία αποκαλούμενοι compressor και limiters. Καταρχήν μια συνάρτηση μεταφοράς είναι η σχέση της εισόδου σε σχέση με την έξοδο αναφορικά με το gain ενός ενισχυτή ή επεξεργαστή. Ένας ενισχυτής ο οποίος παράγει στην έξοδο ένα σήμα με πλάτος όμοιο με αυτό που λαμβάνει στην είσοδο του ή unity - gain¹⁴ θα αποδώσει μια ευθεία διαγώνια γραμμή κατά μήκος των 45°, που ονομάζεται γραμμή unity gain. Μια ομάδα γραφικών παραστάσεων μπορούν να σχεδιαστούν όπως φαίνεται στο σχήμα 3.2.1.1 που δίδεται.



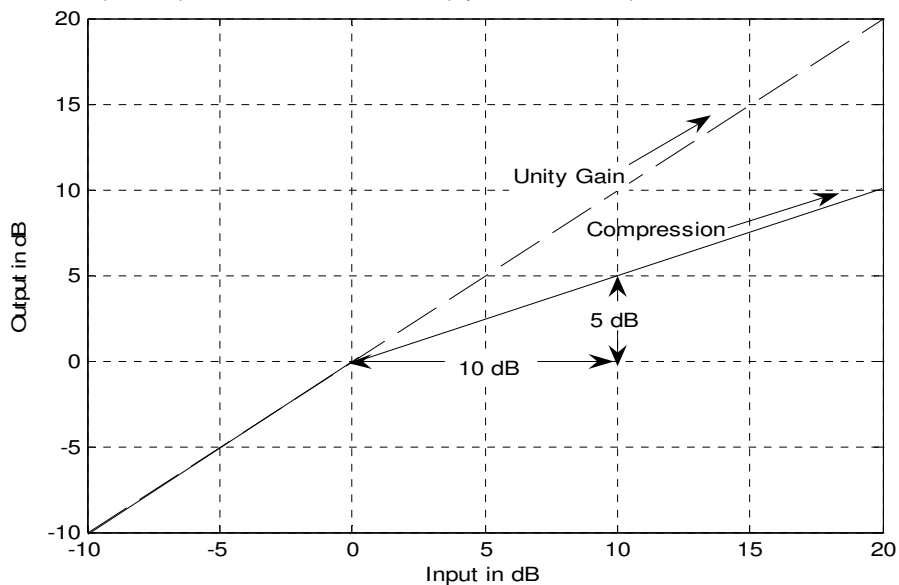
Σχ. 4.2.1.1: Συναρτήσεις μεταφοράς ενός ενισχυτή. (α) Η unity gain του ενισχυτή. (β) Ο ίδιος ενισχυτής με κέρδος (gain) 10 dB στην έξοδο. (γ) Αυτή τη φορά με εξασθένηση του σήματος κατά 10 dB στην έξοδο. (δ) Οι τρεις ευθείες στην ίδια γραφική απεικόνιση.

¹⁴ Σχέση output προς input 1, ή αλλιώς 0dB.

Το level εισόδου βρίσκεται στον άξονα Χ και το level εξόδου στον Υ. Στο διάγραμμα (α) είναι ένας ενισχυτής του οποίου η συνάρτηση μεταφοράς είναι η unity - gain, μετά, στο διάγραμμα (β) ένας με κέρδος 10 dB και στη συνέχεια, στο διάγραμμα (γ) με μείωση 10 dB. Όσο υπάρχει μια ευθεία γραμμή (όχι καμπύλη) στις 45°, οι ενισχυτές είναι γραμμικοί. Αυτό όμως, δε σημαίνει πως η τιμή του level του σήματος στην έξοδο του ενισχυτή θα είναι όμοια με αυτή στην είσοδο του. Σε κάθε μία από τις τρεις γραφικές παραστάσεις του παραπάνω σχήματος 4.2.1.1 υποδηλώνεται πως η επεξεργασία είναι γραμμική εφόσον μία αλλαγή στην είσοδο θα επιφέρει μία αντίστοιχη αλλαγή στην έξοδο. Παρατηρούμε στο διάγραμμα (β) ότι το για κάθε σήμα εισόδου πάνω από -10dBFS προκαλείται παραμόρφωση στην έξοδο.

Τα παραπάνω έχουν σημασία διότι ουσιαστικά ο compressor είναι ένας ενισχυτής ο οποίος, βάσει των ιδιομορφιών του σήματος εισόδου, λειτουργεί με συνεχώς μεταβαλλόμενο gain, ώστε να δίνει στην έξοδο του το εν λόγω σήμα με μειωμένη δυναμική περιοχή.

Το threshold ενός compressor ορίζεται ως το level αυτό, πάνω από το οποίο αρχίζει να λαμβάνει χώρα η συμπίεση του σήματος. Το ratio του compressor, τώρα, είναι ο λόγος του σήματος στην είσοδο του ενισχυτή (input change, Δx) σε σχέση με το ίδιο σήμα στην έξοδο του (output change, Δy), αυτό θα συμβαίνει μόνο όταν το σήμα εισόδου υπερβαίνει το όριο (threshold) που έχει τεθεί από τα πριν. Στο σχήμα 4.2.1.2 παρουσιάζεται με σαφήνεια το πώς λειτουργεί ένας compressor.



Σχ. 4.2.1.2: Ένας compressor με ορισμένο όριο πάνω από το οποίο το σήμα συμπίεζεται τα 0 dB και ratio 2:1. Η διακεκομμένη γραμμή είναι η unity gain του compressor.

Παρατηρούμε στο σχήμα ότι ο compressor δίνει στην έξοδο του ένα ποσοστό του αντίστοιχου level εισόδου. Έτσι προκύπτει συνεπώς ότι η transfer function του compressor παύει να είναι η unity gain και μετασχηματίζεται στην $y = \frac{x}{r}$. Δηλαδή η

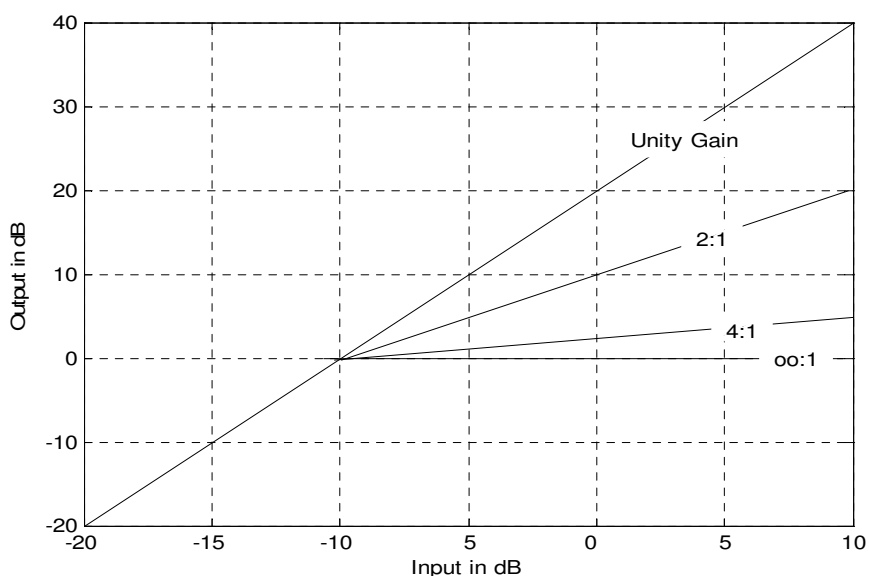
έξοδος είναι το $1/r$ της εισόδου. Συμβολικά το ποσοστό που προκύπτει γράφεται ως

$r:1$ και ορίζει με μαθηματικό τρόπο το ratio του compressor. Παρατηρούμε, επίσης, στο σχήμα 4.2.1.2 ότι $\frac{\Delta y}{\Delta x} = \frac{1}{r} \Rightarrow \frac{\Delta y}{\Delta x} = \frac{5}{10} = \frac{1}{2}$, δηλαδή $r=2$, συνεπώς ο compressor λειτουργεί με ratio = 2:1.

Είναι φανερό ότι η ελάττωση της δυναμικής περιοχής, επιτυγχάνεται μέσω της απώλειας gain (κοινώς gain reduction). Αυτό συμβαίνει κάθε φορά που εμφανίζεται τιμή εισόδου μεγαλύτερη από αυτή του threshold. Φαίνεται για παράδειγμα στο σχήμα ότι στο level εισόδου $x = 10 \text{ dB}$ υπάρχει απώλεια 5 dB εφόσον $y = \frac{x}{2} = 5 \text{ dB}$, ενώ δίχως τον compressor θα είχαμε $y = 10 \text{ dB}$.

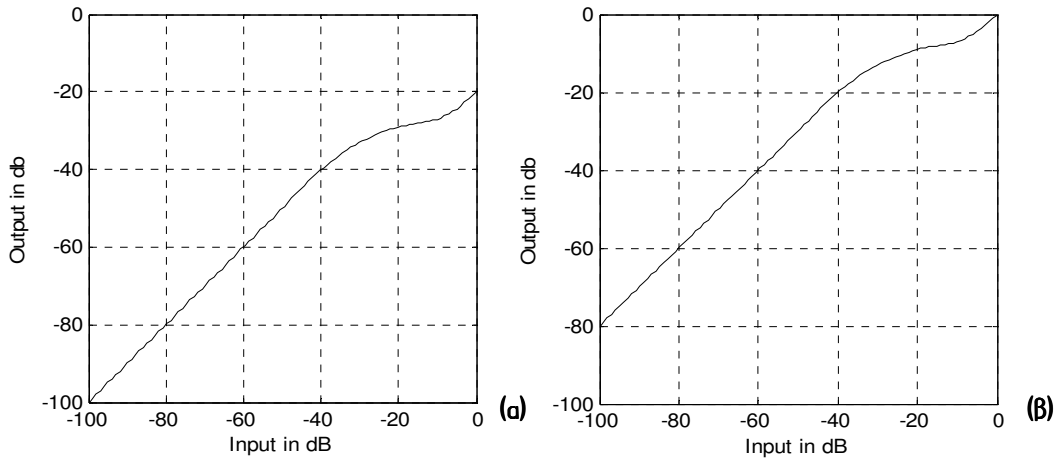
Οι τιμές threshold και ratio είναι μεταβλητές και συνεπώς οι δυνατότητες δημιουργικού πειραματισμού είναι μεγάλες.

Τέλος, το σχήμα 4.2.1.3 που ακολουθεί δείχνει την εξάρτηση του gain reduction από το ratio.



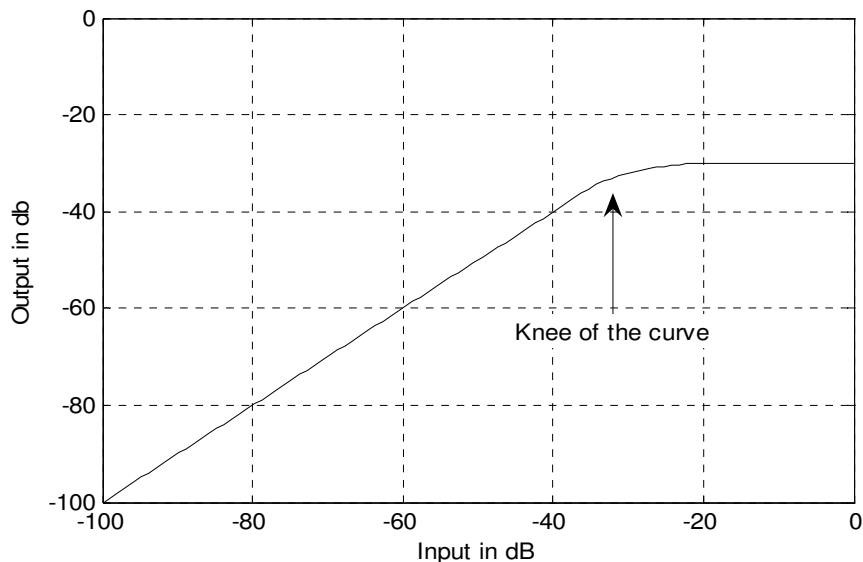
Σχ. 4.2.1.3: Γραφική απεικόνιση της εξάρτησης του gain reduction από το ratio.

Στο σχήμα 4.2.1.4 που δίδεται στην επόμενη σελίδα και στο διάγραμμα (α) παρουσιάζεται ένας απλός compressor με ένα σχετικά ευαίσθητο ratio 2.5:1 και με threshold στα -40 dBFS (που είναι αρκετά χαμηλά και θα επέφερε ισχυρή συμπίεση στα δυνατά σήματα). 2.5:1 και με βάση τα όσα έχουμε αναφέρει πιο πάνω, σημαίνει ότι για μια αύξηση επιπέδου της πηγής κατά 2.5dB, η έξοδος θα ανέβει μόνο 1dB, ή για μια άνοδο 5dB η έξοδος θα είναι μόνο 2dB, ή, όπως φαίνεται στο σχήμα, μια αλλαγή εισόδου κατά 20dB επιφέρει μια αλλαγή εξόδου λίγο μικρότερη από 10dB (όταν η καμπύλη φτάσει στη μέγιστη κλίση της). Ένας τέτοιος compressor θα έκανε πραγματικά τα δυνατά μέρη απαλότερα, γιατί η έξοδος είναι μικρότερη από την είσοδο πάνω από το threshold (όπως είναι και το λογικό άλλωστε). Αυτό συμβαίνει πάντα εκτός αν ο compressor ακολουθείται από ένα gain make-up ενισχυτή.



Σχ. 4.2.1.4: (α) Ένας απλός compressor με σχέση συμπίεσης 2,5:1 και threshold -40 dBFS. (β) Ο ίδιος compressor με τα ίδια χαρακτηριστικά μόνο που μετά τη βαθμίδα ενίσχυσης το σήμα οδηγείται σε έναν gain makeup amplifier ο οποίος και ενισχύει το σήμα κατά 20 dBFS.

Στο διάγραμμα (β) παρατηρούμε πως χρησιμοποιώντας ένα gain makeup (έναν απλό gain amplifier δηλαδή μετά το τμήμα συμπίεσης) μπορούμε να αποκαταστήσουμε το gain σε ένα σήμα εισόδου του οποίου το level ακουμπά το όριο του ψηφιακού audio σήματος (0dBFS) και να επιφέρουμε και πάλι ένα σήμα πλήρους level (0dBFS) στην έξοδο (με αυτό τον τρόπο κατανοούμε ότι ο compressor δεν είναι απαραίτητο πως το σήμα που εξάγει είναι χαμηλότερου level, απλά ο compressor έχει συμπίεσει το σήμα σε επίπεδο δυναμικής περιοχής). Παρατηρούμε, επίσης, ότι ο ενισχυτής δίδει ένα μεγάλο makeup gain, 20 dB, που θα ενίσχυε σημαντικά τα απαλά μέρη (ακόμη και κάτω από το threshold). Σε συνηθισμένη χρήση makeup gains είναι σπάνιο το να δίνουμε περισσότερα από 3 ή 4 dB. Δυνατά τμήματα του σήματος εισόδου, από περίπου -40 έως περίπου -15 dBFS, ενισχύονται ακόμα σε αυτό το σχήμα, αλλά πάνω από τα περίπου -10 dBFS η καμπύλη επανέρχεται και πάλι πίσω στη unity gain και προσομοιάζει αυτή ενός γραμμικού ενισχυτή. Πολύ κάτω από το threshold είναι ένας σχετικά γραμμικός ενισχυτής και μπορεί να έχει αρκετά χαμηλή παραμόρφωση διότι δεν υπάρχει δραστηριότητα μείωσης του gain. Στην κλίμακα των dBFS (dB Full Scale) 20dB gain makeup προστίθεται με 20dB μείωσης του gain επιφέροντας συνολικό level 0 dBFS. Ο συγκεκριμένος compressor λειτουργεί γραμμικά πάνω από ένα συγκεκριμένο ποσό συμπίεσης, ώστε το ratio να λειτουργεί και κατά συνέπεια η συμπίεση του σήματος να λαμβάνει χώρα μόνο για τα πρώτα 15-20dB πάνω από το threshold. Άλλοι compressors συνεχίζουν την απότομη κλίση τους, διατηρώντας έτσι το ratio τους πολύ πάνω από το κατώφλι. Υπάρχουν τόσες ποικιλίες καμπυλών απόκρισης του τρόπου με τον οποίο θα συμπεριφερθεί στο σήμα ο compressor όσες και οι εταιρίες που κατασκευάζουν τέτοιου είδους συσκευές και όλες δίνουν διαφορετικούς ήχους. Για να επιτευχθεί το καλύτερο αισθητικά αποτέλεσμα από κάθε compressor, αυτός πρέπει να έχει λάβει τις ρυθμίσεις με τέτοιο τρόπο ώστε η περισσότερη από τη μουσική δράση να συμβαίνει γύρω από το σημείο



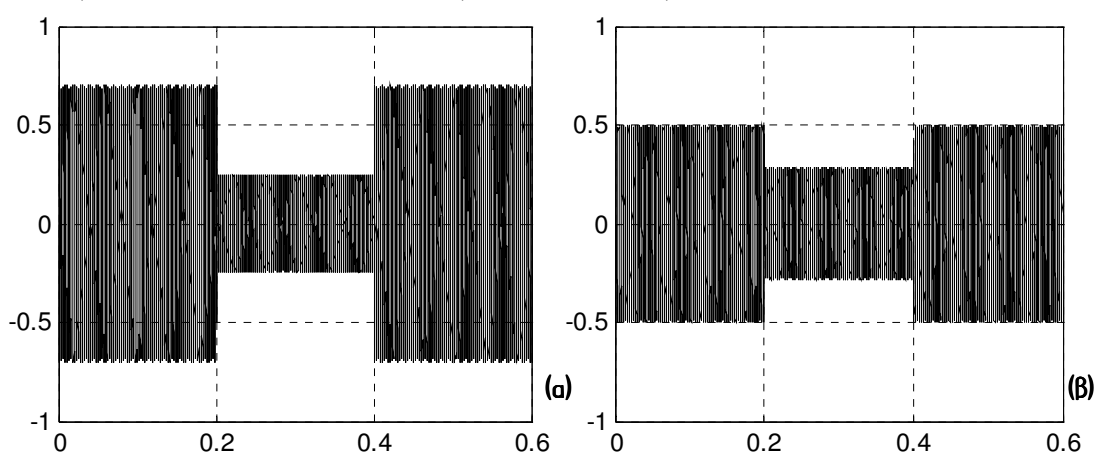
Σχ. 4.2.1.5: Ένας compressor με ratio 10:1, του οποίου το κατώφλιο (threshold) έχει οριστεί στα -35 dBFS, δίχως makeup gain.

του threshold, όπου αλλάζει το σχήμα της καμπύλης. Έτσι είναι πιθανό ότι το κατώφλι ενός compressor στον πραγματικό κόσμο θα ήταν κοντά στα -20 με -10dBFS, εκεί που λαμβάνει χώρα το μεγαλύτερο κομμάτι της μουσικής διαδρομής.

Το σχήμα 4.2.1.5 δείχνει έναν compressor με ένα πολύ υψηλό ratio 10:1 δίχως gain makeup. Παρατηρούμε ότι η έξοδος του compressor είναι μια σχεδόν οριζόντια γραμμή πάνω από το threshold. Οι περισσότεροι τεχνικοί ήχου αποκαλούν κάθε compressor με ratio 10:1 ή μεγαλύτερο limiter. Υπάρχουν πολύ λίγοι αναλογικοί compressor στους οποίους να μπορεί να βρεθεί μεγαλύτερο ratio απο το 10:1, όμως μερικοί ψηφιακοί limiter παρέχουν ακόμη και 1000:1 προκειμένου να προλάβουν και το παραμικρό έστω peak του σήματος ή για να αποτρέψει το clipping όταν το σήμα υπερβαίνει τα 0dBFS. Το μέρος της καμπύλης πάνω ή κοντά στο threshold αποκαλείται knee και είναι η μετάβαση μεταξύ της περιοχής που καταλαμβάνει η unity gain και της περιοχής που καταλαμβάνει η συμπίεση. Το σχήμα που λαμβάνει η καμπύλη στο σημείο αυτό (σημείο knee) μπορεί να κάνει την μετάβαση απο τη μία περιοχή στην άλλη «μαλακή» ή «σκληρή». Ο όρος soft knee αναφέρεται σε ένα στρογγυλεμένο σχήμα της καμπύλης, και ο όρος hard knee σε ένα μυτερό σχήμα, όπου η συμπίεση ή το limiting συμβαίνει γρήγορα πάνω από το threshold. Κατά συνέπεια η αλλαγή από τη unity σε ratio 10:1 θα μπορούσε να είναι στιγμιαία, στην περίπτωση αυτή το knee της καμπύλης θα ήταν μια οξεία γωνία αντί για στρογγυλό, παράγοντας μια οξεία ηχητική αλλαγή, δηλαδή ένα αποτέλεσμα limiting. Η ανάγκη για ένα soft knee εξαρτάται πολύ από το πόση μουσική δραστηριότητα ή κίνηση συμβαίνει στο threshold. Αν υπάρχει πολύ μουσική δραστηριότητα ή κίνηση γύρω από το threshold το σχήμα του γόνατου μπορεί να είναι κρίσιμο. Για compressors που δεν έχουν ρυθμίσεις knee, κάποια από τα ηχητικά αποτελέσματά που προσδίδονται είτε με το hard knee, είτε με το soft knee μπορούν να επιτευχθούν αν πειραματιστούμε με το ratio και/ ή με το threshold.

4.2.2 Χρόνοι Attack και Release:

Η ουσία της επέμβασης του compressor βρίσκεται στο ότι πρέπει αυτόματα να μπαίνει σε λειτουργία αλλά και να βγαίνει από αυτήν, παρακολουθώντας τις αυξομειώσεις του σήματος. Προφανώς, επειδή τίποτα δε συμβαίνει σε μηδενικό χρόνο, αναδεικνύεται σε μείζονος σημασίας ζήτημα το πώς αποκρίνονται χρονικά τα κυκλώματα του compressor. Δηλαδή η ταχύτητα με την οποία ο εν λόγω επεξεργαστής μπαίνει σε λειτουργία και βγαίνει από αυτή. Οι παράμετροι που δείχνουν, αλλά και που ορίζουν αυτές τις ταχύτητες απόκρισης του compressor είναι οι attack και release time. Η παράμετρος attack time ορίζεται ως ο χρόνος μεταξύ της onsets φάσης ενός σήματος που υπερβαίνει το threshold και της πλήρους μείωσης του gain με βάση το ratio που έχει οριστεί. Μπορεί να μετρηθεί σε micro ή millisecond αν και μπορεί να διαρκεί έως και ένα ή δύο δευτέρα. Οι τυπικές τιμές attack time των compressor που χρησιμοποιούνται στη μουσική, κυμαίνονται από 50 ms ως 300 ms με συνήθη χρήση περίπου 100ms. Η παράμετρος, τώρα, release time, γνωστή και ως recovery time λόγω της λειτουργίας της, ορίζεται ως ο χρόνος από τη στιγμή που ένα σήμα σταματά να υπερβαίνει το threshold έως ότου το gain να επανέλθει στην unity τιμή. Οι τυπικές τιμές release time των compressor που χρησιμοποιούνται στη μουσική, κυμαίνονται από 50 ms μέχρι 500 ms ή όσο ένα ή δύο δευτέρα, με συνήθη χρήση μάλλον στα 150 – 250¹⁵ ms. Οι όροι σύντομο (short) ή γρήγορο (fast) για το attack και το release time μπορούν να χρησιμοποιηθούν εναλλακτικά, σημαίνουν το ίδιο πράγμα. Παρόμοια αργό (slow) και μακρείς (long) χρόνοι attack και release σημαίνουν το ίδιο. Στο διάγραμμα (α) του ακόλουθου σχήματος 4.2.2.1 παρουσιάζεται η περιβάλλουσα (envelope) ενός απλού τόνου, από υψηλό level σε χαμηλό και το

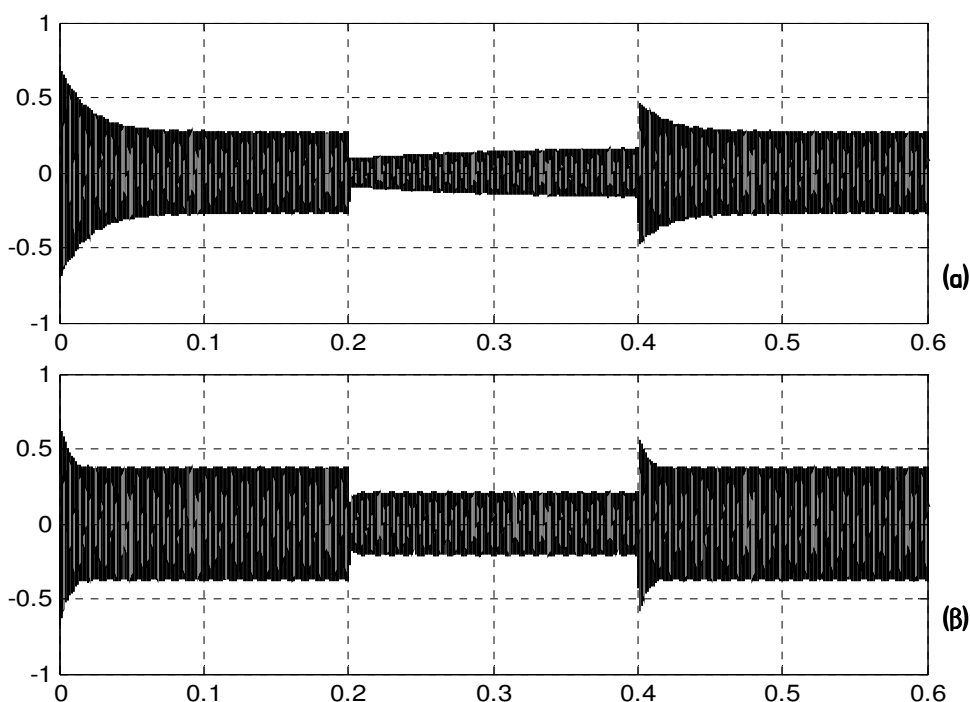


Σχ. 4.2.2.1: (α) Απλός τόνος με περιβάλλουσα από το υψηλό στο χαμηλό level και το αντίστροφο. (β) Ο ίδιος απλός τόνος αφού έχει επεξεργαστεί από έναν compressor με γρήγορο attack, υψηλό ratio και μικρό release time.

¹⁵ Μία κατασκευάστρια εταιρία, η dBx, θέτει στις συσκευές της ως μονάδα μέτρησης του release time όχι το msec, αλλά το dB/sec που πιθανότατα είναι περισσότερο δόκιμος όρος αλλά δύσκλητος.

αντίστροφο. Στο διάγραμμα (β) παρουσιάζεται ο ίδιος καθαρός τόνος αφού πρώτα έχει υποστεί την επεξεργασία από έναν compressor με πολύ γρήγορο attack, υψηλό ratio και γρήγορο release, το threshold του οποίου, είναι στη μέση μεταξύ των δυνατών και απαλών σημάτων. Παρατηρούμε ότι τα δυνατά μέρη ασθενούν άμεσα, τα απαλά μέρη ενισχύονται άμεσα και υπάρχει λιγότερη απόλυτη δυναμική διακύμανση, κρίνοντας από τον κάθετο άξονα Υ του πλάτους.

Στο επόμενο σχήμα 4.2.2.2 και στο διάγραμμα (α) παρουσιάζεται η επεξεργασία μέσω ενός compressor με χαμηλό ratio, αργό attack time και αργό release time.



Σχ. 4.2.2.2: (α) Ένας compressor με χαμηλό ratio, μικρό χρόνο ανόδου (attack) και μικρό χρόνο καθόδου (release). (β) Με υψηλότερο ratio, γρηγορότερο attack και με πολύ γρήγορο release.

Παρατηρούμε πως το αργό attack του compressor επιτρέπει σε μέρος του αρχικού σήματος της πηγής να παραμείνει, έως ότου ο compressor αντιδρά, και στο σημείο αυτό η μείωση του gain μειώνει το level. Τότε, όταν το σήμα πέσει κάτω από το threshold, χρειάζεται ένα κάποιος χρόνος (η διάρκεια του release), και το level παραμένει χαμηλό, μετά σιγά - σιγά επανέρχεται. Μεγάλο μέρος του αποτελέσματος της συμπίεσης (αυτό που κοινώς δηλαδή αποκαλούμε ο «ήχος» του compressor) συμβαίνει κατά την κρίσιμη περίοδο του release, καθώς όπως μπορούμε να δούμε, εκτός του αρχικού χρόνου του attack, ο compressor έχει μειώσει πραγματικά το υψηλού level σήμα.

Αντιπαραβάλλοντας το με τον compressor του διαγράμματος (β) που έχει πολύ ταχύτερο attack και πολύ γρήγορο release time. Με το γρήγορο release, μόλις το σήμα πέσει κάτω από το threshold, ο χρόνος του release ανεβάζει άμεσα το level.

Αυτός ο τύπος γρήγορης δράσης μπορεί να κάνει τη μουσική να ακούγεται ισχυρά συμπιεσμένη γιατί αποδυναμώνει τα δυνατά μέρη και ενισχύει γρήγορα τα απαλά.

4.2.3 Limiting Εναντίον Compression στο Audio Mastering:

Το audio mastering απαιτεί την ανάπτυξη νέων ικανοτήτων, αφού γενικά εργαζόμαστε σε ολοκληρωμένες μίξεις αντί για μεμονωμένα όργανα. Στο mastering όπως και στην ηχοληψία το compression και το limiting αλλάζουν το εύρος της δυναμική περιοχής. Οι περισσότεροι τεχνικοί mastering (που γνωρίζουν τι κάνουν) χρησιμοποιούν compressors για να επηρεάσουν σκόπιμα τον ήχο και limiters για να αλλάξουν τον ήχο όσο το δυνατόν λιγότερο, αλλά απλά να του επιτρέψουν να είναι δυνατότερος. Γι αυτό οι limiters χρησιμοποιούνται συχνότερα στο mastering απ' ότι στην ηχοληψία. Δεν υπάρχει ο τέλεια αόρατος limiter, αλλά το compression επηρεάζει τον ήχο πολύ περισσότερο απ' ότι το limiting. Φανταστείτε την συμπίεση ως εργαλείο για να αλλάξουν οι εσωτερικές δυναμικές της μουσικής. Καθώς μειώνει την δυναμική περιοχή, μπορεί να επιβαρύνει με περίσσιο όγκο ή να προσθέσει περαιτέρω ενέργεια στα μουσικά αυτά μέρη χαμηλού και μεσαίου level για να δημιουργήσει ένα δυνατότερο μουσικό μήνυμα. Χρησιμοποιώντας limiting όμως, με αρκετά γρήγορο attack (1 ή 2 δείγματα - samples), και ένα προσεκτικά ελεγχόμενο γρήγορο release, ακόμα και αρκετά dB limiting μπορεί να είναι διαφανή στο αυτί. Το τρικ είναι να σκεφτόμαστε το limiting όταν θέλουμε να αυξήσουμε το φαινομενικό level του υλικού χωρίς να επηρεάσουμε σημαντικά τον ήχο του. Σκεφτόμαστε το compression ή το upward expansion όταν από το υλικό φαίνεται να λείπει η δύναμη ή η ρυθμική κίνηση συναρτήσει πάντοτε του μουσικού είδους του ηχητικού υλικού.

Όλοι οι ψηφιακοί limiters επηρεάζουν ως ένα βαθμό το ήχο, απαλύνοντας τα μεταβατικά του ηχητικού υλικού και ακόμα «παχαίνοντας» ελαφρά τον ήχο, καθώς μας επιτρέπουν να ανεβάσουμε την RMS τιμή του level και κατά συνέπεια την ακουστότητα του ηχητικού υλικού. Όσο λιγότερο limiting χρησιμοποιούμε τόσο καθαρότερος και πιο «κοφτός» θα είναι ο ήχος, εκτός και αν ψάχνουμε για ήχο με απαλότερα μεταβατικά. Σε ένα ιδανικό mastering project ο limiter θα έπρεπε να δρα μόνο σε τυχόν μη ακουόμενα peaks (και αυτό γιατί τα πολύ σύντομα peaks, τα οποία μπορεί να είναι δύσκολο να γίνουν αντιληπτά, συνήθως, υπερβαίνουν το headroom του μέσου αποθήκευσης. Επίσης μπορεί να καταστούν ιδιαίτερος επιβλαβή για τα tweeters των ηχείων). Η παραμόρφωση, τώρα, του limiting γίνεται ιδιαίτερα ακουστή σε υλικό που έχει λίγα peak και αυτό γιατί ένας limiter δεν έχει σχεδιαστεί να δουλεύει στο RMS πεδίο του μουσικού σήματος και οι limiters μπορούν να ακούγονται πολύ άσχημα όταν «πιέζονται» μέσα στην περιοχή RMS. Όταν χρησιμοποιούμε limiters θα πρέπει να δίνουμε ιδιαίτερη προσοχή σε τυχόν σημαντική παραμόρφωση των χαμηλών συχνοτικών περιοχών του ηχητικού υλικού γιατί οι χρονικές σταθερές ενός limiter είναι πολύ γρήγορες για την καλύτερη δυνατή συμπίεση και αυτό μπορεί να προκαλέσει

διάφορα είδη διαμόρφωσης που προκύπτουν από τη συσχέτιση των γρήγορων εναλλαγών στη λειτουργία του limiter και της περιβάλλουσας ενός ήχου με χαμηλό μήκος κύματος.

4.2.4 Ο πιο διαυγής ψηφιακός limiter:

Ο πιο διαυγής limiter είναι... να μη χρησιμοποιούμε καθόλου limiter! Όταν προσπαθούμε να κάνουμε ένα τμήμα του μουσικού κομματιού δυνατότερο, και αν υπάρχει σε αυτό το τμήμα ένα πολύ σύντομο peak που κλιράρει π.χ. σε ένα τμήμα με έντονη την παρουσία των ντραμς, τότε ένας ικανός τεχνικός mastering μπορεί να λειτουργήσει μειώνοντας το level σε αυτό το πολύ μικρής διάρκειας peak και μόνο, χρησιμοποιώντας τις δυνατότητες του DAW editor. Αυτό θα έχει καλύτερα αποτελέσματα από το να χρησιμοποιούσαμε έναν limiter. Αυτή η τεχνική χειροκίνητου limiting μας επιτρέπει να ανεβάσουμε το φαινομενικό level ενός τραγουδιού χωρίς την επακόλουθη παραμόρφωση που προκύπτει από έναν ψηφιακό limiter, και έτσι είναι η πρώτη διαδικασία που πρέπει να σκεφτούμε όταν δουλεύουμε με μουσική που ακούγεται «ανοιχτά» και μπορεί να καταστραφεί από την υπερβολική επεξεργασία. Συχνά μπορούμε να εφαρμόσουμε την τεχνική του χειροκίνητου limiting για μείωση του level από 1 ως 3dB για διάρκειες, όμως, των peak μικρότερες ή το πολύ έως 3ms. Σε μεγαλύτερης διάρκειας peak η μείωση του level με αυτόν τον τρόπο θα επηρεάσει τον ήχο όσο ή και περισσότερο από έναν καλό ψηφιακό limiter. Επίσης προσπαθούμε να χρησιμοποιούμε όσο το δυνατόν χαμηλότερα το gain reduction όταν προσπαθούμε (στην τελική φάση της επεξεργασίας) να κάνουμε το υλικό δυνατότερο, ενώ παράλληλα προσπαθούμε να επιτύχουμε το καλύτερο δυνατό level χωρίς να εμφανίζεται το clipping, δηλαδή προσπαθούμε να ενισχύσουμε το συνολικό level του μουσικού κομματιού δίχως να το συμπιέσουμε υπερβολικά (για την ακρίβεια το στοίχημα κάθε φορά είναι να επιτύχουμε τον μικρότερο βαθμό συμπίεσης). Μία τέτοια προσέγγιση βοηθά να κρατάμε αόρατη την εφαρμογή του limiting.

4.2.5 Compressors εν ώρα εργασίας:

Η επεξεργασία ενός μουσικού σήματος χρησιμοποιώντας compressors απαιτεί, όπως είναι φυσικό, προσεκτική ρύθμιση του threshold, του attack time του compressor και του release time. Αν το attack time είναι πολύ σύντομο, τότε η onset φάση του snare drum για παράδειγμα μπορεί να αλλοιωθεί χάνοντας τον αρχικό τονισμό του και αναιρώντας όλα τα στάδια της πρωτύτερης εργασίας καθώς και τον σκοπό της σύνθεσης. Αν το release time είναι πολύ μεγάλο, τότε ο compressor δεν θα επανέλθει αρκετά γρήγορα από την μείωση του gain. Αν το release time είναι πολύ γρήγορο αυτή τη φορά, ο ήχος θα αρχίσει να παραμορφώνεται. Αν ο συνδυασμός του attack time και του release δεν είναι ο ιδανικός για τον ρυθμό της μουσικής, ο ήχος που θα προκύψει

θα είναι θολός, υπερβολικά συμπιεσμένος και δίκως «νεύρο». Είναι μια ευαίσθητη διαδικασία που απαιτεί χρόνο, πείρα, ικανότητα και ένα εξαιρετικό σύστημα μόνιτορ. Το καλύτερο σημείο για την έναρξη της ρύθμισης ενός compressor είναι να βρεθεί πρώτα το threshold έστω και κατά προσέγγιση. Στη συνέχεια ορίζουμε το ratio του compressor σε μια σχετικά υψηλή τιμή και θέτουμε ένα γρήγορο (fast) release time. Προσαρμόζουμε το threshold μέχρι το μετρητικό όργανο μείωσης του gain που διαθέτει κάθε compressor να αρχίσει να αναπηδά καθώς περνούν οι «συλλαβές» που θέλουμε να επηρεάσουμε. Αυτό εξασφαλίζει ότι το threshold θα τοποθετηθεί κατά τον καλύτερο τρόπο γύρω από τους μουσικούς τονισμούς που θέλουμε να χειριστούμε, το «σημείο δράσης» δηλαδή της μουσικής. Μετά μειώνουμε το ratio αρκετά χαμηλά και θέτουμε το release time περίπου στα 250ms. Από εκεί και πέρα είναι ζήτημα μικρορυθμίσεων του attack, του release και του ratio, με πιθανή μια επαναρύθμιση του threshold. Ο στόχος είναι να θέσουμε το threshold μεταξύ των χαμηλότερων και υψηλότερων δυναμικών, ώστε να υπάρχει διαρκής εναλλαγή μεταξύ χαμηλής και υψηλής (ή καθόλου) συμπίεσης αναφορικά με το μουσικό σήμα. Ένα πολύ χαμηλό threshold θα αντικρούσει τον σκοπό για τον οποίο προορίζεται, που είναι η διαφοροποίηση των «συλλαβών» της μουσικής, επίσης με ένα πολύ χαμηλό threshold τα πάντα θα κονσερβοποιούνται, κάτι που σε κανένα απολύτως επίπεδο δεν αντικατοπτρίζει το επιδιωκόμενο αποτέλεσμα.

4.2.6 Τυπικές τιμές του ratio και του threshold:

Όταν ασχολούμαστε με τις μικροδυναμικές μεταβολές ενός μουσικού κομματιού κατά τον παραπάνω τρόπο, οι σχέσεις συμπίεσης που χρησιμοποιούνται συνηθέστερα στο audio mastering είναι από περίπου 1,5:1 μέχρι περίπου 3:1, και τα τυπικά threshold εκτείνονται σε ένα εύρος μεταξύ -20 και -10dBFS. Αλλά δεν υπάρχει κανόνας: μερικοί τεχνικοί πετυχαίνουν σπουδαία αποτελέσματα με ratio 5:1, ενώ μια εκλεπτυσμένη «πινελιά» στο μουσικό υλικό μπορεί να επιτευχθεί με ένα ratio μικρό ως και 1,01:1 ή ένα threshold στα -3dBFS. Μερικές φορές μια ηχογράφιση απαιτεί μια απαλή *αόρατη* συμπίεση χωρίς να προσπαθούμε να αλλάξουμε τις ενσωματωμένες δυναμικές της. Ένα κόλπο για να συμπιέσουμε όσο το δυνατόν απαρατήρητα είναι η χρήση μιας εξαιρετικά ελαφριάς σχέσης συμπίεσης, ας πούμε 1,01:1 ή 1,1:1 και ενός πολύ χαμηλού threshold, στα -30 ή -40dBFS, αρχίζοντας πού χαμηλότερα από εκεί που λαμβάνει χώρα η δράση του μουσικού υλικού. Μπορεί να διαλέξουμε ένα χαμηλό ratio για να ελέγξουμε ελαφρά μια ηχογράφιση πολύ νευρική ή να δώσουμε σε μια ηχογράφιση κάποιο από τον απαιτούμενο όγκο. Δεν συνηθίζεται τόσο χαμηλές σχέσεις να χρησιμοποιούνται στην ηχογράφιση ή στη μίξη, είναι όμως πολύ συνηθισμένες κατά τη διάρκεια του mastering μιας ολοκληρωμένης μίξης, κυρίως γιατί με υλικό πλήρους προγράμματος (ολοκληρωμένες δηλαδή μίξεις) οι μεγαλύτερες τιμές στο ratio του compressor

μπορούν να τραβήξουν την προσοχή στη μαγεία πίσω από τη σκηνή ή να αποκαλύψουν ανάσες, το τρίξιμο του δοξαριού πάνω στη χορδή και άλλα.

Επισημάναμε προηγουμένως ότι κάθε εταιρία παραγωγής δυναμικών επεξεργαστών (compressors, expanders, limiters κ.ο.κ) έχει τα δικά της μοναδικά χαρακτηριστικά και ήχο. Ένα μέρος της διασκέδασης στο mastering (αλλά και στη μίξη) είναι η ανακάλυψη των ειδικών χαρακτηριστικών των διαφόρων compressors. Ακόμα και με τις ίδιες ρυθμίσεις κάποια είναι απαλά, αλλά χτυπητά, μερικά αναδεικνύουν τα κρουστά καλύτερα από τα άλλα. Αυτό δεν οφείλεται σε αυτούς καθαυτούς τους χρόνους attack και release, αλλά μάλλον στην καμπύλη (transfer curve) του compressor ή την επιτάχυνση των χρονικών σταθερών του, είτε η συσκευή ανακάμπτει γραμμικά από την μείωση του gain, είτε το gain επιστρέφει στη unity μορφή του σχετικά αργά ή σχετικά γρήγορα. Οι μηχανικοί σχεδιασμού αυτών των συσκευών καταναλώνουν πολύ χρόνο έρευνας στην ανάλυση αυτών των ιδιαίτερων χαρακτηριστικών και το καλύτερο που μπορούμε να κάνουμε εμείς οι κοινοί θνητοί είναι να ακούμε και να κρίνουμε τι είναι αυτό που μας αρέσει.

4.2.7 Ορισμένες επιπλέον και χρήσιμες ρυθμίσεις των compressors:

Μερικοί compressors παρέχουν μία επιπλέον ρύθμιση με το όνομα crest factor (συντελεστής κορυφής)¹⁶, ο οποίος συνήθως εκφράζεται είτε σε dB, είτε σε απόλυτες τιμές. Επίσης, ορισμένοι compressors διαθέτουν στον χρήστη την επιλογή να λειτουργεί ο επεξεργαστής είτε στην RMS τιμή του σήματος, είτε στην quasi peak τιμή του σήματος δηλ. στο μέσο του υψηλότερου peak του σήματος, είτε τέλος στην full peak τιμή του. Αυτό σημαίνει ότι ο compressor δρα είτε στα σημεία του μέσου όρου του σήματος είτε στα σημεία peak ή κάπου ανάμεσα. Φαινομενικά οι compressors με χαρακτηριστικά RMS ακούγονται πιο φυσικά καθώς ανταποκρίνονται στον τρόπο με τον οποίο αντιλαμβάνεται το ανθρώπινο αυτί τον ήχο, αλλά υπάρχουν compressors που λειτουργούν με τις peak to peak τιμές και ακούγονται καλύτερα. Είναι θέμα σχεδιασμού και υλοποίησης της συσκευής.

Το μοντέλο της Weiss DS1 - Mk2, για παράδειγμα, είναι ο πρώτος επεξεργαστής δυναμικών που διαθέτει δυο διαφορετικές χρονικές σταθερές του release time, μία γρήγορη και μία αργή. Ο χρήστης καθορίζει ένα όριο του μέσου όρου διάρκειας του release, όπως π.χ. είναι τα 80ms, πάνω από το οποίο μια μουσική κίνηση αποκαλείται αργή και κάτω από αυτό γρήγορη. Έτσι σε στιγμιαία μεταβατικά του ηχητικού υλικού μπορεί να δοθεί ένα γρηγορότερο release time, αλλά σε παρατεταμένους ήχους ένα πιο αργό release time, πράγμα που έχει ως αποτέλεσμα μια συμπίεση που ακούγεται πιο φυσική, ειδικά στις περιπτώσεις που το ποσοστό της συμπίεσης είναι υψηλό.

¹⁶ Για περισσότερες λεπτομέρειες αναφορικά με τον συντελεστή κορυφής και το πώς αυτός ορίζεται, την RMS τιμή και το πώς αυτή ορίζεται, όπως ομοίως και για την peak to peak τιμή βλέπε το παράρτημα Δ.

4.2.8 Compression και monitoring:

Σχετικά με τη σημασία του συστήματος monitoring κατά τη διαδικασία της επεξεργασίας του compression (και που φυσικά δεν περιορίζεται μόνο σε αυτό το πεδίο) ο Bob Katz σε ένα άρθρο που διαθέτει στην προσωπική του ιστοσελίδα (www.digido.com) αναφέρει τα ακόλουθα:

«θυμάμαι την μίξη μιας ηχογράφησης καθαρής τζαζ όπου χρησιμοποίησα εξαιρετικά ενισχυτικά, καλώδια με εξαιρετικές ωμικές αντιστάσεις συναρτήσεως του μήκους και της διαμέτρου τους καθώς και εξαιρετικά μόνιτορ, εξοπλισμένα με ένα κύκλωμα driver protection, το οποίο ήταν ανενεργό στο μεγαλύτερο μέρος του σήματος και λειτουργούσε μόνο στα πολύ υψηλά peak. Όμως όταν έφτασα στο mastering studio, ανακάλυψα ότι η δυναμική περιοχή της ηχογράφησης ήταν υπερβολικά "jump out" και απαιτούσε κάποια συμπίεση, ένα γεγονός που δεν φαινόταν κατά την μίξη και που αισθάνομαι ότι θα είχε παραμείνει κρυφό αν είχα παρακολουθήσει την μίξη με χαμηλής ισχύος ενισχυτές λυχνίας, π.χ. ένα ζευγάρι set μονομπλόκ (διότι έχουν την ιδιότητα όταν το σήμα ξεπερνά τη δυναμική περιοχή του ενισχυτή, να συμπιέζει ο ίδιος ο ενισχυτής το σήμα. Εξ ου και αυτό που πολλοί αντιλαμβάνονται ως γλυκό ήχο από τη συγκεκριμένη κατηγορία ενισχυτών).

Τελικά παρήγαγα ένα εξαιρετικό, με κοφτό ήχο, master όπου ανησυχούσαμε μήπως οι άνω δυναμικές του ηχητικού υλικού είχαν κάπως υπερβολική επίδραση. Αλλά όταν το master αξιολογήθηκε σε ένα τυπικό δωμάτιο με έντονα στάσιμα κύματα ή όταν, επίσης, αξιολογήθηκε σε ένα τυπικό - συμβατικό σύστημα αναπαραγωγής, οι δυναμικές του ηχητικού υλικού ήταν πνιγμένες σε σχέση με αυτές που αντιλαμβανόμουν στο mastering studio και μάλλον θα "κέρδισαν" περισσότερα αν τους προσφέραμε λίγη περισσότερη "αίσθηση". Έτσι έμαθα ότι αν «ξεχωρίζει λίγο παραπάνω» σε ένα high-headroom σύστημα monitoring, τότε πιθανά θα είναι εντάξει όταν θα παίζεται σε ένα υποδεέστερο σύστημα. Όμως δεν θα μάθετε ποτέ αν κάτι χρειάζεται λίγο παραπάνω compression ή είναι ήδη πολύ συμπιεσμένο ακούγοντας σε ένα σύστημα monitoring, το οποίο δεν έχει υψηλό headroom και κατά συνέπεια συνθλίβει τον ήχο». [19]

4.2.9 Επεξεργασία της δυναμικής περιοχής με Multiband Compression:

Το multiband compression είναι πιθανά η πιο ισχυρή και η εν δυνάμει πιο καταστρεπτική ηχητική διαδικασία που έχει ποτέ εφευρεθεί. Βασικά ένας multiband compressor διαχωρίζει την ηχητική πληροφορία σε δυο, τρεις ή περισσότερες περιοχές βάσει του συχνοτικού περιεχομένου του σήματος, ώστε η δράση του compression σε μια συχνοτική περιοχή του σήματος να μην επηρεάσει τις άλλες. Για παράδειγμα αν τα φωνητικά απαιτούν κάποια μείωση του gain, αυτή η εργασία δεν θα επηρεάσει και το bass drum (ή το αντίστροφο), κάτι που μπορεί να συνέβαινε αν χρησιμοποιούσαμε ένα

full-band compressor. Αυτή είναι ή αξία και η δικαιολόγηση του λόγου για τον οποίο εργαζόμαστε χρησιμοποιώντας multiband compressors οι οποίοι διαχωρίζουν το σήμα σε πολλαπλές συχνотικές περιοχές. Όμως, όπως πολλοί mastering engineers πιστεύουν, το multiband compression έχει, κατά τη γνώμη τους, χρησιμοποιηθεί και προβληθεί υπερβολικά. Και αυτό διότι μπορεί εύκολα να παράγει μη-μουσικό ήχο ή να οδηγήσει μια μίξη σε μονοπάτια που δεν χρειάζεται να βρεθεί. Όπως και στην πραγματική ζωή, σημασία δεν έχει μόνο το χαρακτηριστικό απο μόνο του (εν προκειμένω το multiband compression), αλλά κυρίως ο τρόπος με τον οποίο αυτό αξιοποιείται.

Το multiband compression παρουσιάστηκε πρώτη φορά μάλλον από την TC Electronic στο μοντέλο της M5000 και έπειτα στο πλατιά διαδεδομένο Finalizer. Πιο εξελιγμένο τεχνικά (και με πολύ καλύτερη ποιότητα ήχου) παρουσιάζεται στις μέρες μας στο μοντέλο System6000 και πάλι της TC Electronic. Η Tubetech, επίσης, παράγει έναν multiband compressor λυχνίας τριών συχνотικών περιοχών. Αλλά οι πολλαπλές συχνотικές περιοχές μάλλον είναι περιττές, διότι μια ή δυο δέσμες είναι συνήθως αρκετές. Σπάνια ακόμα και σε μουσικά είδη των οποίων οι ηχογραφήσεις είναι ιδιαίτερες, όπως π.χ. το hip - hop χρειάζονται πάνω από δυο συχνотικές περιοχές κατά την επεξεργασία τους απο έναν multiband compressor για να ακούγονται συμπαγείς και «δυνατές». Γενικότερα χρησιμοποιούμε πάνω από δυο συχνотικές περιοχές κατά την επεξεργασία ενός ηχητικού υλικού μόνον όταν αυτό κρίνεται απολύτως απαραίτητο. Τις περισσότερες φορές χρησιμοποιούμε multiband compression (καθώς και expansion) για να διορθώσουμε κακές μίξεις που δεν μπορούν να επαναμιξαριστούν. Γιατί το κλειδί για ένα σπουδαίο master είναι να αρχίζεις με μια καλή μίξη!

Πότε, όμως, σκεφτόμαστε σοβαρά να επεξεργαστούμε το μουσικό υλικό κάνοντας χρήση του multiband compression;

- Όταν υπάρχει ένα «βαρύ» και κάπως απομονωμένο bass drum και/ ή μπάσο, ο διαχωρισμός της επεξεργασίας σε δύο συχνотικές περιοχές εμποδίζει να υπάρξει συσχέτιση της επεξεργασίας που θα λάβει χώρα στη χαμηλή περιοχή με όλο το υπόλοιπο φάσμα.
- Όταν θέλουμε να αφήσουμε τις κορυφές του audio σήματος (π.χ. τους ήχους των κρουστών) ανεπηρέαστες ενώ επιδιώκουμε να επηρεάσουμε σημεία του audio σήματος τα οποία δεν περιέχουν έντονα μεταβατικά. Οι κρουστικοί ήχοι περιέχουν περισσότερη ενέργεια στις υψηλές συχνότητες από τους συνεχείς ήχους, έτσι χρησιμοποιώντας το multiband compression και χωρίζοντας το φάσμα σε μία χαμηλή και μία υψηλή συχνотική περιοχή μπορούμε, με αυτό τον τρόπο, να επιτύχουμε την χρήση απαλότερης ή και καθόλου συμπίεσης στις υψηλές συχνότητες (πχ υψηλό threshold, χαμηλό ratio).
- Όταν υπάρχουν πολλά συριστικά. Ένας τρόπος για να ελεγχθούν τα συριστικά σε ένα μουσικό κομμάτι μπορεί να επιτευχθεί χρησιμοποιώντας επιλεκτική συμπίεση στο εύρος των 3 ως 9 kHz (η ακριβής συχνότητα ανά περίπτωση

θα πρέπει να επιλεχθεί ακούγοντας προσεκτικά τα φωνητικά). Δοκιμάζουμε ένα πολύ γρήγορο attack, θέτουμε στο release μία μέση τιμή και επιλέγουμε το bandwidth του compressor για τη συγκεκριμένη συχνотική περιοχή να είναι όσο στενό χρειάζεται (συνήθως πολύ στενό).

- Όταν η μίξη είναι κακή ή συγκεκριμένα χαρακτηριστικά στοιχεία της παρουσιάζονται ασθενικά ή υποτονικά, η επεξεργασία κάνοντας χρήση του multiband compression μπορεί να σώσει την κατάσταση, αν υποθέσουμε ότι ένα remix δεν είναι δυνατό. Για παράδειγμα εν ένα ραπ μουσικό κομμάτι έχει μιξαριστεί με πολύ χαμηλά τα φωνητικά και πάρα πολύ δυνατά τα κρουστά, και δεν ήταν δυνατό να γίνει remix, τότε συμπιέζοντας και μετά αυξάνοντας το level των συχνοτήτων στα φωνητικά (περίπου στα 250 Hz) μπορούμε να κάνουμε «remix» το κομμάτι και πολύ απλά να ανεβάσουμε την ένταση των φωνητικών στο κομμάτι.

Όμως, πριν καταφύγουμε στη χρήση του multiband compression, καλό είναι πρώτα:

- Να δούμε αν απλά η αύξηση του attack time σε έναν full - range compressor επιτρέπει σε επαρκή ενέργεια των κορυφών του σήματος να περάσει. Ή αντ' αυτού, μπορούμε να δοκιμάσουμε τη χρήση ενός upward expander (περιγράφεται σε επόμενη παράγραφο).
- Δοκιμάζουμε να χρησιμοποιήσουμε όσο το γίνεται λιγότερες συχνοτικές περιοχές, μόνο δύο αν είναι δυνατόν. Αυτό για να αποφευχθεί μια πιθανή ολίσθηση φάσης (phase shift) και αφύσικες συνέπειες μεταξύ των στοιχείων του μουσικού κομματιού, που μπορεί να αποβούν επικίνδυνες.

4.2.10 Equalization ή multiband compression:

Όταν η επεξεργασία με τη χρήση multiband compression είναι διαθέσιμη, η διαχωριστική γραμμή μεταξύ του equalization και της δυναμικής επεξεργασίας γίνεται ιδιαίτερα θολή, γιατί το συνολικό output level κάθε συχνοτικής περιοχής του multiband compressor, θα μπορούσε να πει κανείς απλοποιώντας τη συνθήκη, ότι σχηματίζουν ένα απλό equalizer.

Ο τρόπος με τον οποίο συμπεριφερόμαστε συνήθως σε πρακτικό επίπεδο είναι να χρησιμοποιούμε equalization όταν τα μουσικά όργανα χρειάζονται τροποποίηση σε ότι αφορά τις φασματικές τους συνιστώσες· απο κει και πέρα χρησιμοποιούμε το multiband compression για να παρέχουμε την αναγκαία φασματική ισορροπία (spectral balancing) στο μουσικό υλικό αν αυτό κρίνεται απαραίτητο. Για παράδειγμα, ένα μουσικό κομμάτι μπορεί να ακούγεται «σκληρό» σε υψηλές εντάσεις (κάτι που εξηγείται απο τις καμπύλες ίσης ακουστότητας των Fletcher & Munson), σε αυτή, λοιπόν, την περίπτωση κάνοντας χρήση των δυνατοτήτων που μας παρέχει το multiband compression μπορούμε να συμπιέσουμε λίγο περισσότερο τις υψηλές

συχνότητες και πιθανότητα με αυτό τον τρόπο να επιφέρουμε τα επιθυμητά αποτελέσματα.

Παρόλα αυτά, δεν υπάρχει κάποια πολύ συγκεκριμένη συνταγή, υπάρχουν μόνο γενικές προσεγγίσεις όπως αυτές που αναφέρθηκαν παραπάνω. Για αυτό το λόγο, σε ένα project δεν είναι απίθανο να χρησιμοποιείται τόσο το equalization, όσο και το multiband compression. Για την ακρίβεια στις περισσότερες περιπτώσεις συμβαίνει ακριβώς αυτό. Συνεπώς το ζήτημα που συνήθως προκύπτει είναι με ποια σειρά θα λάβουν χώρα οι εργασίες κατά τη διάρκεια του project. Στην πράξη η απάντηση συνήθως δίδεται χρησιμοποιώντας πρώτα το equalization και στη συνέχεια το multiband compression. Και αυτό όμως δεν είναι δεδομένο για αυτό πολλές φορές πειραματιζόμαστε τόσο με τη σειρά με την οποία θα λάβουν χώρα οι διεργασίες, όσο και με το γεγονός του αν η προβληματική που έχουμε εντοπίσει επιλύεται πλήρως με μία από τις δύο διαδικασίες. Αν αυτό συμβαίνει τότε δεν υπάρχει λόγος να προχωρήσουμε και στην άλλη.

4.2.11 Compression, Στερεοφωνική Εικόνα και Depth μιας ηχογράφησης:

Ένας σίγουρος τρόπος για να καταστρέψουμε το βάθος (depth), αυτό που αποκαλούμε χωρικό στοιχείο μιας ηχογράφησης είναι το υπερβολικό compression. Το compression, γενικά, αναδεικνύει τις «εσωτερικές φωνές» του μουσικού υλικού. Όργανα που ήταν στην πίσω μεριά της «ορχήστρας» έρχονται μπροστά, και το ηχητικό περιβάλλον, το βάθος, το πλάτος και γενικότερα τα χωρικά στοιχεία μιας μίξης υποβαθμίζονται. Αλλά δεν πρέπει κάθε όργανο να βρίσκεται μπροστά. Θα πρέπει να δίνουμε ιδιαίτερη προσοχή, συγκρίνοντας το original – αρχικό ηχητικό αρχείο με το master που προκύπτει. Αυτό θα πρέπει να συμβαίνει για δύο λόγους, αφενός για να μπορούμε να διαπιστώσουμε αν πράγματι έχουμε διορθώσει τα προβληματικά σημεία που έχουμε εντοπίσει, και αφετέρου για να είμαστε σίγουροι ότι αυτό που έχει επιτύχει ο ηχολήπτης στη μίξη (και δεν υπάρχει λόγος εμείς να το τροποποιήσουμε) παραμένει ως έχει, π.χ. τα χωρικά στοιχεία στο προκείμενο.

4.3 Upward:

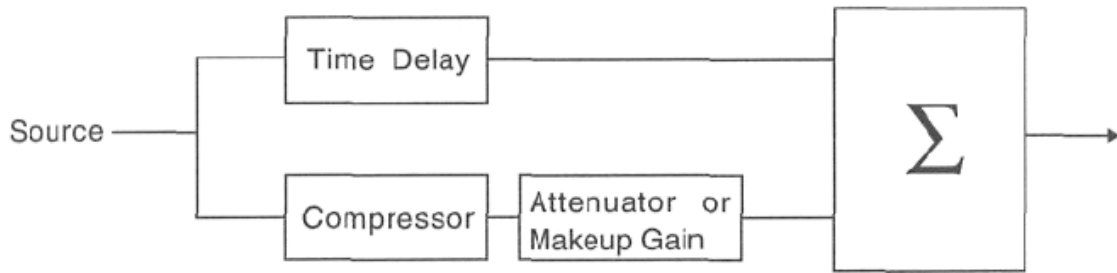
Σε αυτή την παράγραφο παρουσιάζεται μία διαδικασία που θα έπρεπε να είναι μέρος του λεξιλογίου κάθε μηχανικού ήχου. Για να επιτύχουμε, όμως με αυτήν πρέπει να μάθουμε να σκεφτόμαστε με αντίθετο τρόπο, από αυτόν που μέχρι τώρα έχει κατατεθεί στις σελίδες αυτού του κεφαλαίου, αλλά αξίζει τον κόπο.

4.3.1 Upward Compression:

Η υπερβολική συγκέντρωση στη χρήση των downward compressors, καθιστά εύκολο το γεγονός του να παραβλέψουμε το ψυχοακουστικό φαινόμενο ότι το αυτί συγχωρεί πολύ ευκολότερα την αύξηση του level των απαλών τμημάτων του audio σήματος παρά την αδέξια συμπίεση των δυνατών τμημάτων του. Το δεύτερο έχει την αίσθηση μιας τεχνητής απώλειας, ενώ το πρώτο μπορεί να έχει πολύ φυσική αίσθηση.

Σε αυτό το σημείο θα παρουσιάσουμε μια αξιολογική τεχνική συμπίεσης που έχει ωριμάσει μέσα στο χρόνο. Ο νεοζηλανδός μηχανικός ραδιοφώνου Richard Huise ανέπτυξε την πρακτική του parallel compression (παράλληλης συμπίεσης), που επιτυγχάνει με εξαιρετική ηχητική διαύγεια αυτό που είναι γνωστό ως upward compression. Ο Huise χρησιμοποιούσε αναλογικές συσκευές στην ανάπτυξη της συγκεκριμένης τεχνικής και είχε αποδεκτά αποτελέσματα, πίστευε όμως πως η ψηφιακή εφαρμογή της συγκεκριμένης τεχνικής θα τα κατάφερνε καλύτερα και συνεπώς τη δοκίμασε και σε αυτό το πεδίο. Προέκυψε πως η ψηφιακή εκδοχή αυτής της τεχνικής επέφερε καλύτερα αποτελέσματα και πλέον (δηλ. στις μέρες μας) χρησιμοποιείται περισσότερο από την αναλογική της εκδοχή. Ο τρόπος με τον οποίο υλοποιείται η συγκεκριμένη τεχνική είναι σχετικά απλός: Έστω ότι εφαρμόζουμε τη συγκεκριμένη τεχνική σε αναλογικό περιβάλλον, χρησιμοποιούμε δε για το λόγο αυτό δύο από τα modules μιας αναλογικής κονσόλας και τα τροφοδοτούμε με το ίδιο σήμα από το πολυκάναλο. Στη συνέχεια και για το ένα από τα δύο κανάλια (δεν έχει σημασία ποιο) στέλνουμε σε σειρά το σήμα στην είσοδο ενός compressor και έχοντας θέσει τη ρύθμιση PRE στο auxiliary το επαναφέρουμε στο ίδιο κανάλι μέσω του INSERT RETURN. Στη συνέχεια καθορίζουμε το level των δύο modules και τα στέλνουμε στη STEREO έξοδο για να μιξαριστούν. Στην ουσία δηλαδή αυτό που κάνουμε είναι να αθροίζουμε το αρχικό σήμα μαζί με μία παραλλαγή του που προκύπτει από την έξοδο του compressor, έχοντας παράλληλα όμως τη δυνατότητα να καθορίζουμε το ποσοστό του level κάθε σήματος ξεχωριστά. Στον ψηφιακό χώρο, τώρα, είναι δυνατή η άθροιση της πηγής με τον compressor χωρίς επιπτώσεις στην ποιότητα του ηχητικού σήματος, χρησιμοποιώντας μια επακριβώς ορισμένη χρονική καθυστέρηση μέσω ενός delay για το «dry» σήμα έτσι ώστε να μην υπάρχει καμία απολύτως χρονική διαφορά μεταξύ του «dry» σήματος με αυτό που λαμβάνουμε από την έξοδο του compressor. Η παραπάνω διαδικασία για ένα μόνο από τα δύο κανάλια του στερεοφωνικού σήματος παρουσιάζεται στο σχήμα 4.3.1.1 της επόμενης σελίδας.

Σε αυτό το σημείο είναι σημαντικό να αναφέρουμε πως η παραμόρφωση που προκύπτει από το parallel compression συνήθως είναι πολύ χαμηλότερη από αυτή που προκύπτει από το στάνταρ downward compression, καθώς το μεγαλύτερο μέρος του σήματος απολαμβάνει της γραμμικής οδού (linear path) αναφορικά με την επεξεργασία



Σχ. 4.3.1.1: Η τεχνική του parallel compression δοσμένη σε διάγραμμα ροής. Παρατηρούμε την ύπαρξη ενός τμήματος καθυστέρησης, ο λόγος για τον οποίο αυτό συμβαίνει είναι διότι το σήμα μετά την έξοδο του απο τον compressor θα έχει μία μικρή χρονική καθυστέρηση σε σχέση με την ανεπεξέργαστη εκδοχή του, συνεπώς με τη χρήση μιας μικρής καλά καθορισμένης χρονικής καθυστέρησης στο «dry» σήμα θα αποφύγουμε φαινόμενα ολίσθησης φάσης και comb filtering.

του, και η μη γραμμική οδός (nonlinear path) προστίθεται στο κυρίως σήμα¹⁷. Το ποσό της συμπίεσης, τώρα, ελέγχεται από τον attenuator ή το makeup gain του compressor αντίστοιχα. Στόχος της τεχνικής είναι ο compressor να συμβάλει όσο το δυνατόν λιγότερο στον συνολικό ήχο όσο το level του σήματος αυξάνει. Αυτό επιτυγχάνεται με τη χρήση πολύ χαμηλού threshold, ώστε ο compressor να μπαίνει στη διαδικασία μείωσης του gain σχεδόν καθόλη τη διάρκεια του σήματος και συνεπώς με αυτό τον τρόπο να συνεισφέρει λιγότερό από το «dry» σήμα στο τελικά παραγόμενο ηχητικό υλικό.

Εν συνεχεία προτείνονται οι καταλληλότερες ρυθμίσεις για την τεχνική του parallel compression, όπως αυτές προέκυψαν από τα πειράματα του Richard Hulse:

- Το threshold ορίζεται περίπου στα -50dBFS. Ένα πολύ χαμηλό threshold εξασφαλίζει ότι το parallel compression θα είναι στη διαδικασία του gain reduction καθόλη τη διάρκεια των δυνατών, τουλάχιστον, τμημάτων του ηχητικού υλικού. Επειδή το level του σήματος στην έξοδο του compressor έχει συμπιεστεί θα συμβάλει μόνο αμελητέα στο συνολικό level. Είναι σα να προσθέτουμε ένα δεύτερο σήμα με level μικρότερο κατά 20dB ή περισσότερο, στο πρώτο σήμα. Το σήμα άθροισης λοιπόν δεν θα συμβάλει με αντιληπτό τρόπο όπως είναι φυσικό.
- Το attack time θα πρέπει να είναι όσο το δυνατόν ταχύτερο. Ένα millisecond ή λιγότερο αν είναι δυνατόν. Αυτό εξασφαλίζει ότι τα μεταβατικά στοιχεία του ηχητικού υλικού θα διατηρηθούν, γιατί μόλις προκύπτει μια δυνατή μετάβαση ο compressor θα μειώνει το gain. Βοηθητικό στοιχείο στο πλαίσιο της συγκεκριμένης τεχνικής είναι ο compressor να διαθέτει ρύθμιση look ahead, που σημαίνει ότι έχει μια ενσωματωμένη χρονική καθυστέρηση η οποία του

¹⁷ Αυτός ο τρόπος επεξεργασίας αποτελεί επίσης και τη βασική αρχή του συστήματος Dolby A/SR για την αποθρομβοποίηση σημάτων. Δηλ. χρησιμοποιείται το αρχικό σήμα αποθρομβοποιημένο με ασυμπίεστο, καθώς και μία συμπιεσμένη εκδοχή του. Αυτό συμβαίνει με τη λογική του να γίνεται η επεξεργασία αντιληπτή κατά το μικρότερο δυνατό βαθμό.

επιτρέπει να «βλέπει» τα εισερχόμενα σήματα και συνεπώς με αυτό τον τρόπο μπορεί να ανταποκριθεί καλύτερα, διότι ουσιαστικά είναι σα να έχουμε έναν compressor με μηδενικό attack.

- Το ratio του compressor πρέπει να κυμαίνεται απο 2:1 έως 2,5:1 (προτιμητέα είναι η τιμή 2,5:1). Παρόλα αυτά, όμως, υπάρχουν περιθώρια πειραματισμού αναφορικά με το ποσοστό της συμπίεσης και το ratio και αυτό κάθε φορά καθορίζεται απο το είδος του ηχητικού υλικού και τα ιδιαίτερα χαρακτηριστικά του. Σε κάθε περίπτωση οι παραπάνω τιμές αποτελούν ενδεικτικό οδηγό και για αυτό ο τελικός κριτής δε μπορεί να είναι άλλος πέρα απο τα αυτιά μας.
- Το release time του compressor κυμαίνεται στο μέσο όρο των τιμών που χρησιμοποιούμε για τις πλείστες περιπτώσεις. Τα πειράματα του Huise έδειξαν ότι τιμές μεταξύ των 250 και 350 millisecond είναι οι καταλληλότερες για να αποφευχθούν οι αναπνοές του τραγουδιστή/τριας ή τα συριστικά, αν και σε περιπτώσεις που τα χωρικά στοιχεία του μουσικού κομματιού είναι πολύ έντονα, ιδιαίτερα για a capella κομμάτια, μπορεί να χρειάζονται ως και τα 500ms στο release για να αποφευχθεί η υπερβολική έμφαση στις ουρές της αντήχησης (reverb tails).
- Το output level ή το makeup gain ρυθμίζονται κατά βούληση και σύμφωνα με αυτό που εξυπηρετεί καλύτερα το μουσικό κομμάτι.

Για να καθοριστεί ο χρόνος του delay που χρειάζεται έτσι ώστε να αντισταθμίσουμε αυτόν που προκύπτει απο τον compressor, ρυθμίζουμε αρχικά τον compressor με μία τιμή του ratio 1:1 και έξοδο τη unity gain του compressor. Είναι ευνόητο πως για να ορίσουμε τη σωστή τιμή του delay για να ευθυγραμμίσουμε τα δύο σήματα, αυτά πρέπει να έχουν το ίδιο ακριβώς level (θα δούμε στη συνέχεια γιατί). Οι τυπικές τιμές καθυστέρησης κυμαίνονται απο 5 έως 10 samples (που χρονικά ισούνται με 0,11 έως 0,22 msec), αλλά μπορεί να είναι και πολύ μεγαλύτερες αν κάνουμε χρήση του look ahead του compressor. Στην περίπτωση που έχουμε τηρήσει όλα τα παραπάνω, και το κυριότερο έχουμε ορίσει τη σωστή τιμή στο delay, τότε θα προκύψει απο την άθροιση των δύο σημάτων ένα νέο σήμα το οποίο θα είναι όμοιο με τα προηγούμενα αλλά θα έχει κατά 6 dB υψηλότερο level (συμφασικές ηχητικές πηγές ίδιου πλάτους, εξ ου και το γιατί πρέπει τα σήματα να έχουν το ίδιο πλάτος). Πέρα απο απλά ημιτονοειδή σήματα, ο ροζ θόρυβος ενδείκνυται επίσης για τη διεξαγωγή αυτού του test. Έτσι είμαστε σίγουροι πως κατά τη διάρκεια της επεξεργασίας του ηχητικού υλικού με τη συγκεκριμένη τεχνική του parallel compression δε θα προκύψουν ανεπιθύμητα φαινόμενα, όπως για παράδειγμα comb filters.

Η συγκεκριμένη τεχνική μπορεί να εφαρμοστεί με επιτυχία και σε ψηφιακούς σταθμούς εργασίας (DAW) όπως είναι τα Pro Tools, το SADiE, και το Sonic Studio HD. Υπάρχουν επίσης και ψηφιακές hardware συσκευές που περιλαμβάνουν το parallel

compression στο ρεπερτόριο τους. Η Weiss το ενσωμάτωσε στο DS1 - MK2 και η TC Electronic στο System 6000.

Έως αυτό το σημείο εξετάσαμε την τεχνική του parallel compression με τη χρήση full range compressor. Προεκτείνοντας της δυνατότητας της συγκεκριμένης τεχνικής, το parallel compression μπορεί να εφαρμοστεί κάνοντας χρήση των δυνατοτήτων των multiband compressors για να προσδώσει «όγκο» ή να τονώσει την παρουσία μιας συγκεκριμένης συχνοτικής περιοχής ξεχωριστά, ή ακόμα και σε ένα μεμονωμένο μουσικό όργανο σε κάποιες περιπτώσεις. Ακόμα για να τονώσει την παρουσία σε χαμηλού level μέρη του μουσικού σήματος. Απο τα παραπάνω, προκύπτει πως η τεχνική του parallel compression με τη χρήση multiband compressors αφορά περισσότερο, στην ουσία της όλης διαδικασίας μια τονική ισορρόπηση του ηχητικού υλικού, παρά μια πρακτική εφαρμογή συμπίεσης του. Συνεπώς θα μπορούσε να πει κανείς πως αποτελεί και μια ιδιότυπη μορφή equalization.

Το καλό με τον ηχητικό χαρακτήρα που προσδίδει αυτή η τεχνική συμπίεσης είναι ότι βοηθά τη χαμηλή περιοχή (κυρίως) να αποκτήσει μια επιπλέον αίσθηση, χωρίς ωστόσο αυτό να αποβαίνει εις βάρος του ευρύτερου τονικού ισοζυγίου της μίξης και δίχως να καταστρέφει τις δυναμικές τις διακυμάνσεις. Δηλαδή αυτό που συμβαίνει είναι ότι το σώμα του ήχου «παχαίνει» χωρίς όμως να καταστρέφονται τα μεταβατικά χαρακτηριστικά του μουσικού υλικού. Ακόμα και με αυστηρές ρυθμίσεις, η τεχνική του parallel compression ακούγεται καλύτερα και πιο ευχάριστη απο κάθε αυστηρό downward compression το οποίο τείνει να συνθλίβει τον ήχο.

Όπως κάθε επεξεργασία, αν το upward compression (και εν προκειμένω το parallel compression) «πουσαριστεί» πολύ θα τραβήξει την προσοχή με σίγουρα όχι θετικό τρόπο. Το πρώτο ακουστικό σύμπτωμα θα είναι αυξημένες παρατάσεις της διάρκειας των μουσικών οργάνων και η εμφατική αντήχηση. Στο τέλος οι ανάσες απο τα φωνητικά ή διάφορα pops θα κάνουν ιδιαίτερα αισθητή την παρουσία τους. Αυτά τα συμπτώματα μπορούν μερικές φορές να μειωθούν με την αύξηση του release time του compressor. Όμως, αν το ηχητικό υλικό είναι τόσο «ανοιχτό» ώστε αυτού του τύπου η επεξεργασία συνεχίζει να τραβάει την προσοχή, τη μόνη λύση αποτελεί η εγκατάλειψη αυτού του είδους επεξεργασίας και αυτό διότι σημαίνει πως πολύ απλά δεν αποδίδει.

**Noise Reduction
&
Audio Mastering**

5. Noise Reduction - Εισαγωγή:

Ο ανθρωπολόγος Benjamin Whorf παρατήρησε ότι οι Εσκιμώοι έχουν πολλές λέξεις για το χιόνι. Κατά παρόμοιο τρόπο, οι μηχανικοί ήχου διακρίνουν ένα μεγάλο αριθμό κατηγοριών αυτού που συνολικά αποκαλούμε ως θόρυβο. Οι μη γνωρίζοντες, γενικά, δεν διακρίνουν αυτό που αποκαλούμε παραμόρφωση από αυτό που αποκαλούμε θόρυβο, αλλά για μας μια τέτοια κατηγοριοποίηση είναι χρήσιμη: Η παραμόρφωση είναι μια υποκατηγορία της γενικής κατηγορίας που ονομάζουμε θόρυβο, είναι ένα είδος θορύβου που σχετίζεται με το σήμα. Η παραμόρφωση μπορεί να έχει χαμηλό level και να δρα όπως αυτό που συνήθως ονομάζουμε θόρυβο, ή μπορεί να έχει υψηλό level και να γίνεται ιδιαίτερα ενοχλητική εμφανιζόμενη στα peaks του audio σήματος. Ο θόρυβος από μόνος του μπορεί να είναι συνεχής ή διακοπτόμενος, τυχαίος ή ψευδοτυχαίος, να περιέχει αναγνωρίσιμους παράγοντες συχνότητας ή να μην περιέχει, να ηχεί σαν crack, click, tik (κυρίως στην υψηλή συχνότητα περιοχή) ή ως pop (κυρίως στη χαμηλή συχνότητα περιοχή). Κάθε είδος ενοχλητικού θορύβου αναζητά την δική του τεχνική θεραπεία, αλλά η καλύτερη θεραπεία είναι απλά να αγνοήσουμε τον θόρυβο! Συχνά οι μηχανικοί ήχου τείνουν να ξεχνούν ότι το αυτί έχει ένα φυσικό μηχανισμό μείωσης του θορύβου που μας δίνει την δυνατότητα να ξεχωρίζουμε το σήμα από τον θόρυβο, και να ακούμε τις θαμμένες στον θόρυβο πληροφορίες. Έτσι το κλειδί για τη μείωση του θορύβου που επιφέρει καλά (ακουστά) αποτελέσματα, δεν είναι να απομακρύνουμε όλο τον θόρυβο αλλά να αποδεχτούμε μια μικρή βελτίωση στο σηματοθορυβικό λόγο του σήματος ως νίκη. Θα πρέπει να θυμόμαστε ότι τα δυνατότερα σήματα μασκάρουν το θόρυβο, και επίσης ότι το κοινό δεν επικεντρώνεται στον θόρυβο ως πρόβλημα. Οι ακροατές προσέχουν την μουσική και αυτό θα έπρεπε να κάνει και ο μηχανικός. Άρα πριν αναλογιστούμε για μια τεχνική μείωσης του θορύβου, πρέπει να κρίνουμε αν ένας θόρυβος λειτουργεί πραγματικά αποπροσανατολιστικά αναφορικά με την ηχητική πληροφορία.

Οι μέθοδοι μείωσης του θορύβου που περιγράφονται σε αυτό το κεφάλαιο είναι όλοι «ενός τέλους» (single - ended) αντίθετα από τους συμπληρωματικούς (complementary). Το σύστημα Dolby είναι ένα παράδειγμα συμπληρωματικού ή «δύο βημάτων» σύστημα μείωσης του θορύβου το οποίο εφαρμόζει μια διαδικασία κατά την ηχογράφηση και μια αντίστροφη διαδικασία κατά την αναπαραγωγή. Σημαντικό δεδομένο: κανένα single ended σύστημα μείωσης θορύβου δεν είναι τέλειο, όλα τα συστήματα μείωσης θορύβου παίρνουν ένα ποσό του επιθυμητού σήματος μαζί με τον θόρυβο. Επίσης, παράγωγα υπέρ-επιθετικής αποθορυβοποίησης (εννοούμε πως έχουμε θέσει ακραίες τιμές στον επεξεργαστή που επιτελεί την αποθορυβοποίηση. Αυτό μπορεί να συμβεί αρκετά συχνά και, δυστυχώς, πολλές φορές αυτό συμβαίνει δίχως καν (!) να το αντιλαμβάνονται πολλοί μηχανικοί), περιλαμβάνουν: comb-filtering ή phasing noises, γνωστά και ως space monkeys, και χαμηλού level thumps (pops).

Πολύ επιθετικό noise reduction μπορεί, επίσης, να αφαιρέσει το κλίμα και την ατμόσφαιρα από μια ηχογράφηση, δηλαδή να της στερήσουν βασικά χαρακτηριστικά της στοιχεία.

«Η δυσκολία έγκειται στο γεγονός ότι η αντήρηση τείνει να εξασθενεί συν το χρόνο και να συσχετίζεται αναφορικά με το πλάτος που λαμβάνει με το θόρυβο καθιστώντας τα πλάτη αντήρησης και θορύβου συγκρίσιμα. Όμως πολλές από τις πληροφορίες χωρικού καθορισμού των μουσικών οργάνων καθώς και της ατμόσφαιρας του μουσικού κομματιού που λαμβάνουμε προέρχονται από την αντήρηση. Γι' αυτό, αν αφαιρέσουμε ποσοστό της αντήρησης στην προσπάθειά μας να απαλλαγούμε απο το θόρυβο, στην πραγματικότητα είναι σα να αφαιρούμε τους τοίχους, το πάτωμα και το ταβάνι ενός δωματίου».¹⁸ [20]

Η εταιρία Sonic Solutions με το προϊόν της No Noise και η εταιρία Cedar με το De-Noise επιτρέπουν λεπτές ρυθμίσεις της συχνотικής απόκρισης (δηλ. της ενέργειας ανά συχνότητα) της καμπύλης που θα εφαρμοσθεί στο ηχητικό υλικό για τη μείωση του θορύβου, και ένας ικανός μηχανικός θα φτιάξει την καμπύλη αυτή κατορθώνοντας να επιφέρει τον καλύτερο δυνατό συμβιβασμό μεταξύ του ποσοστού του θορύβου που πρέπει να απομακρυνθεί και του κομβικής σημασίας ζητήματος του να μη γίνει αντιληπτή η διαδικασία μείωσης θορύβου. Ο μηχανικός πρέπει διαρκώς να κρατά σε εγρήγορη την προσοχή του, γιατί όσο περισσότερος θόρυβος απομακρύνεται, τόσο περισσότερος θόρυβος αποκαλύπτεται (ο θόρυβος κρύβει άλλους θορύβους από κάτω)! Είναι σαν το ξεφλούδισμα του κρεμμυδιού. Αν αφαιρέσουμε μερικούς τριγμούς, για παράδειγμα, από το δεξί κανάλι, ξαφνικά μπορεί να ακουστούν κάποια tick που προηγουμένως δεν ήταν αντιληπτά από το αριστερό. Σε όλες τις περιπτώσεις είναι αναγκαία η προσεκτική σύγκριση μεταξύ του αρχικού ηχητικού υλικού και του επεξεργασμένου για να είναι σίγουρο πως η μουσική δεν έχει ζημιωθεί. Κατά ειρωνικό τρόπο όσο πιο «ήσυχη» είναι η αυθεντική ηχογράφηση, τόσο πιο αποτελεσματική μπορεί να είναι η διαδικασία του noise reduction. Με άλλα λόγια όσο πιο διαχωρισμένο είναι το αυθεντικό σήμα από τον θόρυβο, τόσο πιο εύκολα μπορεί το σύστημα μείωσης του θορύβου να μειώσει τον θόρυβο χωρίς να πειράξει το σήμα. Έτσι μια ηχογράφηση με υψηλό ποσοστό θορύβου στο σήμα, πιθανά δεν μπορεί να διορθωθεί χωρίς την παραγωγή ταυτόχρονα ορισμένων artifacts όπως αυτά έχουν αναφερθεί παραπάνω.

¹⁸ Η αναφορά αυτή προέρχεται απο τον Gordon Reid, τεχνικό της εταιρίας Cedar, σε μία διαδικτυακή «συζήτηση» η οποία έλαβε χώρα στο forum Mastering webboard.

5.1 Τεχνικές μείωσης θορύβου / Απλό φιλτράρισμα (simple filtering):

Ένα μουσικό τμήμα, σε κάποιο μέρος του μουσικού υλικού με ενοχλητικό υψηύχνο θόρυβο που δεν περιέχει όργανα υψηλών συχνοτήτων μπορεί να διορθωθεί με ένα απλό equalizer εστιασμένο στην υψηλή συχνοτική περιοχή. Για παράδειγμα, ένα σόλο ηλεκτρικού πιάνου στην εισαγωγή ενός τραγουδιού μπορεί να «σφυρίζει», αλλά αυτός ο θόρυβος θα μασκαριστεί όταν μπουν τα υπόλοιπα όργανα. Αυτή η περίπτωση, μάλλον, χρειάζεται επιλεκτικό equalization, για παράδειγμα ας πούμε 1 ως 4 dB dip περίπου στα 3-5 kHz (σε αυτό το εύρος το αυτί είναι περισσότερο ευαίσθητο), ενεργό μόνο κατά την εισαγωγή του πιάνου. Όμως ακόμα και με αυτό τον τρόπο εφαρμογής του, το φίλτρο θα επηρεάσει τις αρμονικές του πιάνου, άρα πρέπει να αποφασίσουμε για το πώς και το αν θα το χρησιμοποιήσουμε με γνώμονα το ηχητικό αποτέλεσμα που προκύπτει από αυτή την επέμβαση μας.

Εν συνεχεία, τα p-pops είναι ένας τύπος θορύβου που σχετίζεται με το σήμα, δηλαδή μια μορφή παραμόρφωσης, και επειδή είναι κυρίως χαμηλής συχνότητας μπορούν να δουλευτούν με ένα επιλεκτικό high-pass φίλτρο, συνήθως γύρω από τα 100Hz αλλά μερικές φορές ως και τα 400 Hz. Αν το φίλτρο εφαρμοστεί για πολύ μικρό χρονικό διάστημα (όσο δηλαδή διαρκούν συνήθως και τα p - pops), το αποτέλεσμα μπορεί να είναι απαλλαγμένο από ενδεχόμενα artifacts που μπορεί να παραχθούν (ή τέλος πάντων αυτά να είναι ανεπαίσθητα στο αυτί). Σε ένα DAW σύστημα, μπορεί κανείς να «κιάσει» με το EQ ένα μικρό τμήμα, μετά, χρησιμοποιώντας τον crossfade editor, να μικρύνει το μέγεθος εφαρμογής του φίλτρου όσο είναι και η διάρκεια του p-pop. Με λίγη εξάσκηση η τεχνική αυτή μπορεί να είναι πάρα πολύ γρήγορη στην εφαρμογή της. Είναι επίσης δυνατό να αφαιρεθεί με επέμβαση τύπου editing (αν μας το «επιτρέπει» το ηχητικό υλικό) το ενοχλητικό μέρος ενός p-pop.

5.2 Narrow-band Expansion:

Ο τρόπος με τον οποίο χρησιμοποιούνται οι compressors τόσο κατά τη μίξη, όσο και κατά την περίοδο που το ηχητικό υλικό βρίσκεται στη φάση του mastering μπορεί να αναδείξουν θορύβους στο αρχικά ηχογραφημένο ηχητικό υλικό, όπως π.χ. είναι τα συριστικά, το θόρυβο των ενισχυτών από τις κιθάρες και το μπάσο, και όλα αυτά μπορούν να θεωρηθούν προβλήματα ή «μέρος του ήχου». Αυτό είναι που κάνει την εργασία μας πολύ υποκειμενική. Από τη στιγμή που η χρήση των compressors χειροτέρευε τον σηματοθορυβικό λόγο του μουσικού υλικού, τη μόνη λύση αποτελούν οι expanders. Χρησιμοποιώντας έναν multiband (downward) expander, και κάνοντας χρήση μιας από τις δέσμες που διαθέτει επικεντρωμένη στα 3 - 5 kHz, με ελάχιστη, όπως 1 ως 4dB, μείωση του gain μπορεί να είναι πολύ αποτελεσματική και, αν γίνει σωστά, αόρατη στο αυτί. Τυπικά αυτές οι μονάδες έχουν 3 ως 4 συχνοτικές δέσμες, αλλά εμείς θα χρησιμοποιήσουμε μόνο μία. Αρχίζουμε βρίσκοντας το threshold, με ένα

αρχικά υψηλό ratio για το expansion, και γρήγορο attack και release time. Βάζουμε ως στόχο ένα threshold μόλις πάνω από το επίπεδο θορύβου. Αν επιτύχουμε στα παραπάνω θα ακούσουμε άσχημα κροταλίσματα και χτυπήματα του noise floor λόγω των πολύ γρήγορων χρονικών σταθερών που έχουμε θέσει στον expander. Εν συνεχεία μειώνουμε το ratio, κάτω από 1:2, ίσως και 1:1,1, και αυξάνουμε τη διάρκεια του release time έως ότου να υπάρχει λίγη ή και καθόλου μετατροπή του noise floor. Υπερβολικό expansion και θα ακούσουμε την παραγωγή artifacts όπως η μείωση των χωρικών στοιχείων του μουσικού κομματιού. Το attack time συνήθως πρέπει να είναι πολύ γρηγορότερο από το release time, έτσι ώστε να μην επηρεάζονται τα γρήγορα crescendos. Ανάλογα με την μουσική, τα δυναμικά της χαρακτηριστικά, και το αυθεντικό της SNR (Signal to Noise Ratio), αυτή η λεπτή προσέγγιση μπορεί να παράγει απαλλαγμένη από artifacts μείωση θορύβου. Οι άλλες δέσμες του expander θα πρέπει να προσπερνιούνται ή το ratio σε αυτές να καθορίζεται στο 1:1. Ένας καλός expander θα πρέπει να έχει look-ahead delay, έτσι ώστε να διατηρούνται τα μεταβατικά στοιχεία του ηχητικού υλικού. Αν η προσέγγιση με τη χρήση expander δεν λειτουργεί, πρέπει να εφαρμόσουμε πιο εξειδικευμένους, καθορισμένους επεξεργαστές μείωσης θορύβου.

5.3 Σύνθετο φιλτράρισμα (complex filtering):

Ο θόρυβος στον οποίο μπορούμε να αναγνωρίσουμε συγκεκριμένα συχνотικά χαρακτηριστικά (απο δω και πέρα θα τον αναφέρουμε ως τονικό θόρυβο), μπορεί να ελαττωθεί χρησιμοποιώντας επιλεκτικό φιλτράρισμα στενού εύρους ζώνης στην κρίσιμη αυτή συχνότητα που μας απασχολεί. Η Sonic Solutions με το σύστημα No Noise, αναπτυγμένο από τον Dr. J. Andrew Moorer, έχει μια επιλογή σύνθετου φιλτραρίσματος που επιτρέπει την εισαγωγή πολλών στενού bandwidth φίλτρων υψηλής ανάλυσης, κατάλληλων για την απομάκρυνση hum και buzz (αρμονικές του hum). Πριν την εισαγωγή των φίλτρων, είναι χρήσιμο να κάνουμε μια ανάλυση FFT του noise floor για να καθορίσουμε ποιες αρμονικές είναι παρούσες ώστε να εφαρμόσουμε μόνο τα αναγκαία φίλτρα. Στα συστήματα SADiE και Artemis, υπάρχει αρκετή ισχύς DSP για την εισαγωγή πολλών φίλτρων στενής δέσμης σε πραγματικό χρόνο, και υπάρχουν περιπτώσεις όπου για τη ρύθμιση ενός dehumming απαιτήθηκαν περίπου 25 φίλτρα με Q στην τιμή του 40 ή και υψηλότερη ακόμα. Υπάρχει επίσης το TG's Backdrop, αναπτυγμένο από τον Dr. Gilbert Soulodre, πολύ αποτελεσματικό με τον τονικό θόρυβο. Συστήματα όπως τα Backdrop, Cedar, και No Noise πρέπει να πάρουν δείγμα ένα σύντομο κομμάτι θορύβου (ακόμα και ένα δευτερόλεπτο αρκεί) ώστε να το αναλύσουν, να αναγνωρίσουν τις φασματικές του συνιστώσες και στη συνέχεια να τις αφαιρέσουν χωρίς να επηρεάσουν το σήμα. Αυτό αναδεικνύει το γεγονός πως δεν πρέπει να «κόβουμε» σφικτά τις εισαγωγές των μουσικών κομματιών ή να επεμβαίνουμε σε υλικό που στέλνεται για αποθρορυβοποίηση.

5.4 Εξειδικευμένοι επεξεργαστές (specialized processors):

Η GML Labs έχει μια εξειδικευμένη μονάδα καταστολής του θορύβου για συνεχή και ασυνεχή θόρυβο. Η Cedar παράγει μια θαυματουργή διαδικασία που λέγεται Retouch, διαθέσιμη μόνο για SADiE DAWs. Το Retouch μπορεί να απομακρύνει ενοχλητικούς θορύβους που κανένα προηγούμενο σύστημα δεν μπορούσε να χειριστεί, όπως το κλάμα ενός μωρού, τριξίματα από καρέκλες, ακόμα και ομιλίες ανθρώπων. Είναι πολύ ακριβό, αλλά δεν υπάρχει υποκατάστατο όταν το χρειαζόμαστε πραγματικά.

Μερικοί κατασκευαστές εξειδικεύονται σε ένα είδος θορύβου, μερικοί έχουν ξεχωριστά (ακριβά) hardware συστήματα για κάθε είδους θόρυβο. Κάθε τύπος θορύβου, όπως χτυπήματα, τριξίματα, σφυρίγματα, κανονικοί θόρυβοι και ακανόνιστοι, υψηλού και χαμηλού level, χρειάζεται τον δικό του ειδικό αλγόριθμο. Ένας decrackler είναι στην πραγματικότητα ένας πολλαπλός decklicker, που ανιχνεύει και παρεμβάλει τα δεδομένα (samples) του crackle ώστε να το απαλείψει, έτσι χρειάζεται μεγάλη ισχύ DSP. Η Sonic και η Cedar έχουν τα πιο δημοφιλή high - end συστήματα μείωσης θορύβου, με ενδιαφέρουσες προσθήκες από τις Algorithmix, Audiodcube, TC Electronic και Waves. Η προσέγγιση της Sonic στον συνεχή θόρυβο, όπως το σφύριγμα ή το rumble, είναι να χρησιμοποιεί 2048 ατομικά συνεχή φίλτρα, δημιουργώντας έναν expander πολλαπλών δεσμών. Τα παραγόμενα artifacts μειώνονται παρά πολύ καθώς η επεξεργασία πολλαπλών δεσμών αποτρέπει την διάδραση μεταξύ των δεσμών. Η Sequoia έχει ένα εξαιρετικό FIR φίλτρο που σου επιτρέπει οπτικά και εργονομικά να «πιάσεις» κάθε αρμονική και να μειώσεις το πλάτος της. Όταν η πηγή του ήχου ποικίλει συχνοτικά, όπως με τις αναλογικές ταινίες με διαφορετικές ταχύτητες, απαιτείται ένα ειδικό είδος φίλτρου αναζήτησης (tracking filter). Το TG's Backdrop βασίζεται στην ψυχοακουστική για την κάλυψη του θορύβου, και είναι πολύ αποτελεσματικό σε συνεχή ή απόλυτο θόρυβο όπως hum, buzz, hiss και rumble, με ελάχιστα παραγόμενα artifacts όταν ρυθμίζεται σωστά. Δυστυχώς η οικονομία της αγοράς δεν έχει αφήσει ανεπηρέαστη τη «μουσική» βιομηχανία (πως θα μπορούσε άλλωστε;) και κατά συνέπεια παίρνουμε ότι πληρώνουμε, και ένα κριτικό αυτί μπορεί να πει την διαφορά μεταξύ των πιο ακριβών και των φθηνότερων συστημάτων.

5.5 Το φαί του ενός είναι το δηλητήριο του άλλου:

Στο βιβλίο του Mastering Audio: The Art and the Science, ο Bob Katz αναφέρει, σχετικά με το ζήτημα της απαλοιφής του θορύβου και της υποκειμενικής φύσης που αυτή πολλές φορές λαμβάνει, τα ακόλουθα:

«Κάποτε επεξεργάστηκα ένα άλμπουμ πανκ ροκ όπου το άνοιγμα ενός τόνου είχε ένα φανερό ηλεκτρικό τικ πάνω από τη νότα του παίκτη του μπάσου. Αφαίρεσα το τικ και αποκαταστάθηκε η ομορφιά της νότας - έτσι νόμιζα. Αλλά

*μετά ο παραγωγός μου είπε ότι του έλειψε το τικ και έτσι αναγκάστηκα να το βάλω πίσω».*σ[5]

Αποδεικνύεται έτσι ότι η ομορφιά βρίσκεται στο αυτί του ακροατή, και πολλοί θόρυβοι θεωρούνται μέρος της μουσικής. Θα πρέπει να είμαστε σε θέση να γνωρίζουμε κάθε μουσική μορφή και είδος, και σε μερικές περιπτώσεις ίσως να πρέπει να σκεφτούμε να αφήσουμε «βρώμικο» το μουσικό υλικό αντί «καθαρό»!

**Aliasing
&
Audio Mastering**

6.1 Aliasing:

Τα περισσότερα σήματα διακριτού χρόνου προέρχονται από τη δειγματοληψία ενός σήματος συνεχές στο χρόνο, τέτοια σήματα π.χ. είναι τα ηχητικά σήματα. Η διαδικασία μετατροπής των σημάτων αυτών σε ψηφιακή μορφή ονομάζεται analog to digital conversion (A/D Conversion). Η αντίστροφη διαδικασία που συνθέτει ένα αναλογικό σήμα από τα δείγματά του ονομάζεται digital to analog conversion (D/A Conversion).

Γιατί όμως θα πρέπει το αρχικό αναλογικό σήμα να υποστεί αυτή την επεξεργασία;

Η απάντηση, πλέον, είναι πασιφανής και αφορά το γεγονός πως με αυτό τον τρόπο το σήμα επωφελείται από ορισμένα πλεονεκτήματα. Ορισμένα από αυτά είναι ο χαμηλός σηματοθορυβικός λόγος του σήματος, η αναλλοίωτη από το χρόνο ποιότητα του σήματος, η εξοικονόμηση αποθηκευτικού χώρου, οι νέες - ευκολότερες - ταχύτερες μορφές επεξεργασίας του ηχητικού υλικού, κ.ο.κ.

Η διαδικασία της μετατροπής ενός σήματος από αναλογικό σε ψηφιακό, η οποία συχνά αναφέρεται και ως ψηφιοποίηση, περιλαμβάνει δύο βασικές φάσεις. Η πρώτη φάση καλείται δειγματοληψία (sampling) και αφορά τη λήψη και καταγραφή δειγμάτων στάθμης από το αναλογικό σήμα σε τακτά και συγκεκριμένα χρονικά διαστήματα. Η δεύτερη φάση έχει ήδη αναφερθεί και ονομάζεται analog to digital conversion (A/D Conversion) και αφορά τη μετατροπή του κάθε δείγματος σε ψηφιακή λέξη.

Η πρώτη φάση της ψηφιακής μετατροπής είναι και η πλέον κρίσιμη, σύμφωνα με το θεώρημα της δειγματοληψίας που ανέπτυξε ο Claude Shannon το 1948. Σε αυτό το θεώρημα αναφέρεται ότι κάθε αναλογικό σήμα με περιορισμένο εύρος ζώνης, μπορεί να περιγραφεί επακριβώς από μία ακολουθία δειγμάτων της στάθμης του, αρκεί αυτά τα δείγματα να έχουν ληφθεί με συχνότητα τουλάχιστον διπλάσια μιας χαρακτηριστικής συχνότητας του αναλογικού σήματος που ονομάζεται f_m . Η συχνότητα f_m είναι η υψηλότερη συχνότητα που περιέχει το αναλογικό σήμα με περιορισμένο εύρος ζώνης και δεν περιέχει καμία άλλη συχνότητα υψηλότερη αυτής. Μόνο κάτω από τις προϋποθέσεις του θεωρήματος της δειγματοληψίας μπορεί η επανασύνδεση των δειγμάτων στάθμης του αρχικού αναλογικού σήματος να αναδημιουργήσει το αρχικό αυτό σήμα δίχως παραμορφώσεις. Το σήμα θεωρείται περιορισμένου εύρους διότι το εύρος ζώνης του εκτείνεται ως την f_m και τελειώνει εκεί. Έτσι η συχνότητα δειγματοληψίας, η οποία στο εξής θα ονομάζεται f_s ή συχνότητα Nyquist, πρέπει σύμφωνα με το θεώρημα της δειγματοληψίας να είναι ίση ή μεγαλύτερη του διπλάσιου της f_m , δηλαδή να ισχύει η συνθήκη:

$$f_s \geq 2 \cdot f_m \quad (6.1.1)$$

Στο audio mastering τα παραπάνω είναι πολύ σημαντικά γιατί πολλές φορές καλούμαστε να επεξεργαστούμε ηχητικό υλικό του οποίου η συχνότητα δειγματοληψίας μπορεί να είναι 48 kHz, 96 kHz ή ακόμα και 192 kHz. Αυτό έχει ως αποτέλεσμα κατά την τελική επεξεργασία να κατεβάζουμε τη συχνότητα δειγματοληψίας αυτή, στα 44,1 kHz ώστε να μπορεί το ηχητικό υλικό να αποθηκευτεί σε μορφή C.D (44,1 kHz / 16 bit).

Συνακόλουθα και σε σχέση με όλα τα παραπάνω μεγάλη σημασία έχει και το πόσες στάθμες κβαντισμού θα χρησιμοποιήσουμε ώστε να αναπαραγάγουμε το αρχικό μας σήμα. Διότι δεν έχει κανένα νόημα να χρησιμοποιούμε υψηλή δειγματοληψία με λίγες στάθμες κβαντισμού (θα έχουμε πάρα πολλά δείγματα με πολύ μικρή διακύμανση). Όταν κάνουμε κβαντισμό του σήματός μας, αντιστοιχίζουμε την τιμή που αυτό έχει μια συγκεκριμένη χρονική στιγμή, σε ένα από τα επίπεδα κβαντισμού (περισσότερες στάθμες κβαντισμού σημαίνει καλύτερη προσέγγιση του αρχικού αναλογικού σήματος, λιγότερες στάθμες κβαντισμού υποδηλώνουν μία αντιστρόφως ανάλογη σχέση)¹⁹. Η κβαντοποίηση του σήματος, όμως, εισάγει θόρυβο σε αυτό ο οποίος καλείται θόρυβος κβαντοποίησης (Quantization noise). Τα σφάλματα που προκύπτουν από την κβαντοποίηση στην ουσία είναι η διαφορά του αρχικού μας σήματος με αυτό που προέκυψε έπειτα από την κβαντοποίηση.

Γνωρίζουμε, επίσης, πως ο θόρυβος κβαντοποίησης δεν είναι τυχαίος, αλλά εξαρτάται από τη συχνότητα του σήματος μας και από τη συχνότητα δειγματοληψίας. Όταν η συχνότητα Nyquist (f_s) είναι πολλαπλάσιο της συχνότητας του σήματος, τότε προκαλείται αρμονική παραμόρφωση με συνιστώσες οι οποίες είναι ακέραια πολλαπλάσια της συχνότητας του σήματος. Όπου η συχνότητα Nyquist ή αλλιώς συχνότητα δειγματοληψίας δίδεται από την σχέση (6.1.1).

Τέλος, απότοκο το οποίο εκπορεύεται από όλα όσα έχουν αναφερθεί μέχρι τώρα, είναι η εμφάνιση συχνοτήτων ακόμη και πέρα από το όριο Nyquist, οπότε θα εμφανίζεται και το φαινόμενο του aliasing. Οι συχνότητες που προστίθενται με αυτό τον τρόπο στο σήμα παράγουν έναν «περίεργο» ήχο ο οποίος στην αρχή των τεχνικών ήχου καλείται «bird singing» ή «birdies». Το aliasing προσθέτει συχνότητες στο σήμα μας σύμφωνα με τη σχέση:

$$f_a = \pm N \cdot f_s \pm f_j \quad (6.1.2)$$

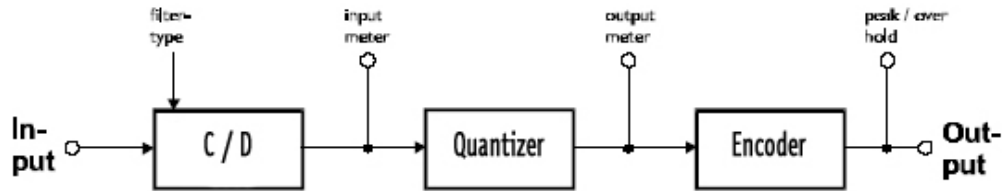
όπου $N = \pm 1, \pm 2, \pm 3, \dots$, κ.ο.κ.

f_a : οι νέες συχνότητες που προστίθενται στο σήμα.

f_s : η συχνότητα δειγματοληψίας.

και f_j : η συχνότητα που έχει τιμή μεγαλύτερη από το μισό της συχνότητας δειγματοληψίας.

¹⁹ Βλ. παράρτημα ΣΤ.



Σχ. 6.1.1: Διάγραμμα ροής υποσυστήματος μετατροπέα A / D.

Η γενική μορφή του συστήματος επεξεργασίας φαίνεται στο σχήμα 1 το οποίο εκφράζει τη διαδικασία μετατροπής ενός σήματος από το αναλογικό στο ψηφιακό πεδίο. Η είσοδος του μετατροπέα A/D είναι μια πραγματική συνάρτηση της συνεχούς μεταβλητής t . Η έξοδος είναι μια ροή δυαδικών ψηφίων (bit stream) που αντιστοιχεί σε μία ακολουθία διακριτού χρόνου, με πλάτος που υφίσταται κβαντισμό σε μία από τις πεπερασμένου πλήθους πιθανές τιμές. Τα υποσυστήματα που συνθέτουν τον εν λόγω converter είναι τρία. Το πρώτο κατά σειρά είναι το σύστημα (κύκλωμα) δειγματοληψίας (sampler), το οποίο πολλές φορές καλείται και Continuous to Discrete Converter (μετατροπέας Συνεχούς σε Διακριτό, C/D). Το δεύτερο είναι ο κβαντιστής (Quantizer), ο οποίος αντιστοιχεί το συνεχές πλάτος σε ένα σύνολο διακριτών τιμών πλάτους. Το τρίτο και τελευταίο είναι ο κωδικοποιητής (Encoder), ο οποίος λαμβάνει το ψηφιακό σήμα και παράγει μία ακολουθία δυαδικών κωδικοποιημένων λέξεων.

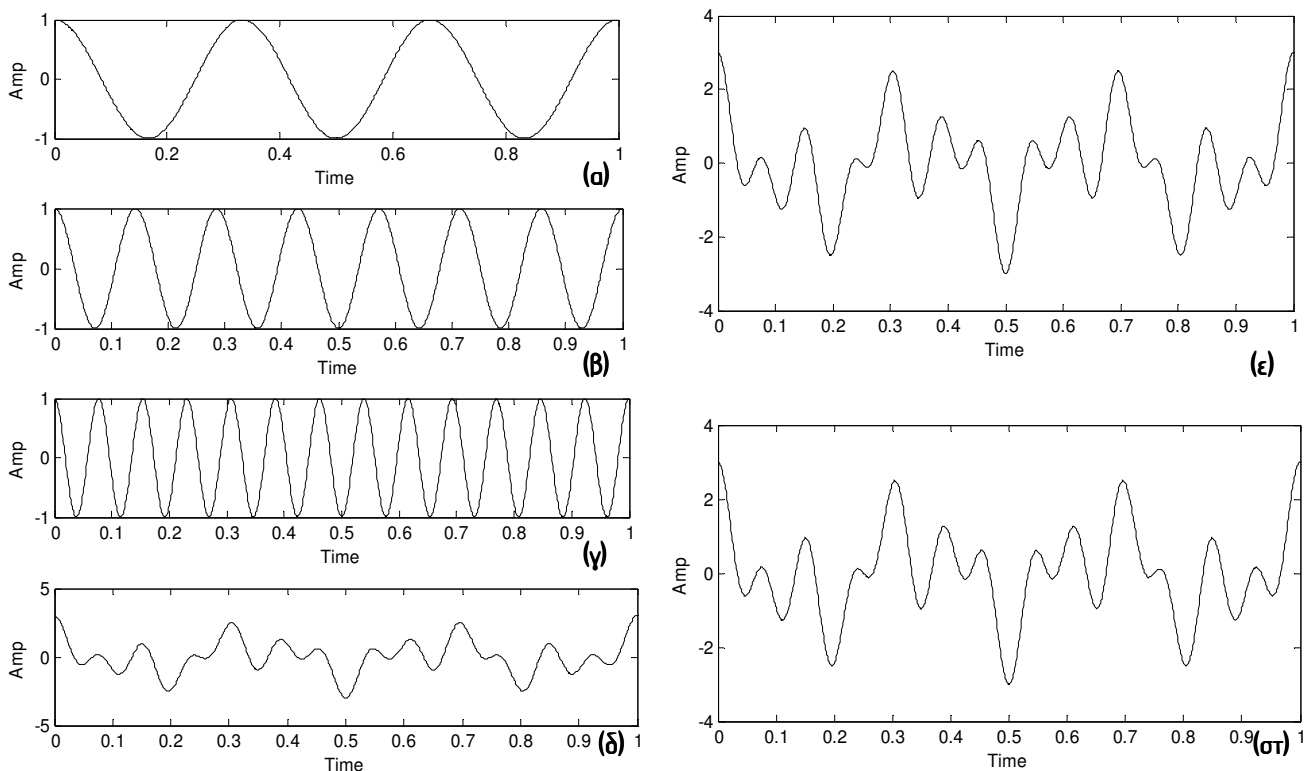
Το παραπάνω σύστημα, όμως, υπαγορεύει (ειδάλλως είναι ελλιπές) την ύπαρξη ενός αντι - aliasing φίλτρου κατά τη διαδικασία της μετατροπής του σήματος από αναλογικό σε ψηφιακό, αλλά και κατά τη διάρκεια της μετατροπής από ψηφιακό σε αναλογικό. Την ανάγκη ύπαρξης του φίλτρου αυτού θα τη δούμε στα παραδείγματα που ακολουθούν παρακάτω. Η ποιότητα των αντι - aliasing φίλτρων στους A/D και D/A μετατροπείς είναι ένα από τα καθοριστικότερα στοιχεία για τη γενικότερη ποιότητα τους. Το θέμα του aliasing είναι εξαιρετικά σημαντικό στο audio mastering και αυτό γιατί επηρεάζει με καθοριστικό τρόπο το ηχητικό υλικό όπως θα δούμε.

Τα παραδείγματα που ακολουθούν σχεδιάστηκαν με μορφή κώδικά²⁰ στο περιβάλλον προγραμματισμού του Matlab.

i. Κώδικας A:

Θεωρούμε αναλογικά ημιτονοειδή σήματα με συχνότητες 3, 7 και 13 Hz. Στα σήματα που έχουμε η υψηλότερη συχνότητα είναι αυτή των 13 Hz. Εμείς παίρνουμε ως συχνότητα δειγματοληψίας το 1 kHz που υπερκαλύπτει τη συχνότητα αυτή. Συνεπώς, προκύπτουν τα ακόλουθα διαγράμματα.

²⁰ Οι κώδικες δίδονται στο παράρτημα Ε.



Σχ. 6.1.2: (α) Σήμα συχνότητας 3 Hz. (β) Σήμα συχνότητας 7 Hz. (γ) Σήμα συχνότητας 13 Hz. (δ) Η υπέρθεση των τριών σημάτων. (ε) Το αρχικό αναλογικό μας σήμα ως η υπέρθεση των τριών αρχικών ημιτονοειδών διαφορετικής συχνότητας. (στ) Το ψηφιοποιημένο σήμα ως η έξοδος ενός ιδανικού A/D converter. Παρατηρούμε πως δεν έχει υποστεί καμία αλλαγή σε σχέση με το αρχικό αναλογικό σήμα.

Στα παραπάνω διαγράμματα παρατηρούμε καταρχήν τα τρία ημιτονοειδή σήματα που συνθέτουν το αναλογικό μας σήμα. Εν συνεχεία παρατίθενται συγκριτικά το αρχικό αναλογικό μας σήμα (διάγραμμα (ε)) με το τελικό ψηφιοποιημένο σήμα (διάγραμμα (στ)), παρατηρούμε σε αυτή την περίπτωση ότι η συχνότητα δειγματοληψίας που έχουμε επιλέξει και είναι το 1 kHz λειτουργεί πολύ καλά συναρτήσει των χαρακτηριστικών του μετατροπέα A/D που έχουμε εισάγει στον κώδικα. Ως αποτέλεσμα, προκύπτει, το ψηφιοποιημένο σήμα το οποίο λαμβάνει διακριτές τιμές στο πεδίο του χρόνου να είναι παντελώς όμοιο με το αναλογικό σήμα το οποίο λαμβάνει συνεχείς τιμές στο πεδίο του χρόνου. Τα παραπάνω επιβεβαιώνουν το θεώρημα δειγματοληψίας του Shannon. Οι γραμμές, τώρα, που χρησιμοποιούνται στον κώδικα και αντιπροσωπεύουν έναν ιδανικό μετατροπέα διακριτού σε συνεχές (ideal discrete to continuous converter) είναι:

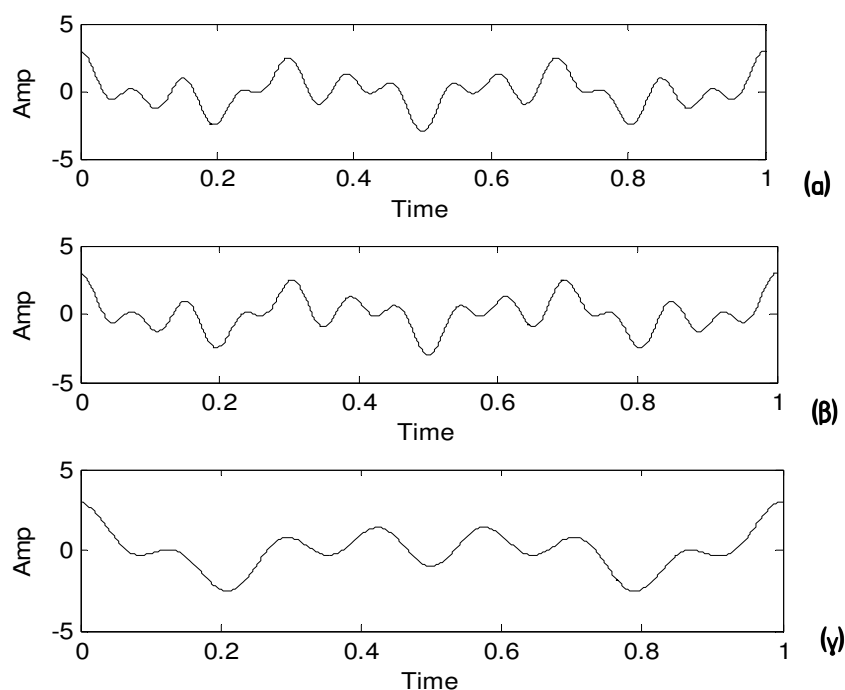
```

for i=1:length(x)
    greconstructed(i)=sum(x.*sinc(fs*(n(i)-n)));
end

```

ii. Κώδικας B:

Στο παράδειγμα αυτού του κώδικα θα κάνουμε δειγματοληψία με συχνότητα μικρότερη από αυτή που απαιτεί το θεώρημα δειγματοληψίας, δηλαδή θα επιλέξουμε συχνότητα μικρότερη από τη συχνότητα Nyquist. Πιο συγκεκριμένα επιλέγουμε ως συχνότητα δειγματοληψίας τα 15 Hz. Συνεπώς τα διαγράμματα που προκύπτουν έχουν τη μορφή:

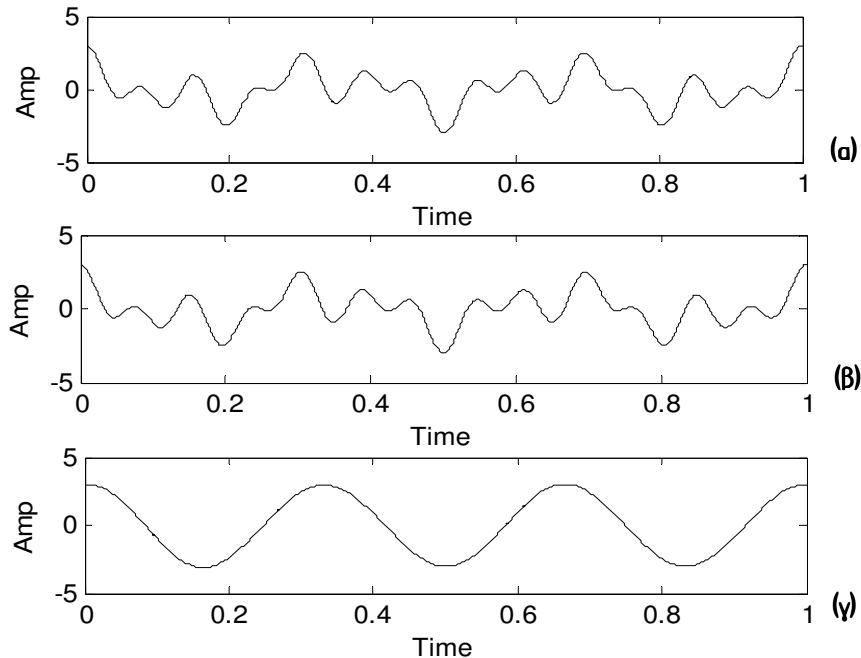


Σχ. 6.13: (α) Το αρχικό αναλογικό σήμα. (β) Το ψηφιοποιημένο σήμα του οποίου η συχνότητα δειγματοληψίας ήταν το 1 kHz. (γ) Το ψηφιοποιημένο σήμα του οποίου η συχνότητα δειγματοληψίας ήταν τα 15 Hz.

Από τον κώδικα αυτό, όπως παρατηρούμε προκύπτουν τρεις γραφικές παραστάσεις. Η πρώτη είναι το αρχικό μας σήμα (διάγραμμα (α)). Η δεύτερη (διάγραμμα (β)) είναι το ανακατασκευασμένο σήμα με συχνότητα δειγματοληψίας μεγαλύτερη από τη συχνότητα Nyquist: αυτή είναι το 1 kHz. Παρατηρούμε ότι το ανακατασκευασμένο σήμα είναι όμοιο με το αρχικό. Η τρίτη (διάγραμμα (γ)) είναι το ανακατασκευασμένο σήμα με συχνότητα δειγματοληψίας μικρότερη από τη συχνότητα Nyquist, πιο συγκεκριμένα η συχνότητα δειγματοληψίας ήταν τα 15 Hz. Αυτό που βλέπουμε είναι ότι με συχνότητα δειγματοληψίας μικρότερη από την απαιτούμενη η μορφή που λαμβάνει το ανακατασκευασμένο σήμα είναι διαφορετική από το αρχικό. Για την ακρίβεια προκύπτει ένα αλλοιωμένο κακέκτυπο του αρχικού σήματος.

iii. Κώδικας Γ (Aliasing στο πεδίο του χρόνου):

Στο παράδειγμα αυτό χρησιμοποιούμε και πάλι τον κώδικα Β, μόνο που αυτή τη φορά εισάγουμε σε δεύτερη συχνότητα δειγματοληψίας τα 10 Hz. Τα διαγράμματα που προκύπτουν δίδονται στο σχήμα 6.1.4 της επόμενης σελίδας.



Σχ. 6.1.4: (α) Το αρχικό αναλογικό σήμα. (β) Το ψηφιοποιημένο σήμα του οποίου η συχνότητα δειγματοληψίας ήταν το 1 kHz. (γ) Το ψηφιοποιημένο σήμα του οποίου η συχνότητα δειγματοληψίας ήταν τα 10 Hz.

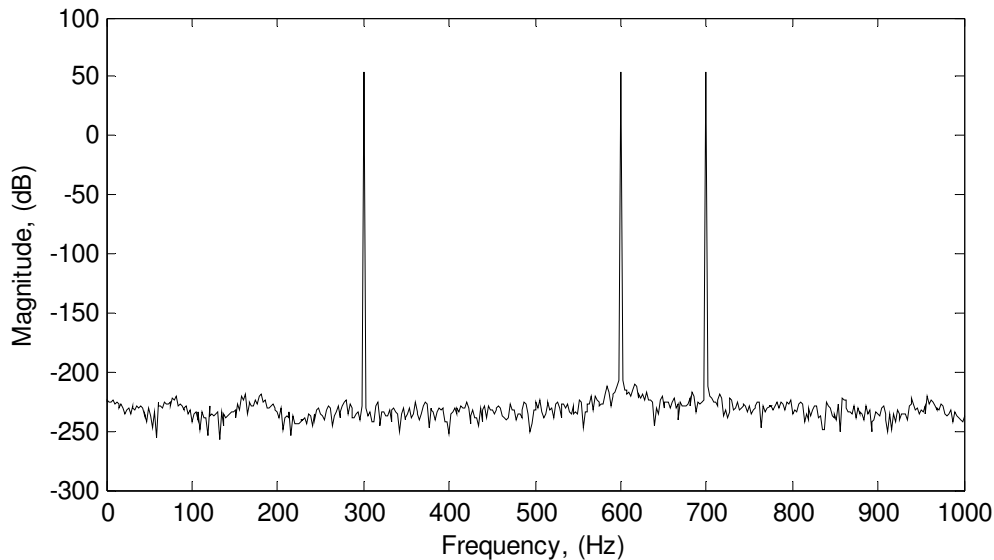
Για τα διαγράμματα (α) και (β) ισχύουν τα όσα έχουμε αναφέρει στην παράγραφο για τον κώδικα Β, μιας και είναι όμοια.

Για το διάγραμμα (γ) όμως, αν προσέξουμε καλύτερα το ανακατασκευασμένο αυτό σήμα, θα διαπιστώσουμε ότι πρόκειται για ένα συνημίτονο με συχνότητα 3 Hz. Παρατηρούμε ακόμα ότι το πλάτος του είναι ίσο με τρία (3). Το φαινόμενο του aliasing συνέβη στην περίπτωση αυτή. Αυτό που έκανε, λοιπόν, η λανθασμένη δειγματοληψία, είναι ότι μετασχημάτισε τα σήματα συχνότητας 7 Hz και 13 Hz και τα έκανε και αυτά σήματα των 3 Hz. Για το λόγο αυτό, άλλωστε, παρατηρούμε ένα συνημίτονο με πλάτος τρία (3), ως υπέρθεση των τριών σημάτων.

Στη γενική, βέβαια, περίπτωση δε θα έχουμε πάντα καταστάσεις τρία προς ένα (3/1), αλλά η ουσία βρίσκεται στο γεγονός ότι λανθασμένη δειγματοληψία παρουσιάζει στο τελικό εξαγόμενο ηχητικό υλικό σήματα με λανθασμένη συχνότητα. Μάλιστα αν το αρχικό μας σήμα αποτελείτο από συνιστώσες των 7 και των 13 Hz μόνο, τότε το σήμα μετά τη δειγματοληψία θα ήταν ένα ημίτονο των 3 Hz με πλάτος δύο (2), μιας συχνότητας δηλαδή που δεν υπήρχε καθόλου στο αρχικό μας σήμα.

iv. Κώδικας Δ (Aliasing στο πεδίο των συχνοτήτων):

Για να δούμε το φαινόμενο του aliasing στο πεδίο των συχνοτήτων θα παρουσιάσουμε τη γραφική παράσταση του σχήματος 6.1.5 (επόμενη σελ.) που προκύπτει από τον κώδικα Δ.



Σχ. 6.1.5: Το φάσμα της ψηφιοποίησης και ανακατασκευής ενός αναλογικού ήχου, ο οποίος αποτελείται από συχνότητες των 300, 700 και 1400 Hz και ο οποίος ψηφιοποιήθηκε με συχνότητα δειγματοληψίας τα 2 kHz.

Στον κώδικα Δ, δημιουργούμε τρία συνημιτονοειδή σήματα συχνοτήτων 300 Hz, 700 Hz και 1400 Hz αντίστοιχα και στη συνέχεια τα προσθέτουμε μεταξύ τους δημιουργώντας έτσι το τελικό αναλογικό μας σήμα. Στη συνέχεια θέτουμε ως συχνότητα δειγματοληψίας τα 2 kHz, συχνότητα η οποία δεν ανταποκρίνεται στο θεώρημα δειγματοληψίας.

Αυτό που παρατηρούμε είναι καταρχήν ότι το εύρος του ηχητικού φάσματος στην έξοδο του converter θα εκτείνεται από τα 0 έως τα 1000 Hz. Αυτό προκύπτει από το γεγονός πως η μέγιστη συχνότητα ηχητικού σήματος που μπορούμε να λάβουμε με συχνότητα δειγματοληψίας τα 2 kHz δίδεται από τη σχέση:

$$f_{\max} = f_s / 2 = \frac{2000 \text{ Hz}}{2} = 1000 \text{ Hz}$$

Συνεπώς προκύπτει πως μία συχνοτική συνιστώσα του αρχικού μας σήματος αυτή των 1400 Hz δεν θα υφίσταται στο ανακατασκευασμένο σήμα που προκύπτει από την έξοδο του converter και κατά συνέπεια προκύπτει αλλοίωση του σήματος.

Παραπέρα, όμως, στο φάσμα του σχήματος 5 παρατηρούμε ότι υπάρχει μία συχνότητα που δεν υπήρχε στο αρχικό μας σήμα και είναι αυτή των 600 Hz. Στην

περίπτωση αυτή έκανε και πάλι την εμφάνιση του το φαινόμενο του aliasing, λόγω της εσφαλμένης συχνότητας δειγματοληψίας.

Η συχνότητα των 600 Hz δεν προέκυψε τυχαία. Σύμφωνα λοιπόν με τη σχέση (6.1.2) και θέτοντας όπου: $\pm N = \pm 1$, $f_s = 2000 \text{ Hz}$ και $f_j = 1400 \text{ Hz}$ προκύπτει τελικά:

$$\begin{aligned} f_a &= \pm N \cdot f_s \pm f_j \Rightarrow \\ f_a &= 1 \cdot 2000 \text{ Hz} - 1400 \text{ Hz} \Rightarrow \\ f_a &= (2000 - 1400) \text{ Hz} \Rightarrow \\ f_a &= 600 \text{ Hz} \end{aligned}$$

Απο τα παραπάνω προκύπτει πως το σήμα που θα προκύψει τελικά απο την έξοδο του converter, δεν θα εμπεριέχει απλώς και μόνο αλλοιώσεις συγκριτικά με το αρχικό αναλογικό σήμα, αλλά παράλληλα θα εμπεριέχει και φασματικές συνιστώσες οι οποίες δεν υπήρχαν στο αρχικό μας σήμα. Κατά συνέπεια το σήμα και όλη η πληροφορία που αυτό μεταφέρει εκφυλίζεται.

6.2 Σύνδεση του Audio mastering με το Aliasing:

Στο mastering καλούμαστε πολύ συχνά να μετατρέψουμε το sampling rate και το bit rate ενός μουσικού κομματιού σε μία άλλη τιμή απο την αρχική που αυτό είχε. Οι λόγοι για τους οποίους γίνεται αυτό αφορούν κυρίως τα format των μέσων αποθήκευσης του ηχητικού υλικού με τα οποία διατίθεται στο ευρύ κοινό (παρόλα αυτά κάτι φαίνεται να αλλάζει σε σχέση με τα format των αποθηκευτικών μέσων τον τελευταίο καιρό). Επίσης αφορούν λόγους μείωσης του όγκου της πληροφορίας δεδομένου του ότι μπορεί το πεδίο διακίνησης του ηχητικού υλικού να αποτελεί ο διαδικτυακός χώρος, κ.ο.κ.

Απο τα παραπάνω προκύπτει η αναγκαιότητα για τους mastering ηχηλήπτες να έχουν την ικανότητα να επιλέγουν την κατάλληλη συχνότητα δειγματοληψίας στην περίπτωση, για παράδειγμα, του transferring απο μία μπομπίνα ενός ηχητικού υλικού αποθηκευμένου σε αναλογική μορφή σε ένα σταθμό εργασίας (DAW) και της ψηφιοποίησης του. Καταδεικνύεται, επίσης, η αναγκαιότητα οι mastering ηχηλήπτες να μπορούν να επιλέξουν και να έχουν στη διάθεσή τους converters και ψηφιακούς αλγόριθμους των οποίων τα anti - alias φίλτρα να είναι απο πολύ καλής ποιότητας ως εξαιρετικής, δεδομένου ότι το aliasing δεν παύει να υφίσταται ακόμα και στις περιπτώσεις που επιλέγουμε υψηλή συχνότητα δειγματοληψίας. Απλά οι παράταιρες συχνότητες μεταφέρονται έξω απο το ηχητικό φάσμα που μας απασχολεί και σε υψηλότερες συχνότητες. Συνεπώς ένα φίλτρο αποκοπής με πολύ καλά χαρακτηριστικά για τις συχνότητες αυτές θεωρείται απαραίτητο.

Jitter
&
Audio Mastering

7.1 Τι είναι το Jitter:

Τα ψηφιακά συστήματα διαφέρουν από τα αναλογικά σε δύο κρίσιμα σημεία.

- Πρώτον το σήμα στην αναλογική του μορφή είναι συνεχές σε αντίθεση με την ψηφιακή του μορφή που καθορίζεται από ένα συγκεκριμένο αριθμό δεδομένων (bytes) με μία συγκεκριμένη διαδοχή.
- Δεύτερον αυτές οι αριθμητικές αξίες αναδημιουργούν το σήμα μόνο με μία συγκεκριμένη διαδοχή στο χρόνο σε αντίθεση με το αναλογικό σήμα που συνεχίζεται κάθε στιγμή στο χρόνο.

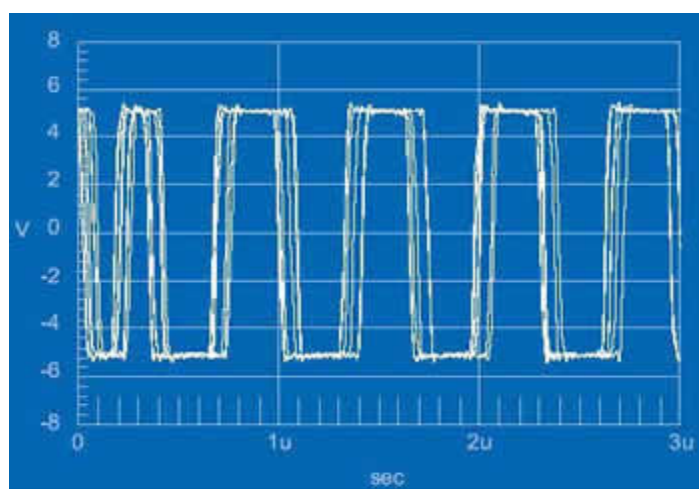
Οι συσκευές που παράγουν και αναπαράγουν τον ψηφιακό ήχο σύμφωνα με τα παραπάνω κριτήρια είναι ο μετατροπέας αναλογικού σήματος σε ψηφιακό (A / D converter) και ο μετατροπέας ψηφιακού σήματος σε αναλογικό (D / A converter).

Σαν jitter θεωρούμε το λάθος χρονισμό στην μεταφορά δεδομένων κατά την σύνδεση δυο ή περισσότερων ψηφιακών συσκευών.

Στην πραγματικότητα όλες οι ψηφιακές συσκευές με μια είσοδο και μια έξοδο εισάγουν ένα μικρό ή μεγαλύτερο ποσό jitter σε μία αλυσίδα.

Αυτό έχει σαν συνέπεια προβλήματα στην μεταξύ τους σύνδεση, επηρεάζει δε την ποιότητα του ήχου κατά την μετατροπή του αναλογικού ήχου σε ψηφιακό (A / D μετατροπή) και κατά την μετατροπή του ψηφιακού ήχου σε αναλογικό (D / A μετατροπή).

Η διάγνωση και οι λύσεις σε αυτό το πρόβλημα δεν είναι πλήρως κατανοητές ακόμα και από γνωστούς κατασκευαστές ψηφιακών συστημάτων, και μόλις τα τελευταία χρόνια έγινε δυνατή η μέτρηση με μηχανήματα του φαινομένου του jitter, εξηγώντας σε μεγάλο βαθμό τις ακουστικές διαφορές που παρατηρούνται από πλήθος ηχοληπτών και δεν είχαν κάποια λογική εξήγηση.



Σχ. 7.1.1: Απεικόνιση του φαινομένου του jitter σε παλμογράφο.

Το ενδιαφέρον με το Jitter είναι ότι δεν μεταφέρεται από ψηφιακή σε άλλη ψηφιακή σύνδεση αλλά είτε γράφεται κατά την αρχική ηχογράφηση του σήματος από τον A / D μετατροπέα είτε το ακούμε στον τελικό D / A μετατροπέα.

Αν για παράδειγμα τεχνικά δημιουργήσουμε τόσο μεγάλο jitter κατά την μεταφορά υλικού από ένα DAT σε ένα σταθμό εργασίας (DAW) ώστε να είναι δύσκολο να κλειδώνει το ρολόι (clock) του σταθμού εργασίας στο εισερχόμενο σήμα και ο ήχος να ακούγεται με διακοπές, ηχογραφήσουμε δε αυτό το σήμα μέσα στο σταθμό εργασίας, παίζοντας το στη συνέχεια μέσα από το σταθμό εργασίας δεν υπάρχουν πλέον διακοπές στο σήμα !!

Μπορούμε να πούμε ότι τα δεδομένα (data) ενός μουσικού κομματιού είναι πάντα τα ίδια (ένας συγκεκριμένος αριθμός 0-1 bytes δεδομένων σε μια συγκεκριμένη σειρά). Όμως το πως θα ακουστεί αυτό το κομμάτι εξαρτάται κάθε φορά από την ποιότητα του τελικού ρολογιού (master clock) που οδηγεί τον τελικό μετατροπέα του ψηφιακού ήχου σε αναλογικό (D / A μετατροπή).

Ιδανικά μόνο αν το τελικό D / A εξαφάνιζε το Jitter (λάθος χρονισμό) θα ακούγαμε πως είναι πραγματικά ηχογραφημένο αυτό το κομμάτι.

Γίνεται έτσι κατανοητή η μεγάλη σημασία που έχει:

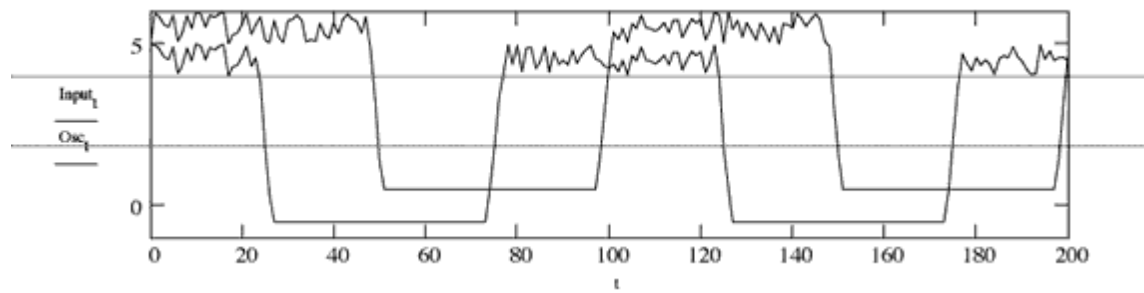
- i. Ο A / D μετατροπέας σε οποιαδήποτε ψηφιοποίηση ήχου, εκεί που ο ήχος "συλλαμβάνεται" για πρώτη φορά γιατί ένας A / D μετατροπέας πρέπει να έχει ένα εξαιρετικά σταθερό ρολόι (master clock) που δεν εισάγει jitter, παραμορφώσεις και θορύβους οι οποίοι από την στιγμή που θα ηχογραφηθούν δεν είναι πλέον δυνατόν να αφαιρεθούν σε μετέπειτα στάδιο.
- ii. Η ικανότητα του τελικού D / A μετατροπέα να μειώνει αν όχι να εξαφανίζει οποίο λάθος χρονισμό εμφανίζεται στο σήμα.

Πώς ακούγεται, όμως, ο ήχος σε υλικό με αυξημένο Jitter;

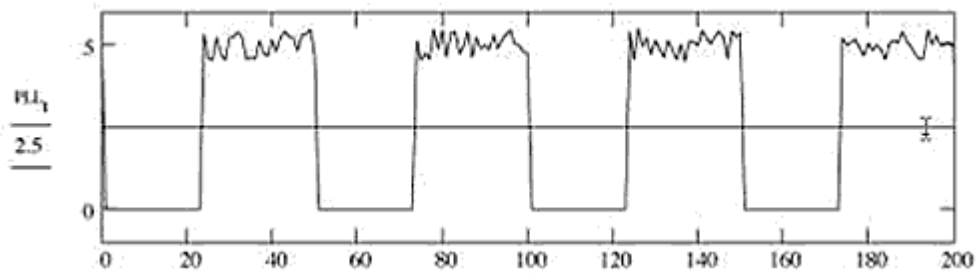
Η βασική αίσθηση είναι ότι ο ήχος είναι στεγνός, φτηνός έχει χάσει δε την έκταση στα άκρα του συχνοτικού φάσματος. Αυτό που αποκαλούμε "αέρα" ή "χώρο" σε μία ηχογράφηση όπως και την έκταση των χαμηλών περιοχών. Έχουν καθεί επίσης τα επίπεδα, η κίνηση και το τρισδιάστατο της ηχογράφησης. Είναι προφανές ότι οι καλές ηχογραφήσεις υποφέρουν περισσότερο από το φαινόμενο του Jitter χωρίς αυτό να σημαίνει ότι δεν ακούγεται η διαφορά σε οποιαδήποτε ηχογράφηση.

Μία από τις πρώτες μεθόδους που προτάθηκαν για την λύση του προβλήματος ήταν η αναδημιουργία (regeneration) και (reclocking) του σήματος (βλ. σχήμα 7.1.2).

Βέβαια τα πράγματα δεν είναι τόσο απλά γιατί τελευταίες έρευνες έδειξαν ότι το jitter εξαρτάται και από το συχνοτικό περιεχόμενο της μουσικής, δηλαδή αλλάζει αναλόγως με την συχνοτική ενέργεια, γεγονός που κάνει πιο σύνθετη την προσπάθεια για τον περιορισμό του.



Σχ. 7.1.2: Δύο κυματομορφές με λάθος χρονισμό και θόρυβο που προέρχεται από το τροφοδοτικό τις οδηγούμε στη συνέχεια σε ένα PLL (Phase Lock Loop) που αναδημιουργεί (regeneration) ότι σήμα βλέπει στην είσοδο του.



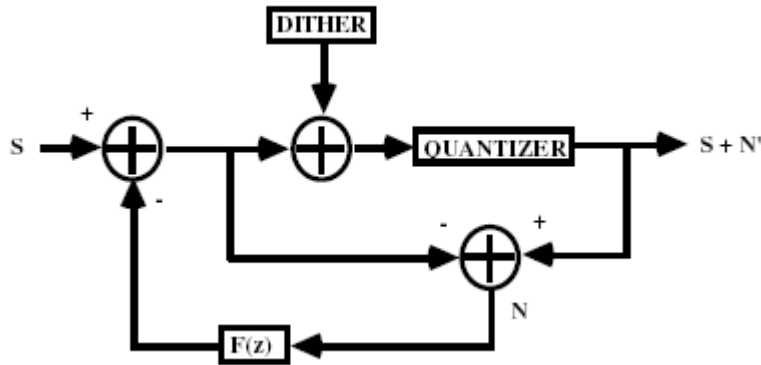
**Dithering
&
Audio Mastering**

8.1 Dithering – Εισαγωγή:

Σήμερα περισσότερο απο ποτέ άλλοτε παράγονται ηχογραφήσεις των οποίων η ανάλυση υπερβαίνει τα 16-bit και το δυναμικό τους εύρος μπορεί και ξεπερνά ακόμα και τα 100 dB. Καθώς όμως το βασικό μέσο αποθήκευσης των περισσότερων εκ των ηχογραφήσεων αυτών εξακολουθεί να παραμένει το Compact Disk (C.D), δημιουργούνται ορισμένα ερωτήματα και προβληματισμοί σχετικά με το πώς μπορεί να επιτευχθεί η μείωση της ανάλυσης του ηχητικού υλικού στα επιβεβλημένα απο το μέσο αποθήκευσης 16-bit. Οι διάφοροι τρόποι κατανομής της ενέργειας ανά συχνότητα (Noise Shaping) καθώς και το dithering χρησιμοποιούνται σαν ικανοποιητικοί τρόποι επίτευξης των παραπάνω. Οι πρακτικές αυτές λειτουργούν ικανοποιητικά και κερδίζουν το ενδιαφέρον καθώς τοποθετούν το μη αποδεκόμενο διόρθωση σφάλμα που προκύπτει απο τη διαδικασία της κβαντοποίησης, σε κάποια περιοχή του συχνοτικού φάσματος στην οποία το ανθρώπινο αυτί είναι λιγότερο ευαίσθητο. Καταδεικνύεται παρακάτω πως η επιτυχία αυτής της πρακτικής εξαρτάται σε πολύ μεγάλο βαθμό απο το τι καλείται να εξυπηρετήσει (μουσικό υλικό, κινηματογραφικό υλικό, κ.ο.κ) καθώς και απο το περιβάλλον ακρόασης. Έτσι γίνεται σαφές πως η επιλογή ενός αλγόριθμου dither απο τον mastering ηχολόγητη αφορά περισσότερο το αισθητικό πεδίο της όλης διαδικασίας και λιγότερο το τεχνικό-επιστημονικό. Συνίσταται λοιπόν στους κατασκευαστές να παρέχουν στο χρήστη ένα ευρύ πλήθος τέτοιων αλγορίθμων έτσι ώστε η διαδικασία να μπορεί να οδηγήσει στα επιθυμητά αποτελέσματα.

8.2 Θεωρητικό υπόβαθρο:

Η μείωση της ανάλυσης ενός ψηφιακού audio σήματος, στα 16-bit ας πούμε, παράγει ένα σφάλμα το οποίο δεν μπορεί να εξαλειφθεί. Σφάλμα το οποίο γίνεται αντιληπτό με ένα ιδιαίτερα ενοχλητικό τρόπο διότι υπόκειται σε διαμόρφωση (modulation) απο το επιθυμητό audio σήμα, παράγει αρμονικές σε πολλαπλάσιες απο αυτό συχνότητες και καταλήγει τελικά να μασκάρει το επιθυμητό σήμα. Υπάρχουν διάφοροι τρόποι έτσι ώστε να επιτευχθεί η μείωση της ανάλυσης ενός ψηφιακού audio σήματος στα 16 π.κ. bit. Σε αυτή την παράγραφο θα συζητήσουμε για τους τρόπους που προκύπτουν απο την επεξεργασία που παρουσιάζεται στο Σχ. 8.2.1 (επόμενη σελίδα).



Σχ. 8.2.1: Block διάγραμμα του τρόπου με τον οποίο παράγεται το dither καθώς και του τρόπου με τον οποίο κατανέμεται το ποσοστό ενέργειας ανά συχνότητα σε αυτό.

Θεωρούμε πώς στο κουτί του block διαγράμματος που ονομάζεται 'DITHER' είναι μία γεννήτρια θορύβου όπου κατανέμει το θόρυβο με τυχαίο τρόπο στο φάσμα και που εκφράζεται είτε από μία συνάρτηση όμοιας κατανομής του πλάτους, η οποία είναι της τάξης του ενός λιγότερου σημαντικού ψηφίου, είτε μέσω μιας τριγωνικής συνάρτησης κατανομής του πλάτους, η οποία είναι διπλάσια αυτής του ενός λιγότερου σημαντικού ψηφίου. Οι συναρτήσεις αυτές είναι γνωστές ως 'UPDF' η πρώτη κι ως 'TPDF' η δεύτερη. Υπάρχουν τέσσερις θεμελιώδεις παραλλαγές του μοντέλου που παρουσιάζεται στο Σχ. 8.2.1 και μπορούν να προσεγγιστούν ως εξής:

- Q1: Απλή μείωση της ανάλυσης (Pure Truncation). Συνοψίζεται στα εξής: Η τιμή του 'DITHER' είναι μηδενική (DITHER=0), καθώς και η τιμή του $f(z)$ είναι μηδενική ($f(z)=0$). Όπου $f(z)$ εννοούμε μια συνάρτηση που περιγράφει την τιμή ενός φίλτρου, το οποίο φίλτρο καθορίζει την καμπύλη του dither, δηλ. το ποσοστό κατανομής της ενέργειας ανά συχνότητα.
- Q2: Επίπεδο dither (Flat dither). Σε αυτή την περίπτωση η τιμή του 'DITHER' είναι μη μηδενική (DITHER \neq 0) και η τιμή του $f(z)$ μηδενική ($f(z)=0$).
- Q3: Noise shaped. Σε αυτή την περίπτωση η τιμή του $f(z)$ είναι μη μηδενική ($f(z)\neq 0$) και η τιμή του 'DITHER' μηδενική (DITHER=0).
- Q4: Το 'DITHER' έχει μη μηδενική τιμή ('DITHER \neq 0') καθώς επίσης και το $f(z)$ έχει μη μηδενική τιμή ($f(z)\neq 0$).

Υπάρχουν διάφοροι λόγοι για τους οποίους θα μπορούσε κάποιος να επιλέξει τη μια μορφή ή την άλλη, κάτι το οποίο κάθε φορά εξαρτάται από τη μορφή του υλικού βάσει του οποίου εργαζόμαστε. Σε κάθε περίπτωση πάντως, η λιγότερο αποδεκτή προσέγγιση φαίνεται να είναι αυτή της απλής μείωσης της ανάλυσης (Pure Truncation, Q1). Στις περιπτώσεις Q1 και Q3 το ποσό της ενέργειας που εμπεριέχεται στο σφάλμα είναι το ίδιο, όμως πάραυτα δε μπορούν να εξαλείψουν το ίδιο το σφάλμα. Στις περιπτώσεις του Q2 και του Q4 το dither προστίθεται στο σήμα με τέτοιο τρόπο ώστε να αυξάνει το απόλυτο πλάτος του κατώτατου ορίου του θορύβου κατά 3 dB. Στην

περίπτωση του Q2 το σφάλμα θα αποκτήσει ένα μεγάλο εύρος συχνοτήτων (θα μετακινηθεί σε πολύ μεγαλύτερο συχνοτικό φάσμα) και με αυτό τον τρόπο θα πάψει να γίνεται αντιληπτό. Οι παραπάνω δεν είναι οι μοναδικοί συνδυασμοί που μπορούν να υπάρξουν, αποτελούν όμως τους πιο συχνά αναφερόμενους στη διεθνή βιβλιογραφία.

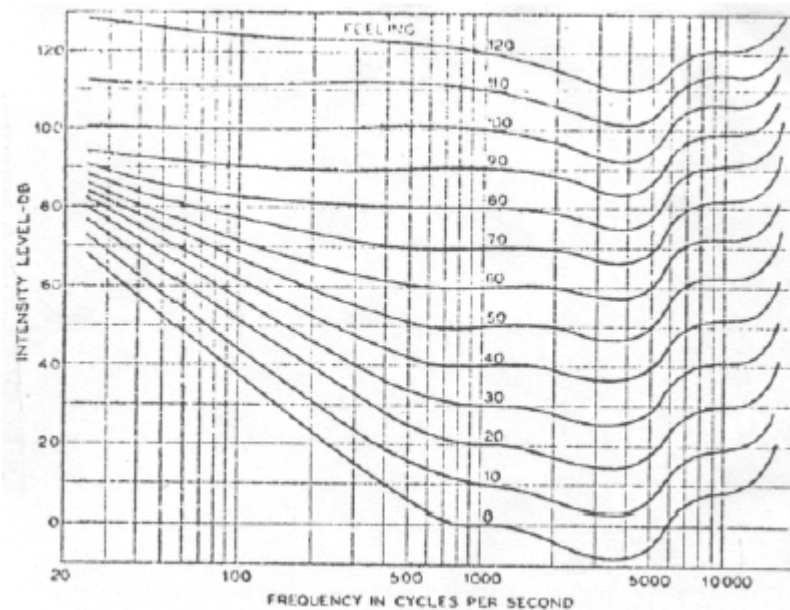
Το σημαντικό με το dithering είναι το ότι με την άθροιση ενός ποσοστού θορύβου παύει να υφίσταται ο συσχετισμός μεταξύ του σήματος και του σφάλματος που έχει προκύψει από τη διαδικασία της κβαντοποίησης.

8.3 Το φάσμα που προκύπτει:

Το νόημα του noise shaping, όπως έχει ήδη αναφερθεί παραπάνω, είναι να μεταφέρει το σφάλμα που προκύπτει στο ψηφιακό σήμα σε μια περιοχή του συχνοτικού φάσματος στην οποία το ανθρώπινο αυτί είναι λιγότερο ευαίσθητο. Παρακολουθώντας αυτό το σκεπτικό προκύπτει πως η καλύτερη κατανομή της ενέργειας στο συχνοτικό φάσμα είναι αυτή που τοποθετεί το θόρυβο ακριβώς με τρόπο αντίστοιχο με αυτόν βάσει του οποίου το ανθρώπινο αυτί αντιλαμβάνεται τη διαφορά έντασης ανά συχνοτική περιοχή. Κατά συνέπεια το πρόβλημα θα λυνόταν με τον απλό σχεδιασμό ενός φίλτρου $F(z)$ που θα παρήγαγε την καμπύλη αυτή. Το ζήτημα όμως γίνεται πολυπλοκότερο ακριβώς σε αυτό το σημείο...

8.4 Όλα άρχισαν με τους Fletcher & Munson:

Το 1932, ένα δίδυμο ερευνητών στα εργαστήρια της Bell αποφάσισαν να ξεκαθαρίσουν κάτι που για εκείνη την εποχή καλυπτόταν από ένα πέπλο μυστηρίου: το πώς το ανθρώπινο αυτί αντιλαμβάνεται απλά ημιτονοειδή σήματα σε διαφορετικές συχνότητες και σε διαφορετικές εντάσεις. Για το πείραμα τους χρησιμοποίησαν ταλαντωτές και ενισχυτές λυχνίας για να παράγουν ηλεκτρικά σήματα απλών τόνων κι ένα και μοναδικό megaphone για να αναπαραγάγουν τους απλούς αυτούς τόνους ως ακουστικό σήμα. Τα υποκείμενα του πειράματος ήταν τοποθετημένα ακριβώς μπροστά από το ηχείο και τους τεινόταν να συγκρίνουν την ακουστότητα των τόνων για διαφορετικές συχνότητες. Το αποτέλεσμα αυτών των πειραμάτων ήταν οι ευρέως διαδεδομένες καμπύλες ίσης ακουστότητας όπως φαίνονται στο Σχ. 8.4.1 (Το Σχ. 8.4.1 αποτελεί αντίγραφο αυτού που δημοσιεύθηκε στο πρωτότυπο άρθρο, επομ. σελίδα.)



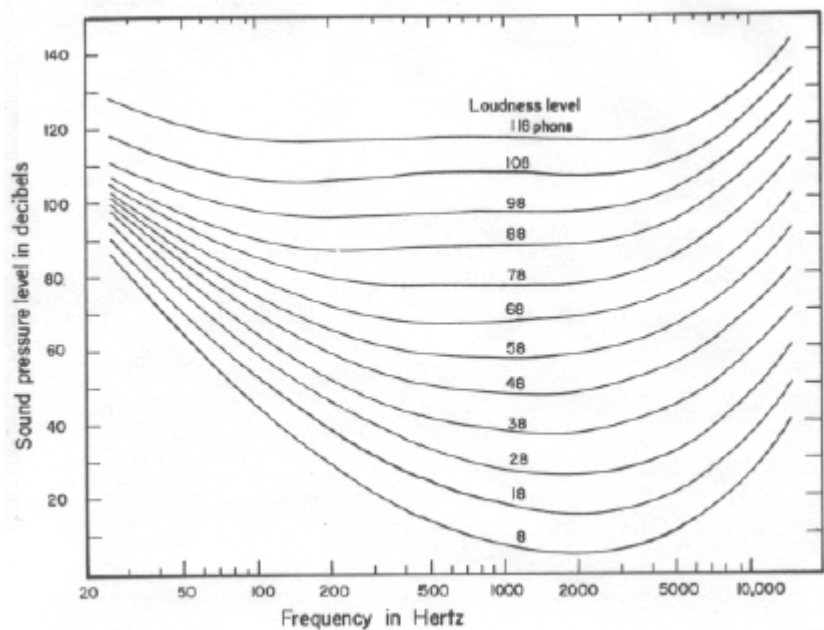
Σχ. 8.4.1: Οι καμπύλες ίσης ακουστότητας των Fletcher & Munson όπως αυτές τυπώθηκαν στο πρωτότυπο άρθρο τους το 1933. [21]

Κάποιος θα μπορούσε να ρωτήσει το αν και κατά πόσον ο εξοπλισμός που ήταν διαθέσιμος το 1932 στους δύο αυτούς ερευνητές, ήταν επαρκής σχετικά με αυτό που καλούνταν να μελετήσουν. Οι ίδιοι περιγράφουν την αρμονική παραμόρφωση που παραγόταν στο σήμα του πειράματος ως εξής:

«Τα στοιχεία ενίσχυσης των συσκευών (λυχνίες) ήταν ηλεκτροδυναμικού τύπου και διαπιστώθηκε πως παράγουν αρμονικές της θεμελίου της τάξης των 50 dB κάτω από το πλάτος αυτής. Στις πολύ υψηλές εντάσεις, η παραμόρφωση που προέκυπτε από τα φίλτρα των συσκευών αποδείχτηκε πως ήταν μεγαλύτερη από αυτήν που προέκυπτε λόγω των λυχνιών. Σε όλες τις περιπτώσεις όμως η ακουστότητα της οποιασδήποτε αρμονικής ήταν τουλάχιστον 20 dB πιο κάτω από αυτή της θεμελίου. Η εμπειρία με τους σύνθετους τόνους έχει δείξει πως κάτω από αυτές τις συνθήκες η συνεισφορά των αρμονικών στη συνολική ένταση του παραγόμενου ήχου είναι ασήμαντη». [21]

Παρόλο του ότι τα συμπεράσματα τους είναι πιθανότατα σωστά (ότι δηλ. οι αρμονικές έτσι όπως περιγράφονται δεν επηρεάζουν το τελικό σήμα), ένα τέτοιο ποσοστό παραμόρφωσης κατά τη διάρκεια ενός ψυχοακουστικού πειράματος δε θα γινόταν σε καμία περίπτωση αποδεκτό στις μέρες μας.

Συνεχίζοντας θα περάσουμε στις καμπύλες ίσης ακουστότητας που φαίνονται στο Σχ. 8.4.2 (επόμενη σελ.). Αυτές προέρχονται από τους Stevens και Davis.



Σχ. 8.4.2: Καμπύλες ίσης ακουστότητας των Stevens και Davis. Αυτές εξήχθησαν μέσω πειραμάτων με τη χρήση ακουστικών. [22]

Είναι χαρακτηριστικό ότι σε πολλά σημεία τους διαφέρουν από αυτές των Fletcher και Munson περισσότερο ακόμα και από 10 dB. Οι καμπύλες αυτές πάρθηκαν με τη χρήση ακουστικών και μόνο στα αυτιά των πειραματικών υποκειμένων. Θα πρέπει να σημειωθεί πως ο κυματισμός που συναντάμε στην περιοχή των 5 ως 10 kHz στις καμπύλες των Fletcher & Munson στις καμπύλες των Stevens & Davis δεν υφίσταται. Από αυτό συμπεραίνουμε πως η μη μονότονη συμπεριφορά στις καμπύλες των Fletcher & Munson για την εν λόγω περιοχή οφείλεται στο ακουστικό περιβάλλον στο οποίο έλαβε χώρα το πείραμα και συνεπώς δεν αποτελεί παράγωγο του μηχανισμού του έσω αυτιού.

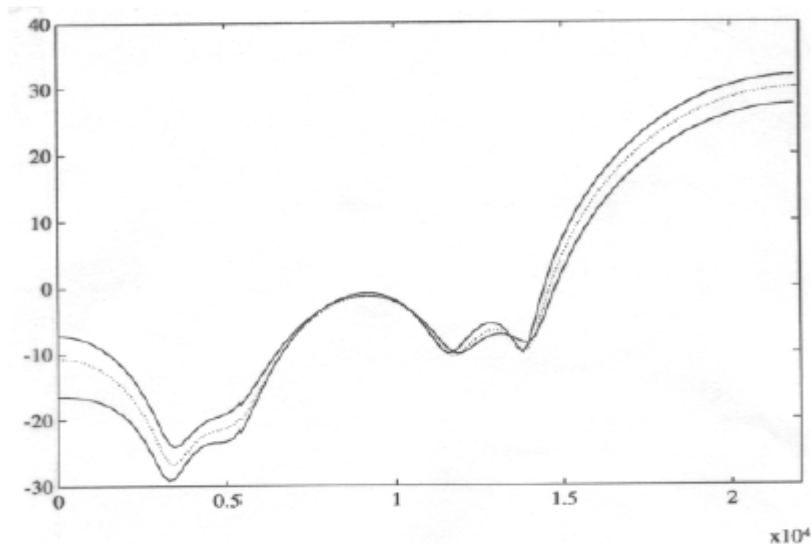
Από τότε έχει λάβει χώρα ένας μεγάλος αριθμός παρόμοιων πειραμάτων με ολόένα και περισσότερο εξελιγμένο πειραματικό εξοπλισμό, των οποίων όμως τα συμπεράσματα παρουσιάζουν μεγάλες διαφοροποιήσεις οι οποίες βασίζονται στην τοποθέτηση του μεγαφώνου και στα ακουστικά χαρακτηριστικά του χώρου που πραγματοποιείται το πείραμα. Συμπεραίνουμε λοιπόν, πως δεν υπάρχει η καμπύλη που περιγράφει το απόλυτο κατώφλι ακουστότητας και πως αυτό εξαρτάται από τον τρόπο με τον οποίο ο ήχος διεγείρει το έσω αυτί. Ακόμα και εξαιρετικά απλές διατάξεις (όπως αυτές που συναντάμε στο καθιστικό ενός διαμερίσματος), οι οποίες αναπαράγουν τον ήχο διαμέσου δύο ηχείων τα οποία είναι τοποθετημένα στα δεξιά και τα αριστερά μιας νοητής ευθείας με ίση απόσταση μεταξύ τους, μπορούν να παρατηρηθούν αλλαγές των καμπυλών ίσης ακουστότητας ακόμη και κατά 6 dB.

Θα πρέπει επίσης να σημειωθεί το εμφανές. Το ότι δηλαδή οι καμπύλες ίσης ακουστότητας αλλάζουν σημαντικά το σχήμα τους συναρτήσει της αλλαγής της έντασης. Στο studio καλούμαστε να επιλέξουμε μία καμπύλη έτσι ώστε να σχεδιάσουμε,

μέσω ενός φίλτρου, τη μορφή που θα έχει ο θόρυβος (dither) που θα χρησιμοποιήσουμε. Αυτό σημαίνει πως πρέπει να αποφασίσουμε όχι μόνο τα ακουστικά χαρακτηριστικά του χώρου ακρόασης (θέση ηχείων - ακροατή, κρίσιμη απόσταση (CD), χρόνο αντήχησης (RT₆₀), κ.α.), αλλά επίσης και το ποίο ακριβώς θα είναι το level στο οποίο θα αναπαράγεται το ηχητικό υλικό. Έτσι προκύπτει πως δεν υπάρχει η καμπύλη που περιγράφει το απόλυτο κατώφλι ακουστότητας και πως αυτό εξαρτάται από το απόλυτο level κατά την αναπαραγωγή. Κατά συνέπεια αυτής της θέσης προκύπτει πως το συχνοτικό φάσμα ενός θορύβου (dither) που δεν είναι ακουστό για μια συγκεκριμένη τιμή της στάθμης του ήχου μπορεί να γίνει τελικά ακουστό αν αυξήσουμε αυτή τη στάθμη.

Λαμβάνοντας υπόψη όλα αυτά τα δεδομένα, προκύπτει το ερώτημα: πως επιλέγουμε τον ένα τύπο dither ή τον άλλο; Μια σύντομη απάντηση θα μπορούσε να είναι πως, φυσικά, δεν μπορούμε να κάνουμε την κατάλληλη επιλογή παρά μόνο εφόσον γνωρίζουμε όλα τα δεδομένα του ήχου που πρόκειται να αναπαραχθεί. Παρόλα αυτά όμως οι περισσότερες ηχογραφήσεις στις μέρες μας κυκλοφορούν χρησιμοποιώντας ως μέσο τα C.Ds και το μουσικό υλικό αυτών των ηχογραφήσεων έχει περάσει από τη διαδικασία του dithering. Οι περισσότεροι άνθρωποι συμφωνούν πως στις περισσότερες περιπτώσεις η χρήση του dithering στο μουσικό υλικό παρέχει πλεονεκτήματα.

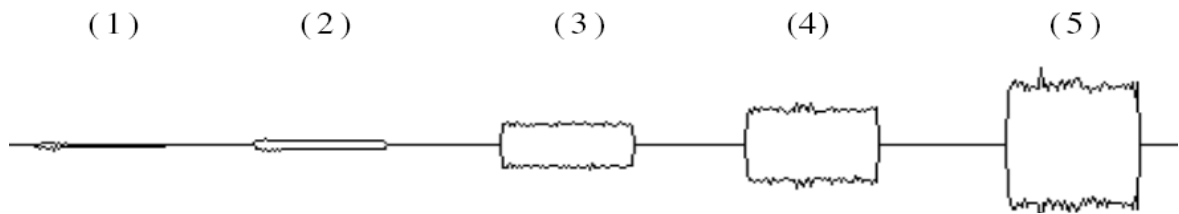
Προχωρώντας την αναλυτική μας ακόμα παραπέρα διαπιστώνουμε ένα είδος συνάφειας στις καμπύλες που περιγράφουν το κατώφλι ακουστότητας (threshold of hearing curves) με τη χρήση του dithering. Μπορούμε όλοι να συμφωνήσουμε πως υπάρχουν κάποιοι γενικότεροι κανόνες που δεν επιδέχονται αμφισβήτηση, όπως το γεγονός πως το ανθρώπινο αυτί είναι περισσότερο ευαίσθητο στην περιοχή από το 1 έως τα 5 kHz περίπου, και πολύ λιγότερο ευαίσθητο στις χαμηλές και στις υψηλές περιοχές του φάσματος. Μιας λοιπόν και η υψηλή περιοχή του φάσματος διαθέτει λιγότερες κρίσιμες μπάντες (critical bands) δε δημιουργεί προβληματισμό το ότι μεταφέρουμε το μεγαλύτερο ποσοστό τις ενέργειας του θορύβου σε αυτή την περιοχή. Κάθε dither του οποίου η κατανομή ενέργειας ανά συχνότητα ακολουθεί μία τέτοιου είδους καμπύλη θεωρείται αποδεκτό. Στο Σχ. 8.4.3 (σελ. 123), βλέπουμε παραδείγματα τέτοιων καμπυλών. Ο τρόπος με τον οποίο τελικά καλούμαστε κάθε φορά να επιλέξουμε ποιο τύπο dither θα χρησιμοποιήσουμε είναι να εφαρμόζουμε κάθε dither ξεχωριστά στο μουσικό μας υλικό και να αποφασίζουμε έπειτα ποιο θεωρούμε πως ταιριάζει καλύτερα. Σήμερα στην παραγωγή των περισσότερων από τους δίσκους που κυκλοφορούν ο άνθρωπος που λαμβάνει αυτές τις αποφάσεις δεν είναι άλλος από τον τεχνικό mastering.



Σχ. 8.4.3: Τρεις καμπύλες απόκρισης συχνότητας ενός dither.

Πρέπει ακόμα να σημειωθεί πως υπάρχει ένα ακόμα σημείο, τεχνικής υφής, που καλούμαστε να λάβουμε υπόψη μας κατά την επιλογή ενός dither. Αυτό είναι η πιθανή απώλεια δυναμικού εύρους²¹ χρησιμοποιώντας κάποιο dither του οποίου η περιβάλλουσα είναι υψηλότερης τάξης. Το σχήμα που ακολουθεί παρακάτω (Σχ. 8.4.4) παρουσιάζει πέντε παραδείγματα εφαρμογής dithering όπου στο κάθε παράδειγμα (με αύξουσα σειρά) έχει χρησιμοποιηθεί κατά μία τάξη μεγαλύτερη καμπύλη περιβάλλουσας. Είναι χαρακτηριστικό πως τα πλάτη των ήχων στους οποίους εφαρμόστηκε dither υψηλότερης τάξης έχουν μεγαλύτερο πλάτος (με ανάλογη κιάλας σχέση), συγκριτικά με τον ήχο στον οποίο έχει εφαρμοστεί dither χαμηλής (πρώτης) τάξης με ομοιόμορφη κατανομή της ενέργειας σε όλες τις συχνότητες. Γενικώς το level του σφάλματος κβαντοποίησης στο συγκεκριμένο παράδειγμα ήταν τόσο μικρό που δεν είχε σημασία ποιο τύπο dither θα χρησιμοποιούσαμε. Με αυτόν τον τρόπο όμως γίνεται σαφές πως πρακτικά, κατά τη διεξαγωγή της εφαρμογής του dithering, υπάρχει ένα άνω όριο σε ότι αφορά την τάξη του dither που χρησιμοποιούμε και της καμπύλης του. Είναι σημαντικό επίσης να ειπωθεί πως ορισμένες audio συσκευές (κυρίως οι χαμηλότερης ποιότητας) παράγουν σοβαρά προβλήματα αντιδρώντας στην παρουσία ενός φάσματος με αρκετή ενέργεια στις υψηλές συχνότητες. Για να μπορούμε να επιλέξουμε το κατάλληλο dither θα πρέπει, λόγω των όσων ειπώθηκαν πιο πριν, ολόκληρη η αλυσίδα των συσκευών που διανύει το audio σήμα να είναι εξαιρετικής ακρίβειας. Η οποιαδήποτε μη γραμμικότητα του συστήματος αναιρεί πλήρως τις αξιώσεις που προϋποτίθενται για την εφαρμογή του dithering.

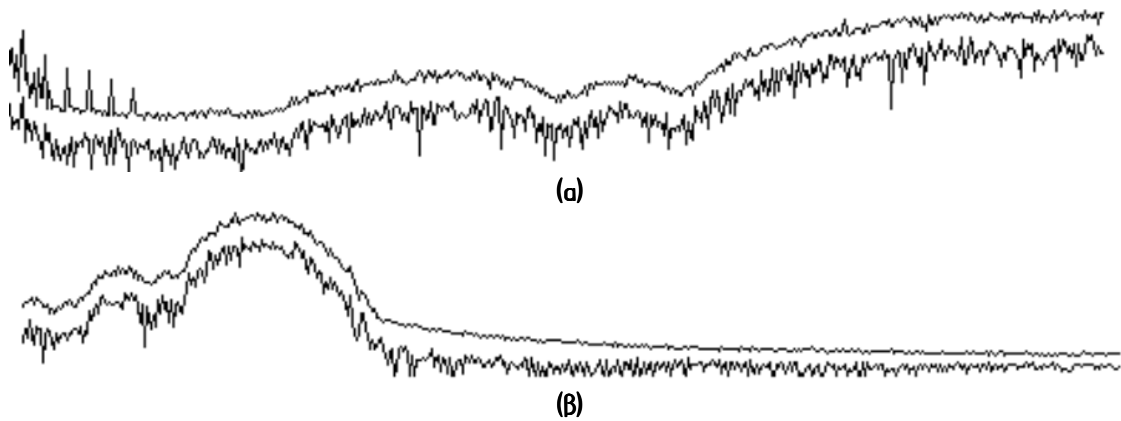
²¹ Όταν αναφερόμαστε στη μείωση του δυναμικού εύρους ενός ήχου δεν εννοούμε πως θα μειωθεί το ολικό του πλάτος (αυτό όπως βλέπουμε παρακάτω μπορεί και να αυξηθεί). Εννοούμε πως μειώνεται η διαφορά στις τιμές μεταξύ της συχνότητας με το μικρότερο πλάτος και αυτής με το μεγαλύτερο (δηλ. ένα είδος compression).



Σχ. 8.4.4: Πέντε παραδείγματα εφαρμογής Dither. (1) Εφαρμογή επίπεδου 'TPDF' dither. (2) Εφαρμογή dither 2^{ης} τάξης. (3) Εφαρμογή dither υψηλής τάξης φιλτραρισμένο στις χαμηλές συχνότητες. (4) Εφαρμογή dither υψηλής τάξης φιλτραρισμένο στις μεσαίες. (5) Εφαρμογή dither υψηλής τάξης φιλτραρισμένο στις υψηλές συχνότητες. Η ολική ηχητική ενέργεια των dither και στις πέντε περιπτώσεις είναι η ίδια. Η αύξηση του πλάτους του σήματος οφείλεται στην καμπύλη του dither.

8.5 Περιπτώσεις χρησιμοποίησης του flat dither:

Έχοντας περιγράψει στις προηγούμενες παραγράφους τα οφέλη που προκύπτουν από την εφαρμογή του noise shaping (της ανομοιόμορφης κατανομής της ενέργειας ανά συχνότητα) στο dithering, για ποιο λόγο κάποιος θα προτιμούσε να χρησιμοποιήσει έναν τύπο dither με ομοιόμορφη κατανομή της ενέργειας στο συχνοτικό του φάσμα; Η απάντηση σε αυτό το ερώτημα εξαρτάται από το είδος του σήματος (μορφή - είδος ηχητικού υλικού) στο οποίο καλείται να εφαρμοστεί το dither. Αν ομιλούμε για ηχογραφήσεις μουσικού υλικού οι οποίες θα «παιχτούν» στο σπίτι, στο αυτοκίνητο ή μέσω ενός ραδιοφωνικού σταθμού, τότε δεν υπάρχει κανένας λόγος να χρησιμοποιήσουμε το flat dither. Εάν όμως το είδος του σήματος στο οποίο θα εφαρμόσουμε το dithering υπάρχει πιθανότητα να υποστεί περαιτέρω επεξεργασία συν το χρόνο, τότε η συνθήκη στην οποία καλούμαστε να λειτουργήσουμε είναι αρκετά διαφορετική. Κάθε χρόνο παράγονται αρκετές χιλιάδες C.Ds τα οποία εμπεριέχουν είτε ηχητικά εφέ, είτε μουσικούς τόνους διαφόρων μουσικών οργάνων (βιβλιοθήκες) ως δείγματα για synthesizers. Και για τις δύο αυτές περιπτώσεις αναμένουμε το ηχητικό υλικό να υποστεί τροποποιήσεις όπως αλλαγή στην ταχύτητα εκτέλεσης του ή να υποβληθεί σε αλλαγή του τονικού του ύψους. Εάν αναπαράγουμε έναν ήχο, στον οποίο έχει εφαρμοστεί dither με noise shaping, σε διαφορετική από την αρχική του ταχύτητα τότε όλα τα χαρακτηριστικά του συχνοτικού φάσματος του ήχου, συμπεριλαμβανομένων αυτών του dither, έχουν μεταβληθεί είτε προς το άνω άκρο, είτε προς το κάτω άκρο, ανάλογα με το αν η ταχύτητα της εκτέλεσης είναι πιο αργή ή πιο γρήγορη από την αρχική. Το σχήμα 8.5.1 (επόμενη σελ.) δείχνει το συχνοτικό φάσμα ενός ήχου στον οποίο έχει εφαρμοστεί dither με noise shaping πριν και μετά την αλλαγή της ταχύτητας αναπαραγωγής του ήχου (δηλ. του τονικού του ύψους). Παρατηρούμε από το σχήμα αυτό πως η ενέργεια του συχνοτικού περιεχομένου έχει μετατοπιστεί προς τα κάτω (αλλαγή του τονικού ύψους προς τα κάτω) και στη συχνοτική αυτή περιοχή όπου το ανθρώπινο αυτί είναι περισσότερο ευαίσθητο. Αν και το ποσοστό της αλλαγής ταχύτητας που ασκήθηκε στο συγκεκριμένο ηχητικό δείγμα



Σχ. 8.5.1: (α) Συχνολικό φάσμα ενός ήχου στον οποίο έχει εφαρμοστεί dither με noise shaping. Παρατηρούμε πως οι κορυφές του σήματος είναι ευδιάκριτες στις χαμηλές & μεσαίες συχνότητες. (β) Το Συχνολικό φάσμα του ίδιου ήχου αφού του έχει γίνει μετατόπιση του τονικού του ύψους προς τα κάτω. Παρατηρούμε πως ενώ η περισσότερη ενέργεια του σήματος βρισκόταν πριν στις υψηλές συχνότητες, τώρα έχει μετατοπιστεί πιο κάτω στο φάσμα με αποτέλεσμα πλέον να είναι ακουστό. (Ο οριζόντιος άξονας στα (α) & (β) είναι σε γραμμική κλίμακα, 0 - 24000 Hz).

ήταν υπερβολικό, αυτό συνέβη για να μπορέσει να γίνει σαφές το τι ακριβώς συντελείται. Απο τα παραπάνω συμπεραίνουμε, όμως, πως οποιαδήποτε μεταβολή του τονικού ύψους ενός ήχου στον οποίο έχει εφαρμοστεί dither με noise shaping έχει οπωσδήποτε αρνητικό αντίκτυπο στην ποιότητα του ηχητικού υλικού. Απο όλα αυτά συμπεραίνουμε πως ήχοι οι οποίοι προορίζονται να χρησιμοποιηθούν ως εφέ ή ως δείγματα συνθετητών (synthesizers) θα πρέπει είτε να μην τους έχει εφαρμοστεί καθόλου dither (Q1), είτε να έχει εφαρμοστεί σε αυτούς flat dither (Q2)²². Σε καμία περίπτωση πάντως δε συνίσταται η χρήση dither με noise shaping.

8.6 Υπάρχει λόγος για τη μη χρήση οποιουδήποτε τύπου dither?:

Υπάρχουν πολλοί A/D converters οι οποίοι δεν εισάγουν κανένα τύπο dither (βλέπε Q3, σελ. 118), βασίζονται όμως στο να υπολογίζουν το σφάλμα κβαντοποίησης και το συχνολικό εύρος που αυτό επηρεάζει κάθε φορά και στη συνέχεια με βάση τον παραπάνω υπολογισμό οδηγούν ένα φίλτρο στην έξοδο του κυκλώματος το οποίο μετασχηματίζει την κατανομή της ενέργειας ανά συχνότητα με τρόπο τέτοιο ώστε να διορθώνονται τα σφάλματα που προκύπτουν απο την κβαντοποίηση (noise shaping). Το επιχείρημα, το οποίο καλείται κάθε φορά για να δικαιολογήσει αυτή την επιλογή σχεδίασης, είναι πως μην αθροίζοντας τίποτα στο αρχικό μας σήμα σημαίνει πως, με αυτό τον τρόπο, δεν αυξάνουμε και το σηματοθορυβικό του λόγο. Αν το σήμα προέρχεται απο έναν converter υψηλής πιστότητας και ανάλυσης το σφάλμα που προκύπτει απο τη στρωγγυλοποίηση των

²² Βλέπε σελ. 118

τιμών (truncation error) επηρεάζει ολόκληρο το ηχητικό φάσμα και είναι απόλυτα επαρκές για να οδηγήσει το noise shaping. Μπορούμε να διακρίνουμε ορισμένα προβληματικά σημεία σε αυτή την προσέγγιση:

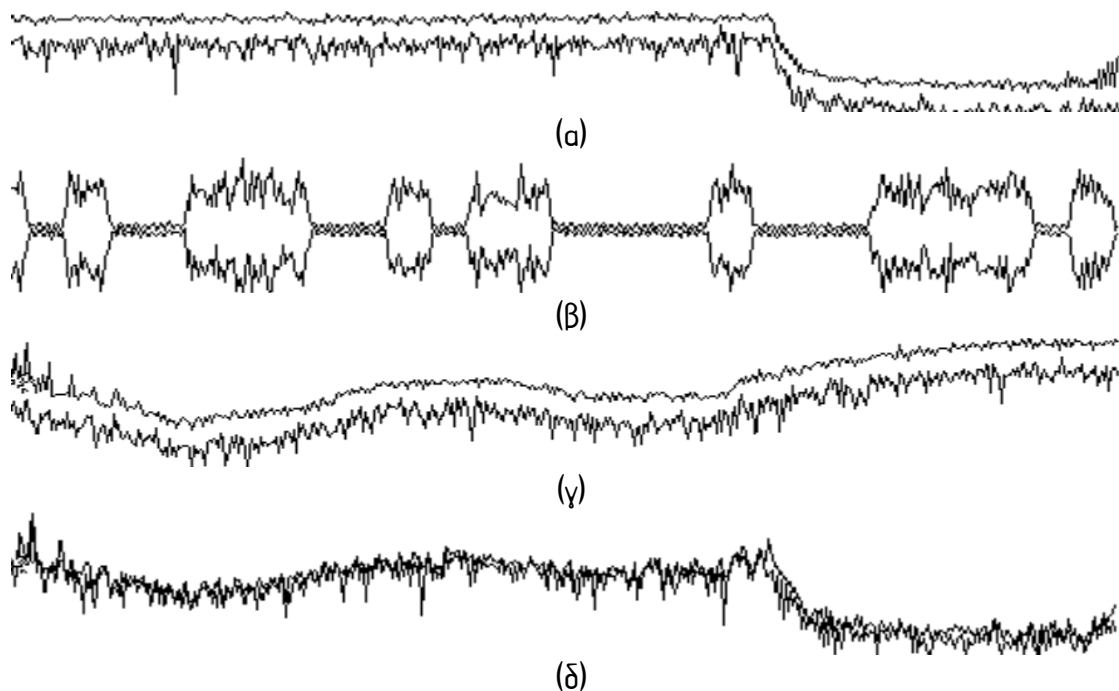
(α) Το σφάλμα που προκύπτει από τη στρογγυλοποίηση του σήματος συσχετίζεται άμεσα με το ίδιο το σήμα και γι' αυτό το λόγο θα μπορούσε να εμφανιστεί διαμόρφωση σήματος. Αυτό παράγει ένα κατώφλι θορύβου (noise floor) για τη συσκευή το οποίο δεν είναι σταθερό και το οποίο εξαρτάται από το ηχητικό υλικό. Αυτό μεταφράζεται στο ότι οι συσκευές αυτές είναι μη αξιόπιστες καθώς είναι αδύνατον να αξιολογήσουμε το σφάλμα κβαντοποίησης να μην είναι σε ποσοστό τέτοιο ώστε να δημιουργεί πρόβλημα, καθώς για ορισμένα μουσικά υλικά θα είναι μικρότερο και για ορισμένα άλλα μεγαλύτερο.

(β) Στο σήμα εμφανίζονται παύσεις (χρονικά κενά) και συνεπώς σε αυτά τα χρονικά σημεία ο converter παράγει μηδενικές τιμές. Παρόλα αυτά όμως ο converter τροφοδοτεί το noise shaping φίλτρο με μία οριακή τιμή. Αυτό σημαίνει πως το noise shaping φίλτρο θα παράγει μια συχνότητα (την κάτω οριακή). Συνεπαγωγή των παραπάνω είναι η παραγωγή μιας συγκεκριμένης συχνότητας και των αρμονικών της στα χρονικά αυτά σημεία του ηχητικού σήματος στα οποία κανονικά έχουμε παύση. Υπό συνθήκες το παραπάνω φαινόμενο γίνεται ιδιαίτερα ακουστό και ενοχλητικό.

Από τους παραπάνω λόγους προκύπτει πως δεν θα έπρεπε ένα οποιοδήποτε σύστημα A/D converter να μην προσθέτει στην έξοδο του κάποιο τύπο dither και αυτό διότι έτσι δημιουργούνται περισσότερα προβλήματα στο ηχητικό σήμα από αυτά που υποτίθεται πως αποφεύγονται.

8.7 Αναφορικά με τη συμπίεση των δεδομένων:

Υπάρχουν κάποιες συσκευές ή προγράμματα οι/τα οποίες/α προκαλούν συμπίεση των δεδομένων του ηχητικού σήματος ώστε να ελαττώσουν το μέγεθος του, πράγμα που έχει ως συνέπεια τον περιορισμό του συχνοτικού ή του δυναμικού του εύρους. Όλες αυτές οι συσκευές χρησιμοποιούν, για να επιτύχουν την καλύτερη δυνατή ποιότητα κατά την αναπαραγωγή, έναν τύπο κωδικοποίησης ο οποίος χωρίζει το ηχητικό υλικό σε συγκεκριμένες μπάντες συχνοτήτων και για το εύρος κάθε μπάντας συχνοτήτων χρησιμοποιείται και ο ανάλογος αριθμός των bit για την κωδικοποίηση της. Ο αριθμός των bit για κάθε μπάνα προκύπτει, όπως είναι φυσικό, από το συχνοτικό εύρος της μπάνας. Οι διαφορές που παρουσιάζονται από συσκευή σε συσκευή και από πρόγραμμα σε πρόγραμμα έχουν να κάνουν με τους αλγόριθμους που χρησιμοποιούνται από το καθένα από αυτά, από τον τρόπο με τον οποίο χωρίζεται και κωδικοποιείται το ηχητικό υλικό στις συχνοτικές αυτές μπάντες και τέλος από την κατανομή των bit στις εν λόγω μπάνες. Θα μπορούσε να τεθεί το ερώτημα του αν υπάρχει κάποια αλληλεπίδραση μεταξύ της μεθόδου κβαντοποίησης που



Σχ. 8.7.1: (α) Το φάσμα ενός ήκου με flat dither το οποίο έχει επεξεργαστεί απο ένα ευρέως διαδεδομένο πρόγραμμα συμπίεσης. Παρατηρείται πως με το συγκεκριμένο αλγόριθμο το ποσοστό της πληροφορίας απο τα 15 kHz και πέρα έχει ελαχιστοποιηθεί. (β) Η χρονική αναπαράσταση ενός ήκου στον οποίο έχει εφαρμοστεί dither υψηλής τάξης (noise shaped dither) το οποίο έχει επεξεργαστεί απο ένα ευρέως διαδεδομένο αλγόριθμο συμπίεσης. Το πλάτος του dither έχει αυξηθεί λόγω της μη ορθής κωδικοποίησης και μάλιστα ιδιαίτερα στις υψηλές συχνότητες. (γ) Το φάσμα του ήκου που φαίνεται στο διάγραμμα (β) το οποίο επιβεβαιώνει την αύξηση του πλάτους στις υψηλές συχνότητες. (δ) Το φάσμα του ήκου που φαίνεται στο διάγραμμα (α) μεγεθυμένο γύρω απο την περιοχή των 15 kHz.

χρησιμοποιήθηκε (P.C.M., Delta/Sigma) και του αλγόριθμου που επιτελεί την συμπίεση των δεδομένων. Ένας πιο πρακτικός τρόπος για να τεθεί το ερώτημα, όμως, είναι να διερωτηθούμε αν υπάρχουν μέθοδοι κβαντοποίησης του ηχητικού υλικού οι οποίες συμπεριφέρονται καλύτερα στην συμπίεση του ηχητικού σήματος απο κάποιες άλλες. Επίσης θα πρέπει να θέσουμε και το ερώτημα της συμπεριφοράς των μέσων αποθήκευσης σε ότι αφορά την συμπίεση του ηχητικού υλικού. Τέλος το επίδικο σε όλη αυτή την αναλυτική εξακολουθεί να παραμένει η διαδικασία του dithering και ο τρόπος με τον οποίο σχετίζεται με όλα τα παραπάνω.

Στο Σχ. 8.7.1 τα διαγράμματα (α) έως (δ) παρουσιάζουν κάποια απο τα αποτελέσματα της συμπίεσης, απο ένα ευρέως διαδεδομένο πρόγραμμα, ενός audio σήματος στο οποίο έχει εφαρμοστεί dither. Το διάγραμμα (α) μας δείχνει πως ο τύπος του TPDF dither (βλ. σελ. 118) όταν εφαρμοστεί παράλληλα με τη συμπίεση του ηχητικού υλικού οδηγεί στην εξασθένηση της μπάντας των υψηλών συχνοτήτων. Το διάγραμμα (β) μας δείχνει πως ένα dither υψηλής τάξης (noise shaped dither) μπορεί να αυξήσει την κατανομή της ενέργειας στις υψηλές συχνότητες. Αν και δεν παρουσιάζουν όλοι οι αλγόριθμοι συμπίεσης αυτού του είδους την προβληματική συμπεριφορά, θα πρέπει να

υπογραμμιστεί πως δεν θα πρέπει να αξιώνεται a priori πως η κβαντοποίηση του σήματος που προκύπτει από την P.C.M. διεργασία μπορεί να υπόκειται σε συμπίεση δίχως εκπτώσεις στην ηχητική ποιότητα. Θα πρέπει κάθε φορά να αξιολογείται το ηχητικό υλικό έπειτα από τη συμπίεση των δεδομένων του και να αποσαφηνίζεται αν είναι αποδεκτό σε αυτή τη μορφή ή όχι.

Από τα παραπάνω προκύπτει πως ο τεχνικός mastering θα πρέπει να εφαρμόσει διαφορετικό τύπο dither στο μουσικό υλικό αν αυτό συμπιεστεί (π.χ. για να διανεμηθεί στο κοινό μέσω του διαδικτύου) απ' ό,τι αν αυτό κυκλοφορούσε σε Compact Disc. Αυτό μας οδηγεί στο συμπέρασμα πως ο τεχνικός mastering θα πρέπει να παράγει διαφορετικές version του ίδιου μουσικού υλικού ανάλογα με το μέσο στο οποίο αυτό καλείται να αποθηκευτεί (π.χ. βινύλιο, Compact Disc, διαδίκτυο). Η δυσκολία στο να συμβεί αυτό όμως είναι το κόστος που αυτό συνεπάγεται, πράγμα που σημαίνει πως φαινόμενα όπως αυτό που παρουσιάζεται στο Σχ. 8.7.1 (α) απαντώνται εξαιρετικά συχνά.

Οι περισσότεροι αλγόριθμοι συμπίεσης ηχητικών δεδομένων έχουν τη δυνατότητα να δέχονται audio σήμα υψηλής ανάλυσης αντί για σήματα των οποίων η ανάλυση έχει «κατέβει» ήδη στα 16 bit. Στην πράξη έχει αποδειχτεί πως η συμπίεση του ηχητικού υλικού υψηλής ανάλυσης προσφέρει πλεονεκτήματα όταν το ηχητικό αυτό υλικό δεν πρόκειται να αποθηκευτεί σε Compact Disc. Αντιθέτως στην περίπτωση που το ηχητικό υλικό αποθηκεύεται σε C.D. είναι προτιμότερη η απλή μείωση της ανάλυσης στα 16 bit χωρίς καμία περεταίρω συμπίεση. Αυτό ισοδυναμεί με το να πούμε πως κάποιος χρησιμοποιεί το σφάλμα που προκύπτει από τη διαδικασία της συμπίεσης και της αποσυμπίεσης του ηχητικού υλικού κατά την αναπαραγωγή για να επιτύχει ακριβώς τα ίδια αποτελέσματα που επιτυγχάνονται με τη χρήση dither και noise shaping. Η παραπάνω πρακτική δεν είναι κατανάγκη καταδικαστέα· η επιτυχία της όμως εξαρτάται καταλυτικά από την ποιότητα συμπίεσης και αποσυμπίεσης του αλγόριθμου που χρησιμοποιείται.

8.8 Κάποια πρώτα συμπεράσματα (πριν συνεχίσουμε):

Είναι σημαντικό να σημειωθούν τα ακόλουθα:

- Το flat dither (ομοιόμορφη κατανομή της ενέργειας ανά συχνότητα, Q2) θα πρέπει να χρησιμοποιείται σε ήχους οι οποίοι προορίζονται για περαιτέρω επεξεργασία (sample rate conversion, αλλαγή του τονικού τους ύψους, κ.ο.κ).
- Τα dither με εφαρμογή noise shaping (ανομοιόμορφη κατανομή της ενέργειας ανά συχνότητα Q4) θα πρέπει να χρησιμοποιούνται για την εφαρμογή σε 16 bit γραμμικό P.C.M αποθηκευτικό μέσο (όπως π.χ. το C.D).

Σε αυτή την περίπτωση ο τεχνικός mastering θα πρέπει να διαθέτει την εμπειρία ώστε να μπορεί να επιτύχει τα επιθυμητά αποτελέσματα.

- Τα audio σήματα τα οποία προορίζονται για αποθηκευτικά μέσα τα οποία απαιτούν τη συμπίεση του ηχητικού υλικού θα πρέπει να παραμένουν σε υψηλή ανάλυση και δίχως την εφαρμογή dithering. Οι αλγόριθμοι συμπίεσης θα πρέπει γι' αυτό το λόγο να μπορούν να δέχονται σήματα υψηλής ανάλυσης και να είναι εξαιρετικής ποιότητας ώστε να μην παράγονται ακουστές παραμορφώσεις.
- Οι κατασκευαστές συσκευών A/D converters καθώς και εκείνοι που παράγουν αλγόριθμους dither θα πρέπει να παρέχουν στο χρήστη πλειάδα επιλογών αναφορικά με τον τύπο του dither και του noise shaping ώστε να είναι εφικτό το «ζευγάρωμα» του dither με το μουσικό υλικό.

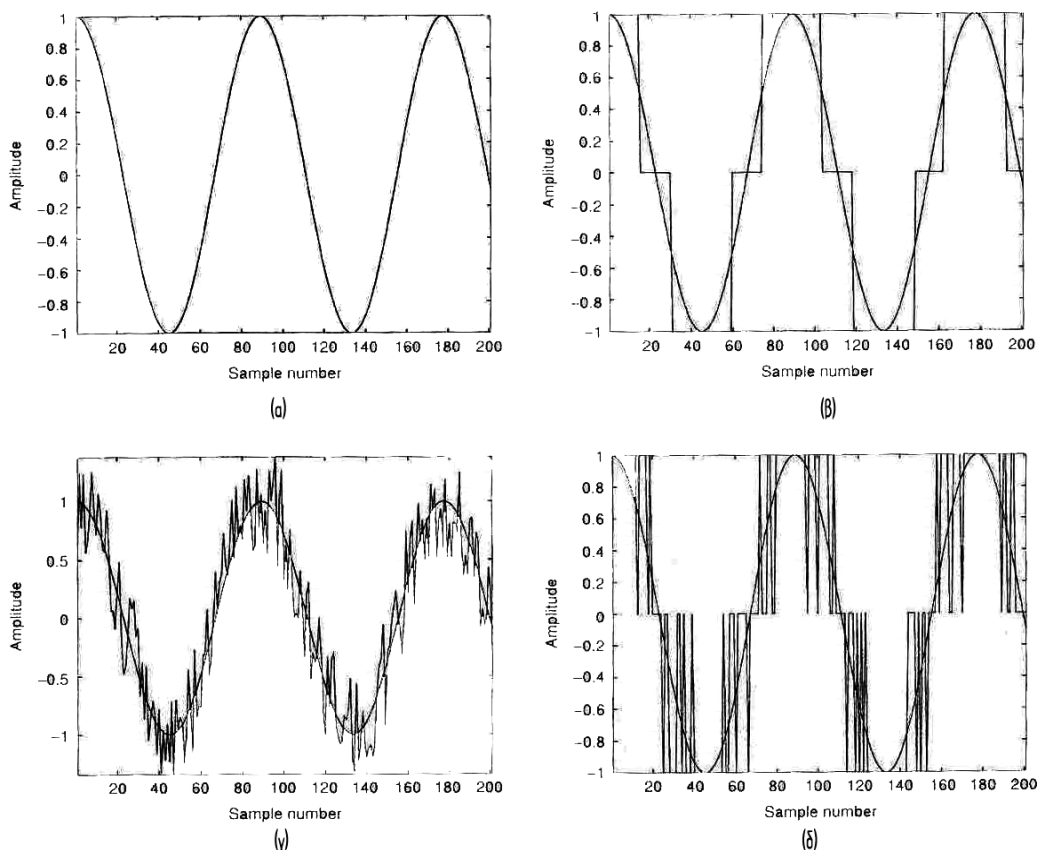
Τα παραπάνω αποτελούν ορισμένα βασικά σημεία εστίασης σε ότι αφορά τη λειτουργία του dithering στο mastering.

8.9 Περαιτέρω εμβάθυνση στο Dither:

Με audio σήματα τα οποία διαθέτουν επαρκές πλάτος έντασης και είναι αρκετά περίπλοκα (π.χ. μια συμφωνική ορχήστρα), το σφάλμα που προκύπτει από τη διαδικασία της κβαντοποίησης επηρεάζει σε πολύ μικρό ποσοστό το ηχητικό υλικό· το σφάλμα αυτό κατανέμεται με τυχαίο τρόπο στο συχνοτικό φάσμα του ήχου και μοιάζει αρκετά με το λευκό θόρυβο των αναλογικών ταινιών. Με σήματα χαμηλού πλάτους ο χαρακτήρας του σφάλματος κβαντοποίησης αλλάζει δραστικά και παρουσιάζονται πιθανές ακουστές παραμορφώσεις του ηχητικού σήματος. Ένα σύστημα ψηφιοποίησης πρέπει να εμποδίζει την εμφάνιση οποιουδήποτε σφάλματος που προκύπτει στην έξοδο του από τη στρωγγυλοποίηση των δειγμάτων του ηχητικού σήματος που δέχεται στην είσοδο του. Προφανώς, ο αριθμός των bit του wordlength²³ μπορεί να αυξηθεί, με αποτέλεσμα για κάθε bit που προστίθεται να έχουμε μείωση στο πλάτος έντασης του παραγόμενου σφάλματος κατά 6 dB. Κάτι τέτοιο όμως δεν είναι πάντοτε εφικτό, όπως επίσης δεν αποτελεί και μια δόκιμη λογική καθώς θα απαιτείτο μεγάλος αριθμός bit για να ικανοποιηθεί η απαίτηση μας για την εξασθένιση του σφάλματος κβαντοποίησης, πράγμα που θα σήμαινε την αύξηση του μεγέθους των αρχείων ήχου και τη χρήση μεγαλύτερων χώρων αποθήκευσης τους, τη χρήση ολοένα και περισσότερης επεξεργαστικής ισχύς για τα DWA συστήματα, καθώς και την αύξηση του κόστους παραγωγής για τις hardware συσκευές.

²³ Εννοούμε το μήκος της ψηφιακής λέξης, δηλ. τον αριθμό των bit που χρησιμοποιείται για να εκφραστεί το ηχητικό υλικό ή αλλιώς την ανάλυση του audio σήματος, bit rate ή bit depth, τεχνικά μιλώντας εννοούμε να αυξήσουμε τις στάθμες κβαντοποίησης με αποτέλεσμα να μειωθεί το βήμα κβαντοποίησης.

Τη λύση σε όλα τα παραπάνω δίνει το dither. Dither ονομάζουμε ένα πολύ μικρό ποσοστό θορύβου το οποίο προστίθεται στο LSB (Last Significant Bit, Λιγότερο Σημαντικό Ψηφίο) του ψηφιακού σήματος και ουσιαστικά εξαλείφει αυτού του τύπου τα σφάλματα και τα παράγωγά τους. Τον τρόπο με τον οποίο λειτουργεί το dither μπορούμε να τον δούμε στο Σχ. 8.9.1 (α) - (δ). Το διάγραμμα (α) παρουσιάζει ένα συνημίτονο με κανονικοποιημένο πλάτος. Το διάγραμμα (β) παρουσιάζει το συνημίτονο του διαγράμματος (α) καθώς και αυτό που προκύπτει, από την ψηφιοποίησή του, χρησιμοποιώντας δύο στάθμες κβαντισμού (δηλ. χρησιμοποιώντας δύο bit για την ανάλυση του σήματος). Επίσης από το διάγραμμα (β) γίνεται σαφές ότι στο συγκεκριμένο παράδειγμα η ψηφιοποίηση αυτή λειτουργεί ως limiter με πολύ χαμηλό threshold, δηλ. παράγει πολύ υψηλό ποσοστό παραμόρφωσης (τετραγωνισμός του σήματος). Όπως όμως φαίνεται στα διαγράμματα (γ) και (δ) το φαινόμενο που παρατηρούμε στην έξοδο διαφοροποιείται ριζικά όταν αθροίσουμε στο συνημίτονο του παραδείγματος ένα ποσοστό dither με πλάτος ίσο με μία στάθμη κβαντοποίησης (ένα bit). Το νέο σήμα εξόδου λοιπόν (διαγρ. (δ)) αποτελείται από ένα πλήθος αυτή τη



Σχ. 8.9.1: Παρουσίαση του τρόπου με τον οποίο το dither μπορεί να εξαλείψει τα ψηφιακά σφάλματα. (α) Ένα συνημίτονο με πλάτος της τάξης των δύο λιγότερων σημαντικών ψηφίων στο οποίο δεν έχει εφαρμοστεί dithering. (β) Το αποτέλεσμα που προκύπτει από την κβαντοποίηση του σήματος εισόδου. Παρατηρούμε τα σφάλματα που προκύπτουν στο audio σήμα, μιας και για την περιγραφή του χρησιμοποιούνται μόνο τρία διακριτά επίπεδα στάθμεων. (γ) Το σήμα εισόδου με την εφαρμογή dither σε αυτό. (δ) Η κβαντοποίηση του σήματος αυτή τη φορά παράγει στην έξοδο ένα σήμα το οποίο αποτελείται από ένα πλήθος τετραγωνικών παλμών που περιγράφουν καλύτερα το αρχικό μας σήμα.

φορά τετραγωνικών παλμών που περιγράφουν με πολύ καλύτερο τρόπο το αρχικό μας συνημίτονο. Αυτό συνέβη διότι λόγω της τυχειότητας της μορφής του dither το αρχικό σήμα εισόδου ενεργοποίησε είτε τη μια είτε την άλλη από τις δύο στάθμες κβαντισμού για περισσότερα samples του σήματος απ' ό,τι στην αρχική περίπτωση (διαγρ. (α) & (β)).

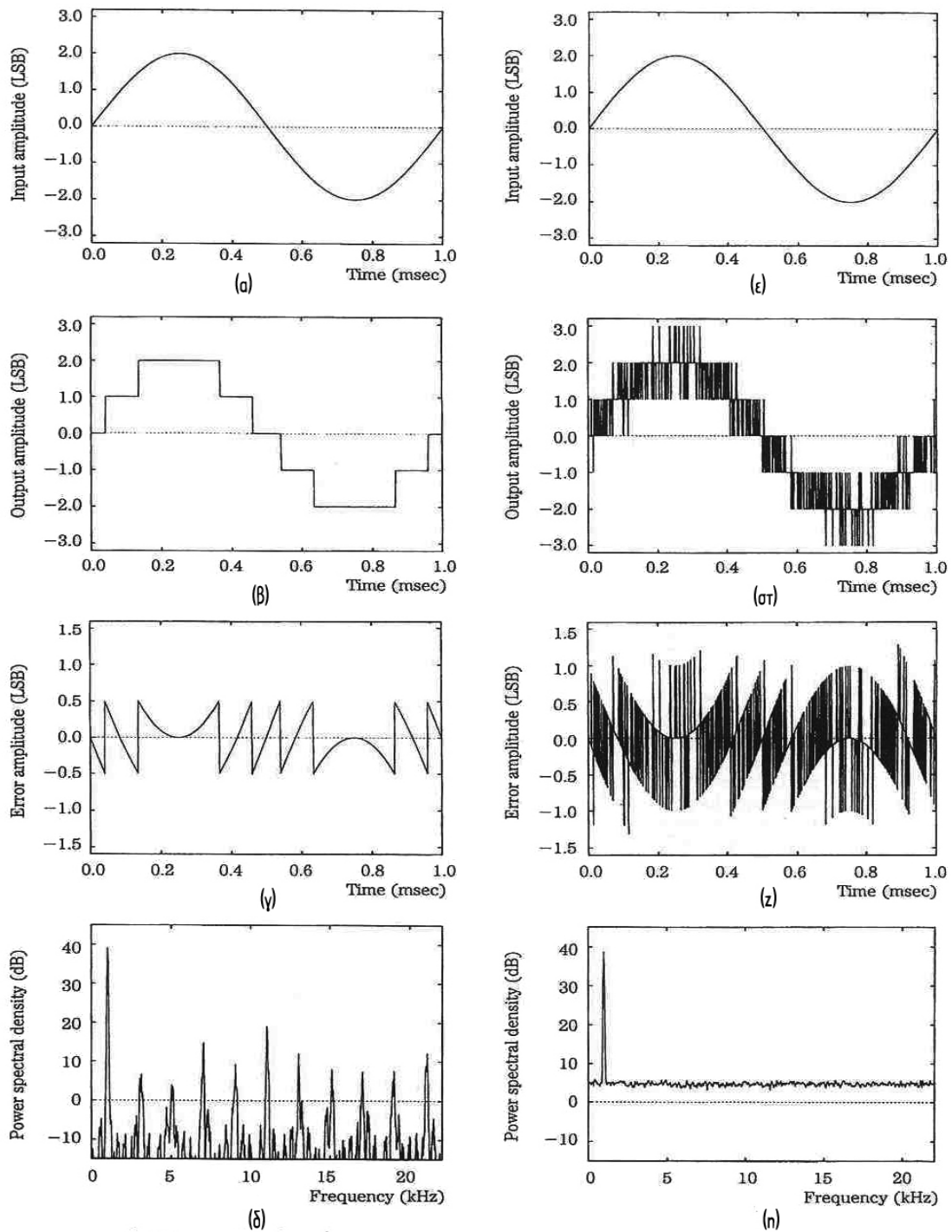
Προχωρώντας περαιτέρω την αναλυτική μας θα επιχειρήσουμε να συνδέσουμε τη λειτουργία του dithering με τον τρόπο λειτουργίας του ανθρώπινου αισθητηρίου ακοής καθώς και τον τρόπο με τον οποίο αυτό αντιλαμβάνεται το dither. Το αυτό, λοιπόν, είναι εξαιρετικά ικανό στο να αναλύει σήματα στενού συχνοτικού εύρους ακόμα και όταν υφίσταται ένα κάποιο ποσοστό θορύβου σε αυτά. Έχει δηλαδή την ικανότητα να συλλαμβάνει ήχους και κάτω από το κατώφλιο θορύβου. Αυτή η ικανότητα του οφείλεται στην ιδιότητα της βασικής μεμβράνης του τυμπάνου να εξαγεί το μέσο όρο (μέση τιμή²⁴) της ενέργειας που εμπερικλείεται στο ακουστικό σήμα. Το αυτό μας, επίσης, συμπεριφέρεται ως ένας φασματικός αναλυτής του 1/3 της οκτάβας με πολύ στενό εύρος· από τα παραπάνω προκύπτει πως το σφάλμα που παράγεται από τη διαδικασία της κβαντοποίησης έχει πάρει, με τη χρήση του dither, τη μορφή λευκού θορύβου πράγμα που με τη σειρά του σημαίνει πως το αυτό παρέχει τη μέση τιμή του σήματος και συνεπώς γίνεται ακουστό το μουσικό περιεχόμενο δίχως όμως να γίνεται αντιληπτό το σφάλμα. Με άλλα λόγια η εφαρμογή του dither μετασχηματίζει τη μορφή του ψηφιακού σφάλματος που προκύπτει από το conversion σε λευκό θόρυβο και επομένως το αυτό μπορεί να συλλάβει (με βάση τη λειτουργία που συνοπτικά σε αυτή την παράγραφο έχουμε αναφέρει) ήχους πλάτους μικρότερους ακόμα και μίας στάθμης κβαντισμού. Αυτός είναι και ο λόγος για τον οποίο με την εφαρμογή του dithering αναδεικνύονται ιδιαίτερα τα fade in & fade out καθώς και οι ουρές των reverbs. Το συμπέρασμα αυτό είναι σημαντικό, διότι με την εφαρμογή του dithering η αναλυτική ικανότητα ενός ψηφιακού συστήματος επεκτείνεται, θεωρητικά, και πέρα από το L.S.B. (Last Significant Bit).

Παραπέρα, η εφαρμογή dither μπορεί να εξαλείψει την παραμόρφωση που προκύπτει υπό τη μορφή clicks μετασχηματίζοντας τα και αυτά σε λευκό θόρυβο και τοποθετώντας τα σε υψηλότερες περιοχές του φάσματος (μέσω του noise shaping) όπου το αυτό μας είναι λιγότερο ευαίσθητο.

Προχωρώντας σε ένα δεύτερο παράδειγμα (από τους Lipshitz, Wannamaker & Vanderkooy, 1992) θα εκφράσουμε τα διαγράμματα που παρουσιάζονται στο Σχ. 8.9.2 (σελ. 132). Ξεκινώντας από το διάγραμμα (δ) παρατηρούμε πως το φάσμα του σήματος στο οποίο δεν έχει εφαρμοστεί dither υποφέρει από αρμονική παραμόρφωση, η οποία

²⁴ Μαθηματικά, η σχέση που εκφράζει τη μέση τιμή είναι η: $x_{RMS} = \sqrt{\frac{1}{T} \int_0^T x^2(t) dt}$.

Όπου x : η μαθηματική περιγραφή του σήματος, και T ο χρόνος ολοκλήρωσης για μία περίοδο του σήματος.



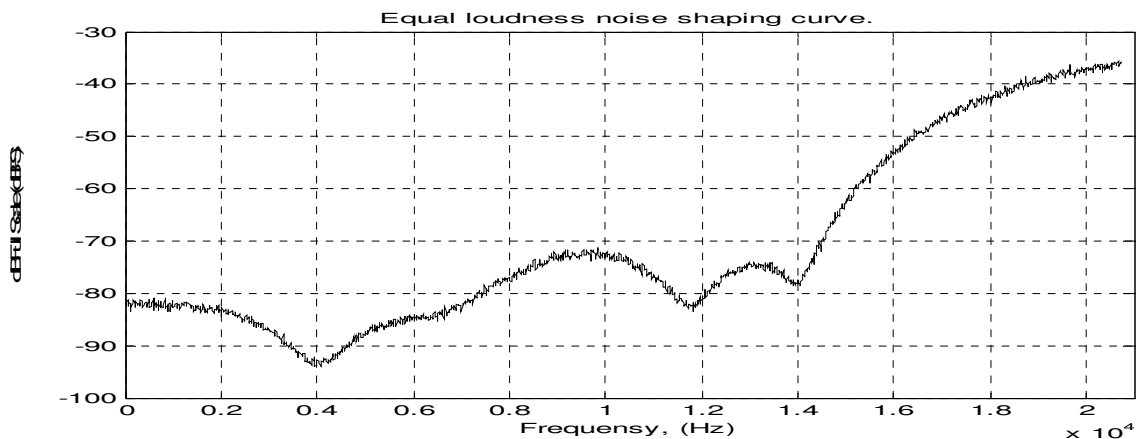
Σχ. 8.9.2: Εξομοίωση σε υπολογιστικό περιβάλλον της κβαντοποίησης ενός καθαρού τόνου στο 1 kHz με και χωρίς dither. (α) Το σήμα εισόδου. (β) Το σήμα εξόδου. (γ) Το ποσοστό σφάλματος στο σήμα δίχως τη χρήση dither. (δ) Το φασματικό περιεχόμενο του σήματος εξόδου δίχως τη χρήση dither. (ε) Το σήμα εισόδου. (στ) Το σήμα εξόδου με τη χρήση TPDF dither (Triangular Probability Density Function). (ζ) Το ποσοστό σφάλματος στο σήμα με τη χρήση dither. (η) Το φασματικό περιεχόμενο του σήματος με dither. (Lipshitz, Wannamaker & Vanderkooy, 1992).

απαρτίζεται απο συχνότητες οι οποίες αποτελούν ακέραια πολλαπλάσια της θεμελιώδους. Παρατηρούμε επίσης ότι υφίσταται και μη αρμονική παραμόρφωση η οποία έχει προκύψει απο το φαινόμενο του aliasing. Το διάγραμμα (z) μας δείχνει το σφάλμα που έχει προκύψει απο την κβαντοποίηση του σήματος στο οποίο (αυτή τη φορά) έχει εφαρμοστεί dither. Παραδόξως το σήμα που παρουσιάζεται στο διάγραμμα (z) ηχεί σαν λευκός θόρυβος, άσχετα αν απ' το διάγραμμα δε φανερώνεται κάτι τέτοιο, και το σήμα εξόδου ηχεί ως ημίτονο με την πρόσθεση του dither όπως φαίνεται στο διάγραμμα (n). Στο διάγραμμα (n) καταδεικνύεται επίσης πως το αρχικό μας σήμα είναι πλέον απαλλαγμένο απο τα παράγωγα του τα οποία προέκυψαν απο το σφάλμα κβαντοποίησης και έχουν περιγραφεί παραπάνω. Τέλος μπορούμε να δούμε πως το dither αυξάνει το πλάτος του σήματος αυξάνοντας παράλληλα και το noise floor.

8.10 Noise Shaping:

Απο το Σχ. 8.9.2 και το διάγραμμα (n) βλέπουμε πως ενώ το αρχικό μας σήμα είναι απαλλαγμένο απο την παραμόρφωση που προέκυψε απο την κβαντοποίηση του, έχουμε παράλληλα αύξηση του πλάτους του θορύβου σε σημείο τέτοιο (και στις συχνотικές αυτές περιοχές) όπου πλέον γίνεται ακουστό. Για το λόγο αυτό χρησιμοποιούμε το noise shaping. Η λειτουργία του noise shaping είναι εξαιρετικά απλή: εφαρμόζεται η καμπύλη ενός φίλτρου υψηλής τάξης στο dither και έτσι, ενώ πριν το dither είχε μία ομοιόμορφη κατανομή της ενέργειας σε όλο το συχνотικό φάσμα, πλέον το μεγαλύτερο ποσοστό της ενέργειας έχει μετακινηθεί στην υψηλότερη συχνотική περιοχή με αποτέλεσμα το noise floor στις περιοχές όπου το ανθρώπινο αυτί είναι περισσότερο ευαίσθητο να μειώνεται και συνεπώς να μη γίνεται ακουστό το dither. Η λογική αυτή εκπορεύεται απο το γεγονός πως η ολική ηχητική ενέργεια του θορύβου δεν περιγράφει κατανάγκη και την ακουστότητα του, διότι εξαρτάται επίσης και απο τα φασματικά χαρακτηριστικά του εν λόγω θορύβου.

Ο τρόπος, όμως, με τον οποίο θα μειωθεί το ποσοστό του θορύβου στις συχνотικές περιοχές που μας ενδιαφέρουν, ενώ παράλληλα θα αυξηθεί το ποσοστό του στις υψηλές συχνотότητες (στις οποίες δεν γίνεται ακουστός), απαιτείται να είναι τέτοιος ώστε η κατανομή της ενέργειας ανά συχνотότητα να είναι η κατάλληλη. Για παράδειγμα ένα πολύ μεγάλο ποσοστό θορύβου στις υψηλές συχνотότητες μπορεί να βλάψει τα tweeters των ηχείων. Όπως επίσης πολλοί ακροατές παρατηρούν πως η άγαρμπη ενίσχυση των υψηλών συχνотότητων του dither παράγει συχνотότητες οι οποίες δεν υπήρχαν αρχικά στο σήμα μας ή ότι πιθανώς μπορεί να μασκάρουν μέρος της ακουστικής πληροφορίας που φέρει το σήμα. Πρακτικά όλα τα παραπάνω εξαρτώνται απο το σχεδιασμό του noise shaping το οποίο θα εφαρμοστεί στο dither. Στο Σχ. 8.10.1 παρουσιάζεται η μορφή ενός dither στο οποίο έχει εφαρμοστεί noise



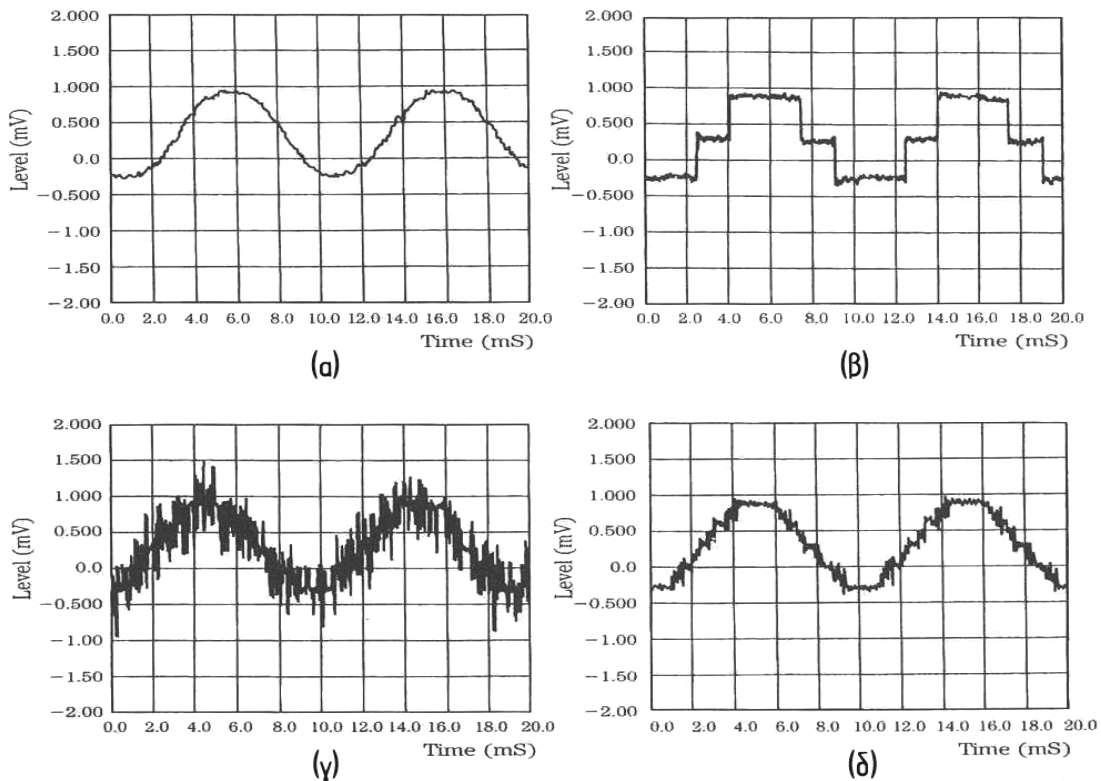
Σχ. 8.10.1: Η κατανομή ενέργειας ανά συχνότητα σε dither στο οποίο έχει εφαρμοστεί noise shaping.

shaping. Στο συγκεκριμένο σχήμα ο άξονας των συχνοτήτων δίδεται σε γραμμική κλίμακα για την καλύτερη παρουσίαση της καμπύλης του noise shaping.

Στις περισσότερες περιπτώσεις η καμπύλη του noise shaping δίδεται ως standard από τον κατασκευαστή και δεν επιδέχεται αλλαγές. Σε κάποιες άλλες, όμως, η καμπύλη του noise shaping μπορεί να μετασχηματιστεί με βάση τα χαρακτηριστικά του audio σήματος στο οποίο επιθυμούμε να την εφαρμόσουμε. Ομοίως, αλλιώς δε θα είχε νόημα το παραπάνω, το ίδιο ισχύει και για το dither ώστε το audio σήμα να μασκάρει το ποσοστό θορύβου που έχει αθροιστεί μέσω του dither. Μία μέθοδος με την οποία μπορούμε να εφαρμόσουμε τα παραπάνω είναι μέσω ενός αναλυτή φάσματος να εξάγουμε τα χαρακτηριστικά του audio σήματος και στη συνέχεια βάσει των εξαγόμενων δεδομένων να ορίσουμε τα χαρακτηριστικά του noise shaping και του dither. Υπάρχουν σουίτες προγραμμάτων που μας επιτρέπουν να προβούμε σε αυτές τις ενέργειες (π.χ. Sonic Studio HD & SADiE v.4 & v.5), όπως επίσης και hardware συσκευές (π.χ. Weiss, Gambit ADC2).

Παρακάτω στο Σχ. 8.10.2 (Pohlmann, 2005) παρουσιάζεται ένας τόνος του 1 kHz με πλάτος στα -90 dBFS. Οι μετρήσεις που παρουσιάζονται έχουν παρθεί με ένα χαμηλοπερατό φίλτρο με συχνότητα αποκοπής στα 16 kHz ώστε να προσομοιωθεί όσο το δυνατόν καλύτερα η συμπεριφορά του ανθρώπινου αυτιού.

Παρατηρούμε πως η ανάλυση του τόνου στα 20 bit (διάγραμμα (α)) είναι αρκετά ακριβής, όταν όμως «κατεβάσουμε» τον τόνο στα 16 bit κάνουν την εμφάνισή τους τα σφάλματα που έχουν προκύψει από την κβαντοποίηση του σήματος (διάγραμμα (β)). Στη συνέχεια εφαρμόζοντας TPDF dither της τάξης του ±1 LSB στο σήμα των 16 bit (διάγραμμα (γ)) βλέπουμε πως ο θόρυβος που έχει προκύψει από την κβαντοποίηση απαλείφεται και το σήμα είναι στα χαρακτηριστικά του αρχικού των 20 bit (διάγραμμα (δ)). Παράλληλα, όμως, παρατηρούμε πως αυξάνεται το ποσοστό του θορύβου στο σηματοθορυβικό λόγο του σήματος.

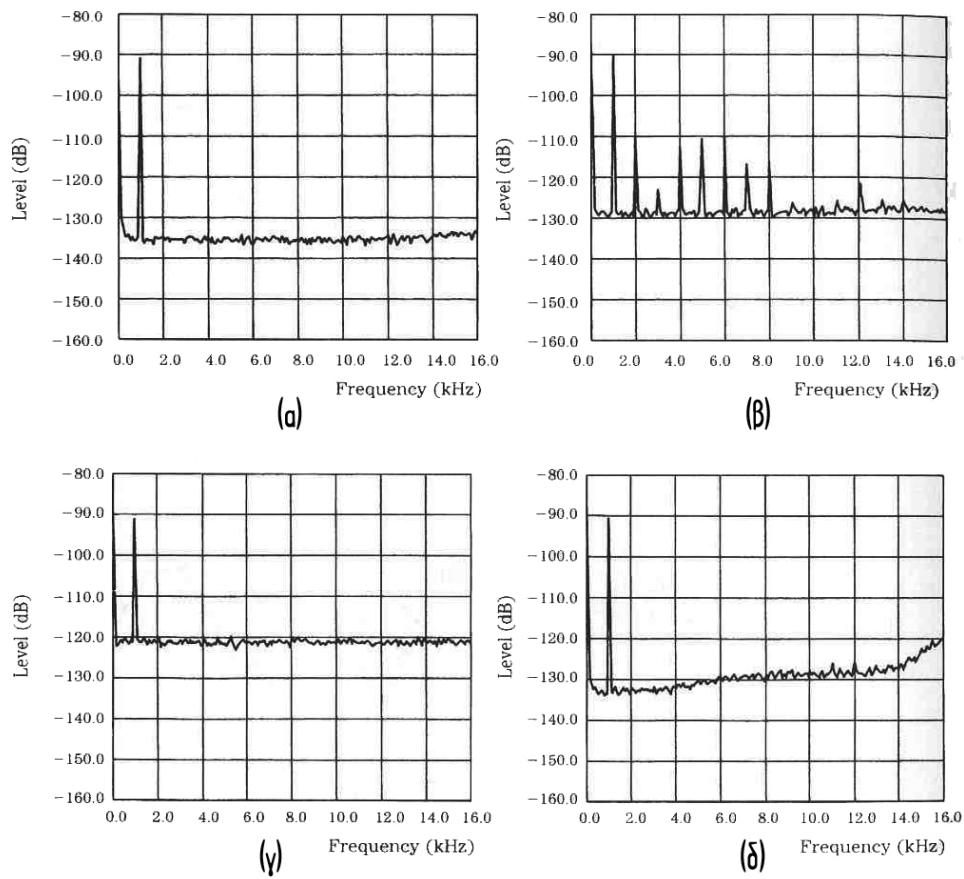


Σχ. 8.10.2: Παράδειγμα εφαρμογής dither με και χωρίς noise shaping σε έναν τόνο του 1 kHz με πλάτος -90 dBFS. (α) Ο τόνος στο αρχικό του bit rate στα 20 - bit. (β) Ο τόνος έπειτα απο το «κατέβασμα» του στα 16 - bit. (γ) Εφαρμογή flat dither στον τόνο των 16 - bit. (δ) Εφαρμογή dither με noise shaping. Παρατηρούμε πως το ποσοστό του θορύβου στο σήμα είναι μειωμένο, πρακτικά αυτό σημαίνει πως διασώζεται η πληροφορία που εμπεριέχουν τα τέσσερα λιγότερο σημαντικά ψηφία.

Η λύση στον παραπάνω αρνητικό αντίκτυπο του dither είναι το noise shaping. Το Σχ. 8.10.3 της επόμενης σελίδας που ακολουθεί, περιγράφει τα φάσματα που προκύπτουν απο τον ίδιο τόνο του 1 kHz και πλάτους -90 dBFS.

Το διάγραμμα (α) μας δείχνει το φάσμα στα 20 bit. Το διάγραμμα (β) μας δείχνει το φάσμα του τόνου μετά το «κατέβασμα» του στα 16 bit. Το διάγραμμα (γ) μας δείχνει τον τόνο στα 16 bit με την εφαρμογή dither, όπου το σφάλμα απαλείφεται αλλά αυξάνεται το ποσοστό του θορύβου στο σηματοθορυβικό λόγο του σήματος. Το διάγραμμα (δ) παρουσιάζει το φάσμα του τόνου στα 16 bit στον οποίο έχει εφαρμοστεί dither με noise shaping. Παρατηρούμε το noise floor να έχει μειωθεί στις περιοχές όπου το ανθρώπινο αυτί είναι περισσότερο ευαίσθητο και να έχει αυξηθεί στις υψηλές συχνότητες στις οποίες δεν είναι ακουστό.

Τέλος παρατηρούμε με αυτόν τον τρόπο πως το noise shaping δεν εφαρμόζεται απλά και μόνο για αισθητικούς λόγους (ταίριασμα του dither με την κατανομή του φάσματος του μουσικού κομματιού), αλλά επίσης και για τεχνικούς λόγους που αφορούν την ποιότητα του ηχητικού σήματος.

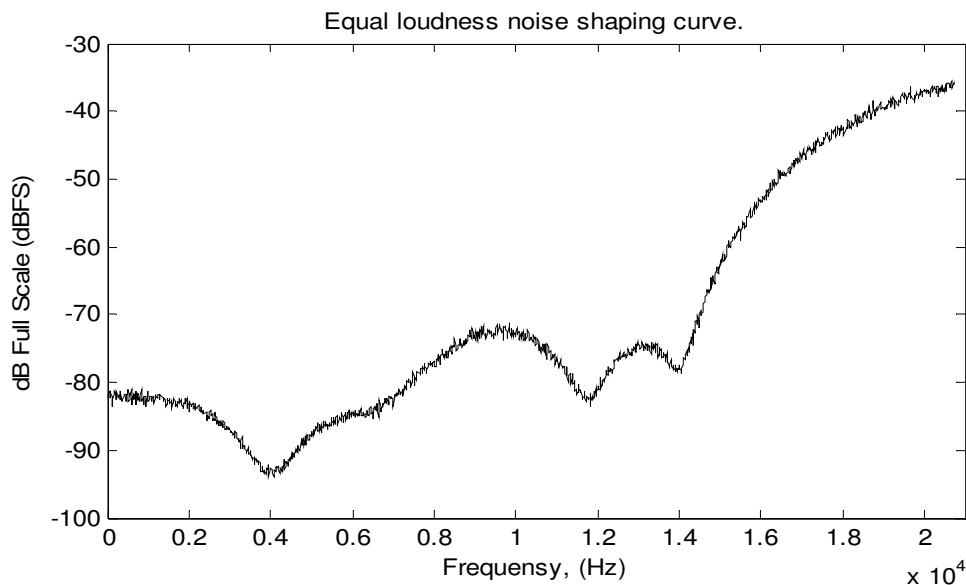


Σχ. 8.10.3: Εφαρμογή dither με noise shaping σε τόνο του 1 kHz και πλάτους -90 dBFS. Εξαγωγή συχνοτικού φάσματος. (α) Το φάσμα του τόνου στα 20 bit. (β) Το φάσμα του τόνου στα 16 bit, παρουσιάζονται τα σφάλματα που προκύπτουν. (γ) Εφαρμογή flat dither στον τόνο των 16 bit και αύξηση του noise floor στο σήμα. (δ) Εφαρμογή dither με noise shaping, παρατηρούμε τη μείωση του noise floor στις χαμηλές και μεσαίες καθώς και την αύξηση του στις υψηλές λόγω του noise shaping.

8.11 Dither So little noise, So much effect:

Η άθροιση ενός dither πλάτους -96 dBFS μοιάζει να είναι ένα πολύ μικρό ποσοστό θορύβου το οποίο αθροίζεται στο μουσικό σήμα. Παρόλαυτα οι ηχοηλίπτες και οι τεχνικοί mastering έχουν την ικανότητα να ακούν το πολύ μικρό αυτό ποσοστό θορύβου ακόμα και σε φυσιολογικές στάθμες έντασης κατά τη διάρκεια της ακρόασης. Το dither αναδεικνύει ορισμένα χαρακτηριστικά του σήματος, όπως τα fade in και fade out καθώς και τις ουρές των reverb, αλλά επίσης έχει την ιδιότητα να αυξάνει το noise floor του audio σήματος με αποτέλεσμα αυτό να γίνεται ακουστό ως ένα είδος hiss.

Για την αποφυγή του παραπάνω φαινομένου χρησιμοποιείται η τεχνική του noise shaping (όπως ήδη έχουμε αναφέρει). Το noise shaping αποτελεί ένα είδος equalization, το οποίο εφαρμόζεται στο συχνοτικό περιεχόμενο του θορύβου του dither ώστε να επιτευχθεί η ανακατανομή της ενέργειας ανά συχνότητα με τρόπο τέτοιο ώστε το dither να καθίσταται μη ακουστό. Με λίγα λόγια αυτό που επιτελείται με το noise shaping είναι η ανακατανομή της ενέργειας στο φάσμα ενώ παράλληλα η συνολική της τιμή παραμένει σταθερή. Το παρακάτω σχήμα (Σχ. 8.11.1) μας δείχνει την εφαρμογή του noise shaping στο dither.



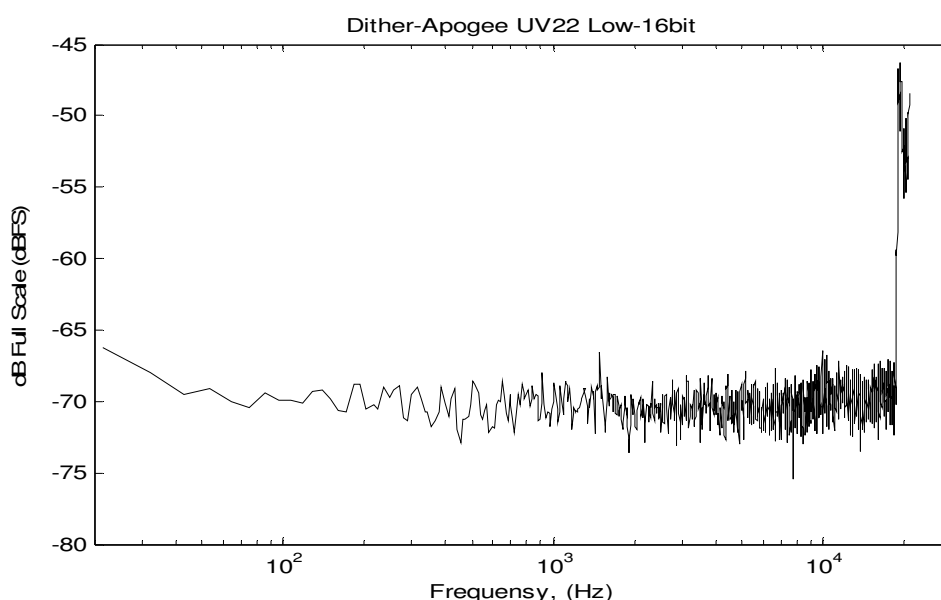
Σχ. 8.11.1: Η κατανομή ενέργειας ανά συχνότητα σε dither στο οποίο έχει εφαρμοστεί noise shaping (Meridian audio, Ltd).

Όπως βλέπουμε, αυτό που προκύπτει είναι ένα dither πολύ υψηλής τάξης με ανομοιόμορφη κατανομή της ενέργειας στο συχνοτικό του φάσμα και πιο συγκεκριμένα με «βυθίσματα» γύρω από τις συχνοτικές αυτές περιοχές στις οποίες το ανθρώπινο αυτί είναι περισσότερο ευαίσθητο (4 kHz, 12kHz, 14 kHz). Παρατηρούμε επίσης την αύξηση της ενέργειας με εκθετικό τρόπο από το 15 kHz έως τα 22 kHz, όπου για το συγκεκριμένο συχνοτικό εύρος δε γίνεται ακουστός ο θόρυβος του dither. Ηχητικά το

αποτέλεσμα που προκύπτει είναι ένα audio σήμα διαυγέστερο στις μικροδυναμικές και μακροδυναμικές του συνιστώσες ακόμα και στην περίπτωση του 16 bit C.D.

Ορισμένες απο τις πιο αξιόλογες hardware και software συσκευές dithering με τη χρήση noise shaping είναι της db Technologies και το μοντέλο 3000 Digital Optimizer, της Meridian (Model 618), τα plug ins όπως αυτό της Sony (Sony Super Bit Mapping), καθώς και αυτά της Waves (L1 & L2 Ultramaximizers), επίσης το ιδιαίτερα διαδεδομένο τύπο dither pow-r (αποτελεί ακρωνύμιο των λέξεων: *Psychoacoustically Optimized Wordlength Reduction*).

Η Apogee electronics παράγει το επίσης ιδιαίτερα διαδεδομένο UV-22. Η Apogee δε χρησιμοποιεί τον όρο dither. Προτιμά να αποκαλεί το προϊόν της «σήμα» (signal) δεδομένου του ότι το dither (καταχρηστικά εδώ) της Apogee αποτελείται απο περιοδικό θόρυβο, σε αντίθεση με τα όσα άλλα γνωρίζουμε τα οποία αποτελούνται απο τυχαία παραγόμενο θόρυβο. Παρόλαυτα ηχεί ακριβώς ως dither. Στο Σχ. 8.11.2 φαίνεται η κατανομή της ενέργειας ανά συχνότητα στο UV-22 της Apogee.

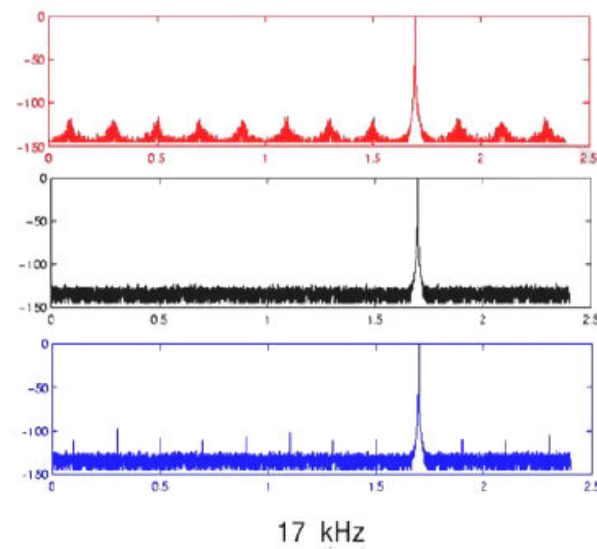
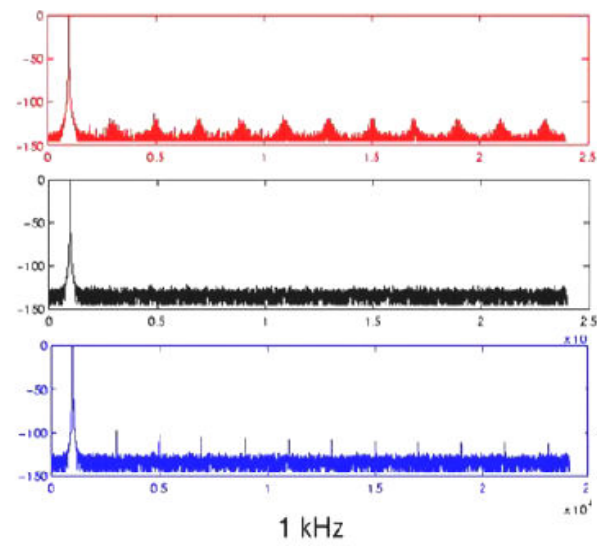
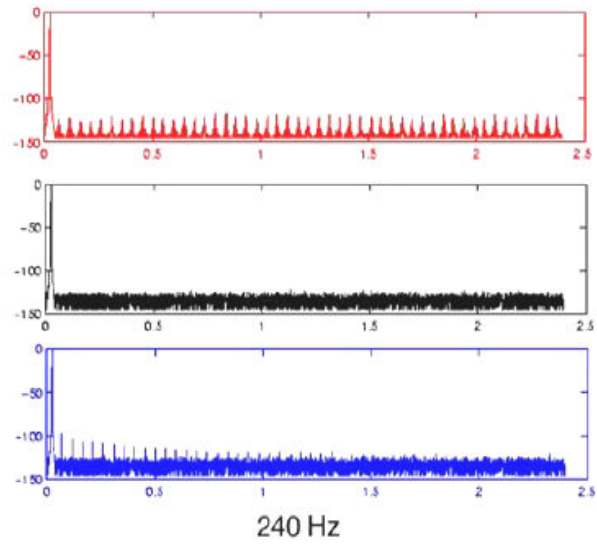


Σχ. 8.11.2: Το «dither» UV-22 της Apogee.

Όπως παρατηρούμε στο Σχ. 8.11.2 η Apogee δεν εφαρμόζει την τεχνική του noise shaping στη δική της πρόταση. Αθροίζει, όμως, ένα υψηλότερο ποσοστό ενέργειας στο θόρυβο στην περιοχή των 22 kHz.

Αξίζει, τέλος, να σημειωθούν κάποια σημαντικά σημεία τα οποία αφορούν την προσέγγιση μας κατά τη διάρκεια της εφαρμογής του dithering στο audio mastering. Για να διατηρηθεί πλήρως η ποιότητα του ηχητικού σήματος θα πρέπει να λάβουμε υπόψη, ότι τα διάφορα plug - ins που έχουμε στη διάθεση μας δεν είναι όλα τους πλήρως αξιόπιστα. Επίσης δε θα πρέπει να εμπιστευόμαστε τη BYPASS κατάσταση ενός plug - in, εκτός αν είμαστε απολύτως πεπεισμένοι ότι ο κλώνος που παράγεται απο το plug - in στη συγκεκριμένη κατάσταση είναι όμοιος με αυτόν του αρχικού

σήματος. Στο παρακάτω σχήμα (Σχ. 8.11.3, σελ. 140, Jim Johnston / AT&T research) παρουσιάζονται τα φάσματα έπειτα απο FFT ανάλυση ορισμένων καθαρών τόνων (240 Hz, 1 kHz & 17 kHz). Για κάθε τόνο δίδεται μία τριάδα διαγραμμάτων, το πρώτο κάθε φορά απεικονίζει το ημιτονικό σήμα δίχως την εφαρμογή dither στα 16 bit· το δεύτερο απεικονίζει, κάθε φορά, το ημιτονικό σήμα με την εφαρμογή ενός flat dither· τέλος το τρίτο διάγραμμα απεικονίζει, κάθε φορά, την έξοδο που παράγεται απο ένα δημοφιλές plug - in (στη βιβλιογραφία μας δεν αναφέρεται ποιο είναι αυτό) στην κατάσταση BYPASS. Παρατηρούμε πως για το συγκεκριμένο plug - in, στην έξοδο του (και στην BYPASS κατάσταση) έχουμε ξανά την εμφάνιση αρμονικής παραμόρφωσης η οποία, λογικά, θα έπρεπε να είχε απαλειφθεί με τη χρήση του dither. Κατανοούμε, λοιπόν, πως θα πρέπει να είμαστε εξαιρετικά επιλεκτικοί στη χρήση των διάφορων plug - ins και επεξεργαστών (και αυτό δεν αφορά μόνο την περίπτωση του dithering).

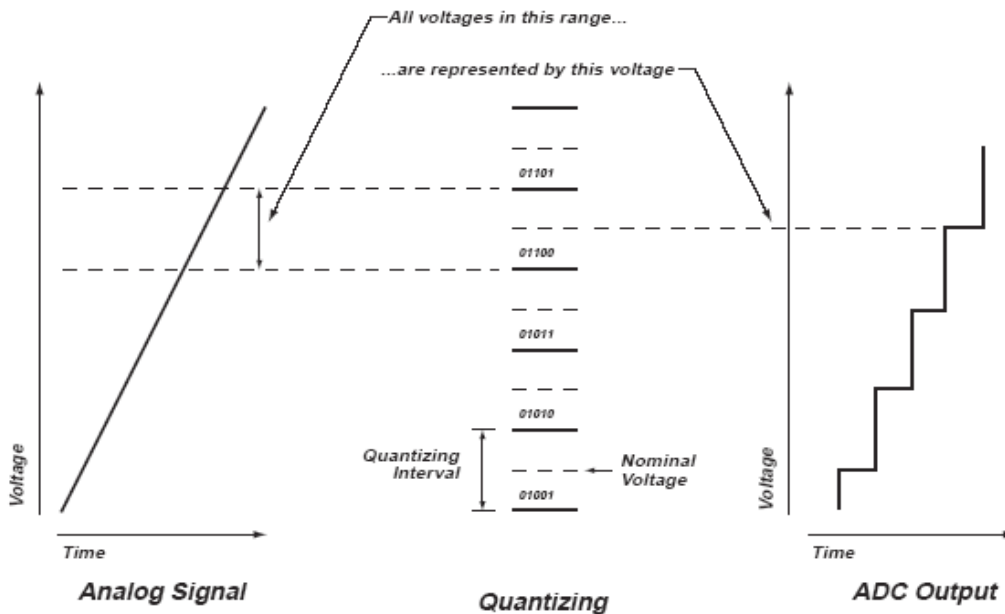


Σχ. 8.11.3: Παρουσίαση της λειτουργίας ενός plug - in εφαρμογής dither στην κατάσταση BYPASS, για τρεις καθαρούς τόνους - 240 Hz, 1 kHz & 17 kHz - (Jim Johnston / AT&T research, 1998).

8.12 Μια μικρή επανάληψη κι ένα παράδειγμα στο περιβάλλον του Sonic Studio:

Το Dither είναι θόρυβος ο οποίος αθροίζεται στο ψηφιακό audio σήμα ώστε να εξαλείψει τα ψηφιακά σφάλματα και τις παραμορφώσεις που προκύπτουν από την κβαντοποίηση του αναλογικού σήματος. Τα σφάλματα αυτά προκύπτουν λόγω της στρογγυλοποίησης που υφίσταται το σήμα, δεδομένου του ότι οι στάθμες κβαντισμού (bit depth) δεν είναι άπειρες.

Όταν ένας ήχος μεταλλάσσεται και από το αναλογικό περνά στο ψηφιακό πεδίο, τότε το συνεχές αναλογικό σήμα αναπαρίσταται από σειρές διακριτών βημάτων οι οποίες κωδικοποιούνται ως ψηφιακοί αριθμοί. Αυτού του τύπου η επεξεργασία καλείται κβαντοποίηση. Επειδή κάθε ψηφιακή τιμή αναπαριστά ένα σύνολο αναλογικών τιμών, η εκ νέου κβαντοποίηση από το ψηφιακό στο αναλογικό πεδίο παράγει σφάλματα. Ανεξάρτητα από το ακριβές πλήθος των τιμών που εισέρχονται σε έναν A/D converter, στην έξοδο αυτού ένας ορισμένος αριθμός και μόνο ψηφιακών τιμών εμφανίζεται. Δηλαδή άσχετα από το ότι στην είσοδο του εμφανίζονται άπειρες τιμές (αναλογικό σήμα), στην έξοδο οι τιμές θα είναι πάντα ίδιες (π.χ. 44100, ψηφιακό σήμα).



Σχ. 8.12.1: Σχηματική αναπαράσταση του σφάλματος κβαντοποίησης.

Εάν το σήμα εισόδου στον converter δεν είναι ένα απλό ημίτονο, αλλά ένα μουσικό σήμα (όπως π.χ. μια συμφωνική ορχήστρα) και διαθέτει επαρκές πλάτος ώστε να χρησιμοποιεί τα περισσότερα από τα bit του converter, τότε το σφάλμα κβαντοποίησης που προκύπτει έχει τη μορφή θορύβου και μασκάρεται από το κατά πολύ δυνατότερο μουσικό σήμα. Στην περίπτωση όμως που το σήμα που εισέρχεται είναι σχετικά απλό και/ή έχει χαμηλό πλάτος έντασης, τότε τα σφάλματα που προκύπτουν από την κβαντοποίηση παράγουν ακουστές παραμορφώσεις οι οποίες πέρα από το ότι είναι από

μόνες τους μη επιθυμητές, μπορούν επίσης μέσω του φαινομένου του aliasing να παράγουν διαμόρφωση στο audio σήμα η οποία ηχεί ιδιαίτερα δυσάρεστα.

Προσθέτοντας όμως στο LSB (Last Significant Bit, Λιγότερο Σημαντικό Ψηφίο) του ψηφιακού σήματος έναν ιδιαίτερα χαμηλού πλάτους θόρυβο, τότε ουσιαστικά εξαλείφουμε αυτού του τύπου τα σφάλματα και τα παράγωγά τους. Η διαδικασία που το επιτυγχάνει αυτό ονομάζεται Dithering. Η πρόσθεση του dither θέτει το σηματοθορυβικό λόγο ενός 16-bit συστήματος περίπου στα 96 dB, επειδή όμως το ανθρώπινο αυτί έχει την ικανότητα να ακούει και κάτω από αυτό το level θορύβου, το δυναμικό εύρος ενός 16-bit συστήματος που έχει υποστεί την κατάλληλη επεξεργασία με dither μπορεί να είναι ακόμη και της τάξης των 115 dB. Συνεπώς προκύπτει πως η διαδικασία του Dithering, μέσω τις οποίας αντιμετωπίζουμε τις παραμορφώσεις και τα ψηφιακά σφάλματα, έχει επίσης την ικανότητα να αναδεικνύει τις λεπτομέρειες ενός ηχητικού κομματιού, όπως τις ουρές των reverb και τα fade ins, fade outs.

Το Dither θα πρέπει επίσης να εφαρμόζεται κάθε φορά που το bit depth ενός ψηφιακού audio σήματος αλλοιώνεται. Αυτό θα πρέπει να λαμβάνεται υπόψη όχι μόνο κατά τη μετατροπή ενός 24-bit υψηλής ανάλυσης audio σήματος σε 16-bit, αλλά κάθε φορά που εκτελούμε κάποια εργασία επεξεργασίας σήματος: όπως την αλλαγή του gain, κάποιο crossfade, την εφαρμογή equalizer, κ.α.

Μολονότι όμως αυξάνει το δυναμικό εύρος, το Dither προσθέτει θόρυβο σε ένα δυνητικά ακουστικό level για ορισμένες συχνотικές περιοχές. Γι' αυτό το λόγο, υπάρχουν διάφοροι τύποι Dither, καθώς και διάφορες τεχνικές που μας επιτρέπουν να μειώσουμε το πλάτος του στις επιθυμητές συχνотικές περιοχές. Μια από αυτές τις τεχνικές (η πιο απλή) είναι να εφαρμόσουμε equalizer για τη συχνотική περιοχή που μας ενδιαφέρει καθώς και για το εύρος της. Αυτές οι συχνотότητες είναι οι μεσαίες και μεσαιονψηλές όπου το αυτί μας είναι περισσότερο ευαίσθητο.

8.12.1 Τύποι Dither στο Sonic Studio HD:

Όποτε ενεργοποιείται το dither, το Sonic Studio HD εφαρμόζει την επιλογή αυτή σε όλες τις εργασίες ενός project που είτε εκτελούνται εκείνη τη στιγμή, είτε είναι αποθηκευμένες. Τέσσερις τύποι Dither είναι διαθέσιμοι:

- Flat: Στο flat dither, το ποσοστό ενέργειας για κάθε συχνотότητα είναι το ίδιο (λευκός θόρυβος). Το flat είναι ο περισσότερο ακουστός τύπος dither.
- Second Order: ή dither δεύτερης τάξης. Αυτός ο τύπος dither είναι κατά 3 dB χαμηλότερος από το flat dither στις χαμηλές συχνотότητες, και κατά 3 dB δυνατώτερος από το flat στις υψηλές. Η κατανομή αυτή του πλάτους ανά συχνотικό εύρος κάνει το συγκεκριμένο dither να γίνεται λιγότερο αντιληπτό.
- Turbo Bitmapping: Το Turbo Bitmapping είναι ένας αλγόριθμος κατασκευασμένος από τη Sonic Solutions ο οποίος δίνει τη δυνατότητα στο

χρήστη να «κουρδίσει» το dither κατάλληλα, ώστε η κατανομή του πλάτους ανά συχνότητα που θα διαθέτει να ταιριάζει με την κατανομή του πλάτους στις συχνότητες του μουσικού υλικού. Το συγκεκριμένο dither μπορεί να χρησιμοποιηθεί μόνο σε στερεοφωνικό αρχείο ήχου και εφόσον η ανάλυση του είναι από 24-bit και πάνω.

- Sony Super Bitmapping: Το S.B.M. είναι ένας αλγόριθμος ο οποίος είναι σχεδιασμένος από την εταιρία Sony. Δε δίνει καμία δυνατότητα ρύθμισης των χαρακτηριστικών του και αυτά συνεπώς δε μεταβάλλονται (όπως π.χ. στο Turbo Bitmapping). Το S.B.M. μπορεί να χρησιμοποιηθεί μόνο σε στερεοφωνικό αρχείο ήχου και εφόσον η ανάλυση του είναι από 16-bit και πάνω. Η μόνη επιλογή που δίνει στο χρήστη είναι η επιλογή του ενός εκ των δύο τύπων dither (S.B.M.1 & S.B.M.2) με τους οποίους διατίθεται.

Κατά τη διάρκεια του dithering θα πρέπει να είμαστε πολύ προσεκτικοί καθώς το κάθε τύπος dither δε λειτουργεί το ίδιο αποτελεσματικά για κάθε διαφορετικό μουσικό πρόγραμμα. Πράγμα που σημαίνει πως κάποιος ίσως να χρειαστεί να πειραματιστεί αρκετά με τους διάφορους τύπους dither ώστε να επιτύχει τα επιθυμητά αποτελέσματα.

8.12.2 Εφαρμόζοντας το Dither στο Sonic Studio:

Στο Sonic Studio HD το dither είναι διαθέσιμο στην καρτέλα HDSP plug-in. Για να εφαρμόσουμε το dither:

1. Επιλέγουμε Windows --> HDSP Manager, ή πληκτρολογούμε COMMAND+6. Τότε εμφανίζεται η καρτέλα HDSP Manager (Σχ. 2).



Σχ. 8.12.2.1: Το παράθυρο HDSP Manager.

Τα διαθέσιμα plug-ins παρουσιάζονται σε λίστα στο HDSP Processing πάνελ. Στο Σχ. 8.12.2.1 το μόνο διαθέσιμο είναι το Global Dither.

2. Κάντε διπλό κλικ στο Global Dither ή σύρετε το στο Active Global Plug-ins.
3. Όταν το Global Dither εμφανιστεί στο πάνελ των ενεργών plug-in, κάντε του διπλό κλικ και επιλέξτε τον τύπο του dither που επιθυμείτε να χρησιμοποιήσετε.

Το παρακάτω σχήμα (Σχ. 8.12.2.2) παρουσιάζει το παράθυρο με τις επιλογές dither.



Σχ. 8.12.2.2: Το παράθυρο απο το οποίο καλείται ο χρήστης να επιλέξει το dither που επιθυμεί.

Για να απενεργοποιήσετε την επιλογή του dither plug-in:

1. Πηγαίνετε Windows --> HDSP Manager, ή πληκτρολογήστε COMMAND+6.
2. Επιλέξτε το plug-in απο την περιοχή Active Global Plug-ins.
3. Επιλέξτε διαγραφή ή μετακινήστε σέρνοντας το plug-in στην περιοχή Global plug-in area.

8.12.3 Παράμετροι του Dither:

Τα dither flat, Second Order και Sonic Turbo Bitmapping παρουσιάζουν ρυθμιζόμενες παραμέτρους οι οποίες και καθορίζουν το πώς τα συγκεκριμένα dithers επενεργούν στο ηχητικό υλικό. Αυτές οι παράμετροι είναι:

- Η ολική ενέργεια του dither (Total dither energy).
- Η χρήση ή όχι delay στο dither (Turn-off delay).
- Το εύρος των bit του dither (Bit width).

Το Turbo Bitmapping προσφέρει επίσης και μία τέταρτη ρύθμιση: Τη μετακίνηση μέρους της ενέργειας του dither απο τις χαμηλότερες συχνοτικές περιοχές στις υψηλές συχνότητες (High frequency weight).



(α)



(β)



(γ)



(δ)

Σχ. 8.12.3.1: Τα παράθυρα των dither που εμφανίζονται στο Sonic Studio με τις ρυθμίσεις που διαθέτουν. (α): Flat, (β): Second Order, (γ): Sonic Turbo Bitmapping & (δ): Sony Super Bitmapping.

Πιο συγκεκριμένα η καθεμιά απο τις παραπάνω ρυθμίσεις αφορά:

Η ολική ενέργεια του dither (Total dither energy):

Αυτή η παράμετρος ορίζει το επίπεδο του θορύβου του dither. Μηδενική τιμή θέτει το dither εκτός λειτουργίας, τιμή της τάξης του 100% ανταποκρίνεται στο μισό του βήματος κβαντοποίησης (θεωρητικά η ιδανικότερη τιμή). Η μέγιστη τιμή που μπορεί να λάβει η συγκεκριμένη παράμετρος είναι 1000%.

Η χρήση ή όχι delay στο dither (Turn-off delay):

Εφόσον το dither είναι ένα ηχητικό σήμα θορύβου, κατανοούμε πως σε περιπτώσεις ηχητικών παύσεων στη διάρκεια ενός μουσικού κομματιού, μπορεί υπό συγκεκριμένες προϋποθέσεις να γίνει ακουστό. Το Sonic Studio HD θέτει, by default, εκτός λειτουργίας το dither απο τη στιγμή που θα ανιληφθεί περιόδους σιωπής. Αυτό

συμβαίνει μετά απο κάποια σχετική χρονική καθυστέρηση (delay) που εισάγει ο χρήστης απο αυτή την παράμετρο (Turn-off delay). Οι τιμές που λαμβάνει η συγκεκριμένη παράμετρος εμπεριέχουν ένα εύρος απο 0 έως 999 δευτερόλεπτα.

Το εύρος των bit του dither (Bit width):

Αυτή η παράμετρος θέτει την ανάλυση του dither. Η default τιμή είναι αυτή των 16 bit. Το εύρος των τιμών κινείται πάντως απο τα 8 έως τα 16 bit.

High frequency weight:

Αυτή η ρύθμιση διατίθεται μόνο για ένα συγκεκριμένο τύπο dither: το Sonic Turbo Bitmapping. Η παράμετρος αυτή ρυθμίζει την κατανομή της ενέργειας ανά συχνότητα του συγκεκριμένου dither. Η default τιμή είναι το μηδέν (0) και είναι ελάχιστα αντιληπτή απο το ανθρώπινο αυτί. Η μέγιστη τιμή που μπορεί να λάβει είναι το ένα (1) και σε αυτή την περίπτωση μετακινείται το φασματικό περιεχόμενο του dither κατά τέτοιο τρόπο ώστε να μη γίνεται αντιληπτό στη συχνοτική περιοχή που μας ενδιαφέρει. Ο κατάλληλος χειρισμός αυτής της παραμέτρου μας επιτρέπει να «χρωματίσουμε» με τέτοιο τρόπο το dither ώστε να επιτύχουμε τα καλύτερα δυνατά αποτελέσματα.

Audio Mastering
Απο τη θεωρία στην πράξη

9. Εισαγωγή:

Στο συγκεκριμένο κεφάλαιο θα επιχειρήσουμε να αποτυπώσουμε τις εργασίες που έλαβαν χώρα, από μέρους μας, σε μία πλειάδα μουσικών κομματιών διαφορετικών ειδών. Απαραίτητη προϋπόθεση για να γίνει αυτό, είναι να αποτυπωθούν τα προβληματικά σημεία που εντοπίστηκαν στα μουσικά αυτά κομμάτια.

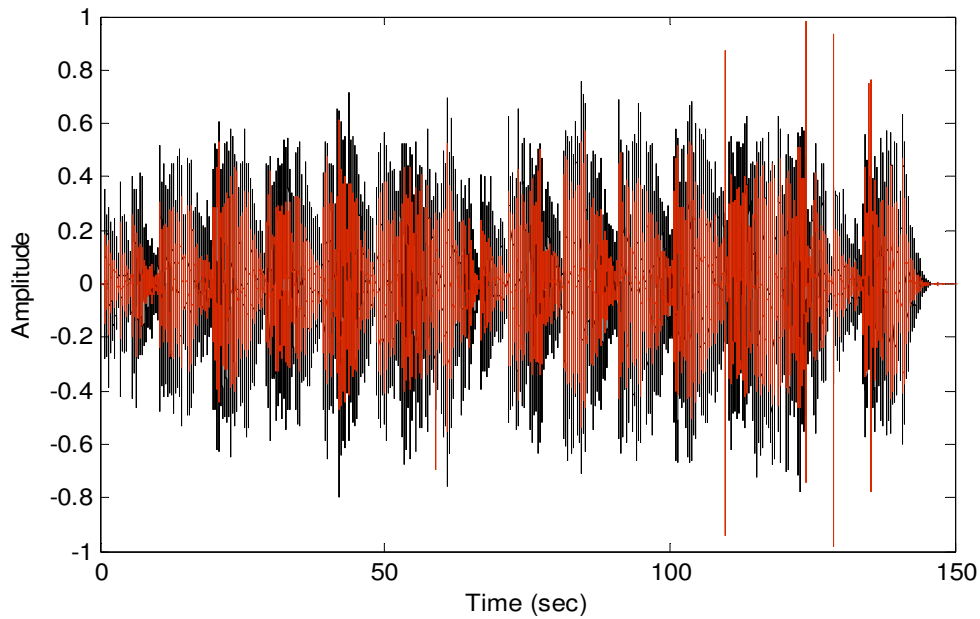
Στο κεφάλαιο αυτό παρατίθενται συγκριτικά πλήθος διαγραμμάτων τα οποία αποτυπώνουν τις διαφοροποιήσεις σε ότι αφορά την προγενέστερη έναντι της μεταγενέστερης μορφής του κάθε μουσικού κομματιού. Τα διαγράμματα αυτά αφορούν τις δυναμικές του σήματος, τη συχνотική του απόκριση, όπως επίσης δίνονται και πίνακες οι οποίοι αποτυπώνουν με ξεκάθαρο τρόπο τις δυναμικές επίσης διαφοροποιήσεις. Από τα παραπάνω προκύπτουν χρήσιμα συμπεράσματα τα οποία και θα εξετάσουμε στο κεφάλαιο 10.

9.1 Σαβίνα Γιαννάτου:

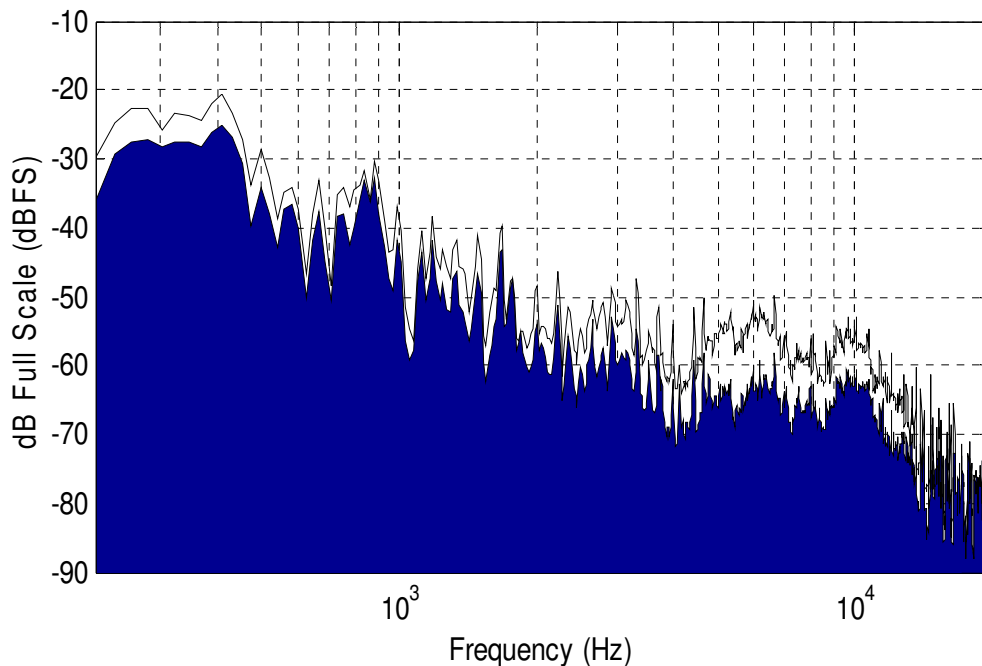
Το συγκεκριμένο μουσικό κομμάτι πρόκειται, σίγουρα, για μια εξαιρετική μουσική σύνθεση με πολύ καλές ερμηνείες τόσο απο τη μεριά των φωνητικών, όσο και απο τη μεριά των μουσικών οργάνων.

Παρατηρούμε, όμως, πως το αρχικό ηχητικό υλικό πάσχει απο πολλά και υψηλού level clicks, καθώς και απο ακόμη περισσότερα crackles (πολύ χαμηλού level clicks). Επίσης παρατηρούμε πως η αρχική ηχογράφιση τείνει να είναι λίγο άτονη, παρά τις πολύ καλές ερμηνείες, και πως επίσης ο αρμονικός πλούτος των οργάνων και ιδιαίτερα της υψηλής περιοχής του φάσματος είναι κατά μία έννοια «πνιγμένος».

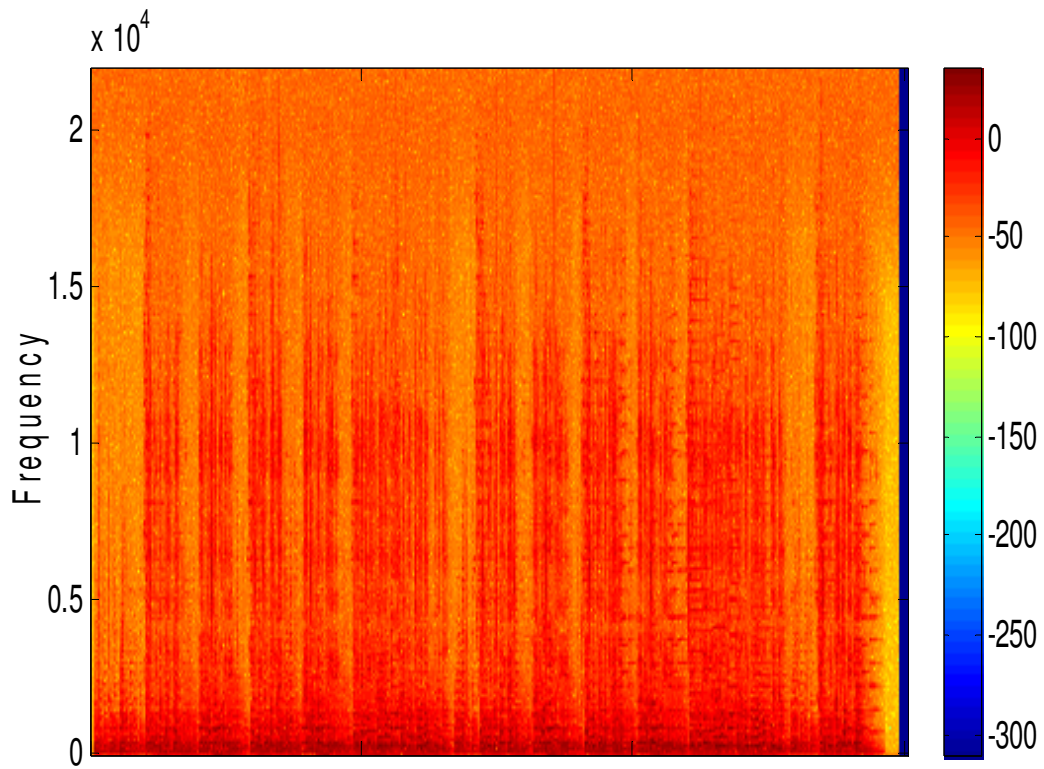
Τα έντονα clicks αφαιρέθηκαν κάνοντας χρήση του manual declicker, μαρκάροντας κάθε φορά το click που μας απασχολούσε και τελικά το απομακρύναμε απο το μουσικό υλικό. Τα crackles τα αντιμετωπίσαμε κάνοντας χρήση του plug-in x-crackle της Waves. Τη ζωντάνια της υψηλής περιοχής την επαναφέραμε κάνοντας χρήση του EQ και συγκεκριμένα της τεχνικής τύπου Baxandall. Επίσης καταφέραμε να «δέσουμε» τη μίξη κάνοντας χρήση της τεχνικής του parallel compression. Συνακόλουθα με τη χρήση και πάλι του equalization προσπαθήσαμε να δώσουμε λίγη παραπάνω αίσθηση στην μεσαιούψηλή περιοχή και σε συνδυασμό με το multiband compression πιστεύουμε πως τελικά το κατορθώσαμε. Τέλος κάναμε συντηρητική χρήση του limiter L2 της Waves έτσι ώστε να αυξήσουμε λίγο το συνολικό level του μουσικού κομματιού χωρίς όμως να προκύψει καμία είδους συμπίεση του μουσικού κομματιού καθώς επίσης και των μικροδυναμικών και μακροδυναμικών του συνιστωσών. Τέλος και λόγω της λιπής καθώς και του τρόπου της ενορχήστρωσης δεν προέκυψε καμία ανάγκη σε ότι αφορά την ενασχόληση με την στερεοφωνική εικόνα του συγκεκριμένου μουσικού κομματιού. Παρακάτω ακολουθούν τα γραφήματα και ο πίνακας με περαιτέρω επεξηγήσεις.



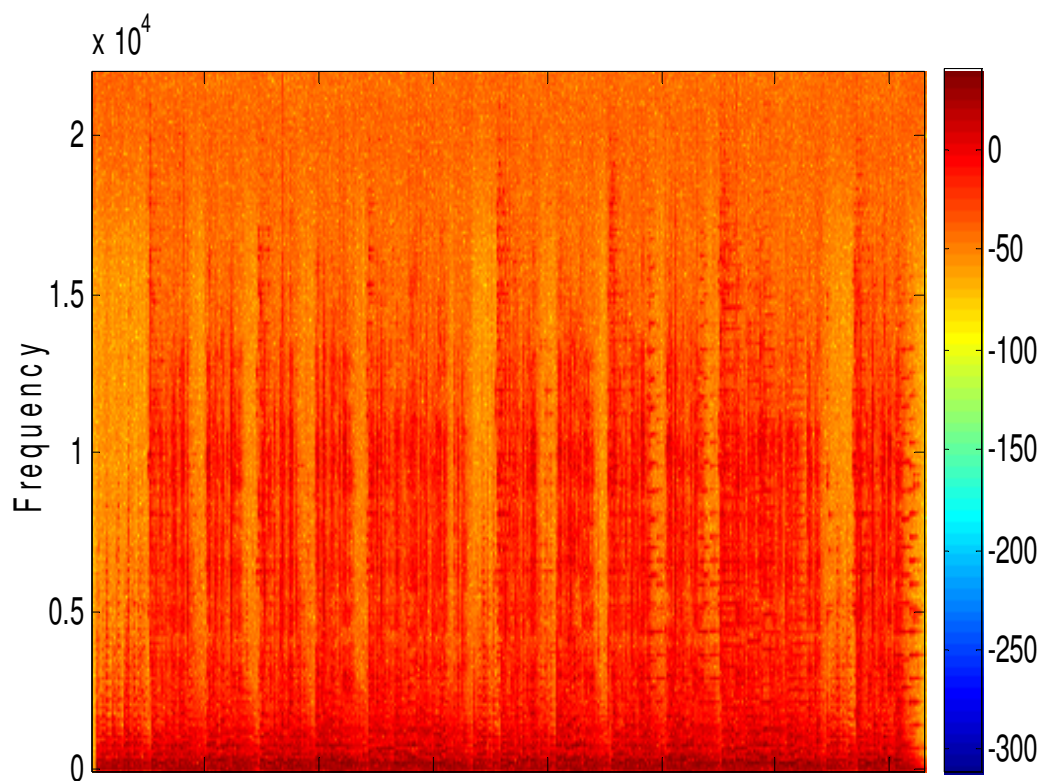
Σχ. 9.1.1: Το αριστερό κανάλι στο πεδίο του χρόνου. Φαίνονται ξεκάθαρα τα clicks που υπήρχαν αρχικά καθώς και η απομάκρυνση τους στο τελικό ηχητικό υλικό (Master). Καταδεικνύεται επίσης και η διαφορά level μεταξύ του original και του master.



Σχ. 9.1.2: Το ίδιο κανάλι (αριστερό) στο πεδίο των συχνοτήτων. Παρατηρούμε και πάλι τη συνολική αύξηση του level στο master. Παρατηρούμε, επίσης, ότι στην υψηλή περιοχή (18 - 20 kHz) το level του master είναι χαμηλότερο από αυτό του original. Αυτό οφείλεται στο γεγονός της απομάκρυνσης των clicks δεδομένου ότι αυτά περιέχουν το μεγαλύτερο ποσοστό της ενέργειας τους στην υψηλή περιοχή.



Σχ. 9.1.3: Το φασματογράφημα που προκύπτει απο το αρχικό ηχητικό υλικό (original).



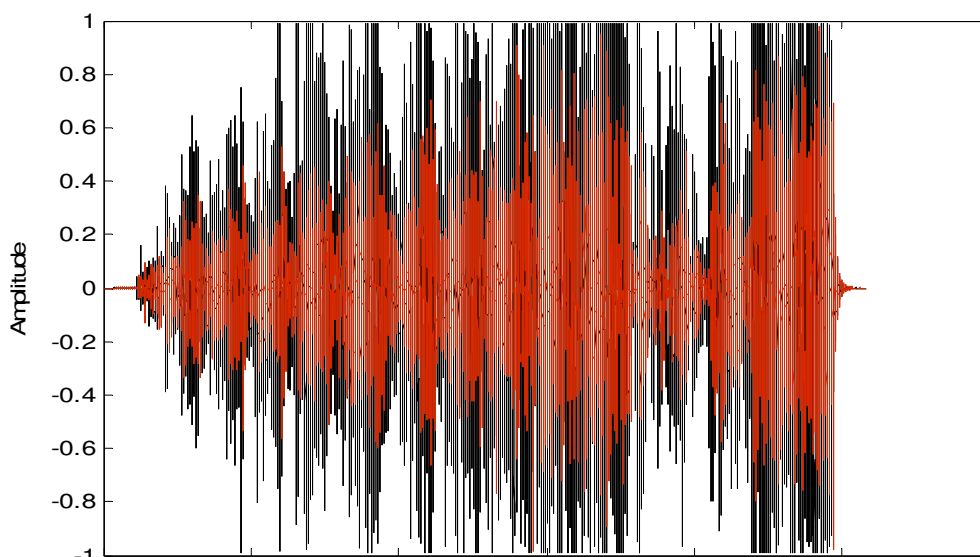
Σχ. 9.1.4: Το φασματογράφημα που προκύπτει απο το τελικά παραγόμενο ηχητικό υλικό (master). Παρατηρούμε και σε αυτό το σχήμα την αύξηση του συνολικού level του μουσικού κομματιού.

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	-0.12 dBFS	-0.09 dBFS	-15.03 dBFS	-15.02 dBFS	-21.42 dBFS	-21.35 dBFS
Master wav	-1.95 dBFS	-1.71 dBFS	-12.15 dBFS	-12.15 dBFS	-16.92 dBFS	-16.84 dBFS

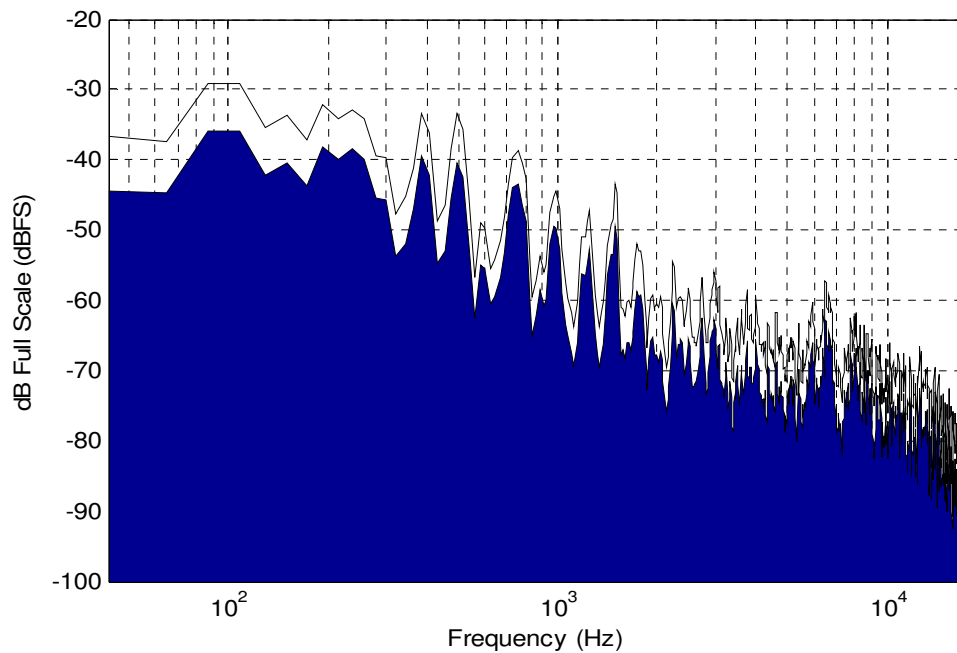
Πιν. 9.1.1: Σε αυτόν τον πίνακα βλέπουμε τις τιμές των μέγιστων peak του σήματος, όπως και των RMS τιμών του, τόσο για το δεξί όσο και για το αριστερό κανάλι. Η τιμή RMS (max) αναφέρεται στο γεγονός του ότι ένα μουσικό σήμα δεν έχει μια καθορισμένη περίοδο και συνεπώς δεν μπορούμε να εξάγουμε την RMS τιμή του με έναν και μόνο τρόπο. Πιο συγκεκριμένα το μουσικό σήμα έχει σημεία κατά τη διάρκεια του τα οποία είναι υψηλότερου level από ορισμένα άλλα. Συνεπώς το μέγεθος RMS (max) εκφράζει τη μέση τιμή του σήματος για τα σημεία του σήματος με υψηλό level και ουσιαστικά υφίσταται λόγω των μακροδυναμικών συνιστωσών του εκάστοτε ηχητικού υλικού. Η τιμή RMS (aver) από την άλλη εκφράζει τη μέση τιμή του σήματος για όλη τη διάρκεια του ανεξάρτητα από τις μακροδυναμικές του. Από τον παραπάνω πίνακα, λοιπόν προκύπτουν αβίαστα τα συμπεράσματα μεταξύ του αρχικού υλικού και του τελικά παραγόμενου και συνεπώς δεν χρειάζεται να πούμε εμείς κάτι. Στο μόνο, ίσως, σημείο στο οποίο θα άξιζε να σταθούμε αφορά το γεγονός του ότι στο original wav τα peaks είναι υψηλότερα απ' ό τι στο master. Αυτό οφείλεται και πάλι στα clicks και συνεπώς εφόσον στο master αυτά έχουν αφαιρεθεί έχουμε μικρότερες τιμές που αφορούν όμως αποκλειστικά το μουσικό σήμα.

9.2 Θανάσης Γκαϊφίλιας & Μίλτος Πασχαλίδης:

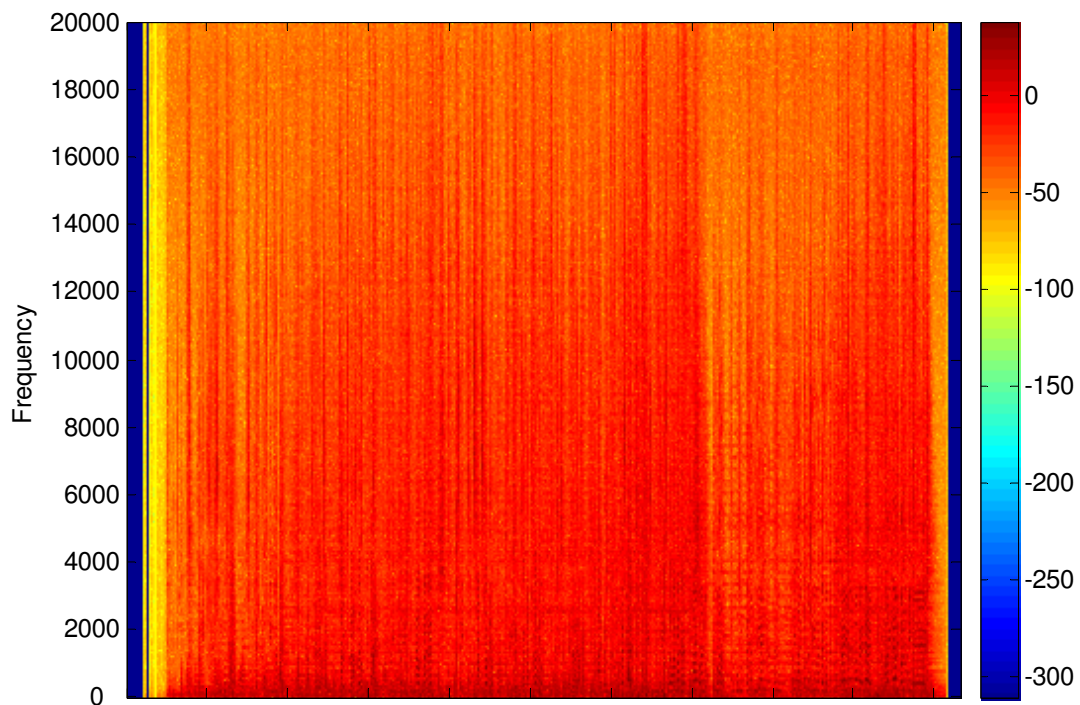
Εδώ πρόκειται για άλλη μία πολύ καλής ποιότητας μίξη. Ουσιαστικά σε αυτή την περίπτωση δεν είχαμε και πολλά πράγματα να διορθώσουμε παρά μόνο να αναδείξουμε ορισμένες εσωτερικές λεπτομέρειες του μουσικού κομματιού. Σε ότι αφορά το συχνотικό φάσμα των μουσικών οργάνων αυτό κρίθηκε ως επαρκέστατα ισορροπημένο και συνεπώς δεν χρειάστηκε να ασχοληθούμε με αυτό. Το μόνο που έγινε απο τη μεριά μας είναι ότι τονίστηκε φειδωλά η υψηλή περιοχή του φάσματος με ένα φίλτρο τύπου *Baxandall* ώστε η μίξη να αποκτήσει λίγο παραπάνω «αέρα». Σε ότι αφορά τις δυναμικές του μουσικού κομματιού αυτές κρίθηκαν επαρκείς ποιοτικά, αλλά με περιθώρια βελτίωσης. Για αυτό το λόγο και επεξεργαστήκαμε το μουσικό υλικό με την τεχνική του *parallel compression* ώστε να δέσει λίγο παραπάνω και επιπροσθέτως για να τονισθούν λίγο περισσότερο τα χωρικά στοιχεία του. Μία προβληματική που εντοπίσαμε ήταν ότι τα χορωδιακά μέρη του κομματιού, καθώς και τα φωνητικά του Μίλτου Πασχαλίδη σε ορισμένα σημεία ακούγονταν κάπως πέραν του δέοντος «τονισμένα», αυτά τα σημεία αντιμετωπίστηκαν με επιτυχία κάνοντας χρήση του *multiband compressor*. Τέλος ένα ακόμα σημείο που διορθώθηκε ήταν πως στο τέλος του μουσικού κομματιού ή ουρά της αντήχησης είχε «κοπεί» κατά τη μίξη απότομα (τέτοια λάθη συναντώνται πολύ συχνά). Αυτό αντιμετωπίστηκε πολύ εύκολα κάνοντας χρήση ενός *fade out* με την κατάλληλη ρύθμιση για τη συγκεκριμένη περίπτωση. Επίσης για να ενισχύσουμε τη φυσική δομή του μουσικού κομματιού κάναμε χρήση της τεχνικής του *gain riding* όπως αυτή έχει αναλυθεί σε προηγούμενη παράγραφο. Παρακάτω ακολουθούν τα διαγράμματα και για αυτή την περίπτωση.



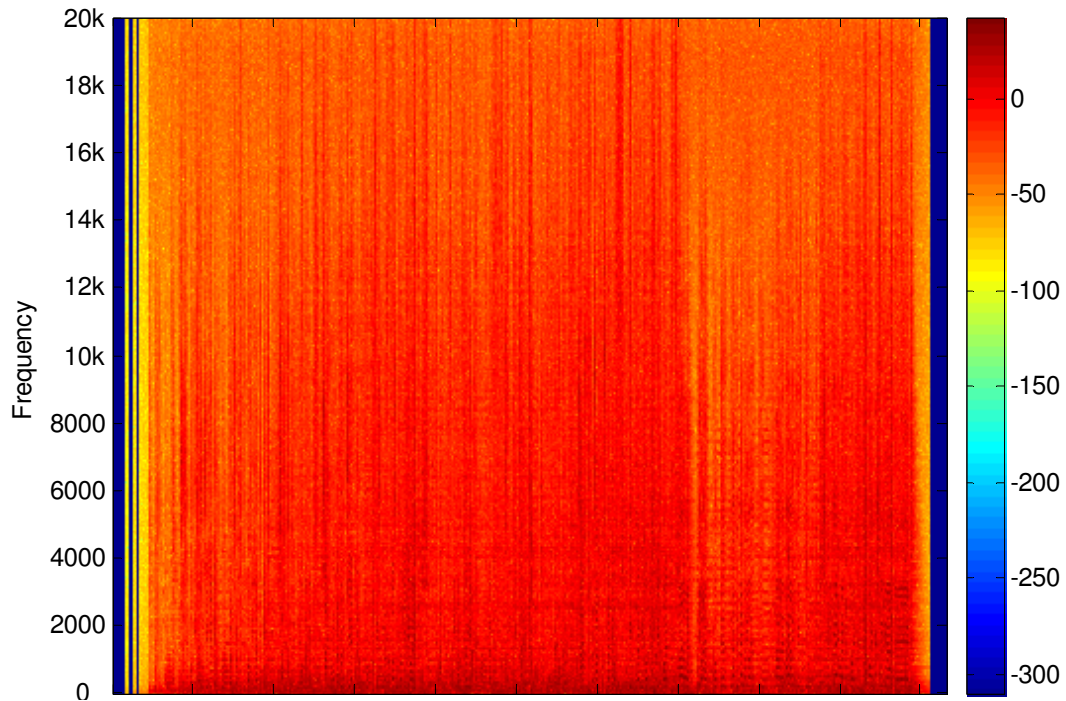
Σχ. 9.2.1: Το *original* και το *master wav* στο πεδίο του χρόνου. Παρατηρούμε την ενίσχυση του συνολικού *level* του σήματος στο *master*. Παρατηρούμε επίσης, την εφαρμογή του *fade out* στο *master* μιας και βλέπουμε ότι στο σημείο του τέλους του γραφήματος το *original* (κόκκινο) διαρκεί χρονικά κάποια *millisecond* παραπάνω.



Σχ. 9.2.2: Το σήμα στο πεδίο των συχνοτήτων. Παρατηρούμε την αύξηση της ενέργειας σε όλο το συχνοτικό φάσμα. Η ασάφεια στην υψηλή περιοχή του γραφήματος προέρχεται από το γεγονός του ότι για την απεικόνιση του συχνοτικού περιεχομένου των δύο σημάτων χρησιμοποιήθηκε λογαριθμικός άξονας και ως εκ τούτου έχουμε συμπίεση στην υψηλή περιοχή.



Σχ. 9.2.3: Το φασματογράφημα του αρχικού ηχητικού υλικού (original).



Σχ. 9.2.4: Το φασματογράφημα του τελικά παραγόμενου ηχητικού υλικού (master). Παρατηρούμε και πάλι την αύξηση του level.

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	-0.14 dBFS	-0.5 dBFS	-9.78 dBFS	-9.87 dBFS	-21.79 dBFS	-22.36 dBFS
Master wav	-0.1 dBFS	-0.1 dBFS	-7.68 dBFS	-7.53 dBFS	-15.76 dBFS	-16.28 dBFS

Πιν. 9.2.1: Όπως και στον πίνακα 1.1 τα μεγέθη δεν αλλάζουν. Αυτό που αλλάζει είναι οι τιμές των μεγεθών μιας και τώρα αναφερόμαστε σε διαφορετικά αρχεία ήχου. Οι τιμές είναι και πάλι εύγλωπτες και τα χαρακτηριστικά τους πασιφανή. Το μόνο που αξίζει να αναφερθεί σε αυτή την περίπτωση είναι ότι η μέγιστη τιμή των peaks του σήματος είναι τα -0,1 dB και αυτό γιατί εμείς θέσαμε ως μέγιστη τιμή εξόδου στον limiter τα -0,1 και όχι τα 0 dB. Αυτό συνέβη για να αποφύγουμε τυχόν πολύ γρήγορα μεταβατικά περάσματα του σήματος (τα οποία όσο γρήγορο attack time κι αν έχει ένας limiter υπάρχει πάντοτε η πιθανότητα να μην τα προλάβει) τα οποία και θα μπορούσαν να δημιουργήσουν clicks.

9.3 Wish Upon a Star – Frank lines:

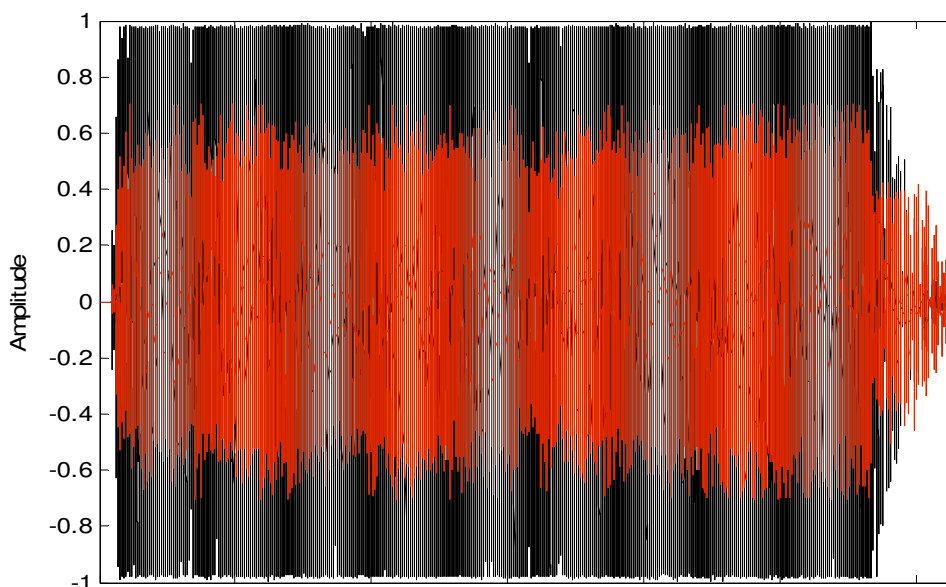
Οι Wish Upon a Star είναι μία punk – rock μπάντα· δύο εκ των μελών της μάλιστα είναι (ήταν?) και φοιτητές της σχολής μας. Η ηχογράφιση του κομματιού τους Frank lines έγινε στο studio του τμήματος και συγκεκριμένα στα Pro Tools. Αν και πρόκειται για μια αρκετά καλή ηχογράφιση και για μια αξιοπρεπή μίξη του ηχογραφηθέντος υλικού, παρόλα αυτά υπάρχουν ορισμένα προβληματικά σημεία σε αυτή. Τόσο σε επίπεδο τεχνικών λεπτομερειών, όσο και στο επίπεδο της αισθητικής. Καταρχήν και σε ότι αφορά το audio mastering θα πρέπει ένας ηχολόγητος να γνωρίζει πως το ηχητικό υλικό που προορίζεται για το mastering studio θα πρέπει να έχει περιθώρια, ώστε να μπορεί να το επεξεργαστεί ο τεχνικός mastering. Στη συγκεκριμένη περίπτωση αυτό δε συνέβη· πιο συγκεκριμένα οι άνθρωποι που έκαναν τη μίξη αυτού του κομματιού είχαν τοποθετήσει limiter στο master fader των pro tools ώστε να «ανεβάσουν» το συνολικό level του κομματιού (αυτή είναι μια γενικότερα αποδεκτή πρακτική όταν πρόκειται για παραγωγές τύπου demo. Σε καμία περίπτωση όμως δε θεωρείται αποδεκτή όταν μιλάμε για επαγγελματικού τύπου παραγωγές). Συν τις άλλους είχαν αθροίσει και dither στο μουσικό υλικό και μάλιστα υψηλής τάξης, ενώ η συχνότητα δειγματοληψίας του ήταν στα 48000 Hz και το bit rate του ήταν στα 24 bit. Τα παραπάνω έχουν ως αποτέλεσμα να μην μπορεί καταρχήν ο τεχνικός mastering να επεξεργαστεί το ηχητικό υλικό ούτε καν με τη χρήση ενός απλού EQ καθώς η οποιαδήποτε ενίσχυση της οποιασδήποτε συχνοτικής περιοχής σε ένα ήδη «πουσαρισμένο» ηχητικό υλικό θα έχει ως συνέπεια τη δημιουργία clicks καθώς και άλλων παθογόνων καταστάσεων. Το dither δε σε υλικό του οποίου η συχνότητα δειγματοληψίας είναι στα 48000 Hz και το bit rate του στα 24 bit είναι ακραίο λάθος καθώς, αν υποθέσουμε ότι το ηχητικό υλικό θα αποθηκευτεί σε μορφή c.d, θα πρέπει να το κατεβάσουμε στα 44100 Hz και στα 16 bit και συνεπώς θα πρέπει για ακόμη μία φορά να του αθροίσουμε dither πράγμα που σημαίνει ότι θα χαθεί μεγάλο μέρος, τουλάχιστον, της ηχητικής ποιότητας του μουσικού κομματιού.

Για όλους αυτούς τους λόγους, λοιπόν, οι εννέα στους δέκα τεχνικούς mastering θα απαιτούσαν το επαναμιξάρισμα του συγκεκριμένου κομματιού. Κάτι τέτοιο στην περίπτωση μας δεν ήταν δυνατό να συμβεί, συνεπώς έπρεπε να βαδίσουμε με βάση αυτά τα δεδομένα.

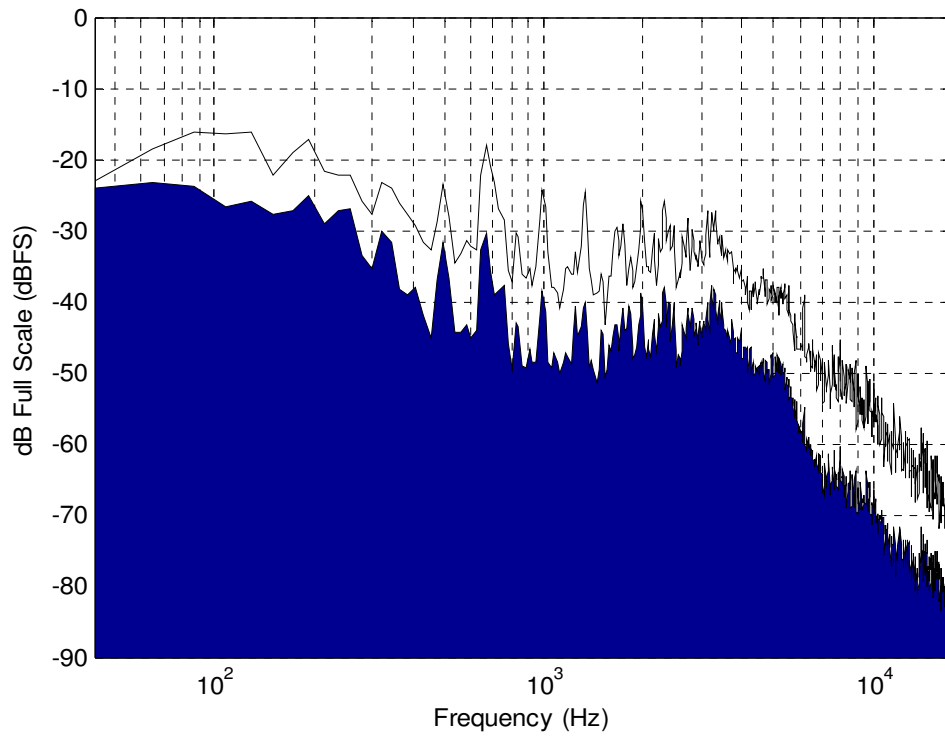
Για να μπορέσουμε να επεξεργαστούμε, λοιπόν, το ηχητικό υλικό καταρχήν του κάναμε ένα normalize της τάξης των 3 dB προς τα κάτω (με όλες τις αρνητικές συνέπειες που μπορεί να έχει αυτό). Σε ότι αφορά το truncation του κομματιού και την εκ νέου χρήση dither, χρησιμοποιήσαμε για αυτό το σκοπό τη σουίτα του Sonic Studio και κατά πρώτο λόγο κατεβάσαμε το ηχητικό υλικό από τα 48000 στα 44100 Hz και από τα 24 στα 16 bit. Εν συνεχεία αθροίσαμε στο σήμα ένα dither 1^{ης} τάξης και όσο το δυνατόν χαμηλότερου level (κάτι το οποίο δεν είναι τόσο εύκολο όσο ακούγεται και απαιτήσε μια τρομερή δαπάνη χρόνου από τη μεριά μας).

Σε ότι αφορά τα αισθητικά χαρακτηριστικά του μουσικού αυτού κομματιού και σε συνάρτηση πάντα με το μουσικό είδος που αυτό πρεσβεύει, οι προβληματικές που εντοπίσαμε αφορούσαν το γεγονός του ότι θα θέλαμε τα φωνητικά λίγο πιο «ψηλά», στο ότι θα θέλαμε τη στερεοφωνική εικόνα πιο «ανοιχτή», το συχνотικό φάσμα πιο «ισορροπημένο» γενικά και τη μίξη περισσότερο «δεμένη» συνολικότερα. Έτσι, λοιπόν, και για να «δέσουμε» περισσότερο τη μίξη κάναμε χρήση της τεχνικής του parallel compression. Για να «ισορροπήσουμε» το τονικό ισοζύγιο του μουσικού κομματιού κάναμε χρήση της τεχνικής Gerzon με τη χρήση EQ, καθώς και χρήση του multiband compression με την έννοια του δυναμικού equalization. «Διαπλατώναμε» τη στερεοφωνική εικόνα με τη χρήση του plug-in S1 - shuffler της Waves. Τα φωνητικά προσπαθήσαμε να τα «ανεβάσουμε» κάνοντας χρήση της τεχνικής του upward expansion με πενιχρά ωστόσο αποτελέσματα και αυτό διότι περεταίρω εφαρμογή της τεχνικής αυτής απο εκείνη που συντελέστηκε είχε αρνητικό αντίκτυπο στις υπόλοιπες ηχητικές συστατώσεις του μουσικού κομματιού. Τέλος εφαρμόσαμε ένα στερεοφωνικό reverb κάνοντας χρήση του plug-in TC Reverb της TC Electronic και αυτό γιατί θεωρήσαμε ότι χρειαζόνταν μια μικρή ενίσχυση τα χωρικά στοιχεία του συγκεκριμένου μουσικού υλικού.

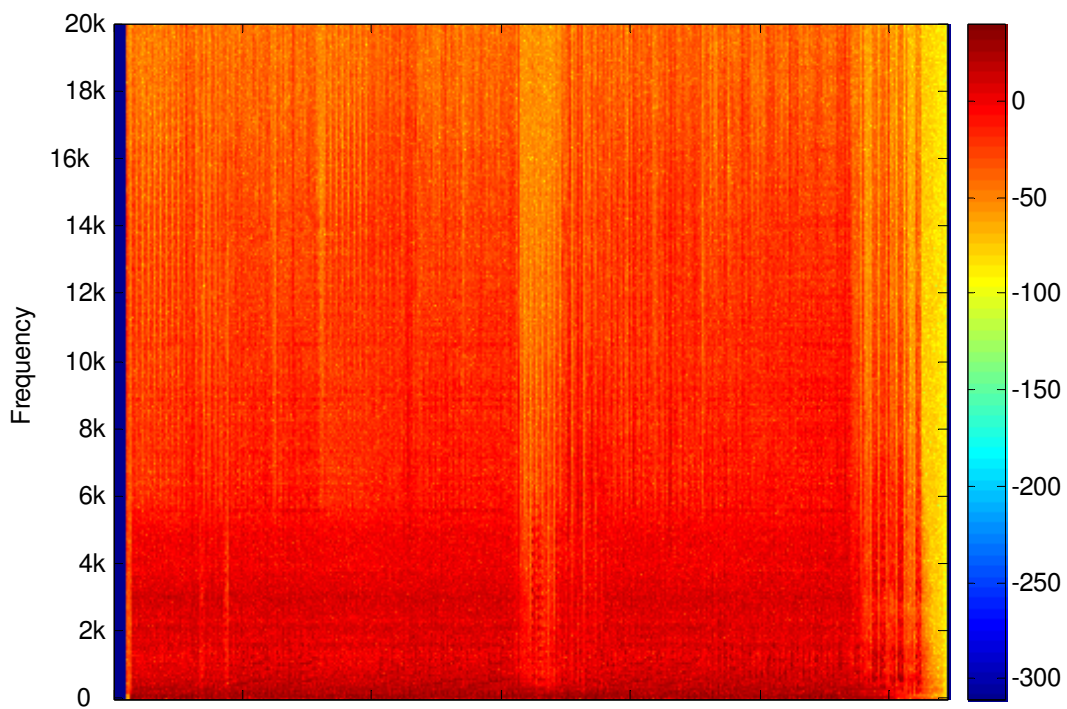
Και πάλι ακολουθούν τα διαγράμματα και για αυτή την περίπτωση.



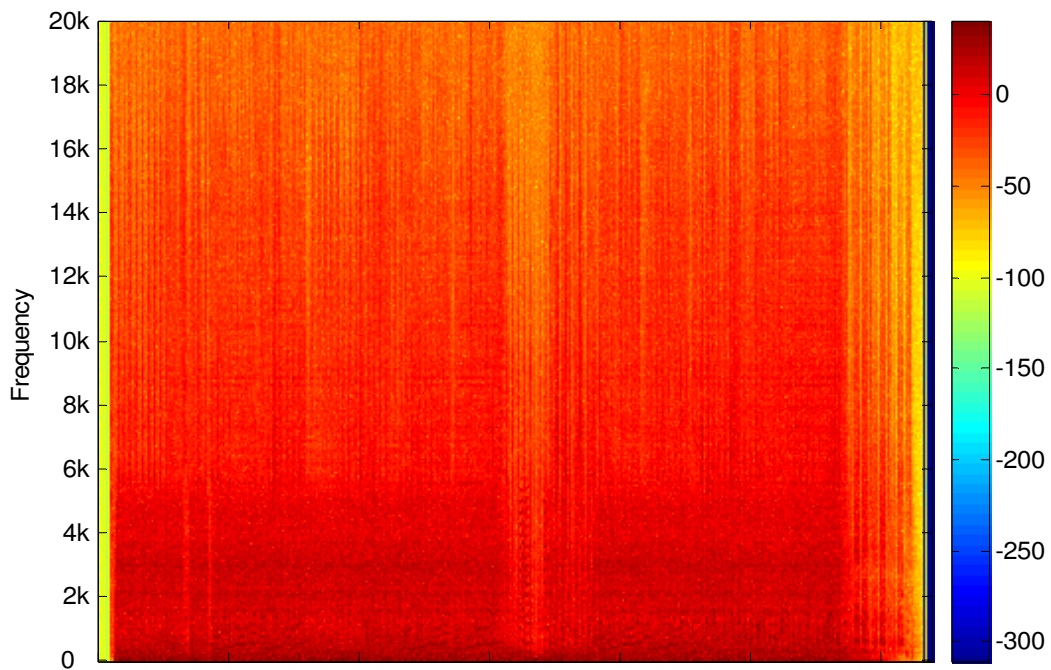
Σχ. 9.3.1: Τα δύο σήματα (original & master) στο πεδίο του χρόνου. Στο συγκεκριμένο γράφημα παρατηρούμε κατά πρώτο λόγο πως το ποσοστό της ενίσχυσης στο συνολικό level του master είναι αρκετά υψηλό. Για την ακρίβεια έχουμε δημιουργήσει αυτό που αγγλιστί αποκαλείται the wall of sound. Αυτό συνέβη για δύο κυρίως λόγους: πρώτον και κύριων διότι αισθητικά αρμόζει με το ύφος και με το είδος του κομματιού· δεύτερον διότι με αυτό το λόγο έχουμε καλύτερο signal to noise και αντιμετωπίζουμε καλύτερα την προβληματική που ήδη έχει αναφερθεί και αφορά την εκ νέου άθροιση dither. Το δεύτερο συνεπαγόμενο απο το συγκεκριμένο γράφημα είναι πως στο τέλος του μουσικού κομματιού το original έχει υψηλότερο πλάτος απο το master. Αυτό συνέβη γιατί πολύ απλά στο master εφαρμόσθηκε fade out.



Σχ. 9.3.2: Το αρχικό ηχητικό υλικό και το τελικά παραγόμενο στο πεδίο των συχνοτήτων.



Σχ. 9.3.3; Το φασματογράφημα του αρχικού ηχητικού υλικού (original).



Σχ. 9.3.4: Το φασματογράφημα του τελικά παραγόμενου ηχητικού υλικού (master).

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	-0,16 dBFS	-0,16 dBFS	-9,31 dBFS	-8,87 dBFS	-14,35 dBFS	-14,58 dBFS
Master wav	-0,01 dBFS	-0,01 dBFS	-6,93 dBFS	-6,97 dBFS	-10,45 dBFS	-10,64 dBFS

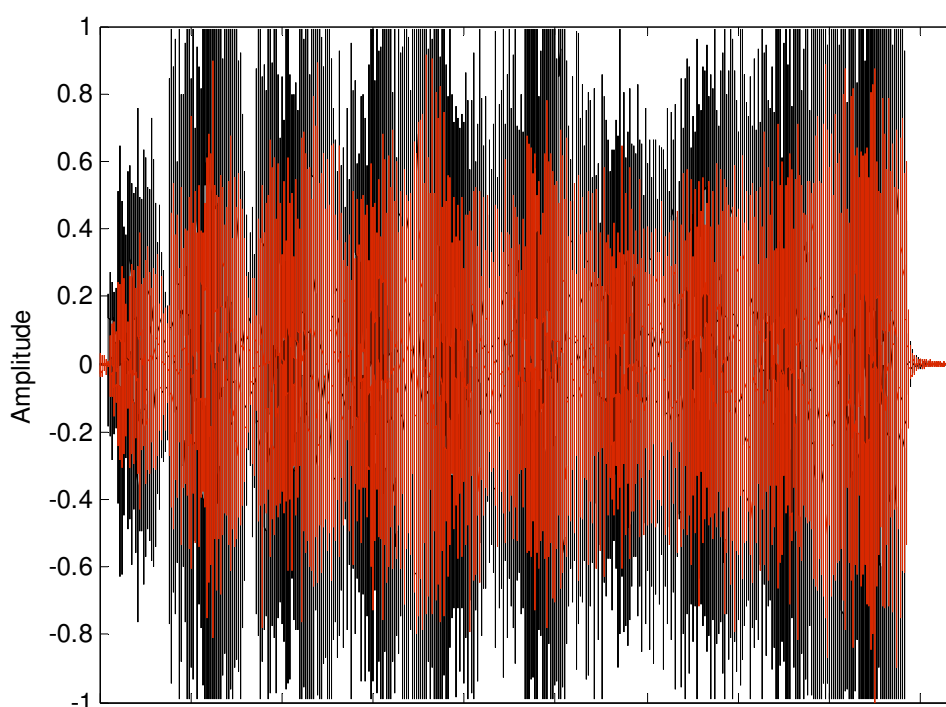
Πιν. 9.3.1: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το original και το master αντίστοιχα του μουσικού κομματιού των Wish Upon a Star.

9.4 Μίκης Θεοδωράκης - 1:

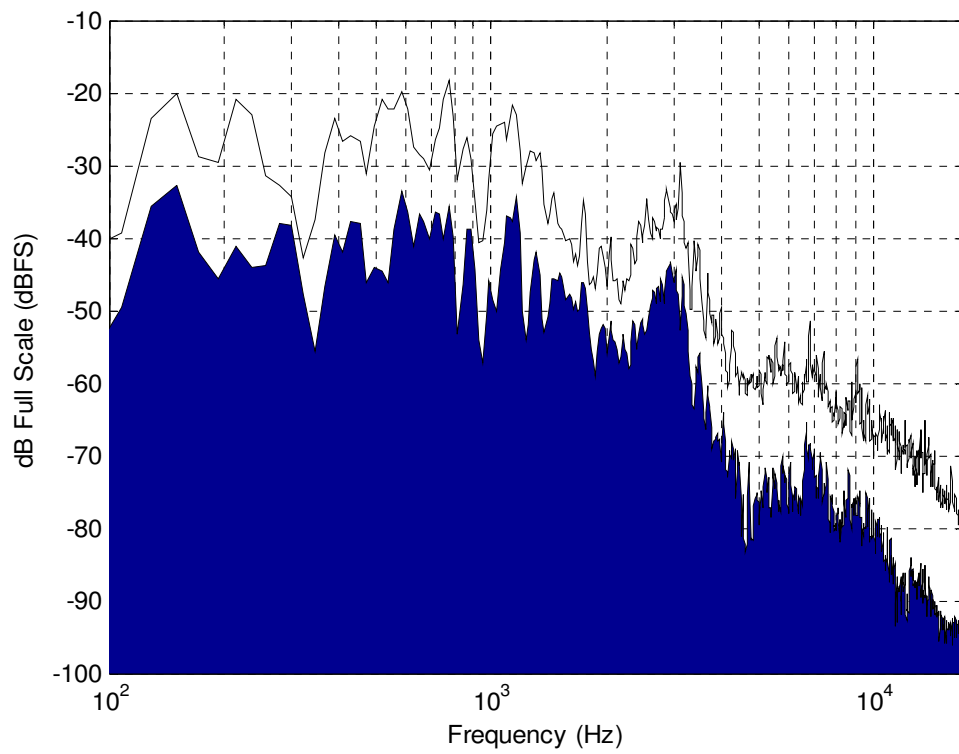
Αυτή η ηχογράφιση, καθώς και αυτή που εξετάζεται στην επόμενη παράγραφο (9.5) μπορούν να θεωρηθούν ως κορυφαία δείγματα ηχοληπτικής δεινότητας. Και αυτό γιατί ηχογραφήθηκαν σε πραγματικές συνθήκες υπό την παρουσία κοινού σε ένα και μόνο φορητό οκτακάναλο.

Σχετικά με το συγκεκριμένο μουσικό κομμάτι που πραγματευόμαστε σε αυτήν εδώ την παράγραφο όμως τα προβληματικά σημεία που εντοπίσαμε αφορούν: α) την ισορροπία του συχνωτικού φάσματος, β) τις μικροδυναμικές διακυμάνσεις του ηχητικού υλικού.

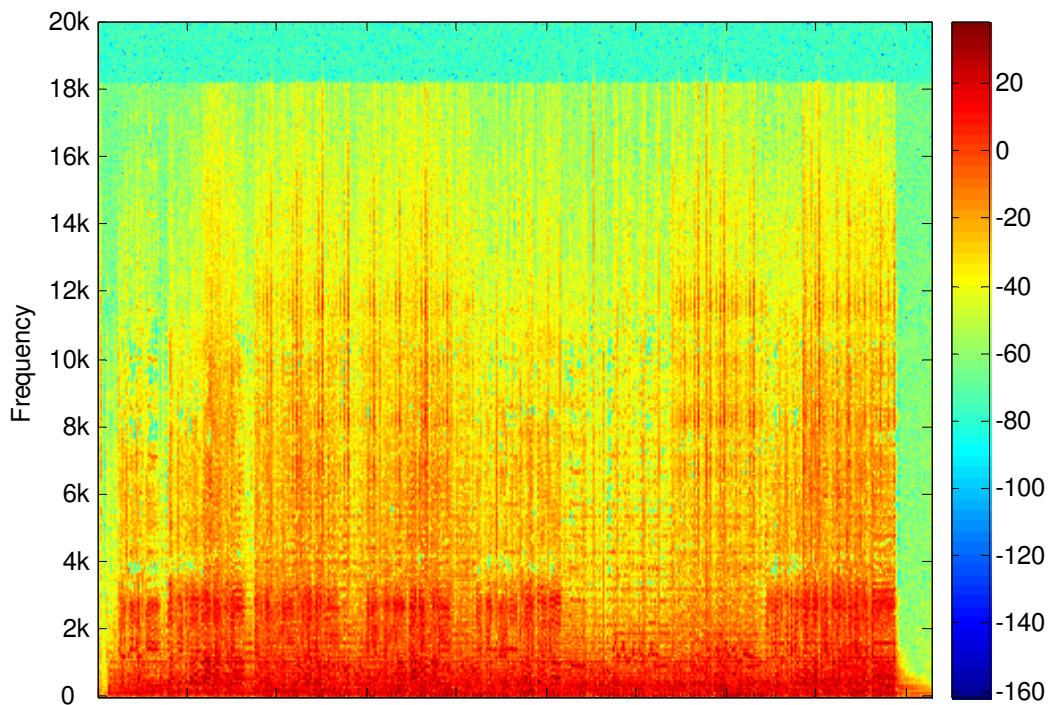
Για να φτιάξουμε την ισορροπία του συχνωτικού φάσματος κάναμε χρήση του equalization, που σε συνδυασμό με τη χρήση του harmonic exciter (ένας τρόπος enhancement ο οποίος τονίζει τις αρμονικές μιας δοθείσας συχνότητας ή μιας ολόκληρης επιλεγείσας συχνωτικής περιοχής. Επίσης προσθέτει και ένα μικρό ποσοστό delay στις συχνότητες στις οποίες επεμβαίνει ώστε να μη χάνεται η φυσική τους διαύγεια) επέφερε τα επιθυμητά αποτελέσματα. Τις μικροδυναμικές διακυμάνσεις του συγκεκριμένου μουσικού κομματιού τις βελτιώσαμε κάνοντας χρήση του multiband compression. Τέλος ακολούθησε και εδώ, όπως και σε όλες τις άλλες περιπτώσεις που εξετάζουμε, το limiting. Ακολουθούν τα διαγράμματα για αυτό το μουσικό κομμάτι.



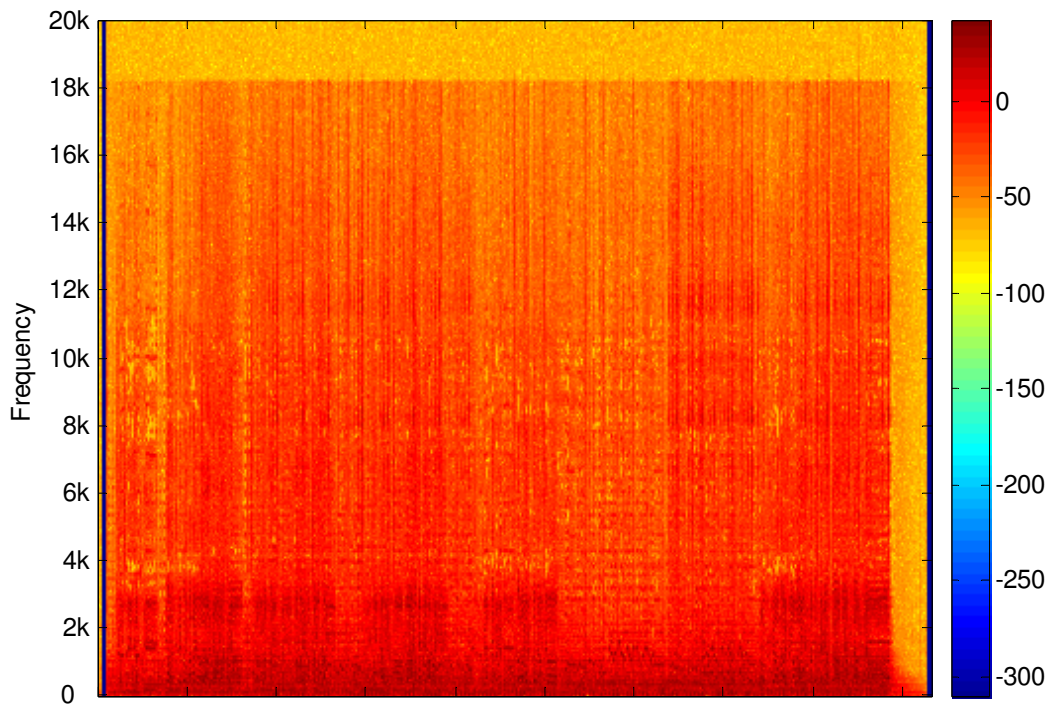
Σχ. 9.4.1: Το αρχικό αρχείο και το master στο πεδίο του χρόνου. Παρατηρούμε ότι αν και έχει εφαρμοσθεί στο master η τεχνική του limiting οι δυναμικές διακυμάνσεις του κομματιού παραμένουν (πράγμα που ήταν και το ζητούμενο). Παρατηρούμε επίσης πως και σε αυτή την περίπτωση το original αρχείο έχει μεγαλύτερο πλάτος στο τέλος του κομματιού από το master. Αυτό οφείλεται και πάλι στο fade out που εφαρμόστηκε στο master.



Σχ. 9.4.2: Το αρχικό ηχητικό υλικό και το τελικά παραγόμενο στο πεδίο των συχνοτήτων.



Σχ. 9.4.3: Το φασματογράφημα του αρχικού ηχητικού υλικού.



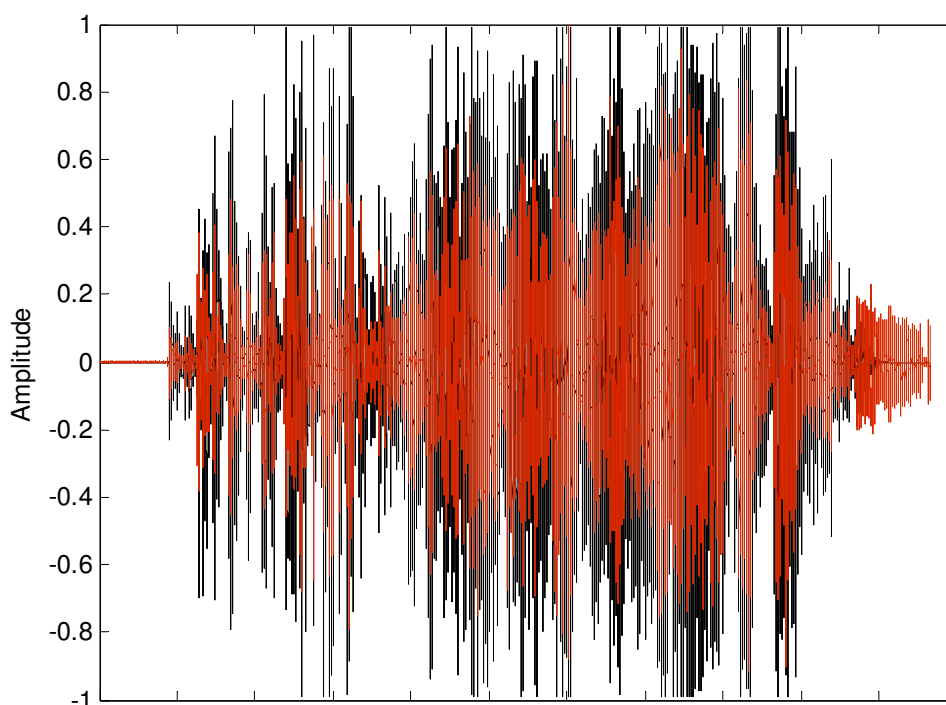
Σχ. 9.4.4: Το φασματογράφημα του τελικά παραγόμενου ηχητικού υλικού.

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	0,00 dBFS	-2,01 dBFS	-10,41 dBFS	-10,47 dBFS	-18,30 dBFS	-19,32 dBFS
Master wav	-0,10 dBFS	-0,10 dBFS	-6,48 dBFS	-6,32 dBFS	-12,89 dBFS	-13,77 dBFS

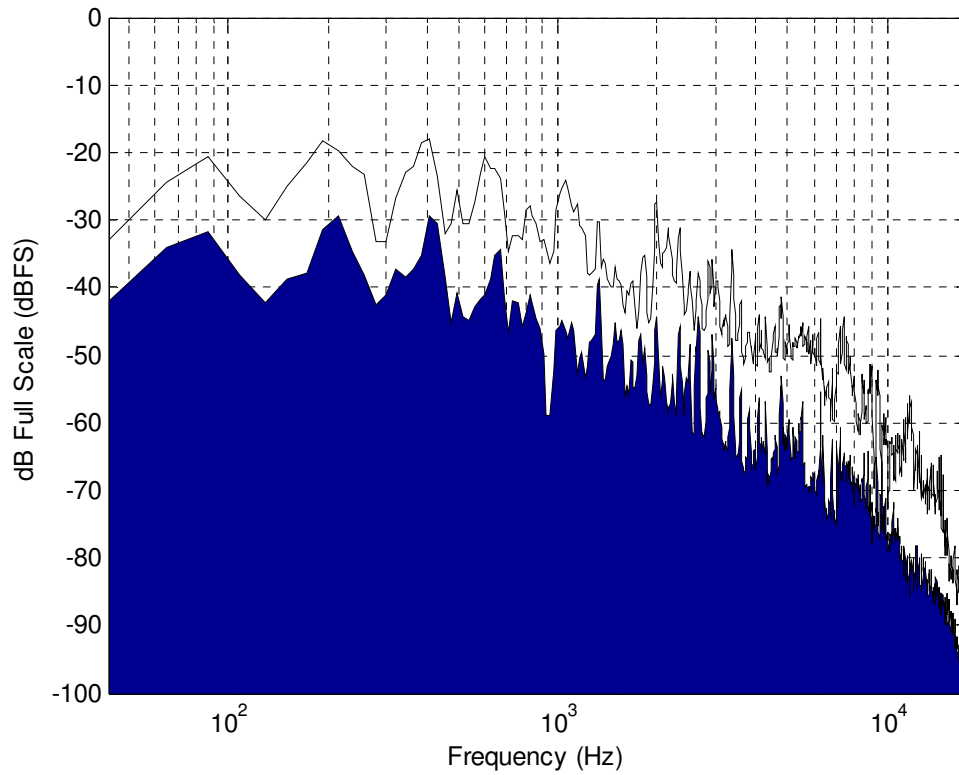
Πιν. 9.4.1: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το original και το master αντίστοιχα του συγκεκριμένου μουσικού κομματιού.

9.5 Μίκης Θεοδωράκης - 2:

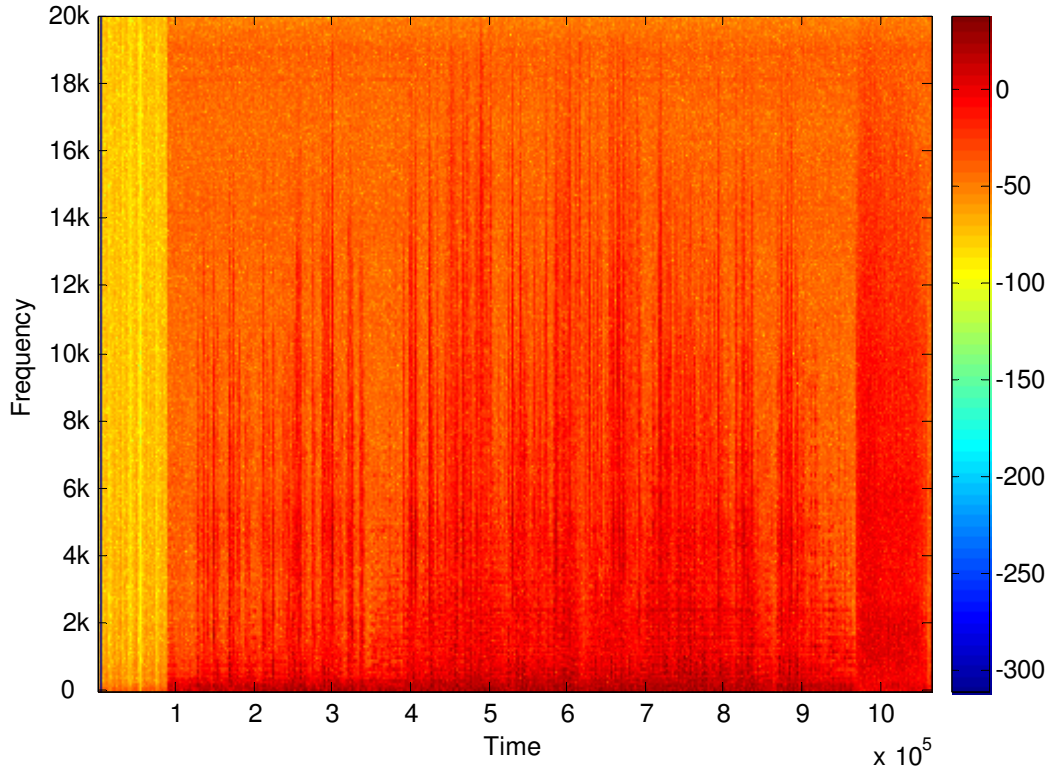
Η συγκεκριμένη μίξη αντιμετώπιζε τα ίδια ακριβώς προβλήματα που αντιμετώπιζε και αυτή της παραγράφου 9.4. Οι λύσεις συνεπώς που δόθηκαν και για αυτή την περίπτωση ήταν οι ίδιες. Ένα πρόβλημα όμως που αντιμετώπιζε η συγκεκριμένη μίξη το οποίο δε συναντήθηκε σε αυτήν της παραγράφου 9.4, ήταν το ότι η μίξη που εξετάζουμε εδώ είχε ένα ποσοστό θορύβου στο σήμα το οποίο και έπρεπε να αφαιρεθεί. Αυτό το πετύχαμε κάνοντας χρήση του plug-in X - noise της Waves. Ακολουθούν τα διαγράμματα



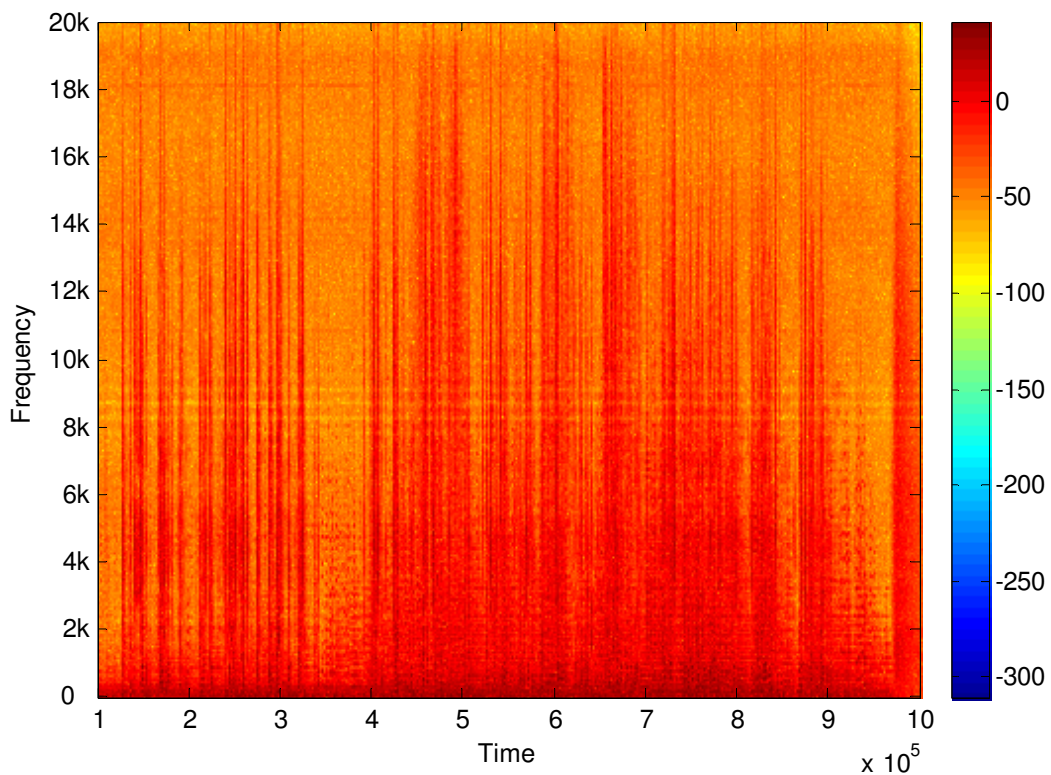
Σχ. 9.5.1: Το αρχικό αρχείο και το master στο πεδίο του χρόνου. Παρατηρούμε πως και σε αυτή την περίπτωση το original αρχείο έχει μεγαλύτερο πλάτος στο τέλος του κομματιού από το master. Αυτό οφείλεται και πάλι στο fade out που εφαρμόστηκε στο master. Αυτό συνέβη διότι στο τέλος του αρχικού αρχείου ήταν πολύ έντονο το χειροκρότημα του κοινού.



Σχ. 9.5.2: Τα σήματα των δύο αρχείων ήχου στο πεδίο των συχνοτήτων.



Σχ. 9.5.3: Το φασματογράφημα του αρχικού ηχητικού υλικού.



Σχ. 9.5.4: Το φασματογράφημα του τελικά παραγόμενου ηχητικού υλικού (master).

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	0,00 dBFS	-1,27 dBFS	-10,41 dBFS	-10,27 dBFS	-21,22 dBFS	-21,08 dBFS
Master wav	-0,10 dBFS	-0,10 dBFS	-6,95 dBFS	-7,09 dBFS	-15,47 dBFS	-15,30 dBFS

Πιν. 9.5.1: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το original και το master αντίστοιχα του μουσικού αυτού κομματιού.

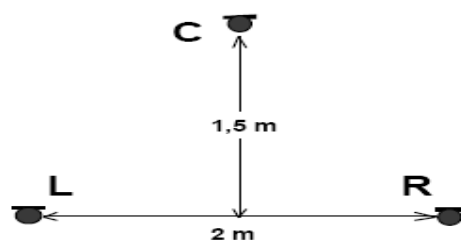
9.6 Kurt Weill - Ballad of Mack the Knife:

Η ηχογράφιση που εξετάζεται σε αυτή την παράγραφο, καθώς και αυτές που εξετάζονται στις παραγράφους 9.7 και 9.8 αποτελούν δημιουργίες του συνθέτη Kurt Weill. Η ηχογράφιση τους έγινε από τον υπεύθυνο καθηγητή της παρούσης πτυχιακής εργασίας. Για την ηχογράφιση του μουσικού συνόλου χρησιμοποιήθηκε μια stereo τεχνική ηχογράφησης και πιο συγκεκριμένα η τεχνική Decca Tree²⁵.

Τα προβληματικά σημεία που εντοπίσαμε στην παρούσα ηχογράφιση του Ballad of Mack the Knife, αφορούσαν κυρίως το συχνοτικό του φάσμα. Πιο συγκεκριμένα τα χάλκινα πνευστά φαίνονταν σε κάποια σημεία να τσιρίζουν. Η προβληματική αυτή αντιμετωπίστηκε κάνοντας χρήση της τεχνικής Gerzon με τη χρήση EQ. Αντιμετωπίστηκε επίσης χρησιμοποιώντας την τεχνική του δυναμικού equalization με τη χρήση multiband compressor.

Μία ακόμη προβληματική που εντοπίστηκε αφορούσε το δυναμικό εύρος του μουσικού κομματιού, δεδομένου ότι αυτό ηχογραφήθηκε με πολύ χαμηλό πλάτος. Στις κλασσικές συνθέσεις γενικότερα, αποτελεί σχεδόν κανόνα το να μην επεμβαίνουμε στο δυναμικό τους εύρος, δεδομένου ότι μεγάλο μέρος των ηχητικών - μουσικών τους χαρακτηριστικών, καθώς και της ομορφιάς τους, βρίσκεται ακριβώς στις μικροδυναμικές και μακροδυναμικές τους διακυμάνσεις. Αυτή η προβληματική, λοιπόν, και με βάση αυτά τα δεδομένα αντιμετωπίστηκε κάνοντας συντηρητική χρήση του limiter (άλλωστε, όπως έχει αναφερθεί ήδη σε προηγούμενο κεφάλαιο, οι limiters στο audio mastering χρησιμοποιούνται για την αύξηση του συνολικού level του σήματος,

²⁵ Decca Tree: Πρωτοεμφανίστηκε το Μάρτιο του 1954 όταν οι μηχανικοί ήχου Roy Wallace και Arthur Haddy στα στούντιο Decca του Λονδίνου ηχογράφησαν την ορχήστρα Mantovani. Ο Wallace έφτιαξε μια σειρά μικροφώνων σε μορφή T, όπως φαίνεται στο παρακάτω σχήμα, από παντοκατευθυντικά μικρόφωνα Neumann M49 τοποθετημένα το καθένα στις τρεις άκρες.



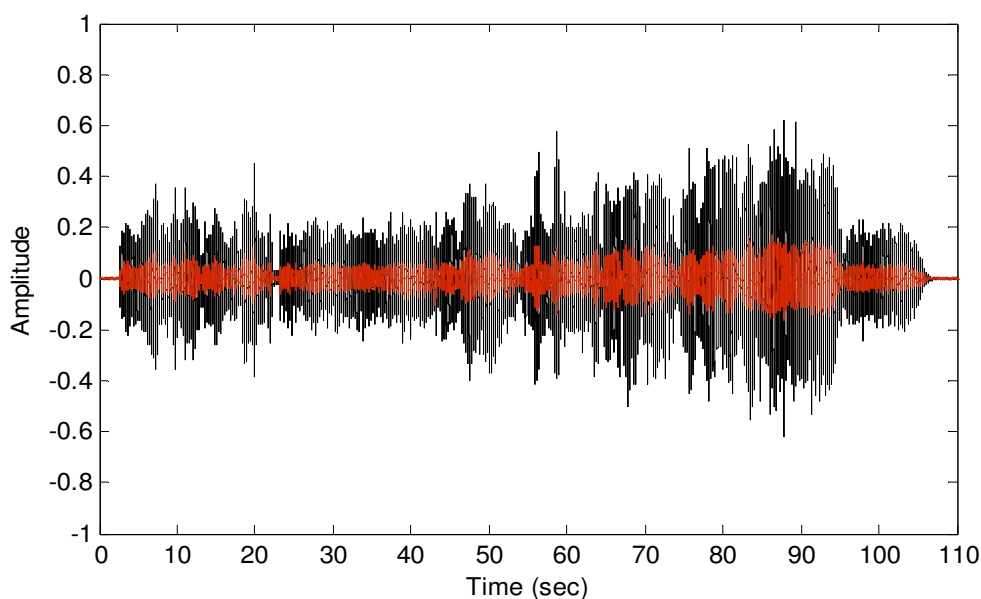
Σχ. σ.26: Σχηματική μορφή του συστήματος Decca Tree.

Όταν ο Haddy πρωτοείδε το σύστημα αυτό, σχολίασε ότι μοιάζει με χριστουγεννιάτικο δέντρο. Έτσι το όνομα έμεινε ως Decca Tree.

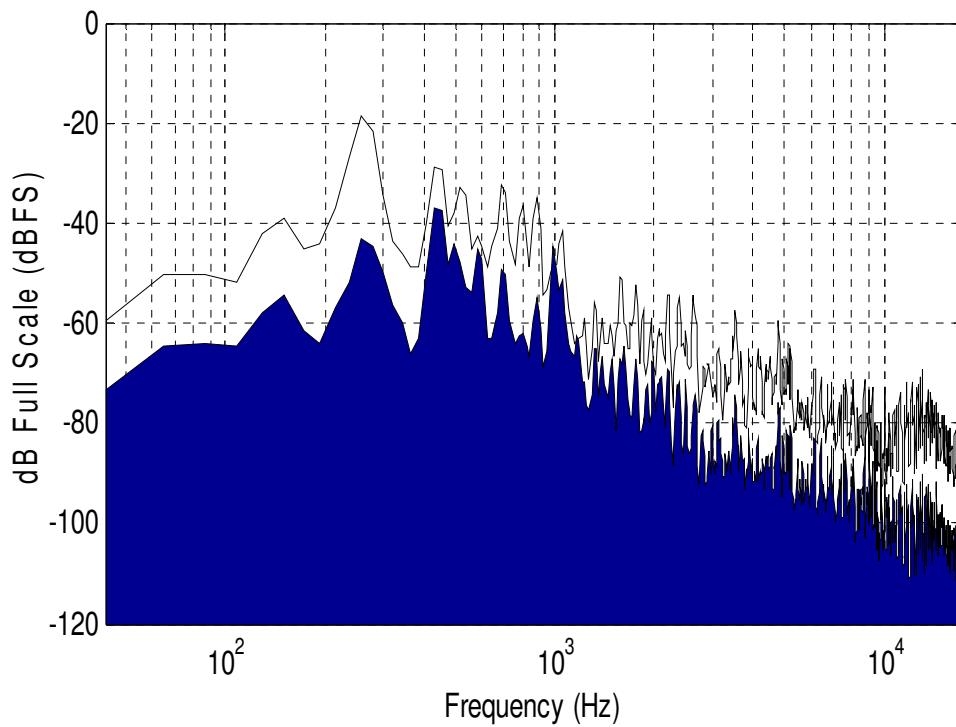
Επειδή ο ήχος φθάνει στο κεντρικό μικρόφωνο νωρίτερα απ' ό,τι στο αριστερό και το δεξί, έχει ως αποτέλεσμα η κεντρική εικόνα να είναι πολύ έντονη και σαφής. Μια παραλλαγή του συστήματος Decca Tree είναι να χρησιμοποιήσουμε ένα ζευγάρι μικροφώνων M-S τεχνικής για το μπροστινό μικρόφωνο, αντί του omni. Στόχος είναι να πάρουμε τη ζωτικότητα από το M-S ζεύγος και σε συνδυασμό με τα πλαινά μικρόφωνα να διευρύνουμε ελαφρώς το ηχητικό πεδίο.

ενώ παράλληλα είναι αυτοί που επηρεάζουν κατά το μικρότερο δυνατό ποσοστό τις δυναμικές διακυμάνσεις του σήματος). Έτσι είχαμε ως αποτέλεσμα τη διατήρηση των δυναμικών χαρακτηριστικών της αρχικής μορφής του κομματιού, ενώ παράλληλα, στην τελική του μορφή, του μουσικό κομμάτι έχει αυξημένο συνολικό level. Επίσης, κάποιοι ήχοι (μη μουσικοί) που προέρχονταν απο μέλη της ορχήστρας, όπως ανάσες, κ.α. επιλέχθηκε να μην «πειραχθούν». Η επιλογή αυτή πάρθηκε για δύο λόγους. Πρώτον γιατί δεν λειτουργούν αποπροσανατολιστικά σε σχέση με τη μουσικότητα του ηχητικού υλικού και δεύτερον γιατί η οποιαδήποτε προσπάθεια απομάκρυνσης τους με τα μέσα που είχαμε στη διάθεση μας, μάλλον αρνητικά αποτελέσματα θα επέφερε στο ηχητικό υλικό, παρά θετικά.

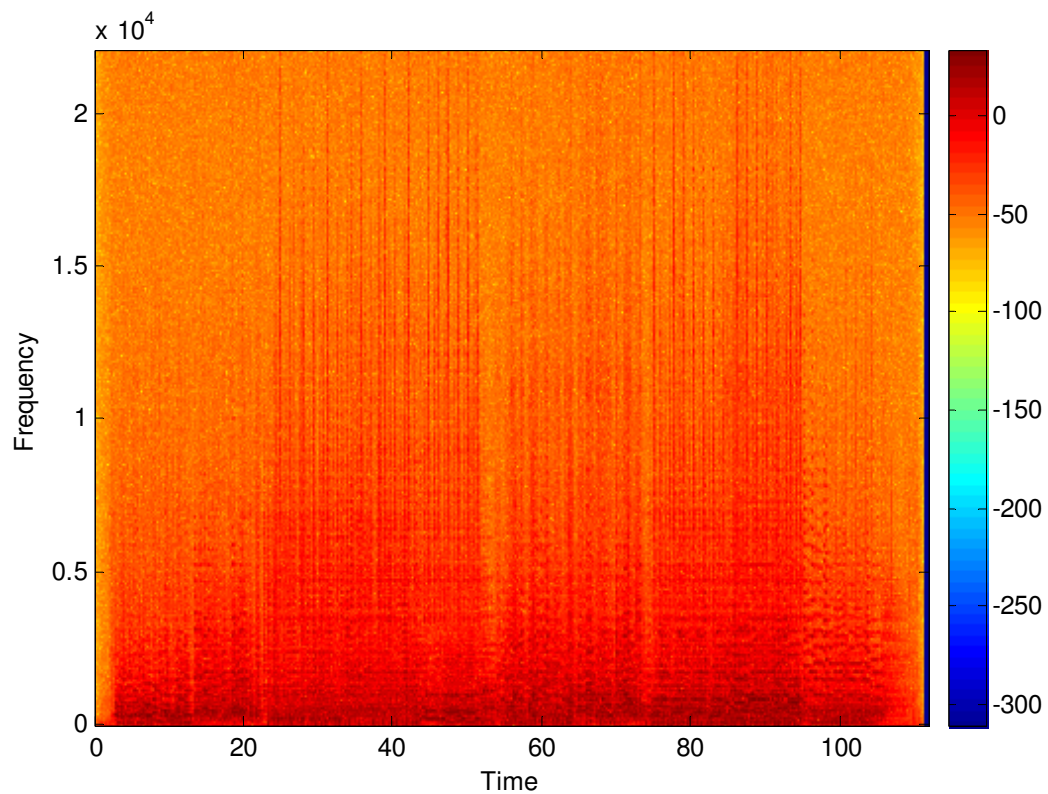
Τέλος αβροίστηκε στο ηχητικό υλικό ένα μικρό ποσοστό reverb για να δοθεί μια κάποια μεγαλύτερη διάρκεια στο σβήσιμο των μουσικών φράσεων. Και σε αυτό το πεδίο όμως η επεξεργασία μας ήταν συντηρητική, δεδομένου του ότι πρόκειται για μία στερεοφωνική ηχογράφιση μιας ολόκληρης ορχήστρας. Ακολουθούν τα διαγράμματα.



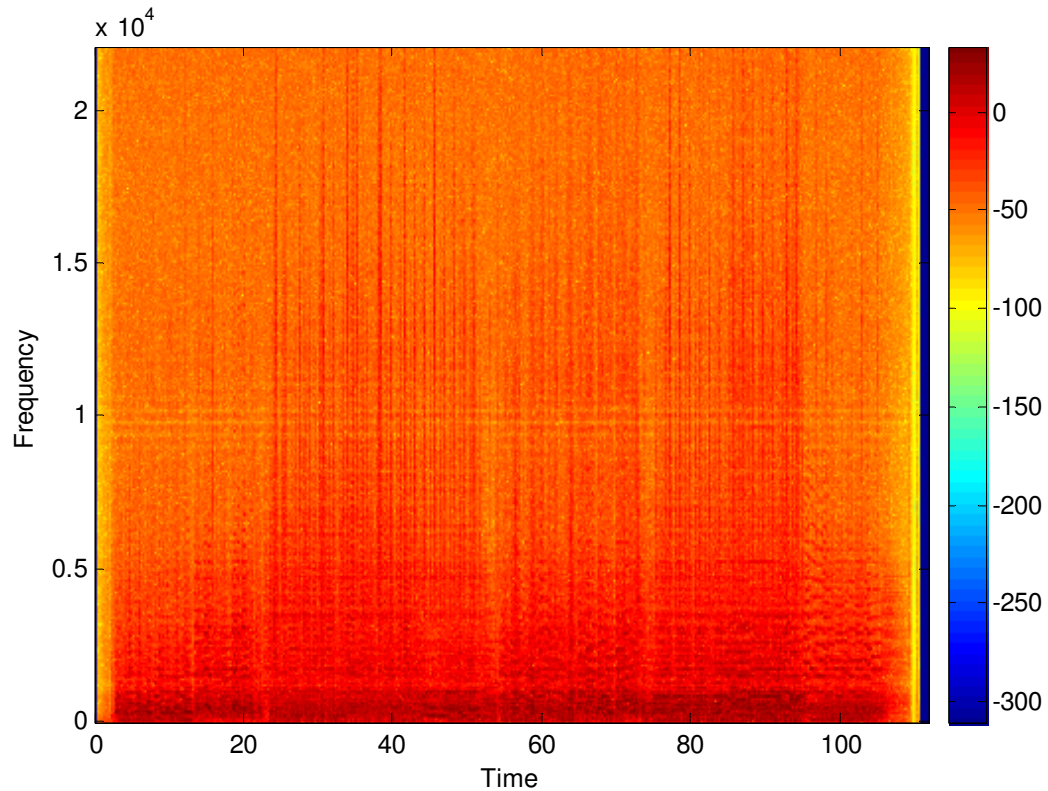
Σχ. 9.6.1: Το αρχικό και το τελικό αρχείο για το κομμάτι *Ballad of Mack the Knife* στο πεδίο του χρόνου.



Σχ. 9.6.2: Τα δύο σήματα (original και master) στο πεδίο των συχνοτήτων.



Σχ. 9.6.3: Το φασματογράφημα του αρχικού αρχείου ήχου (original).



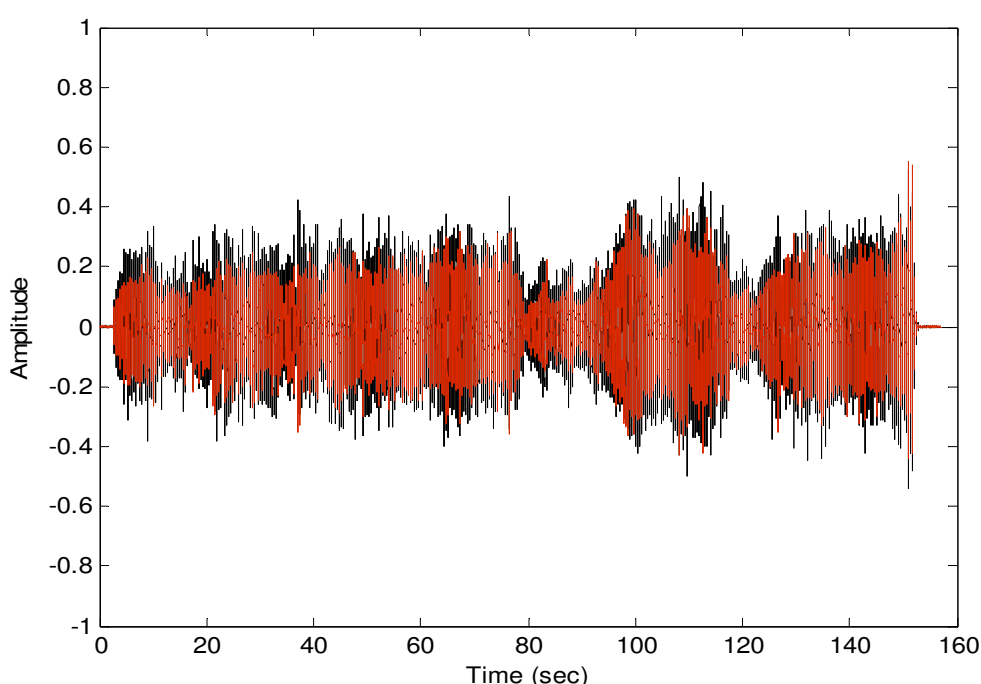
Σχ. 9.6.4: Το φασματογράφημα του τελικά παραγόμενου ηχητικού υλικού (master).

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	-15.43 dBFS	-14.77 dBFS	-25.38 dBFS	-25.43 dBFS	-32.73 dBFS	-34.10 dBFS
Master wav	-4.14 dBFS	-3.66 dBFS	-14.42 dBFS	-14.38 dBFS	-21.70 dBFS	-23.14 dBFS

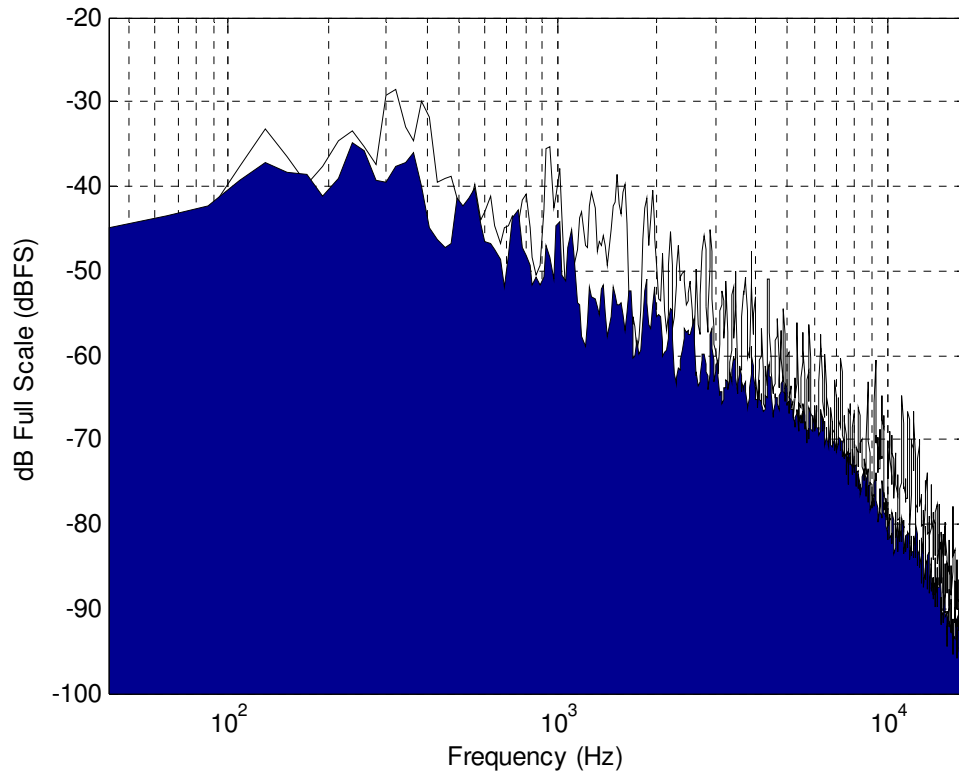
Πιν. 9.6.1: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το original και το master αντίστοιχα του μουσικού κομματιού Ballad of Mack the Knife του Kurt Weill.

9.7 Kurt Weill - Canon Song:

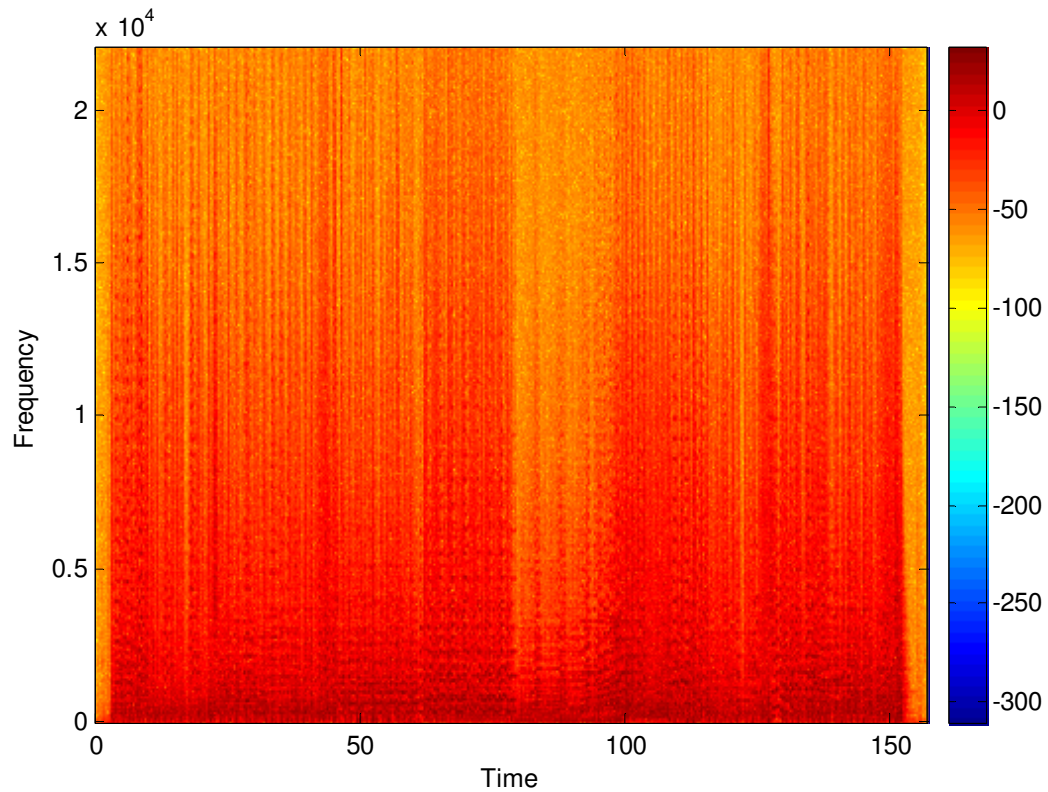
Για το μουσικό κομμάτι Canon Song του Kurt Weill, ισχύουν τα όσα έχουμε αναφέρει στην προηγούμενη παράγραφο για το κομμάτι Ballad of Mack the Knife, μιας και προέρχονται από την ίδια ηχογράφιση. Συνεπώς οι προβληματικές που εντοπίσαμε και οι λύσεις που κατά συνέπεια δόθηκαν ήταν όμοιες και άρα δεν θα είχε νόημα να τις επαναλάβουμε. Επομένως προχωράμε κατευθείαν στα διαγράμματα που προκύπτουν για το μουσικό αρχείο Canon Song. Το μόνο το οποίο αξίζει να σημειωθεί είναι πως στην περίπτωση αυτού του κομματιού η στάθμη ηχογράφησης του ήταν υψηλότερη και συνεπώς το limiting για το συγκεκριμένο μουσικό κομμάτι είχε ακόμη ευγενικότερα χαρακτηριστικά (κάτι που φαίνεται και από το σχήμα 9.7.1).



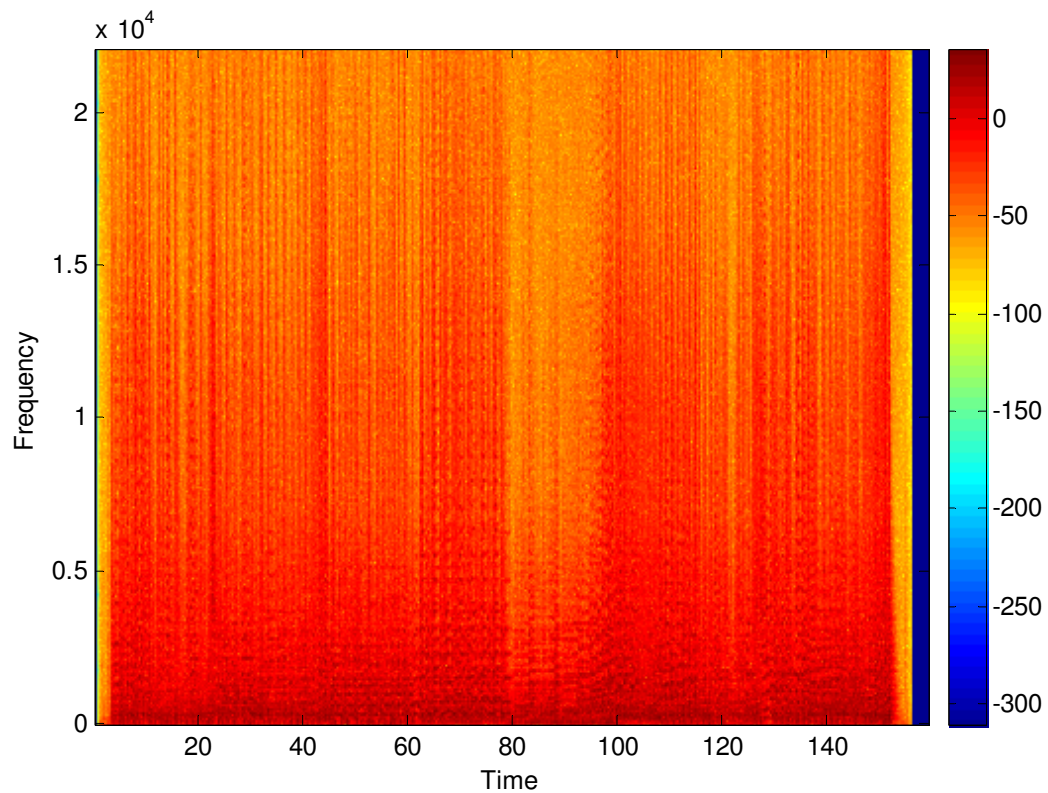
Σχ. 9.7.1: Τα δύο αρχεία για το τραγούδι Canon Song στο πεδίο του χρόνου. Παρατηρούμε πως το limiting στο συγκεκριμένο αρχείο είναι περιορισμένης τάξεως. Παρατηρούμε επίσης στο τέλος του κομματιού το fade out που έχει εφαρμοσθεί στο master.



Σχ. 9.7.2: Τα δύο σήματα στο πεδίο των συχνοτήτων. Παρατηρούμε ότι η διαφορά μεταξύ τους αναφορικά με το ποσοστό της ενέργειας που περιέχεται ανά συχνότητα είναι αμελητέα. Αυτό οφείλεται στη συντηρητική χρήση του limiter. Παρόλα αυτά οι διαφοροποιήσεις που έχουν προκύψει από το equalization και το multiband compression είναι εμφανείς.



Σχ. 9.7.3: Το φασματογράφημα του αρχικού ηχητικού υλικού (original).



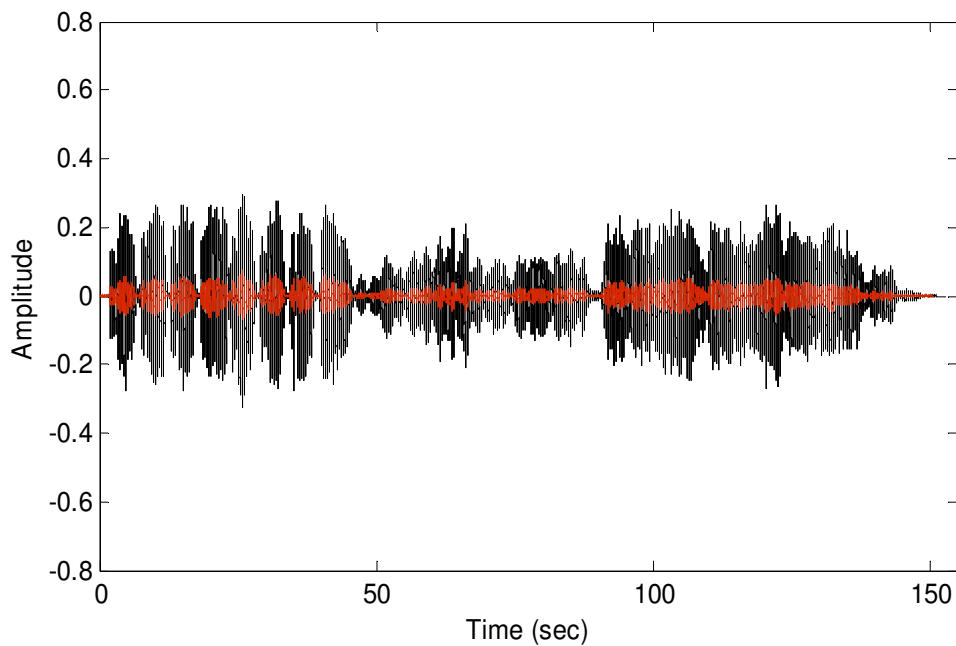
Σχ. 9.7.4: Το φασματογράφημα του τελικά παραγόμενου ηχητικού υλικού (master).

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	-5.11 dBFS	-5.19 dBFS	-16.34 dBFS	-16.27 dBFS	-25.96 dBFS	-26.35 dBFS
Master wav	-5.30 dBFS	-4.37 dBFS	-14.59 dBFS	-14.45 dBFS	-22.93 dBFS	-23.32 dBFS

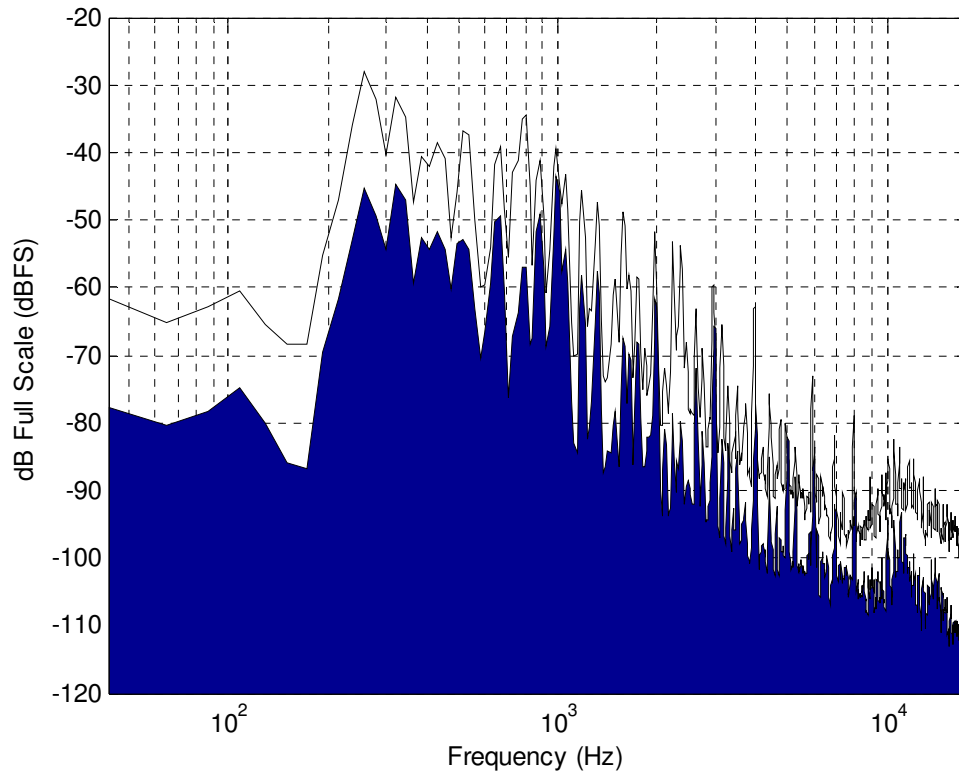
Πιν. 9.7.1: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το original και το master αντίστοιχα του μουσικού κομματιού Canon Song του Kurt Weill.

9.8 Kurt Weill - Polly's Song:

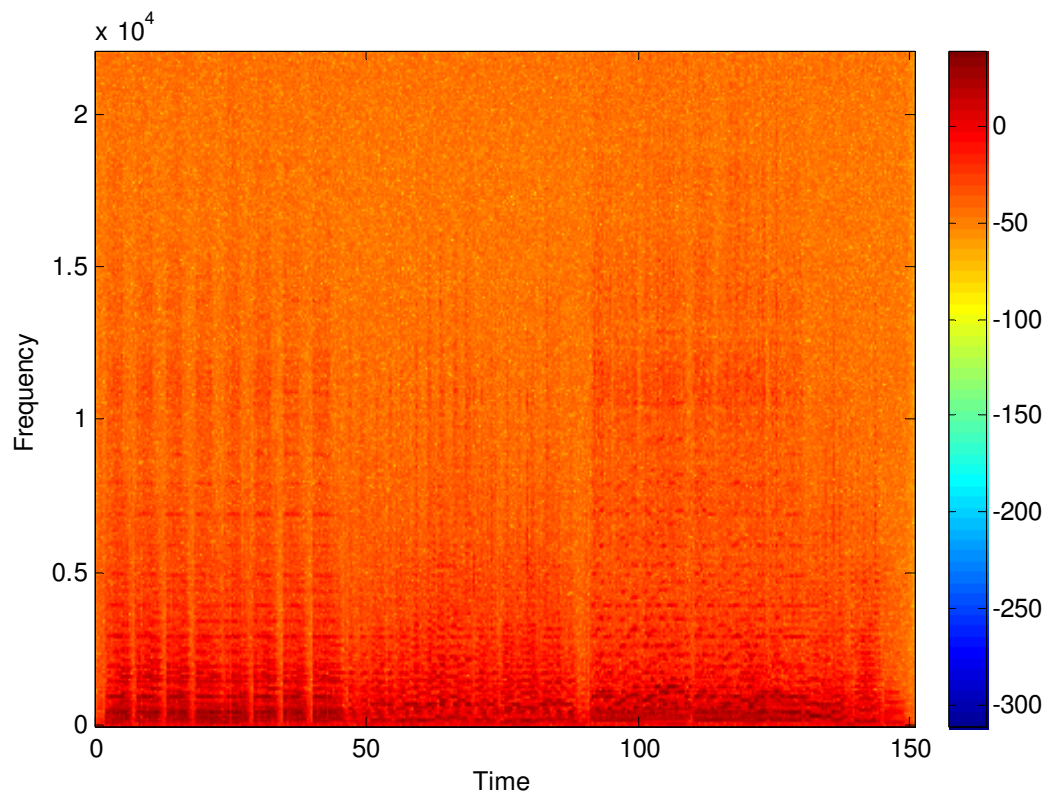
Για το μουσικό κομμάτι Polly's Song του Kurt Weill, ισχύουν τα όσα έχουμε αναφέρει στις δύο προηγούμενες παραγράφους. Το πλάτος εγγραφής του συγκεκριμένου μουσικού κομματιού ήταν όμοιο με αυτό για το τραγούδι Ballad of Mack the Knife και άρα οι ρυθμίσεις στη χρήση του limiter ήταν οι ίδιες. Σε ότι αφορά όλες τις άλλες μορφές επεξεργασίας του μουσικού υλικού, εφαρμόστηκαν αυτές που έχουν αναφερθεί στην παράγραφο 9.6. Προχωράμε, λοιπόν, στα διαγράμματα για το τραγούδι Polly's Song.



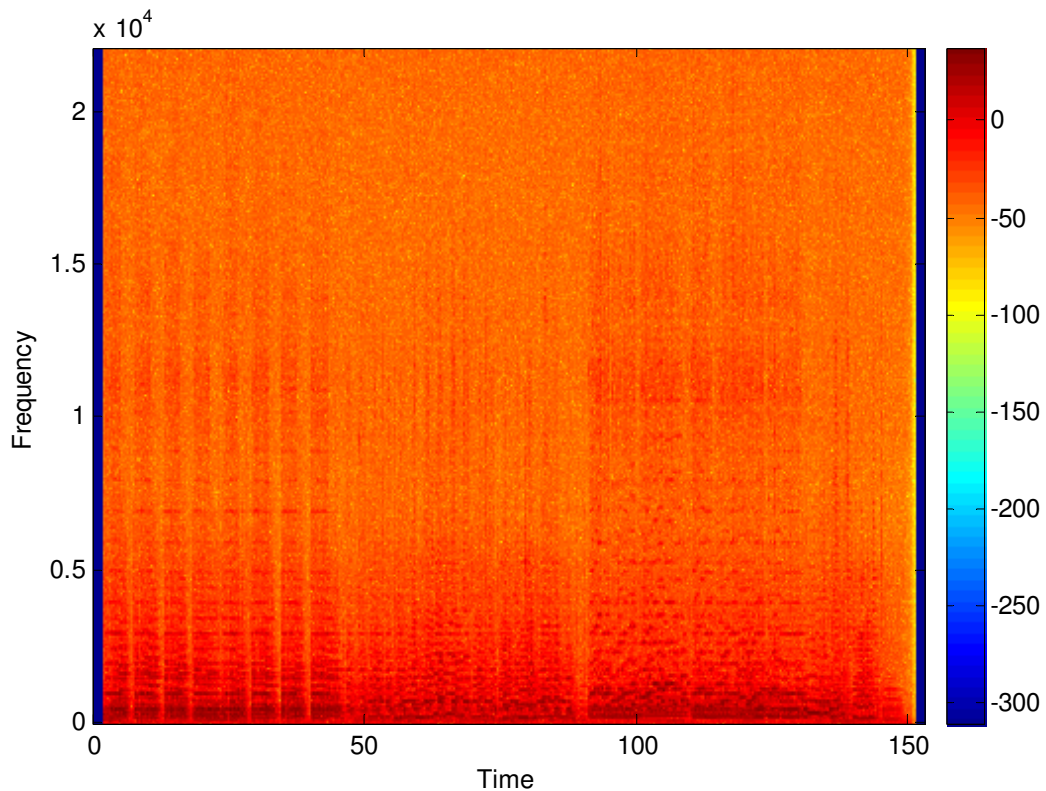
Σχ. 9.8.1: Τα δύο σήματα στο πεδίο του χρόνου.



Σχ. 9.8.2: Τα δύο σήματα στο πεδίο των συχνοτήτων.



Σχ. 9.8.3: Το φασματογράφημα του αρχικού ηχητικού υλικού (original).



Σχ. 9.8.4: Το φασματογράφημα του τελικά παραγόμενου ηχητικού υλικού (master).

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	-22.55 dBFS	-22.74 dBFS	-28.54 dBFS	-28.51 dBFS	-38.99 dBFS	-39.67 dBFS
Master wav	-9.67 dBFS	-10.46 dBFS	-17.69 dBFS	-17.71 dBFS	-25.74 dBFS	-26.93 dBFS

Πιν. 9.8.1: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το original και το master αντίστοιχα του μουσικού κομματιού Polly's Song του Kurt Weill.

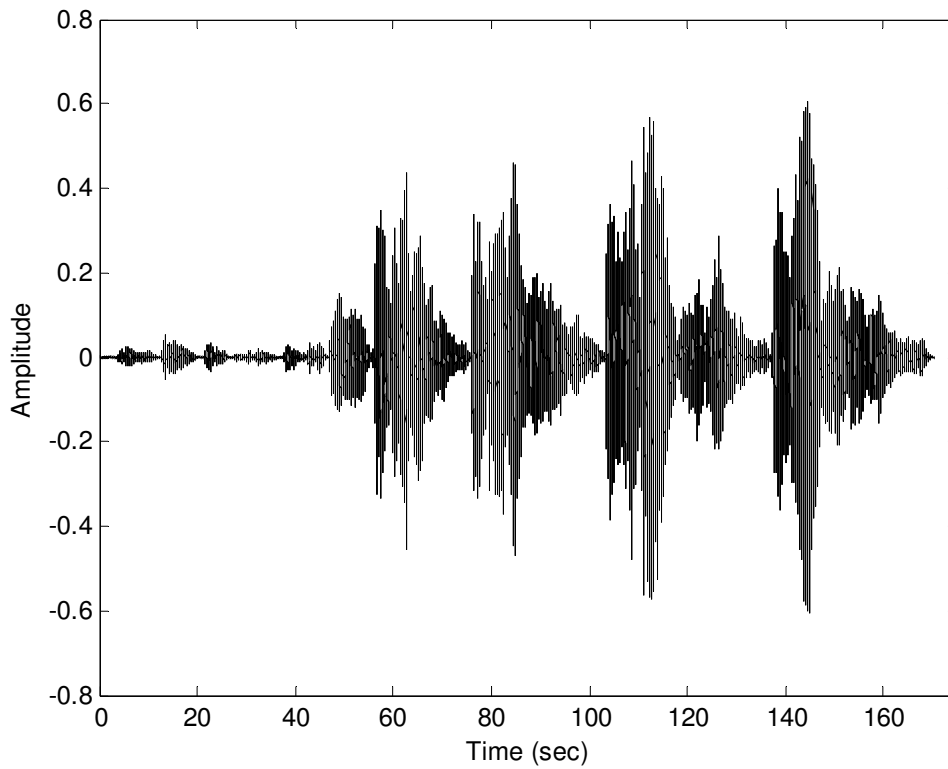
9.9 Sergei Rachmaninov – Come let us Worship:

Η συγκεκριμένη σύνθεση του Rachmaninov αποτελεί μέρος ενός ολόκληρου κύκλου συνθέσεων υπό τον τίτλο ολονυκτία, η οποία αναφέρεται στο θρησκευτικό αίσθημα.

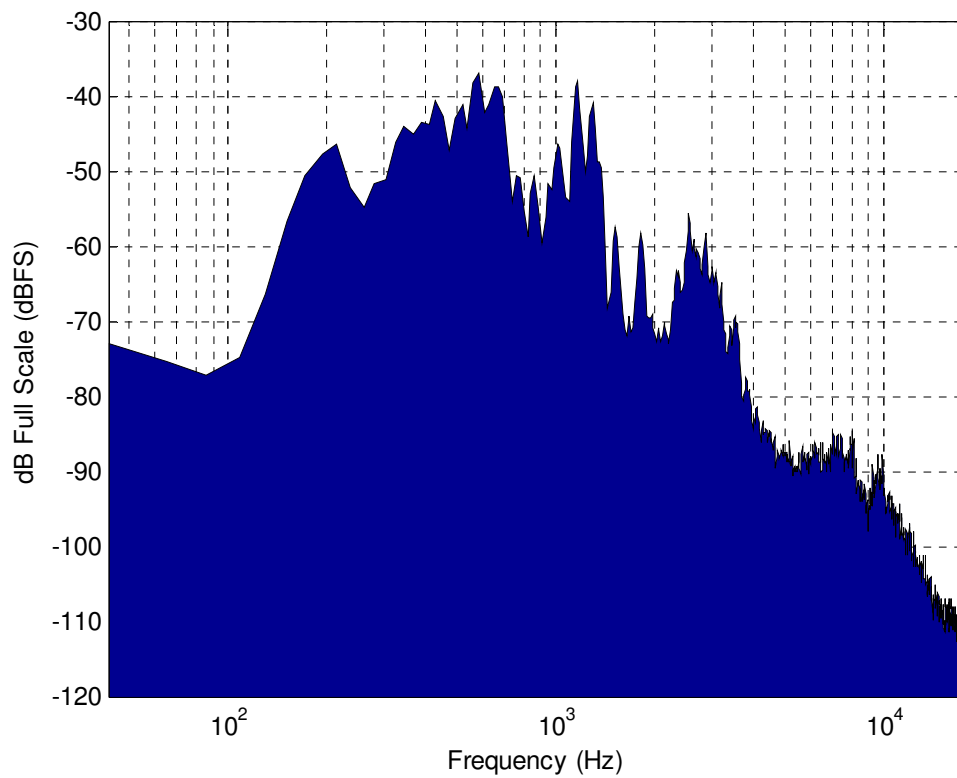
Στο προκείμενο αυτό που εμάς μας αφορά είναι η ίδια η σύνθεση και τα ηχητικά της ποιοτικά χαρακτηριστικά. Πριν περάσουμε όμως σε αυτά θα πρέπει να πούμε ότι στη συγκεκριμένη ηχογράφιση έλαβε μέρος η φιλαρμονική χορωδία της Εσθονίας, η οποία και θεωρείται μια από τις καλύτερες στον κόσμο στο είδος της, υπό την διεύθυνση του Paul Hillier. Η ηχογράφιση έγινε στις 24 Μαΐου του 2004 στην εκκλησία του αγίου Παύλου η οποία βρίσκεται στο Haapsalu της Εσθονίας. Ηχολήπτης της συγκεκριμένης ηχογράφησης ήταν ο Brad Michael. Η ηχογράφιση έγινε με ένα σύστημα Decca Tree συνεπικουρούμενο από μικρόφωνα χώρου.

Επανερχόμενοι στα ποιοτικά χαρακτηριστικά του ήχου αναφορικά με τη συγκεκριμένη ηχογράφιση, η αλήθεια είναι πως δεν έχουμε να πούμε και πολλά πράγματα. Είναι εξαιρετική σε όλα τα επίπεδα. Η κατανομή της ενέργειας στο συχνοτικό φάσμα, δηλ. το τονικό ισοζύγιο του μουσικού κομματιού είναι ιδανικό για το συγκεκριμένο μουσικό είδος. Τα μακροδυναμικά και μικροδυναμικά χαρακτηριστικά της συγκεκριμένης ηχογράφησης είναι πέρα από κάθε αμφιβολία εξαιρετικά. Η στερεοφωνική εικόνα είναι ανοικτή και εξαιρετικά φυσική (δίχως αυτή την τεχνική αίσθηση που εκλαμβάνουμε στις close mic. ηχογραφήσεις). Η αίσθηση του βάθους είναι φυσικότητα και τα χωρικά στοιχεία αναδεικνύουν το συνολικό έργο.

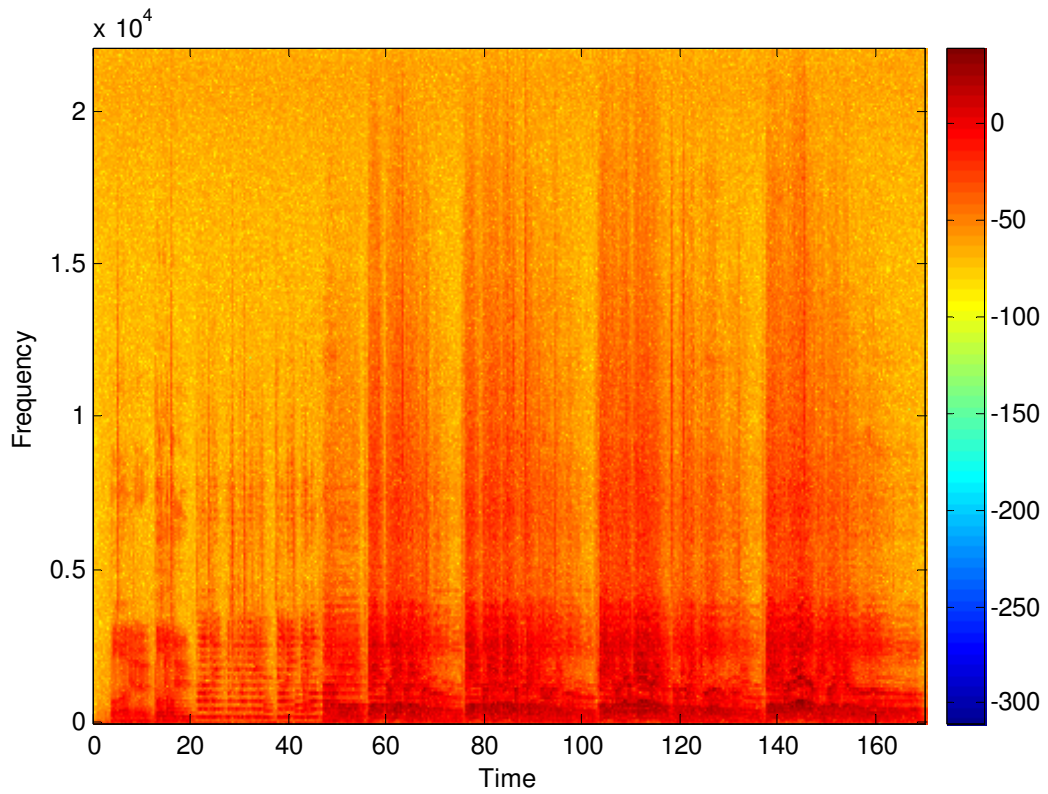
Με βάση, λοιπόν, όλα τα παραπάνω η διαδικασία του audio mastering για τη συγκεκριμένη ηχογράφιση θεωρείται περιττή. Η οποιαδήποτε περαιτέρω επεξεργασία θα είχε περισσότερο αρνητικά παρά θετικά αποτελέσματα. Είναι αυτό που είχαμε αναφέρει νωρίτερα σε μια παράγραφο όταν μιλούσαμε για τη χρήση του equalization: μία πολύ καλή ηχογράφιση μπορεί να μην χρειαστεί καθόλου mastering. Έτσι λοιπόν προχωράμε κατευθείαν στα διαγράμματα για να πάρουμε μια εικόνα της συγκεκριμένης ηχογράφησης.



Σχ. 9.9.1: Το μουσικό κομμάτι Come let us worship του Rachmaninov στο πεδίο του χρόνου.



Σχ. 9.9.2: Το ίδιο μουσικό κομμάτι στο πεδίο των συχνοτήτων.



Σχ. 9.9.3: Το φασματογράφημα του συγκεκριμένου μουσικού αρχείου.

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	-4,36 dBFS	-4,54 dBFS	-11,76 dBFS	-11,80 dBFS	-24,69 dBFS	-26,25 dBFS

Πιν. 9.9.1: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το μουσικό κομμάτι Come let us worship του Sergei Rachmaninov.

9.10 Jean Claude Risset – Sud La mer le matin:

Πριν ασχοληθούμε με τη συγκεκριμένη ηλεκτροακουστική σύνθεση του Jean Claude Risset, καλό θα ήταν να αναφέρουμε κάποια πράγματα σχετικά με την ηλεκτροακουστική μουσική γενικότερα και τη σχέση της με το audio mastering.

Είναι, λίγο - πολύ, σε όλους μας γνωστό το πώς πρέπει να ηχεί μια fender telecaster ή μια Les Paul, όπως επίσης μας είναι γνωστό το πώς θα έπρεπε να ηχεί μια στιβαρή και γεμάτη μπότα ή ένα στακάτο ταμπούρο. Όλα αυτά όμως δεν ισχύουν στην περίπτωση της ηλεκτροακουστικής μουσικής και αυτό γιατί κανείς δεν μπορεί να πει εκ των προτέρων πως θα πρέπει να ηχεί μια ηλεκτροακουστική σύνθεση για να είναι ακουστικά ενδιαφέρουσα. Όλα εξαρτώνται από αυτό που θέλει να αποδώσει ο συνθέτης, από το είδος της τεχνικής ή των τεχνικών που χρησιμοποίησε και φυσικά από τις δεξιότητες του. Όλα αυτά είχαν, μέχρι λίγο καιρό πριν, ως συνέπεια ελάχιστες ηλεκτροακουστικές συνθέσεις να φτάνουν στο στάδιο της επεξεργασίας του audio mastering. Παρόλα αυτά στο audio mastering μπορούν να βελτιωθούν τυχόν προβληματικά σημεία που μπορεί να υπάρχουν ακόμη και σε ηλεκτροακουστικές συνθέσεις.

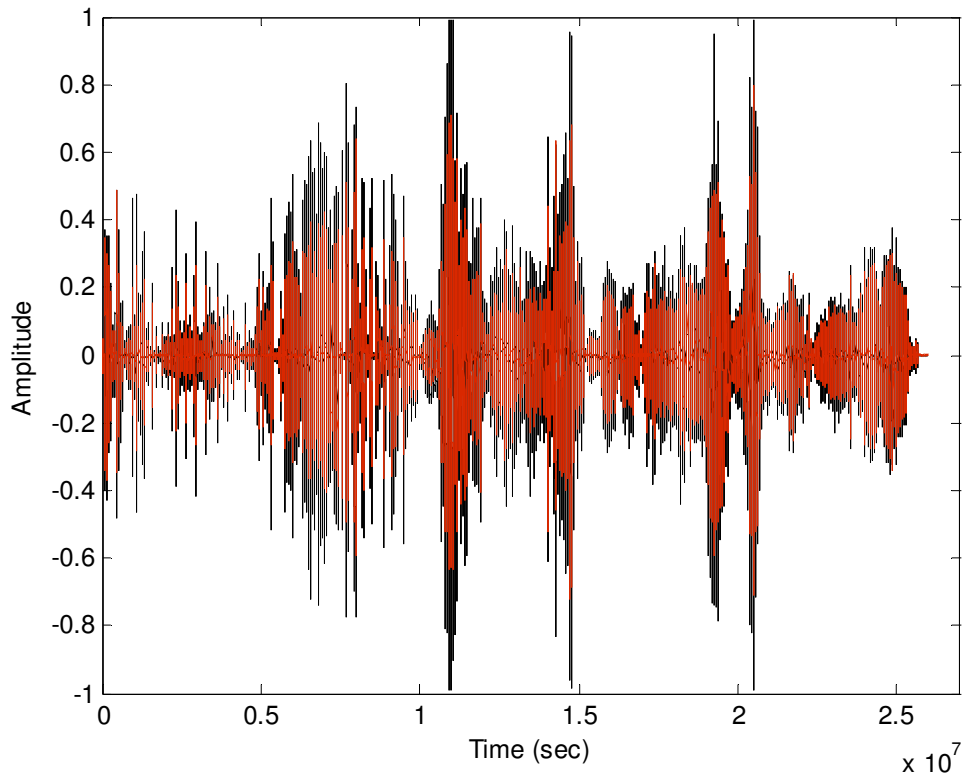
Όπως και σε όλα τα άλλα είδη της μουσικής, έτσι και στην ηλεκτροακουστική τα πάντα εξαρτώνται από την ποιότητα του αρχικού υλικού. Αυτό θα υποδείξει το αν και το κατά πόσον, μέχρι ποιου σημείου και σε ποια έκταση θα πρέπει ο τεχνικός να επεξεργαστεί το ηχητικό υλικό.

Στο προκείμενο τώρα, και αναφορικά με τη συγκεκριμένη σύνθεση του Jean Claude Risset η οποία φέρει τον τίτλο Sud – La mer le matin θα πρέπει να πούμε καταρχήν πως είναι ένα εξαιρετικό δείγμα του είδους της καθώς και της εποχής της.

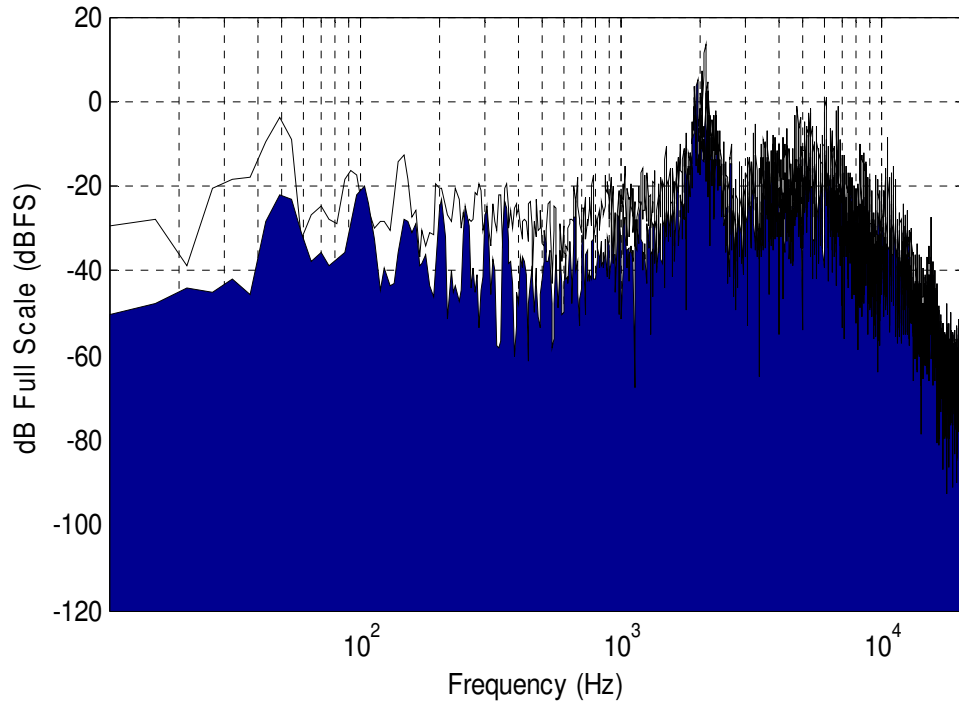
Συνεχίζοντας παραπέρα και αναφορικά με τα προβληματικά σημεία που εντοπίσαμε σε αυτό το μουσικό κομμάτι μπορούμε να πούμε ότι αυτά είχαν να κάνουν πρώτον με το συχνοτικό φάσμα του κομματιού, πιο συγκεκριμένα αισθανόμασταν πως η κάτω συχνοτική περιοχή ήταν αρκετά ελλειπής και πως η υψηλή συχνοτική περιοχή ήταν κατά μία έννοια υπερτονισμένη. Τα παραπάνω αντιμετωπίστηκαν με τη χρήση EQ καθώς και κάνοντας χρήση των δυνατοτήτων που μας προσφέρει το multiband compression. Το δεύτερο προβληματικό σημείο που εντοπίσαμε αφορούσε τη στερεοφωνική εικόνα του συγκεκριμένου μουσικού κομματιού. Πιο συγκεκριμένα διαπιστώσαμε πως σε κάποια σημεία, λόγω του έντονου pan – pot που λαμβάνει χώρα, υπήρχε μια ασάφεια στο κέντρο της στερεοφωνικής εικόνας, σαν κάποιος σε αυτά τα σημεία να άνοιγε μια τρύπα. Η προβληματική αυτή αντιμετωπίστηκε κάνοντας χρήση του plug-in: S1 – Imager της Waves· ουσιαστικά αυτά που κάναμε ήταν να σμικρύνουμε το εύρος της στερεοφωνίας. Αυτό έγινε με εξαιρετική φειδώ ώστε να μη χαθούν τα χαρακτηριστικά της σύνθεσης του Risset. Παρόλα αυτά (δηλ. αν και με συντηρητική χρήση της παραπάνω τεχνικής) το πρόβλημα αμβλύνθηκε και η στερεοφωνική εικόνα παρέμεινε «ανοιχτή». Ένα τρίτο πράγμα που κάναμε ήταν να προσθέσουμε ένα μικρό ποσοστό

reverb. Αυτό έγινε γιατί πιστεύουμε ότι λειτουργεί βοηθητικά στο συγκεκριμένο κομμάτι. Τέλος εφαρμόστηκε ένα μικρό, επίσης, ποσοστό limiting έτσι ώστε να αυξηθεί το συνολικό level του κομματιού δίχως παράλληλα να ζημιωθούν οι δυναμικές του διακυμάνσεις.

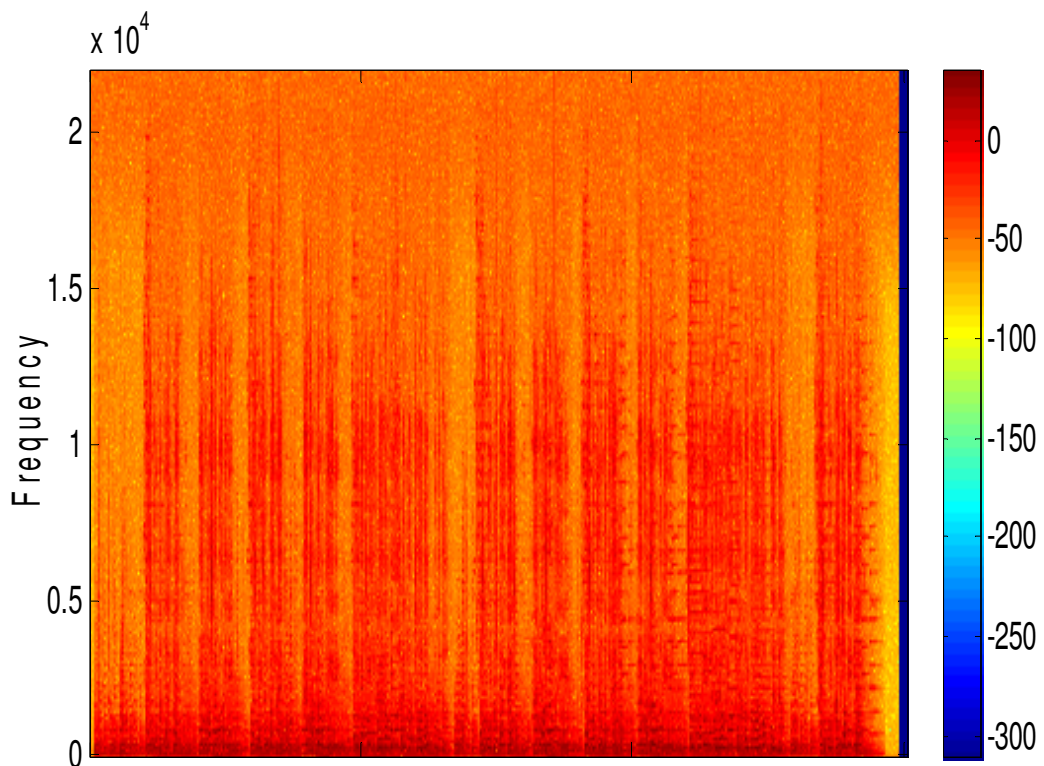
Περνάμε τώρα στα γραφήματα.



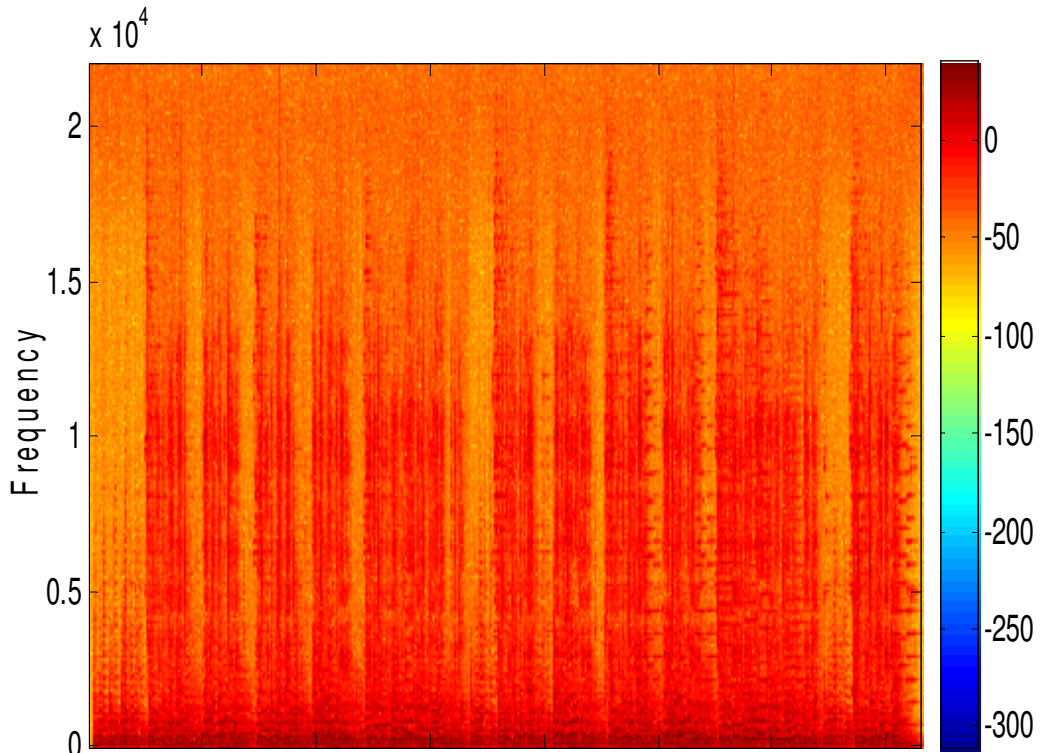
Σχ. 9.10.1: Τα δύο σήματα στο πεδίο του χρόνου για το κομμάτι Sud La mer le matin.



Σχ. 9.10.2: Τα δύο σήματα στο πεδίο των συχνοτήτων για το κομμάτι Sud La mer le matin.



Σχ. 9.10.3: Το φασματογράφημα του αρχικού ηχητικού υλικού (original) για το κομμάτι Sud La mer le matin.



Σχ. 9.10.4: Το φασματογράφημα που προκύπτει από το τελικά παραγόμενο ηχητικό υλικό (master).

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	0,00 dBFS	-2,01 dBFS	-10,41 dBFS	-10,47 dBFS	-18,30 dBFS	-19,32 dBFS
Master wav	-0,10 dBFS	-0,10 dBFS	-6,48 dBFS	-6,32 dBFS	-12,89 dBFS	-13,77 dBFS

Πιν. 9.10.6: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το original και το master αντίστοιχα του συγκεκριμένου μουσικού κομματιού.

9.11 P. Kokoras - Response:

Η σύνθεση αυτή του Παναγιώτη Κόκορα χρονολογείται κατά το έτος 2002. Ο τρόπος με τον οποίο συνέθεσε το συγκεκριμένο κομμάτι είναι χρησιμοποιώντας διάφορους χάρακες, διαφορετικού μήκους και υλικού ο καθένας, τους οποίους χτυπούσε σε μία σταθερή επιφάνεια, ενώ παράλληλα ηχογραφούσε τον παραγόμενο από την παραπάνω διαδικασία ήχο. Στη συνέχεια επεξεργάστηκε το αρχικό ηχητικό υλικό τόσο ως προς τις ακουστικές του συνιστώσες, όσο και ως προς την τοποθέτηση των διαφορετικών του πεδίων αναφορικά με - και στο - χρόνο.

Είναι ευνόητο, πως και γι' αυτή τη σύνθεση, ισχύουν τα όσα έχουμε αναφέρει περί ηλεκτροακουστικής μουσικής στην προηγούμενη παράγραφο. Λαμβάνοντας, λοιπόν, όλα τα παραπάνω υπόψη μας μπορούμε να περάσουμε στην εξέταση και της συγκεκριμένης ηλεκτροακουστικής σύνθεσης.

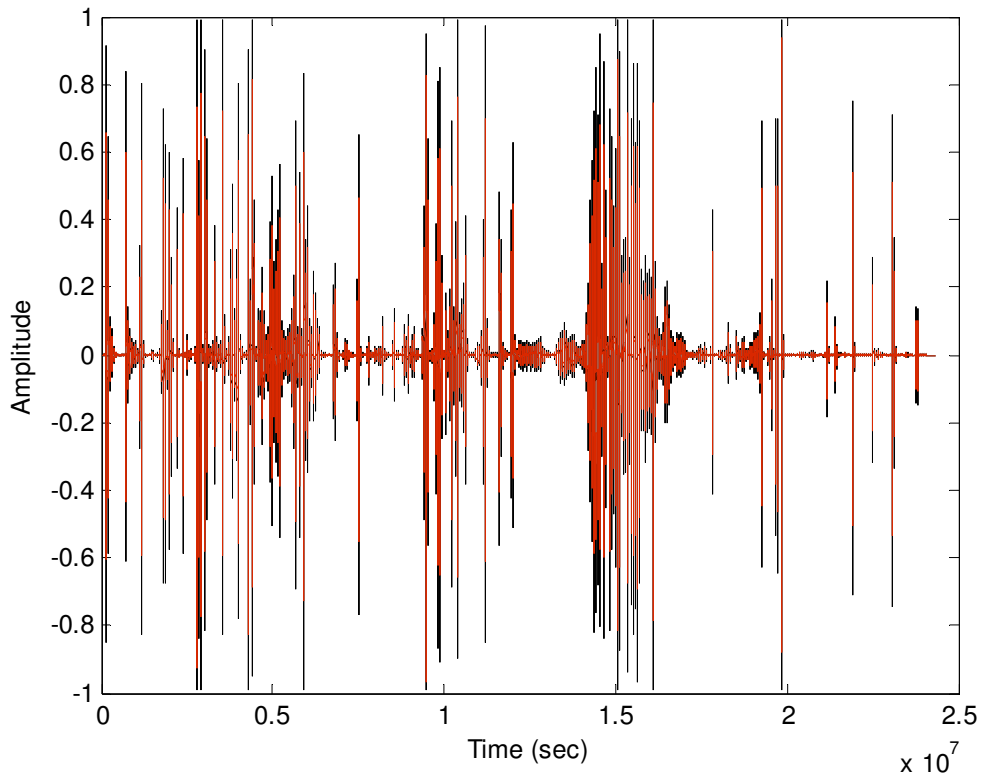
Το συγκεκριμένο μουσικό κομμάτι μπορούμε να πούμε πως δεν εμπεριείχε κανένα προβληματικό σημείο - πέρα από κάποια μεμονωμένα clicks που υπήρχαν, τα οποία όμως δεν γίνονταν αντιληπτά δεδομένης της πολύ μικρής τους διάρκειας και δεδομένης, επίσης, της μορφής των ήχων στους οποίους εμπεριέχονταν (κρουστικοί ήχοι) - συνεπώς το μεγάλο διακύβευμα και ο στόχος στην συγκεκριμένη περίπτωση δεν ήταν τόσο η αντιμετώπιση συγκεκριμένων προβληματικών σημείων, όσο ήταν η ανάδειξη των «δυνατών» σημείων της συγκεκριμένης σύνθεσης.

Στο πρακτικό, τώρα, επίπεδο τα clicks αντιμετωπίστηκαν με τη χρήση του manual declicker, κάτι το οποίο δεν απαίτησε ούτε ιδιαίτερο χρόνο, ούτε ιδιαίτερη προσπάθεια από τη μεριά μας. Τα «δυνατά» σημεία τώρα που έπρεπε να αναδείξουμε αφορούσαν δύο κυρίως επίπεδα: Το συχνοτικό εύρος του μουσικού κομματιού και τα χωρικά του στοιχεία. Ένα τρίτο επίπεδο αφορούσε το δυναμικό εύρος. Για να επιτευχθούν λοιπόν όλα τα παραπάνω επικεντρωθήκαμε αρχικά στη χρήση του equalizer και του multiband compressor για να φτιάξουμε κατά πρώτο λόγο μια στιβαρή μεσαία συχνοτική περιοχή και κατά δεύτερο για να τονίσουμε λίγο την περιοχή των χαμηλών συχνοτήτων. Τα παραπάνω ήταν αρκετά για να αναδειχθεί πλήρως ο ηχοχρωματικός πλούτος της σύνθεσης. Στην υψηλή περιοχή δεν υπήρξε καμία παρέμβαση μιας και δεν κρίναμε σκόπιμο κάτι τέτοιο.

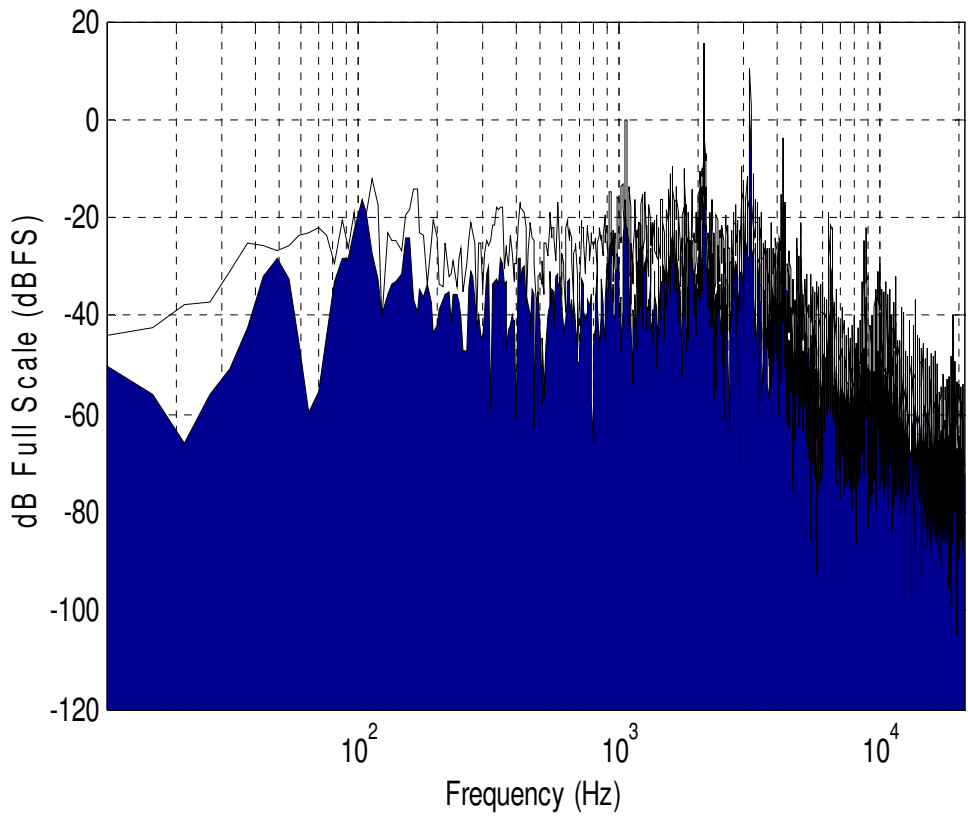
Αναφορικά με τα χωρικά στοιχεία του μουσικού κομματιού και την περαιτέρω, αλλά με μέτρο, ανάδειξη τους, κάναμε χρήση της τεχνικής του parallel compression όπως αυτή έχει περιγραφεί σε προηγούμενο κεφάλαιο. Θεωρούμε πως η επιλογή της συγκεκριμένης τεχνικής για τις ανάγκες του συγκεκριμένου κομματιού, μας δικαιώνει πλήρως.

Αναφορικά, τώρα, με το δυναμικό εύρος σκοπός μας ήταν να αποκτήσει μια περαιτέρω «αίσθηση» πράγμα που θα μπορούσε να αποτελέσει ένα από τα κύρια ποιοτικά χαρακτηριστικά, συνολικά, της σύνθεσης. Κάτι τέτοιο κατορθώθηκε κάνοντας φειδωλή χρήση του limiter.

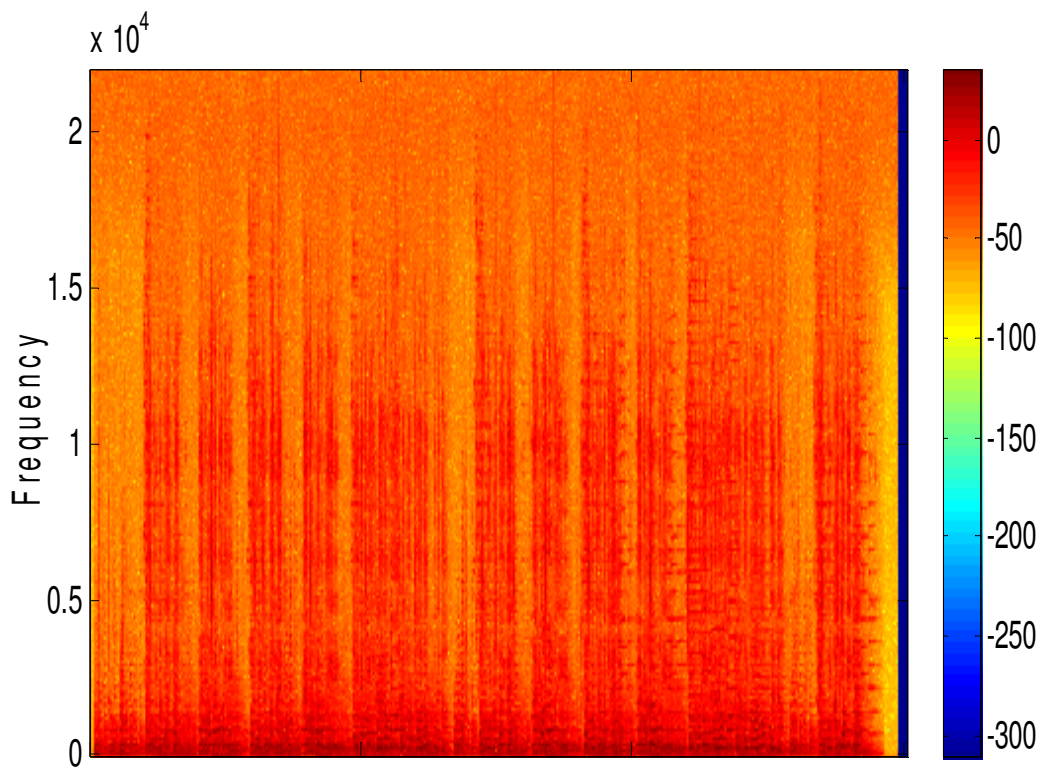
Αξίζει α σημειωθεί, για μια ακόμη φορά πως κύριο μέλημα μας, αναφορικά με τη σύνθεση του Π. Κόκορα, δεν ήταν τόσο η απαλοιφή τυχών προβληματικών της σημείων, όσο η ανάδειξη των ποιοτικών αυτών της χαρακτηριστικών που θα μπορούσαν να ενσταλάξουν μια περαιτέρω δυναμική στην ίδια τη σύνθεση. Παρακάτω ακολουθούν τα διαγράμματα και για αυτό το μουσικό κομμάτι.



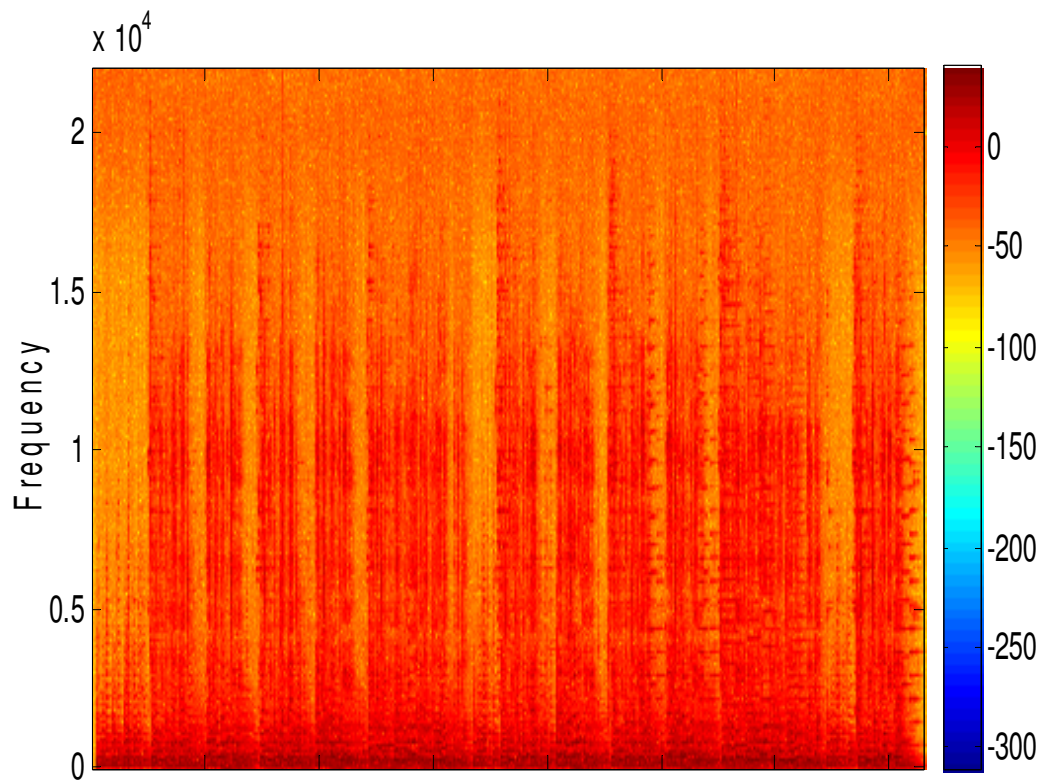
Σχ. 9.11.1: Τα δύο σήματα (original & master) στο πεδίο του χρόνου.



Σχ. 9.11.2: Τα δύο σήματα (original & master) στο πεδίο των συχνοτήτων.



Σχ. 9.11.3: Το φασματογράφημα που προκύπτει απο το αρχικό ηχητικό υλικό (original).



Σχ. 9.11.4: Το φασματογράφημα που προκύπτει απο το τελικά παραγόμενο ηχητικό υλικό (master).

	Peaks (Left)	Peaks (Right)	RMS(max) (Left)	RMS(max) (Right)	RMS(aver) (Left)	RMS(aver) (Right)
Original wav	0,00 dBFS	0,00 dBFS	-17,41 dBFS	-16,47 dBFS	-22,30 dBFS	-23,00 dBFS
Master wav	-0,10 dBFS	-0,10 dBFS	-14,48 dBFS	-14,32 dBFS	-17,09 dBFS	-17,77 dBFS

Πιν. 9.11.6: Οι τιμές των peaks καθώς και της ενεργούς τιμής (RMS) για το original και το master αντίστοιχα του συγκεκριμένης μουσικής σύνθεσης.

Συμπεράσματα

10. Συμπεράσματα – Αντί επιλόγου:

Στόχος της παρούσας πτυχιακής εργασίας ήταν να δώσει μια πρώτη απάντηση στο γιατί και στο με ποιο τρόπο υιοθετήθηκε το audio mastering από τη μουσική βιομηχανία στο σύνολο της. Επιπλέον στόχο από τελούσε το γεγονός του να δούμε τι επιπτώσεις, θετικές ή αρνητικές, επέφερε τελικά η επιλογή αυτή στον τρόπο που ακούμε και αντιλαμβανόμαστε το μουσικό συμβαινόμενο. Για να επιτύχουμε απαντήσεις στα παραπάνω ζητούμενα κάναμε χρήση της σχετικής με το αντικείμενο βιβλιογραφίας και αρθρογραφίας κατά πρώτο λόγο· και κατά δεύτερο, επιλέξαμε να επεξεργαστούμε μια σειρά μουσικών κομματιών διαφορετικού ύφους το καθένα.

Ο τρόπος με τον οποίο, αρχικά, είχαμε δομήσει τον τρόπο εργασίας μας αφορούσε κυρίως τρία επίπεδα από τα οποία και θα προέκυπταν τα τελικά συμπεράσματα. Αυτά ήταν: πρώτον μία αναφορά στον τρόπο επεξεργασίας που λαμβάνει χώρα κατά τη διαδικασία του mastering· δεύτερον η ίδια η επεξεργασία του ηχητικού υλικού που είχαμε στη διάθεση μας και η αξιολόγηση της μεταγενέστερης έναντι της προγενέστερης του μορφής· και τέλος, τρίτον, η σύναψη ερωτηματολογίου το οποίο θα απευθυνόταν σε ένα συγκεκριμένο πλήθος ανθρώπων με ανομοιογενή μουσικά χαρακτηριστικά όπου, παράλληλα με την ακρόαση του μουσικού υλικού, θα μπορούσαν να εξαχθούν συμπεράσματα κατά πόσο ωφελεί τελικά το mastering το ίδιο το ηχητικό υλικό.

Όμως, όπως κάθε έρευνα και εργασία, έτσι και η δική μας στην πορεία της υπέστη ορισμένες μεταλλαγές, ελπίζουμε προς το καλύτερο. Οι μεταλλαγές αυτές προέκυψαν τόσο από ζητήματα στην πορεία της εργασίας και τα οποία εκ των προτέρων δεν είχαμε συμπεριλάβει, όσο και από άποψη θέσης αναφορικά με το ζήτημα με το οποίο καταπιαστήκαμε.

Ένα από τα πράγματα που θέλαμε να κάνουμε αρχικά και τελικά δεν το κατορθώσαμε ήταν η δημιουργία ενός ερωτηματολογίου βάση του οποίου θα αξιολογούταν το μαστερarisμένο υλικό με βάση τα υποκειμενικά κριτήρια του καθενός. Αυτό δεν κατέστη δυνατό λόγω της δυσκολίας συγγραφής ενός τέτοιου ερωτηματολογίου το οποίο να μπορεί να μας δίνει απαντήσεις και δεδομένα με σαφήνεια. Επίσης δεν ήταν δυνατό να βρεθεί χώρος με την κατάλληλη ακουστική για τις ακροάσεις με βάση αυτό το ερωτηματολόγιο δίκως να παρεμποδίζεται κάποια άλλη διαδικασία στις εγκαταστάσεις του τμήματος και τρίτον γιατί δεν είχαμε τον κατάλληλο εξοπλισμό που απαιτούνταν ώστε να μην εκφυλιστεί το όλο διακύβευμα.

Θέση μας αποτέλεσε το ότι όντως το audio mastering μπορεί να βοηθήσει το ηχητικό υλικό στο σύνολο του (χωρίς αυτό να σημαίνει ότι το κάνει πάντα, πολλές φορές συμβαίνει το αντίθετο μάλιστα). Συνεπώς δεν εκκινήσαμε από μια ουδέτερη αφετηρία, γιατί ως τεχνικοί που ενδιαφέρονται για αυτό που κάνουν δεν θα μπορούσαμε άλλωστε.

Επομένως η πτυχιακή αυτή εργασία πήρε έναν περισσότερο ερευνητικό χαρακτήρα σε ένα, όμως, περισσότερο θεωρητικό επίπεδο προσπαθώντας να ορίσει τα όρια, τους περιορισμούς, αλλά και τις δυνατότητες του audio mastering. Βασική προϋπόθεση για να συμβεί αυτό ήταν η ενασχόληση μας με την ψυχοακουστική και ο τρόπος με τον οποίο αυτή διασυνδέεται με τις διεργασίες του γνωστικού μας αντικειμένου. Πιστεύουμε πως στα κεφάλαια που έχουν προηγηθεί υπήρξαμε εξαντλητικοί σε αυτό το πεδίο, τουλάχιστον για αυτό το αφηγηριακό σημείο στο οποίο βρισκόμαστε.

Συμπέρασμα αποτελεί πως σε τεχνικό επίπεδο οι περιορισμοί τίθενται περισσότερο από τα μέσα αποθήκευσης του ηχητικού υλικού. Το σημαντικό όμως είναι οι περιορισμοί που τίθεντο με βάση το ίδιο το ηχητικό υλικό και το τι πραγματικά χρειάζεται. Σε ένα τέτοιο επίπεδο ενασχόλησης δεν υπάρχουν βασικές κατευθυντήριες γραμμές, τα πάντα εξαρτώνται από τη μορφή, την ποιότητα (ηχητική & μουσική), το μουσικό είδος και πάνω απ' όλα τις «προθέσεις» του ηχητικού υλικού.

Σχηματικά, λοιπόν, θα μπορούσαμε να πούμε ότι η ηχογράφιση ενός κομματιού αποτελεί το 50% της ηχητικής του ποιότητας, η μίξη το 40% και το mastering το 10%. Το ποσοστό αυτό όμως μπορεί να είναι και 50% - 30% - 20% κατά αναλογία, ή ακόμα και 40% - 30% - 30% (όμως θα πρέπει να έχουμε κατά νου αναφορικά με το mastering ότι το θετικό του ποσοστιαίο πρόσθετο μπορεί να είναι και αρνητικό στην περίπτωση που ο σκοπός του είναι το απλά να ακούγεται δυνατά).

Έτσι μπορούμε λοιπόν να κατανοήσουμε ότι το audio mastering μπορεί να παραλληλιστεί με την διαγράμμιση και τη σήμανση ενός οδικού δικτύου. Αν αυτή είναι δόκιμη, μπορεί να κάνει το μουσικό μας ταξίδι πιο όμορφο και πιο ασφαλές, αν δεν είναι μπορεί να επιφέρει ακόμη και θανατηφόρο δυστύχημα. Ή ακόμη και με το μακιγιάζ μιας γυναίκας το οποίο μπορεί είτε να αναδεικνύει με εμφανέστερο τρόπο τα πιο όμορφα χαρακτηριστικά της, είτε να την προσομοιώνει με καρνάβαλο.

Το c.d. επομένως άλλαξε τον τρόπο αποθήκευσης και αναπαραγωγής του ήχου. Έχει μεγαλύτερη διάρκεια από το δίσκο βινυλίου 33 στροφών και μεγαλύτερο εύρος στις δυναμικές του ήχου. Τα μπάσα και τα πρίμα δεν στριμώχνονται όπως στα αυλάκια του βινυλίου. Αλλά οι δυνατότητες αυτές, στα ανθρώπινα χέρια και τις εμπορικές επιδιώξεις, δεν βελτιώνουν πάντα τη μουσική. Ίδιως από τη δεκαετία του '90, με τη χρήση ειδικευμένων προγραμμάτων και λειτουργικών, το ηχητικό προϊόν των ηχογραφήσεων προσαρμόζεται στις ανάγκες του εμπορίου με τρόπο πρωτόγνωρο. Για να ακούγεται πιο δυνατά, χωρίς «κενά», «παύσεις» και «χαμηλόφωνα σημεία» από τους ραδιοφωνικούς σταθμούς και τα μηχανήματα αναπαραγωγής, η μουσική υφίσταται βαθιές αλλοιώσεις.

Ένας απ τους τεχνικούς που διαμορφώνουν αυτό το νέο άκουσμα, ο Βλάντο Μέλερ, υποστηρίζει με ζήλο ότι «η σημερινή γενιά θέλει να ακούει τα c.d δυνατά». Αντιθέτως, ορισμένοι εκ των οποίων πρωτοπόρησαν της τεχνολογίας που εφηύρε αυτά

τα προγράμματα και λειτουργικά επεξεργασίας του ήχου, όπως ο Μάικ Ντόροου, αποσύρθηκαν απογοητευμένοι από τις καταχρήσεις που οδήγησαν στην «*καταστροφή της μουσικής*». Αυτός ο δυνατός ήχος, που έχει επικρατήσει στη συντριπτική πλειονότητα των ηχογραφήσεων, είναι τεχνητός και επιτυγχάνεται ακυρώνοντας όλες τις εσωτερικές ισορροπίες των μουσικών κομματιών. Εισαγωγές, σιγανά σημεία κι ανάσες καταργούνται,. Εξισώνονται ψηφιακά και ακούγονται, αν ακούγονται, το ίδιο δυνατά, ισοπεδώνοντας ακουστικά το σύνολο ενός κομματιού.

Εξαιτίας του τρομερού ανταγωνισμού στα ραδιοσυχνότητες, αυτή η τεχνική εφαρμόστηκε κατά κόρον στο ραδιόφωνο. Με ειδικά προγράμματα ή συσκευές εξισώνονται ακουστικά καθώς μεταδίδονται ραδιοφωνικά. Επίπεδα και δυνατά, χωρίς δυναμικές διακυμάνσεις, Για να μη χάσει ο ακροατής ούτε ένα δευτερόλεπτο το ισχυρό σήμα του ραδιοσταθμού.

Το ίδιο ισχύει και για τις νέες ηχογραφήσεις κατά γενική ομολογία. Συμπιέζονται προς τα κάτω οι υψηλές συχνότητες και ενισχύονται προς τα πάνω οι χαμηλές, μέχρι να δημιουργηθεί ένας νέος ακουστικά μέσος όρος, ο οποίος στη συνέχεια ενισχύεται περαιτέρω μέσω των limiters. Είμαστε περικυκλωμένοι από μουσική που δεν κάνει τίποτε άλλο πέρα από το να φωνάζει.

Οι σημερινοί μουσικοί, σε όλα τα μουσικά είδη, προσαρμόστηκαν σ αυτά τα νέα δεδομένα και πολλοί μάλιστα υπερέβαλλαν, όπως οι Metallika και οι Red Hot Chili Peppers που ανήγαγαν την παραμόρφωση σε επιστήμη, αμφότεροι με παραγωγό τον Ρικ Ρούμπιν.

Άλλοι, όμως, τραγουδοποιοί φρίττουν με τις παρενέργειες της συγκεκριμένης αντίληψης στη μουσική. Αυτή η ευρέως διαδεδομένη έλλειψη δυναμικών διακυμάνσεων δεν έχει προηγούμενο στην ιστορία της μουσικής. Το δυναμικό εύρος, αλλά ιδιαίτερα η δυναμική διακύμανση κατά τη διάρκεια ενός μουσικού κομματιού είναι από τους σημαντικότερους παράγοντες που κάνουν τη μουσική μια μαγική εμπειρία.

Αλλά οι παραγωγοί σε συνεργασία με τους τεχνικούς και με τη συναίνεση των «καλλιτεχνών» που επιζητούν πάση θυσία την εμπορική επιτυχία, διαμορφώνουν τον ήχο των τραγουδιών σύμφωνα με τις προσαγές του λάιφ στάιλ. Τα τραγούδια κατασκευάζονται (κι εδώ οι λέξεις δεν είναι τυχαίες, μιας κι αναφερόμαστε στην τεχνική κατασκευή και όχι στην ανθρώπινη δημιουργικότητα) για να παίζονται σε μπάρ και φουαγιέ ξενοδοχείων, αυτοκίνητα που κινούνται, τρένα και μετρό, κυλιόμενες σκάλες και καφετέριες, χώρους με πάσης φύσεως θορύβους. Όμως, αυτός ο «πόλεμος του δυνατού ήχου» για την επιβολή του στον ακροατή, θυμίζει σε μερικούς από εμάς την περίφημη ρήση του Χίτλερ ότι δεν θα είχε καταφέρει να κατακτήσει τη Γερμανία χωρίς τη βοήθεια των μεγαφώνων!

Γνωρίζουμε πως η αντοχή μιας αλυσίδας ορίζεται από την αντοχή του πιο αδύναμου κρίκου της. Στη αλυσίδα της μουσικής όμως ο πιο αδύναμος κρίκος δεν είναι το audio mastering. Είναι ο τρόπος με τον οποίο αναφερόμαστε και σχετιζόμαστε μαζί της. Ο πιο αδύναμος κρίκος της λοιπόν είναι η εμπορευματοποίηση της και η απεμπόληση του κοινωνικού της χαρακτήρα. Όλα τα υπόλοιπα αποτελούν ψελλίσματα ανοησίας.

Προτάσεις για τη συνέχεια:

α) Μια πρώτη πρόταση αφορά την εκπόνηση αυτού που δεν κατορθώσαμε να κάνουμε εμείς στην παρούσα εργασία. Δηλαδή τη μελέτη για τη συγγραφή ενός ερωτηματολογίου το οποίο θα απευθύνεται σε ένα ευρύ πλήθος ανθρώπων. Το ερωτηματολόγιο αυτό θα αφορά τη σύγκριση μουσικών τραγουδιών πριν αλλά και έπειτα από το στάδιο της επεξεργασίας του audio mastering. Είναι ευνόητο πως για να μπορέσει να πραγματοποιηθεί κάτι τέτοιο θα πρέπει να υπάρξει ιδιαίτερη μέριμνα κατά τη συγγραφή του ερωτηματολογίου, αλλά ιδιαίτερα για τη δημιουργία των υλικών αυτών υποδομών και των αντικειμενικών συνθηκών ώστε να μπορέσει να πραγματοποιηθεί κάτι τέτοιο.

β) Μια άλλη πρόταση μπορεί να αφορά την ενασχόληση με τη διαδικασία του audio mastering πολυκάναλων ηχογραφήσεων. Τόσο για τον κινηματογράφο και τη χρήση ειδικών εφέ, όσο και για τη μουσική. Θα μπορούσε, επίσης, και σε αυτό το πεδίο να αποπειραθεί η σύνδεση της συγκεκριμένης διαδικασίας με το γνωστικό αντικείμενο της ψυχοακουστικής. Τέλος θα μπορούσαν να διερευνηθούν τόσο οι δυνατότητες αναφορικά με αυτό το ζητούμενο, όσο και οι περιορισμοί που, ενδεχομένως, μπορεί να τίθενται.

γ) Μια τρίτη πρόταση αφορά την ενασχόληση με το αντικείμενο των binaural ηχογραφήσεων με τη διαδικασία, καθώς και με τον τρόπο αντιμετώπισης των συγκεκριμένων ηχογραφήσεων, στο audio mastering. Αναφορικά με τη συγκεκριμένη αυτή πρόταση, η ενασχόληση με το πεδίο της ψυχοακουστικής θεωρείται επιβεβλημένη. Αυτή η πρόταση, από τη στιγμή που θα πραγματοποιηθεί ως τέτοια, θα μπορούσε να επιφέρει πολύ ενδιαφέροντα συμπεράσματα καθώς ο πλούτος και η έκταση των εν λόγω ζητημάτων μπορούν να αποτελέσουν πεδίο τριβής, σύνθεσης και πειραματισμού και κατά συνέπεια να δημιουργήσουν έστω και σε ένα πρώτο μοριακό επίπεδο νέους ορίζοντες και κατευθύνσεις σχετικά με τις δυνατότητες και τη δυναμική του audio mastering.

δ) Τέλος μια άλλη πρόταση θα μπορούσε να αφορά τη διεξοδική μελέτη και ανάλυση του τρόπου με τον οποίο το audio mastering διαμορφώνει τον ήχο που ακούμε στις μέρες μας. Με την κατάλληλη επιχειρηματολογία θα μπορούσαν να καταδειχθούν τόσο τα φωτεινά, όσο και τα μελανά του σημεία. Από εκεί και έπειτα μπορούν να αρθρωθούν οι λύσεις αυτές, τόσο στο τεχνικό, όσο και στο αισθητικό επίπεδο, που σύμφωνα με την κρίση του γράφοντα αποτελούν διέξοδο ή που αποτελούν μία ενδιαφέρουσα προσέγγιση σε κάθε περίπτωση.

Βιβλιογραφία:

- [1] Ballou G. M._ Handbook for sound engineers, 3rd Edition, Focal Press, 2002.
- [2] Blauert J._ Spatial Hearing: the psychophysics of human sound localization, M.I.T, 1997.
- [3] Davis D. Patronis E._ Sound system engineering, 3rd Edition, Focal Press, 2006.
- [4] Fastl H. Zwicker E._ Psychoacoustics: facts and models, 3rd Edition, Springer, 2007.
- [5] Katz B._ Mastering Audio: the art and the science, Focal Press, 2002.
- [6] SADiE 4_ Operation manual
- [7] Sonic Studio HD_ No Noise Guide
- [8] Sonic Studio HD_ User Guide
- [9] Manley_ Massive passive stereo tube EQ, owner's manual.
- [10] MathWorks_ Signal Processing Toolbox with Matlab, Version 4.2, 1999.
- [11] Waves_ Q10 Paragraphic EQ, setup guide.
- [12]. Κουζούπης Σ._ Εφαρμοσμένη ακουστική ΙΙ, 2008.
- [13] Ξενικάκης Δ._ Ηχοληψία Ι: Σημειώσεις θεωρίας, 2006
- [14] Ξενικάκης Δ._ Ηχοληψία ΙΙ: Σημειώσεις θεωρίας, 2005.
- [15] Παπαδάκης Ν._ Ψηφιακή Επεξεργασία Σήματος: Σημειώσεις εργαστηρίου, 2008.
- [16] Παπαδάκης Ν._ Σημειώσεις Ψυχοακουστικής, 2008.
- [17] Σκαρλάτος Δ._ Εφαρμοσμένη ακουστική, Φιλομάθεια, 2005.
- [18] Φιτοανάκης Μ._ Mastering: Σημειώσεις θεωρίας, 2006.

[19] <http://www.digido.com/articles-and-demos12/13-bob-katz/19-mixing-tips-and-tricks.html>

[20] <http://www.digido.com/articles-and-demos12/15-other-audio-articles/34-mastering-engineers-discuss-dynamics-in-mixing.html>

[21] H. Fletcher and W.A. Munson, "Loudness, Its Definition, Measurement and Calculation, " J. Acoustical Society of America, Vol. 5, pp 82-108, October, 1933.

[22] S.S. Stevens and H. Davis, "Hearing - Its Psychology and Physiology," New York: Wiley, 1938, p. 124

[23] James A. Moorer Julia C. Wen, WHITHER DITHER: Experience with High-Order Dithering Algorithms in the Studio, Sonic Solutions San Rafael, CA USA.

Παράρτημα

Παράρτημα Α:

Οι παρακάτω πίνακες είναι πολύ χρήσιμοι όταν συναντάμε equalizer των οποίων τα ρυθμιστικά του εύρους ζώνης τους δίδονται με την τιμή του bandwidth (δηλ. σε Hz), αλλά εμείς επιθυμούμε να γνωρίζουμε την τιμή του Q, καθώς επίσης και το αντίστροφο. Οι τιμές του bandwidth είναι εκφρασμένες σε αριθμό οκτάβων και είναι μετρημένες στο σημείο των -3 dB. Η σχέση τώρα που μας δίνει τον αριθμό των οκτάβων από την τιμή των Hz είναι η:

$$\text{Αριθμ. Οκτ. (Α.Ο.)} = \frac{\log_{10}\left(\frac{f_2}{f_1}\right)}{\log_{10} 2} \quad (\text{Π.Α.1})$$

B/W	Q
0.02	72.13
0.03	48.09
0.04	36.07
0.05	28.85
0.06	24.04
0.07	20.61
0.08	18.03
0.09	16.03
0.10	14.42
0.20	7.21
0.30	4.80
0.40	3.60
0.50	2.87
0.60	2.39
0.70	2.04
0.80	1.78
0.90	1.58
1.00	1.41
1.20	1.17
1.40	0.99
1.60	0.86
1.80	0.75
1.90	0.71
2.00	0.67
2.20	0.60
2.40	0.54
2.60	0.49
2.80	0.44
3.00	0.40

(α)

Q	B/W
0.50	2.54
0.55	2.35
0.60	2.19
0.65	2.04
0.70	2.00
0.75	1.80
0.80	1.70
0.85	1.61
0.90	1.53
0.95	1.46
1.00	1.39
1.10	1.27
1.20	1.17
1.30	1.08
1.40	1.01
1.50	0.94
1.60	0.89
1.70	0.84
1.80	0.79
1.90	0.75
2.00	0.71
3.00	0.48
4.00	0.36
5.00	0.29
6.00	0.24
8.00	0.18

(β)

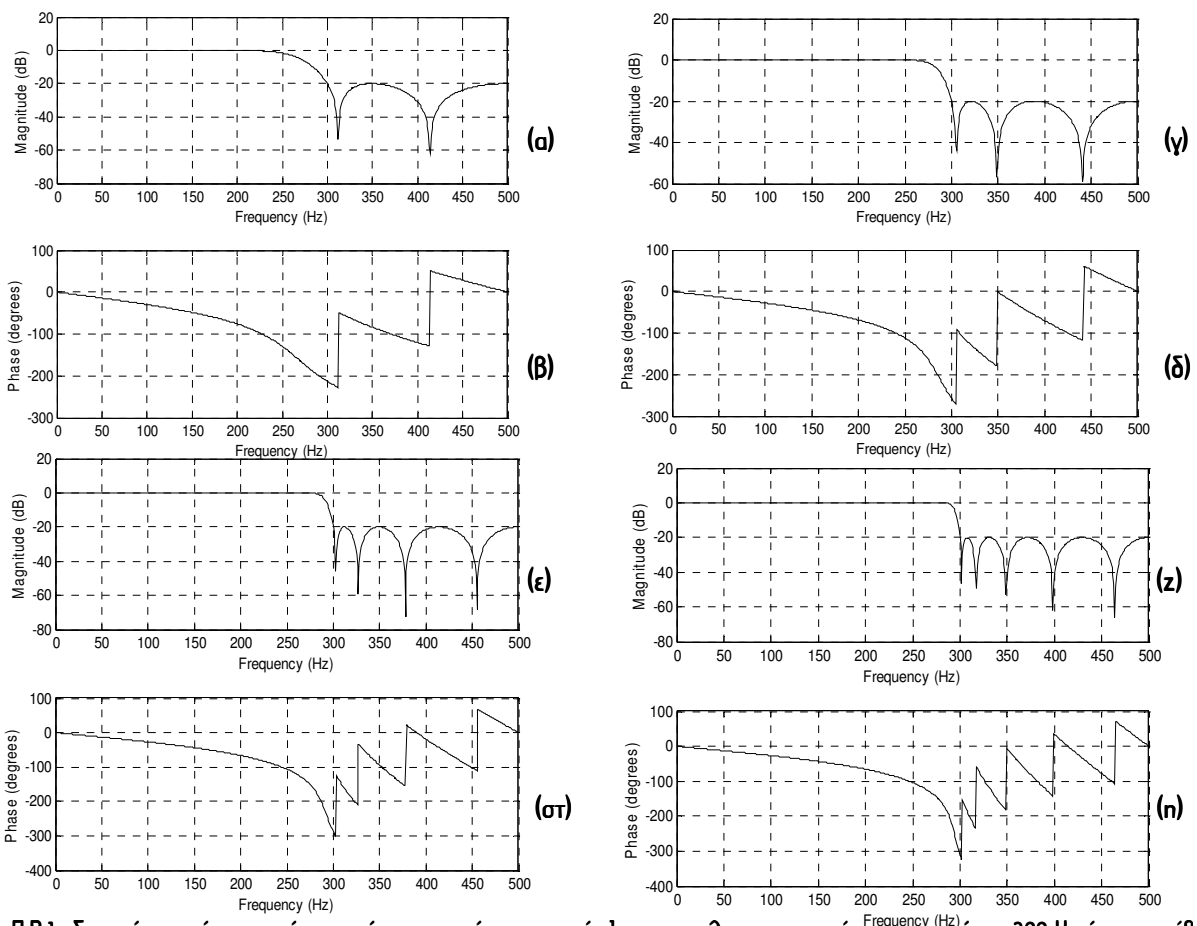
Πιν. Π.Α.1: (α) Μετατροπή του εύρους του φίλτρου από bandwidth σε Q. (β) Μετατροπή του εύρους του φίλτρου από Q σε bandwidth.

Παράρτημα Β:

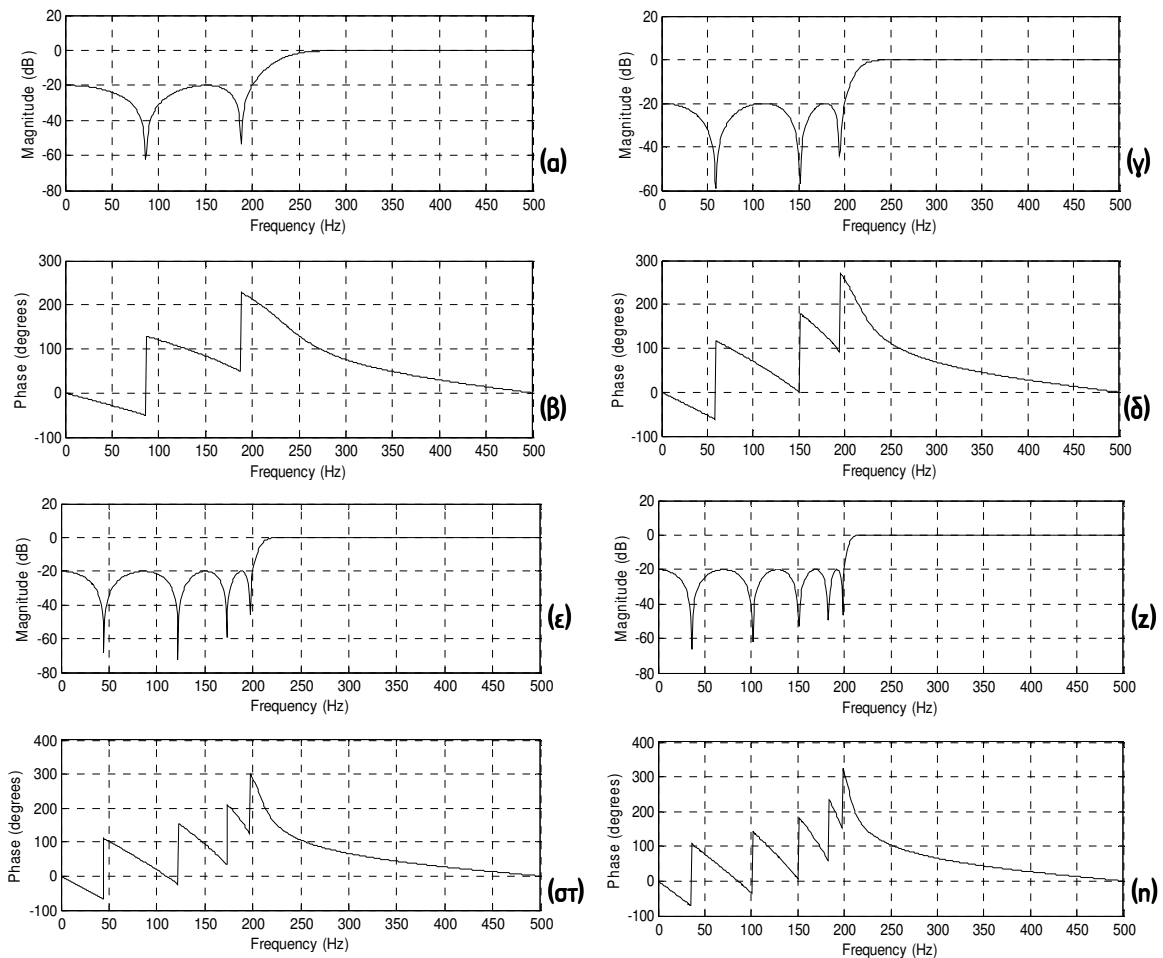
Σε αυτό το παράρτημα θα ασχοληθούμε με τη αναστροφή φάσης που επιφέρει η χρήση των equalizer στο μουσικό σήμα, καθώς και με τα διάφορα επιφαινόμενα, όπως επίσης και με τα αίτια τους, που προκύπτουν από αυτή τη συνθήκη.

Αρχικά θα μελετήσουμε ορισμένα low pass φίλτρα, διαφορετικής τάξης (order), και θα δούμε την απόκριση της αναστροφής φάσης των φίλτρων αυτών συναρτήσει της συχνότητας και της τάξης τους, δηλ. συναρτήσει του πόσο κάθετη ή όχι είναι η καμπύλη αποκοπής τους και κατά συνέπεια την επίδραση που έχει αυτός ο παράγοντας στο μουσικό σήμα. Στη συνέχεια θα πράξουμε το ίδιο και για μια σειρά από high pass φίλτρα. Επίσης το ίδιο, συν την εξαγωγή του delay που προκύπτει από ένα notch equalizer, θα εξετάσουμε ως αντιπροσωπευτικό παράδειγμα της οικογένειας των παραμετρικών EQ.

Τέλος με αφορμή όλα τα παραπάνω θα εξετάσουμε πώς αυτά επιδρούν στο μουσικό υλικό συναρτήσει του πώς το αντιλαμβάνεται η ανθρώπινη ακοή και ως απότοκο αυτού θα αναφέρουμε μία εξειδικευμένη (στη σύλληψη της ως τέτοια και όχι στην εφαρμογή της) τεχνική του audio mastering που κάνει χρήση όλων αυτών.



Σχ. Π.Β.1.: Στο σχήμα αυτό παρουσιάζονται τέσσερα ζευγάρια ψηφιακών low pass φίλτρων με συχνότητα αποκοπής τα 300 Hz όπου το κάθε ένα από αυτά είναι διαφορετικής τάξης (με αύξουσα σειρά από το (α) στο (γ) κ.ο.κ.). Τα (α), (γ), (ε) και (z) παρουσιάζουν το κέρδος του φίλτρου συναρτήσει της συχνότητας. Τα (β), (δ), (στ) και (η) παρουσιάζουν την απόκριση φάσης του φίλτρου συναρτήσει της συχνότητας και αποτελούν το συμπλήρωμα του ζευγαρώματος των (α), (γ), (ε) και (z) αντίστοιχα.



Σχ. Π.Β.2.: Εδώ παρουσιάζονται τέσσερα ζευγάρια ψηφιακών High pass φίλτρων με συχνότητα αποκοπής τα 250 Hz όπου το κάθε ένα από αυτά είναι διαφορετικής τάξης (με αύξουσα σειρά από το (α) στο (γ) κ.ο.κ.). Τα (α), (γ), (ε) και (ζ) παρουσιάζουν το κέρδος του φίλτρου συναρτήσει της συχνότητας. Τα (β), (δ), (στ) και (η) παρουσιάζουν την απόκριση φάσης του φίλτρου συναρτήσει της συχνότητας και αποτελούν το συμπλήρωμα του ζευγαρώματος των (α), (γ), (ε) και (ζ) αντίστοιχα.

Στο σχήμα Π.Β.1 της προηγούμενης σελίδας παρουσιάζονται τέσσερα (4) low pass φίλτρα με τα διαγράμματα τους (κέρδος του φίλτρου συναρτήσει της συχνότητας και φάση του φίλτρου συναρτήσει της συχνότητας). Τα φίλτρα αυτά σχεδιάστηκαν στο Matlab και έχουν συχνότητα αποκοπής τα 300 Hz. Καταρχήν να αναφέρουμε πως η κυμάτωση που παρουσιάζεται από τη συχνότητα αποκοπής και πέρα είναι ένα χαρακτηριστικό που συναντάμε και ορίζεται ως stop band ripple (stopripple). Όπως φαίνεται στα διαγράμματα ανεξάρτητα από την τάξη του φίλτρου, το πλάτος του stop band ripple είναι πάντα κατά 20 dB ασθενέστερο από αυτό των συχνοτήτων των οποίων επιτρέπεται η διέλευση. Παρατηρούμε, επίσης, πως όσο αυξάνεται η τάξη του φίλτρου αυξάνεται και το stopripple, τόσο σε ότι αφορά το level του, όσο και σε ότι αφορά το εύρος των συχνοτήτων στο οποίο αυτό εμφανίζεται.

Πριν περάσουμε στα διαγράμματα φάσης συναρτήσει της συχνότητας ας ξεκινήσουμε με μια μικρή υπενθύμιση ορισμένων θεμελιωδών θεωρητικών γνώσεων.

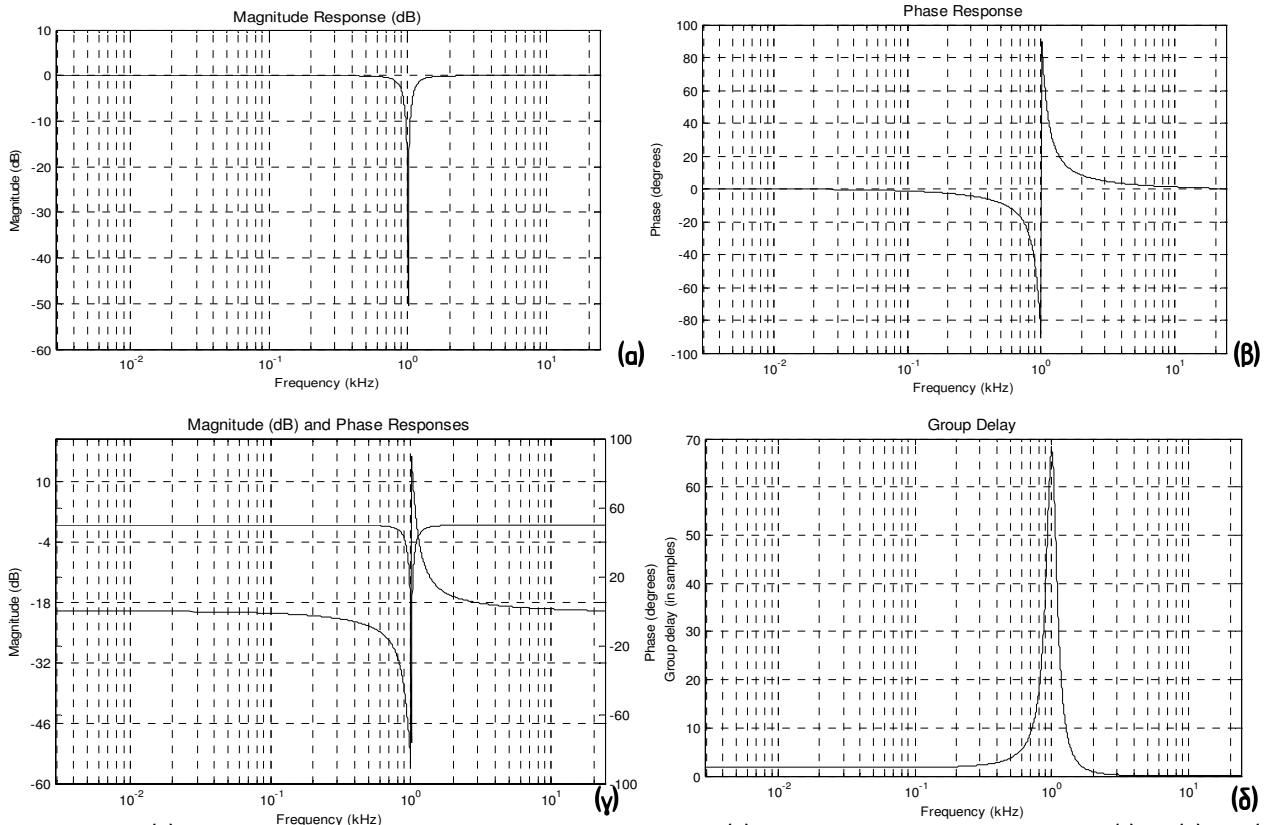
Αποτελεί, λοιπόν, κοινή παραδοχή πως όλα τα αναλογικά equalizer και σχεδόν όλα τα ψηφιακά, παράγουν μία ορισμένη αναστροφή φάσης στο audio σήμα όταν

ενισχύουμε ή αποκόπτουμε με αυτά μια συγκεκριμένη συχνотική περιοχή του φάσματος. Αυτή η αναστροφή φάσης παράγει μια χρονική καθυστέρηση στο σήμα η οποία, σημειωτέον, δεν είναι σταθερή αλλά εξαρτάται από τη συχνότητα καθώς, επίσης, και από το ποσοστό ενίσχυσης ή εξασθένησης που εφαρμόζει το EQ σε αυτή. Τα φίλτρα υψηλής τάξης, καθώς κι αυτά των οποίων η τιμή του Q είναι υψηλή παράγουν υψηλότερο ποσοστό αναστροφής φάσης και κατά συνέπεια μεγαλύτερες τιμές delay. Διαπιστώνουμε, λοιπόν, πως τα equalizer δεν εξαντλούν την επιρροή τους, στο σήμα, μόνο στο πεδίο των συχνοτήτων· αλλά, επίσης, επηρεάζουν και το πεδίο του χρόνου. Για αυτό το λόγο η αναστροφή φάσης, στο audio λεξιλόγιο, αναφέρεται συχνά κι ως φασική παραμόρφωση (phase distortion).

Περνώντας τώρα στα διαγράμματα φάσης του σχήματος Π.Β.1, διαπιστώνουμε πως όσο αυξάνεται η τάξη του φίλτρου αυξάνεται τόσο το stopripple, όσο και η αναστροφή φάσης που αυτό επιφέρει. Διαπιστώνουμε, επομένως, αυτό που και αμέσως πιο πάνω αναφέραμε, ότι δηλ. η αναστροφή φάσης που επιφέρει ένα φίλτρο εξαρτάται και είναι ανάλογη της τάξης (order) του φίλτρου.

Στο σχήμα Π.Β.2 (σελ. 26) δίδονται τέσσερα (4) High pass φίλτρα με τα διαγράμματα τους (κέρδος του φίλτρου συναρτήσει της συχνότητας και φάση του φίλτρου συναρτήσει της συχνότητας). Και αυτά σχεδιάστηκαν στο Matlab, έχουν, όμως, συχνότητα αποκοπής τα 250 Hz. Για αυτά δεν χρειάζεται να πούμε και πολλά πράγματα, μιας και ισχύουν τα όσα έχουμε αναφέρει για τα Low pass, παρά μόνο να αναφέρουμε πως τα διαγράμματα φάσης σε αυτή την περίπτωση είναι παντελώς όμοια με αυτά που δίδονται στα Low pass με μόνη διαφορά του ότι είναι ανεστραμμένα τόσο κατά τον οριζόντιο άξονα, όσο και κατά τον κάθετο άξονα. Αυτό είναι φυσικό να συμβαίνει, δεδομένου ότι στη συγκεκριμένη περίπτωση έχουμε να κάνουμε με high pass φίλτρα και δεδομένου πως πέραν αυτής της διαφοράς τα φίλτρα (low & high) σε όλες τις υπόλοιπες παραμέτρους σχεδίασης τους έχουν όμοιες τιμές.

Για να γίνουν κατανοητά όλα τα παραπάνω σε αυτό το σημείο του κειμένου θα ανοίξουμε μια παρένθεση και θα αναφέρουμε ένα παράδειγμα από τον «κόσμο» των μεγαφώνων. Οι αξιόλογοι, λοιπόν, κατασκευαστές ηχείων δεν είναι τυχαίο ότι χρησιμοποιούν στα κυκλώματα των crossover των μεγαφώνων τους όσο το δυνατόν λιγότερα ηλεκτρονικά στοιχεία (αντιστάσεις, πυκνωτές, πηνία) έτσι ώστε να διατηρούν τη διαδρομή του σήματος σε αυτά όσο το δυνατόν λιγότερο. Το κυριότερο, δεν είναι τυχαίο ότι χρησιμοποιούν πάντοτε ή σχεδόν πάντοτε φίλτρα πρώτης τάξης έτσι ώστε η αναστροφή φάσης που θα υποστεί το σήμα να είναι η μικρότερη δυνατή. Αυτός είναι και ο λόγος που τα ηχεία δύο δρόμων θεωρούνται περισσότερο «μουσικά» από αυτά των τριών ή τεσσάρων δρόμων, μιας και χρησιμοποιούν λιγότερα φίλτρα στα crossover τους για το διαχωρισμό του φάσματος του audio σήματος που θα οδηγήσει, τελικά, το κάθε μεγάφωνο. Επίσης, θα εκμεταλλευτούμε αυτή την παρένθεση για να ασχοληθούμε λίγο με μία διαρκή μέσα στο χρόνο παρερμηνεία των ανθρώπων που ασχολούνται γενικότερα με την παραγωγή ή την επεξεργασία ηχητικού υλικού. Αναφερόμαστε στο ζήτημα της



Σχ. Π.Β.3: (α) Ένα notch φίλτρο με κεντρική συχνότητα το 1kHz. (β) Η απόκριση φάσης του φίλτρου. (γ) Τα (α) και (β) μαζί στο ίδιο διάγραμμα. (δ) Το delay που αθροίζει το φίλτρο στο σήμα μετρημένο σε samples.

διαφοράς της Φάσης σε σχέση με την Πολικότητα. Η Πολικότητα αναφέρεται στη θετική ή στην αρνητική ημιπερίοδο ενός σήματος. Η Φάση αναφέρεται στα όσα έχουμε ήδη αναφέρει και είναι εξαρτώμενη από το χρόνο. Συνεπώς, όταν αναφερόμαστε στη αναστροφή Φάσης, ουσιαστικά αναφερόμαστε σε ένα χρονικό delay. Αν, όμως, αναστρέψουμε την πολικότητα ενός σήματος, κάνουμε απλά και μόνο αυτό, δεν το επηρεάζουμε χρονικά. Για παράδειγμα αν αναστρέψουμε τους ακροδέκτες ενός μεγαφώνου (φυσικά αναφερόμαστε σε ένα και μόνο μέγαςφωνο, χωρίς την ύπαρξη δεύτερου), τότε δεν επηρεάζουμε το πόσο γρήγορα θα πάλλεται ο κώνος του μεγαφώνου για να αναπαραγάγει τη δεδομένη συχνότητα, απλά εκεί που είχε θετική φορά τώρα έχει αρνητική και το αντίστροφο. Συνεπώς κατανοούμε ότι η φάση και η πολικότητα δεν είναι το ίδιο πράγμα. Ένα δεύτερο παράδειγμα μπορεί να βρεθεί στις κονσόλες μίξης των ηχοληπτών. Ο διακόπτης που αναφέρεται ως phase shift, δεν είναι σε καμία περίπτωση αυτό που δηλώνει, διότι οι διακόπτες αυτοί δεν επηρεάζουν χρονικά καθόλου το σήμα (στην περίπτωση που αυτό συμβαίνει η κονσόλα είναι για πέταμα), συνεπώς οι διακόπτες αυτοί δεν θα έπρεπε να ονομάζονται phase shift, αλλά polarity shift. Εδώ κλείνει η παρένθεση.

Στο διάγραμμα Π.Β.3 (α) παρουσιάζεται το κέρδος συνάρτησης της συχνότητας ενός notch EQ με κεντρική συχνότητα το 1 kHz. Το Π.Β.3 (β) παρουσιάζει την απόκριση

φάσης του φίλτρου. Το Π.Β.3 (γ) παρουσιάζει τα (α) & (β) παράλληλα. Τέλος το Π.Β.3 (δ) παρουσιάζει τη χρονική καθυστέρηση που θα αθροιστεί στο audio σήμα απο τη χρήση του συγκεκριμένου φίλτρου, λόγω της αναστροφής φάσης που επιφέρει (που είναι και το σημαντικότερο στοιχείο αυτού που μελετάμε). Απο τα διαγράμματα (β) και (γ) διαπιστώνουμε ότι το ποσοστό της αναστροφής φάσης γύρω απο τη συχνότητα του 1 kHz λαμβάνει τις υψηλότερες τιμές. Το διάγραμμα (δ) μας παρουσιάζει το delay που προκύπτει απο αυτό το equalizer, ο οριζόντιος άξονας είναι ο άξονας των συχνοτήτων και ο κάθετος άξονας μας παρουσιάζει τα samples της καθυστέρησης που προστίθενται στο σήμα. Ο τρόπος με τον οποίο εξάγεται το συγκεκριμένο γράφημα είναι ιδιαίτερα απλώς αρκεί να λάβουμε υπόψη μας ορισμένα βασικά σημεία απο τη θεωρία των ΕQ. Όπως έχουμε ήδη δει, ένα ψηφιακό φίλτρο μπορεί να εκφραστεί απο τη σχέση:

$$Y(n) = \frac{b_1x_1 + b_2x_{n-1} + \dots + b_{N+1}x_{n-N} - a_2y_{n-1} - \dots - a_{N+1}y_{n-N}}{a_1} \quad (\text{Π.Β.1})$$

όπου: a... & b... οι όροι που προσδιορίζουν τη μορφή του φίλτρου.

Η παραπάνω σχέση, όμως, εκφράζει το φίλτρο στο πεδίο του χρόνου. Στο πεδίο των συχνοτήτων το φίλτρο μπορεί να εκφραστεί απο τη σχέση:

$$H(e^{j\omega}) = \frac{b_1 + b_2e^{-j\omega} + \dots + b_{N+1}e^{-j\omega(N)}}{a_1 + a_2e^{-j\omega} + \dots + a_{N+1}e^{-j\omega(N)}} \quad (\text{Π.Β.2})$$

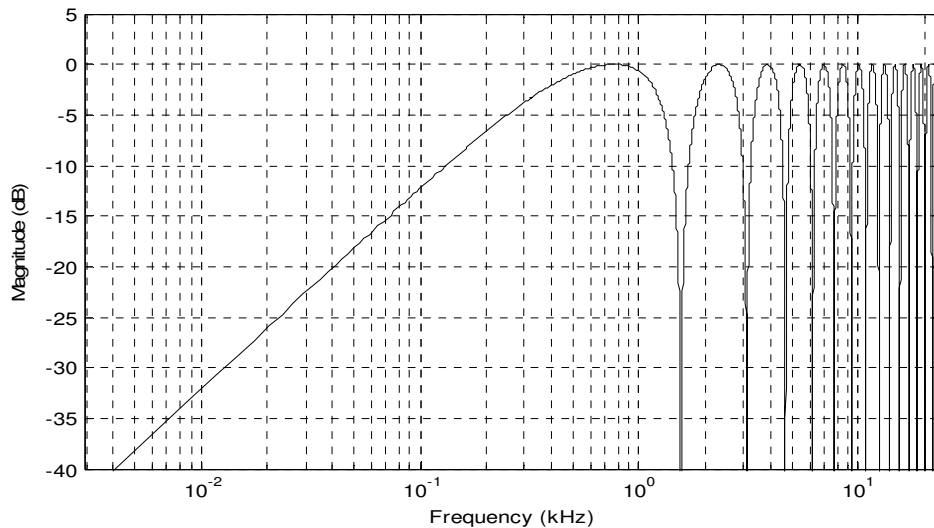
Συνεπώς, μπορούμε να λάβουμε το delay που προκύπτει απο το φίλτρο βρίσκοντας την πρώτη αρνητική παράγωγο της φασικής απόκρισης του φίλτρου, η οποία δίδεται απο τη σχέση:

$$\tau_g(\omega) = -\frac{d\mathcal{G}(\omega)}{d\omega} \quad (\text{Π.Β.3})$$

όπου: \mathcal{G} είναι η γωνία φάσης που προκύπτει απο την $H(e^{j\omega})$.

Παρατηρούμε, λοιπόν, ότι η μεγαλύτερη τιμή του delay υπεισέρχεται στη συχνότητα του 1 kHz και είναι περίπου 70 samples, πράγμα που στο πεδίο του χρόνου μεταφράζεται σε μία τιμή της τάξης του 1.6 msec για ηχητικό υλικό δειγματοληπτημένο στα 44100 Hz. Τι σημαίνει όμως αυτή η τιμή καθυστέρησης για το ηχητικό σήμα, και ποια είναι τα επιμέρους χαρακτηριστικά που του προσδίδει;

Το αποτέλεσμα που προκύπτει είναι το λεγόμενο comb filtering, δηλ. το αποτέλεσμα της άθροισης ενός σήματος με μία δεύτερη εκδοχή του η οποία, όμως, εμφανίζει μια κάποια (χρονική) διαφορά φάσης. Στο σχήμα Π.Β.4 που ακολουθεί στην επόμενη σελίδα παρουσιάζεται η μορφή του comb filter, καθώς και οι μαθηματικές σχέσεις απο τις οποίες προκύπτει.



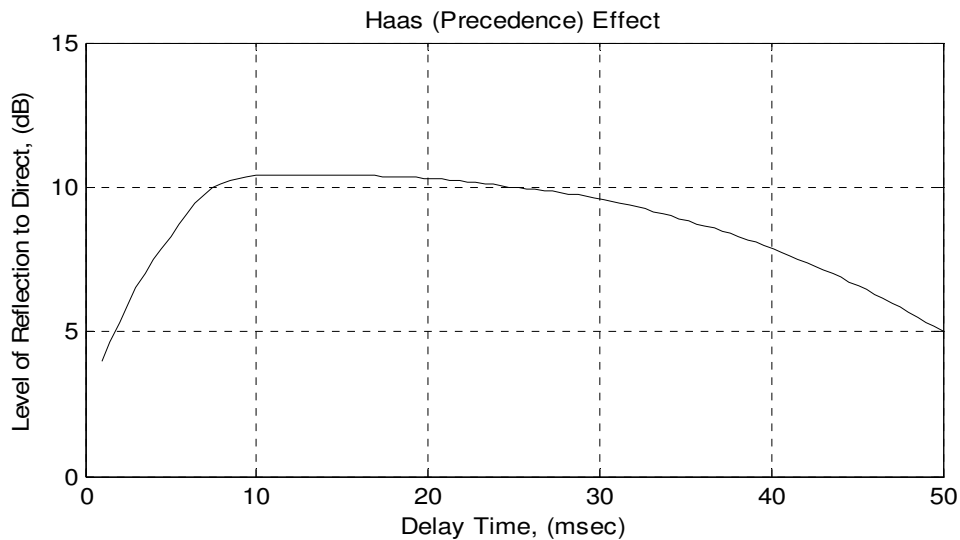
Σχ. Π.Β.4: Παρουσίαση του comb filter που προκύπτει για $\Delta t=1.6$ msec.

Απο όλα τα παραπάνω, λοιπόν, κατανοούμε ότι για τη δημιουργία του comb filtering απαιτούνται δύο σήματα, τα οποία δίδονται απο τις σχέσεις:

$$f_{co} = n \cdot \frac{1}{\Delta t}, \quad n = 1,2,3... \quad (\text{Π.Β.4})$$

$$f_{de} = \frac{2 \cdot n + 1}{2} \cdot \frac{1}{\Delta t}, \quad n = 0,1,2,3... \quad (\text{Π.Β.5})$$

Απο όσα έχουν αναφερθεί, θα μπορούσε κανείς να βγάλει το συμπέρασμα πως η χρήση των equalizer είναι καταστροφική για το ηχητικό υλικό. Τα πράγματα σε καμία περίπτωση δεν είναι έτσι. Το συγκεκριμένο notch EQ δεν επιλέχθηκε τυχαία. Όπως έχουμε αναφέρει πιο πάνω η αναστροφή φάσης του φίλτρου και κατά συνέπεια το delay που προκύπτει, εξαρτάται απο το bandwidth του φίλτρου καθώς και απο το ποσοστό ενίσχυσης ή εξασθένησης που επιφέρει στο σήμα. Συνεπώς ένα notch φίλτρο θα επιφέρει το μεγαλύτερο ποσοστό delay στο σήμα (στην περίπτωση μας 1.6 msec). Συνακόλουθα απο τη συχνότητα του 1 kHz, όπως πολύ εύκολα προκύπτει απο τη σχέση $T = \frac{1}{f}$, η περίοδος του, είναι ίση με 1 msec. Συνεπώς ήταν αναμενόμενο να προκύψει comb filter με τα χαρακτηριστικά που δίδονται στο σχήμα Π.Β.4. Συν τοις άλλοις, θα πρέπει να συμπεριλάβουμε στο σκεπτικό μας, ότι στο δικό μας γνωστικό πεδίο δεν έχουμε να κάνουμε ποτέ με καθαρούς τόνους και ιδιαίτερα στο mastering που το πεδίο εργασίας μας αφορά ολοκληρωμένες μίξεις. Παρά τα όσα, όμως, έχουμε αναφέρει, δεν έχουμε απαντήσει στο ερώτημα που αρχικά έχει τεθεί: ποια είναι η πραγματική συνθήκη που λαμβάνει χώρα μέσω της αναστροφής φάσης και κατά συνέπεια του delay, τι συνεπάγεται για το ηχητικό υλικό αυτή η συνθήκη και πως αντιλαμβάνεται όλα τα παραπάνω η ανθρώπινη ακοή;

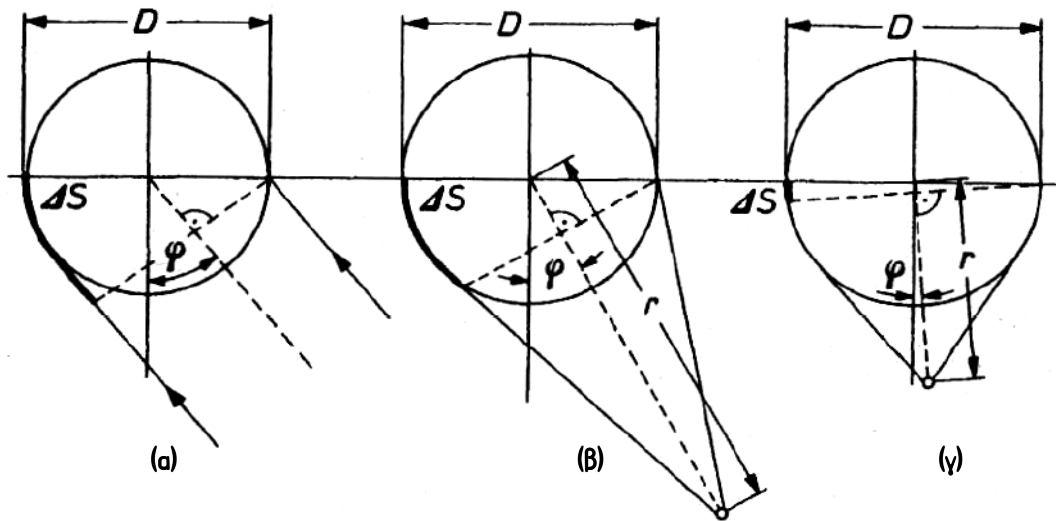


Σχ. Π.Β.5: Η καμπύλη που προέκυψε από τα πειράματα του Helmut Haas και η οποία φέρει το όνομα του (ή αλλιώς Precedence Effect).

Την απάντηση σε όλα τα παραπάνω έχει δώσει ήδη από το 1951 ο Helmut Haas. Διερευνώντας τον προσδιορισμό του ρόλου των ανακλάσεων στη διαμόρφωση του ακουστικού χαρακτήρα ενός κλειστού χώρου. Τα συμπεράσματα του, μπορούν να γενικευτούν και για τον ακουστικό χαρακτήρα και την πρόσληψη ενός οποιουδήποτε ηχητικού υλικού και αυτό γιατί τα πειράματα του δεν αφορούσαν τη μέτρηση χώρων, αλλά αντίθετα αφορούσαν τη διακριτική ικανότητα των υποκειμένων αναφορικά με τον απευθείας ήχο και τις μετέπειτα ανακλάσεις του. Το ζητούμενο σε αυτά τα πειράματα ήταν το πότε τα υποκείμενα αντιλαμβάνονταν την επανάληψη ενώ αυτή ηχούσε ταυτόχρονα με τον απευθείας ήχο αλλά από διαφορετική διεύθυνση - μεγάφωνο. Το σχήμα Π.Β.5 παρουσιάζει την καμπύλη που προέκυψε από αυτά τα πειράματα και η οποία είναι γνωστή ως καμπύλη Haas. Ο κατακόρυφος άξονας του σχήματος δηλώνει τον αριθμό των επιπλέον dB που απαιτούνται για την επανάληψη, ώστε αυτή να αποκτήσει διακριτό χαρακτήρα. Συνεπώς, γίνεται κατανοητό πως, η περιοχή του σχήματος μεταξύ της καμπύλης - όριο και του οριζώντιου άξονα (του χρόνου) καθορίζει τη λεγόμενη «ζώνη Haas».

Από το Precedence effect υποδηλώνονται με σαφήνεια δύο πολύ σημαντικά συνεπαγόμενα. Πρώτον, ότι το delay γίνεται διακριτό ως τέτοιο με τη μορφή επαναλήψεων αν η χρονική τιμή της διαφοράς του είναι μεγαλύτερη από 50 msec περίπου. Αν το $\Delta t < 50 \text{ msec}$ τότε δε γίνεται αντιληπτό το delay. Συνακόλουθα, αυτό που προκύπτει είναι μια κάποια μεγαλύτερη διάρκεια του πρωτογενούς ηχητικού σήματος. Επιπλέον εισπράττουμε κι ερμηνεύουμε μια επιπρόσθετη αίσθηση χώρου. Το ηχητικό υλικό δηλ. λαμβάνει ένα επιπλέον «άπλωμα» στο χώρο (εν προκειμένω στη στερεοφωνική εικόνα).

Το δεύτερο (συνεπαγόμενο) που προκύπτει είναι ότι στη ζώνη του Haas (δηλ. στο Precedence effect) υπάρχει ένα κάτω όριο. Ο λόγος είναι ότι καθώς μικραίνει η



Σχ. Π.Β.6: Απλοποιημένη σχηματική αναπαράσταση της Ι.Τ.Δ καθώς και του υπολογισμού της διαφοράς δρόμου Δs για (α) Παράλληλη πρόσπτωση και μακριά από την πηγή (συνεπώς το ηχητικό κύμα μπορεί να θεωρηθεί επίπεδο) και (β), (γ) για σημεία εκπομπής κοντά στο κεφάλι (σφαιρικό κύμα) και για διαφορετικές γωνίες πρόσπτωσης.

τιμή του delay, από το 1 msec περίπου και κάτω, μπαίνουμε σε μία περιοχή όπου ένας άλλος μηχανισμός αντίληψης επικρατεί. Αυτός της binaural αντίληψης.

Το σχήμα Π.Β.6 μας παρουσιάζει τη χρονική διαφορά που προκύπτει λόγω του πλάτους της ανθρώπινης κεφαλής. Το κεφάλι αποτελεί εμπόδιο στην πορεία του ήχου και συνεπώς προκύπτει το φαινόμενο σκίασης όταν ταιριάζει η χωροθέτηση πηγής - ακροατή. Επίσης στις περιπτώσεις που η πηγή είναι αρκετά κοντά στον ακροατή, η διαφορά μεταξύ των L/R διαδρομών επιφέρει μία ακόμη διαφορά level βάσει του νόμου του αντιστρόφου τετραγώνου. Τις χρονικές - φασικές διαφορές τις ονομάζουμε Interaural Time Delay (I.T.D) και τις διαφορές level τις ονομάζουμε Interaural Level Delay (I.L.D). Δεδομένου, όμως, πως οι διαφορές level δεν είναι ποσοτικά αξιόλογες για φυσιολογικές αποστάσεις από την πηγή δεν θα τις εξετάσουμε παραπέρα και θα ασχοληθούμε μόνο με την Ι.Τ.Δ η οποία στο παραπάνω σχήμα συμβολίζεται με Δs . Επομένως η Ι.Τ.Δ για τα παραδείγματα του σχήματος Π.Β.6 μπορεί να υπολογιστεί από τις σχέσεις που δίδονται παρακάτω.

Για το Π.Β.6.(α), για πρόσπτωση υπό γωνιά φ και με την πηγή μακριά από τον ακροατή για την οποία ισχύει $r \gg D/2$, όπου r η απόσταση από την πηγή και D η απόσταση μεταξύ των αυτιών του υποκειμένου, η διαφορά δρόμου Δs ή αλλιώς Ι.Τ.Δ προκύπτει:

$$\Delta s = \frac{D}{2}(\varphi + \sin \varphi) \quad (\text{Π.Β.6})$$

Για το Π.Β.6.(β), για πρόσπτωση υπό γωνιά φ και με την πηγή κοντά στον ακροατή για την οποία ισχύει $\varphi > \frac{D}{2 \cdot r}$ η διαφορά δρόμου Δs προκύπτει:

$$\Delta s' = D \left[\left(n + \frac{1}{2} \right) \cos(m) + \frac{1}{2} (\phi + m) - \sqrt{n^2 + n + \frac{1}{2} - \left(n + \frac{1}{2} \right) \sin(\varphi)} \right] \quad (\text{Π.Β.7})$$

όπου:

$$n = \frac{r - (D/2)}{D}$$

και όπου:

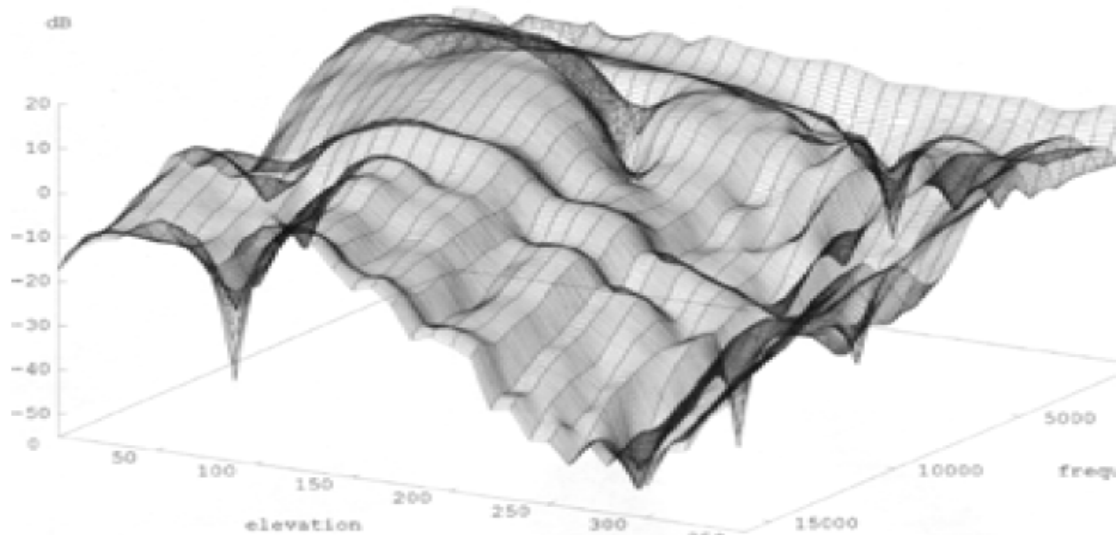
$$m = \arcsin \left(\frac{D}{2 \cdot r} \right) = \arcsin \left(\frac{1}{1 + 2 \cdot n} \right)$$

Για το Π.Β.6.(γ), για πρόσπτωση υπό γωνιά φ και με την πηγή κοντά στον ακροατή για την οποία ισχύει $\varphi \leq \frac{D}{2 \cdot r}$ η διαφορά δρόμου Δs προκύπτει:

$$\Delta s'' = D \cdot \varphi \quad (\text{Π.Β.8})$$

Απο τα παραπάνω προκύπτει ότι η I.T.D οφείλεται στη διαφορά δρόμου, η οποία με τη σειρά της οφείλεται στη γωνία φ . Η μέγιστη, λοιπόν, τιμή της I.T.D αντιστοιχεί στο πλάτος ενός ανθρώπινου κεφαλιού, ονομάζεται *Binaural delay* και είναι της τάξεως των 0.65 msec-650 msec. Συνεπώς προκύπτει ότι $0 \leq \text{I.T.D} \leq 650 \text{ msec}$. Λαμβάνοντας υπόψη τις ιδιαιτερότητες της κυματικής συμπεριφοράς, καθώς επίσης και τις διαφοροποιήσεις του ανθρώπινου κεφαλιού από άτομο σε άτομο προκύπτει ότι $0 \leq \text{I.T.D} \leq 1 \text{ msec}$. Με αυτό τον τρόπο προσδιορίζεται και το κάτω όριο της περιοχής του *Precedence effect* (Haas). Δηλ. το κάτω όριο της είναι το 1 msec περίπου, όπως φαίνεται και στο διάγραμμα του σχήματος Π.Β.5.

Αν, συνακόλουθα, λάβουμε υπόψη μας τα όσα έχουμε αναφέρει, και προσθέσουμε και το γεγονός του ότι όταν ο ήχος προσπίπτει στο ανθρώπινο αυτί με μια κάποια διεύθυνση πρόσπτωσης δημιουργούνται *comb filters* λόγω των μικροανακλάσεων στο πτερύγιο του αυτιού, τότε δημιουργούνται οι λεγόμενες *Head Related Transfer Functions - H.R.T.Fs*. Παράδειγμα των οποίων δίδεται στο σχήμα Π.Β.7 της επόμενης σελίδας.

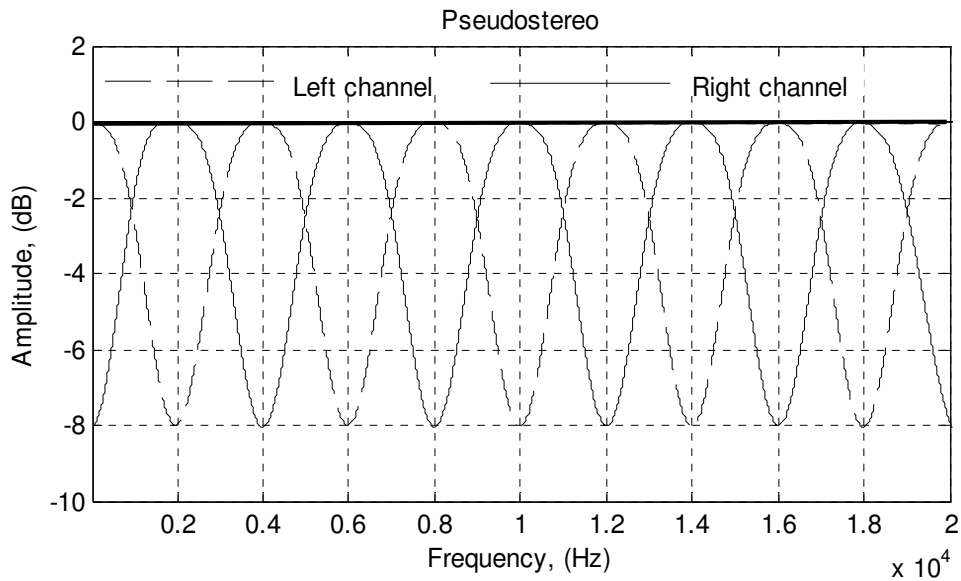


Σχ. Π.Β.7: Τυπική μορφή στις τρεις διαστάσεις των Head Related Transfer Functions – H.R.T.Fs.

Αυτό που προκύπτει από όλη αυτή τη έκθεση και ανάλυση των ψυχοακουστικών ιδιοτήτων των delays, είναι ότι στο «αρχικό» σήμα προστίθενται κάποια χωρικά στοιχεία. Τα χωρικά αυτά στοιχεία, όμως, διαφοροποιούνται μεταξύ τους και λειτουργούν σε διαφορετικές τιμές των delays. Δηλαδή, το απλό φυσικό φαινόμενο της επανάληψης έχει τη σημαντική ιδιότητα, σε συνδυασμό με την ταυτόχρονη λειτουργία με το «πρωτογενές» σήμα, να δημιουργεί κάποιο επιφαινόμενο του οποίου η ταυτότητα, το πως θα ηχεί δηλαδή, εξαρτάται αποκλειστικά από την τιμή του delay. Η ταξινόμηση των συμβαινόμενων (που όπως θα δούμε παρακάτω, τουλάχιστον για την περίπτωση των equalizer, δεν ταυτίζονται με το συμβαίνον) που προκύπτει είναι:

- i) $0 \cong 1 \text{ msec} \rightarrow$ I.T.D , Binaural αντίληψη.
- ii) $1 \cong 10 \text{ msec} \rightarrow$ Precedence effect, Comb filters.
- iii) $10 \cong 50 \text{ msec} \rightarrow$ Precedence effect.
- iv) $> 50 \text{ msec} \rightarrow$ Επαναλήψεις (Echoes).

Αν τώρα επανέλθουμε στο παράδειγμα του σχήματος Π.Β.3 για το notch φίλτρο και δούμε ότι το delay που προκύπτει είναι της τάξης του 1.6 msec θα διαπιστώσουμε ότι βρίσκεται εντός της περιοχής του Precedence effect και των comb filters. Όπως, όμως, έχουμε δει αυτή είναι η μέγιστη τιμή του delay (τηρουμένων των αναλογιών) που μπορούμε να λάβουμε από ένα equalizer και η οποία προκύπτει από τα χαρακτηριστικά του συγκεκριμένου φίλτρου όπως έχουμε καταδείξει. Οι συνήθεις τιμές του delay που προκύπτουν από ένα equalizer κυμαίνονται περίπου από τα 0.5 msec έως το 1 msec είναι δηλαδή εντός της περιοχής στην οποία επικρατεί ο μηχανισμός της Binaural αντίληψης.



Σχ. Π.Β.8: Σχηματική αναπαράσταση της τεχνικής του Pseudostereo με τη χρήση έντεκα (11) παραμετρικών equalizer με διαφορετική κεντρική συχνότητα, με το ίδιο bandwidth της τάξης των 600 Hz και με εξασθένιση του πλάτους του σήματος κατά 8 dB. Με τη διακεκομμένη γραμμή σημειώνονται τα φίλτρα που έχουν εφαρμοστεί στο αριστερό κανάλι της μίξης, ενώ με τη συνεχή αυτά που έχουν εφαρμοστεί στο δεξί.

Γνωρίζοντας τα όσα έχουμε αναφέρει (και προφανώς ακόμη περισσότερα), εμπνευσμένοι ηχολήπτες και τεχνικοί mastering δημιούργησαν την τεχνική του pseudostereo χρησιμοποιώντας μια σειρά από παραμετρικά equalizer. Η σχηματική αναπαράσταση της τεχνικής αυτής, με τη χρήση έντεκα (11) φίλτρων δίδεται στο σχήμα Π.Β.8. Η τεχνική του pseudostereo είναι πολύ απλή στη σύλληψη της, συμπεριλαμβάνει όμως όλη τη ψυχοακουστική ανάλυση που έχουμε παραθέσει.

Γνωρίζοντας πως κάθε EQ δημιουργεί ένα ποσοστό αναστροφής φάσης, καθώς επίσης πως λόγω αυτού προσθέτει και ένα delay στο σήμα, θα μπορούσαμε να χρησιμοποιήσουμε μια σειρά από παραμετρικά equalizer τα οποία θα επενεργούν για διαφορετικές συχνотικές περιοχές του audio φάσματος και για διαφορετικό κανάλι (L/R) το κάθε ένα από αυτά. Το αποτέλεσμα που θα προέκυπτε λόγω των αναστροφών φάσης, των συνεπαγόμενων comb filters και του delay που προστίθενται στο σήμα θα ήταν η δημιουργία της ψευδαίσθησης μιας πιο ευρείας στερεοφωνικής εικόνας.

Το pseudostereo χρησιμοποιείται από τους mastering ηχολήπτες είτε για την απόκτηση μιας στερεοφωνικής αίσθησης σε παλιές mono ηχογραφήσεις, είτε για τη διαπλάτυνση στο οριζόντιο επίπεδο της στερεοφωνικής εικόνας μιας δικάναλης (L/R) μίξης.

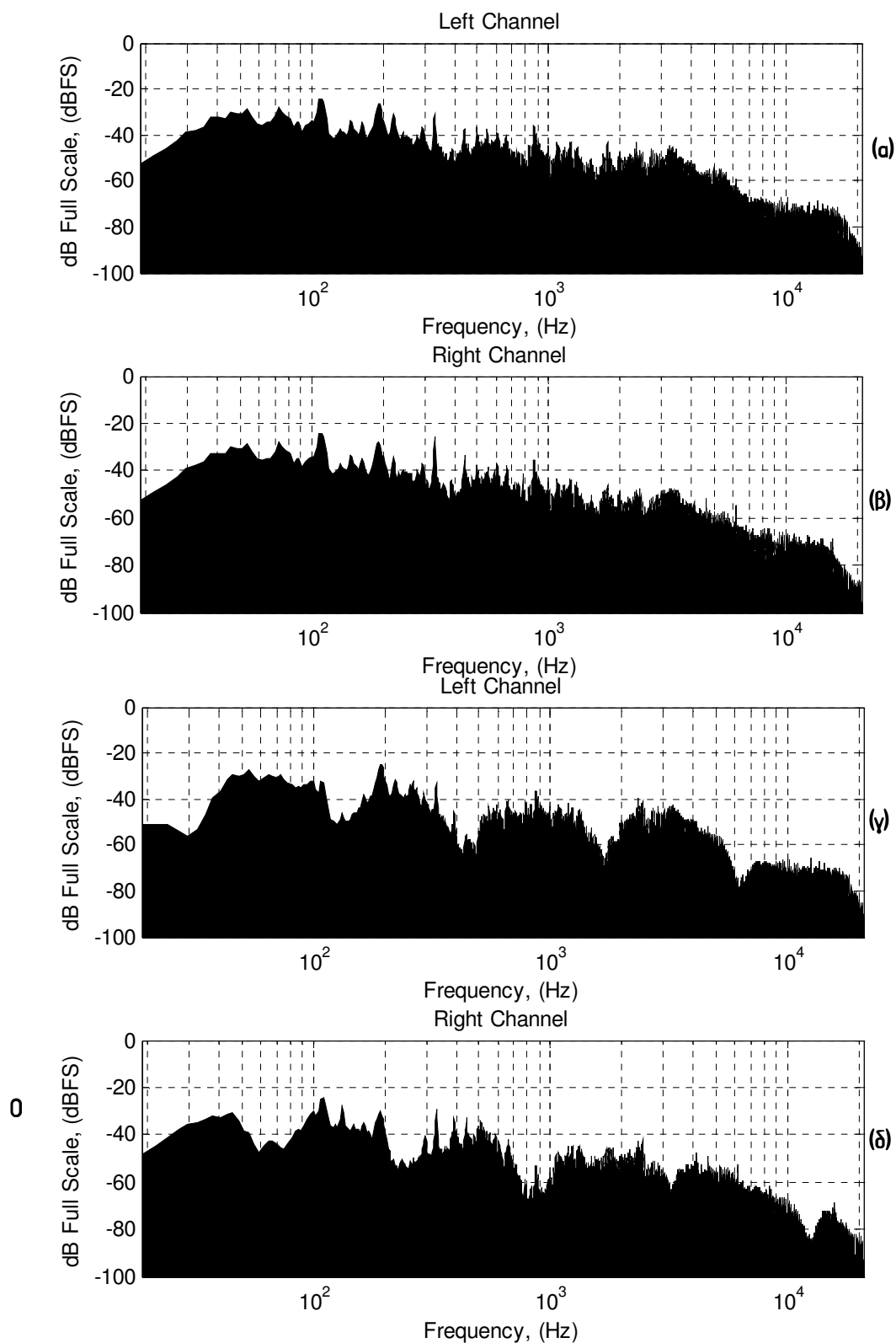
Η τεχνική αυτή (όπως και οι περισσότερες άλλωστε) έχει τα πλεονεκτήματα και τα μειονεκτήματα της. Ορισμένα από τα πλεονεκτήματα της είναι η εξαιρετική μονοφωνική συμβατότητα (mono compatibility) που προσφέρει, καθώς και το γεγονός του ότι μας δίνει τρεις παραμέτρους ρύθμισης (κεντρική συχνότητα - f_c , εύρος ζώνης - BW/Q και το ποσοστό ενίσχυσης και εξασθένισης - gain) για κάθε ένα φίλτρο

ξεχωριστά, ώστε να ορίζουμε την τιμή της αναστροφής φάσης και κατά συνέπεια το delay, πράγμα που σημαίνει πως ουσιαστικά μπορούμε να δημιουργήσουμε άπειρους συνδυασμούς συναρτήσεων του αριθμού των φίλτρων που χρησιμοποιούμε και αναφορικά με το συχνοτικό φάσμα το ηχητικού υλικού.

Στα μειονεκτήματα, τώρα, της τεχνικής αυτής συγκαταλέγονται το γεγονός του ότι απαιτεί πολύ μεγάλο χρόνο ενασχόλησης για την επίτευξη των επιθυμητών ηχητικών χαρακτηριστικών, η εμπειρία που απαιτείται από το χρήστη, το φτωχό ηχητικό αποτέλεσμα σε αρκετές περιπτώσεις και τέλος (το σημαντικότερο) οι αλλοιώσεις που επιφέρει στο φάσμα. Η τεχνική αυτή, και για τα παραπάνω μειονεκτήματα της, έχει αποκτήσει στο πέρασμα του χρόνου μια κακή φήμη. Παρόλα αυτά, με τη χρήση των κατάλληλων ρυθμίσεων, των κατάλληλων hardware ή software συσκευών και με «τεντωμένα» αυτιά μπορούμε να επιτύχουμε αξιοπρεπή αποτελέσματα. Σαφώς και υπάρχουν διάφορες άλλες τεχνικές, που κάνουν χρήση του ίδιου θεωρητικού υπόβαθρου, χρησιμοποιούν όμως διαφορετικού τύπου επεξεργασίες. Η τεχνική αυτή αναφέρεται διότι αποτελεί μία εξαιρετικά ευφυή επινοήση και διότι μας επιτρέπει να κατανοήσουμε σε βάθος τη λειτουργία των equalizer, καθώς και των συνεπειών τους (καλών και κακών) στο ηχητικό σήμα. Στο σχήμα Π.Β.9 της επόμενης σελίδας δίδονται τα φάσματα των Left και Right καναλιών ενός μουσικού κομματιού πριν και μετά από την εφαρμογή του pseudostereo. Για την εφαρμογή της τεχνικής χρησιμοποιήθηκε το plug - in της Waves: Q10 Paragraphic EQ.

Από τα διαγράμματα του σχήματος Π.Β.9 γίνεται σαφές τι είδους αλλοιώσεις παράγει αυτού του τύπου η επεξεργασία στο ηχητικό υλικό και γιατί συνήθως αποφεύγεται. Πιο συγκεκριμένα παρατηρούμε τα deep's και τα peaks που έχει επιφέρει η τεχνική αυτή στο ηχητικό υλικό. Γίνεται σαφές με αυτό τον τρόπο πως δημιουργούνται comb filters στο φάσμα και ακυρωτικές συμβολές· μόνο που στη συγκεκριμένη περίπτωση δε γίνονται αντιληπτά αυτά τα φαινόμενα. Αυτό που γίνεται αντιληπτό από τον ακροατή είναι ένα άπλωμα του ηχητικού υλικού στο χώρο, καθώς και η διεύρυνση της στερεοφωνικής εικόνας. Τέλος φαίνονται και οι διαφορετικές συχνοτικές περιοχές, για το κάθε κανάλι, στις οποίες έχουν εφαρμοστεί τα equalizer.

Κλείνοντας αυτό (το εκτενές ομολογουμένως) παράρτημα, θα πρέπει να πούμε ότι δε θα πρέπει να φοβάται κανείς τη χρήση των equalizer. Τα equalizer είναι από τα πιο χρήσιμα εργαλεία στα χέρια (και τα αυτιά) του ηχολήπτη και του τεχνικού mastering. Η αναστροφή φάσης και το delay που επιφέρουν είναι, για τις συνήθεις περιπτώσεις, άνευ σημασίας και σίγουρα τα πλεονεκτήματα που προσφέρει είναι πολύ περισσότερα από τα μειονεκτήματα του. Χρειάζεται εξάσκηση, καλλιέργεια της ακοής και εμπειρία, αλλά ανταποδίδει στο μέγιστο το χρόνο που του έχει αφιερωθεί. Επίσης, όπως καταδείξαμε με όλη την προηγηθείσα αναλυτική, τα EQ στα χέρια (και τα αυτιά) εμπνευσμένων και ευφυών ανθρώπων μπορούν να προεκτείνουν το λόγω της ύπαρξής τους και το πεδίο εφαρμογής τους.



Σχ. Π.Β.9: Τα συχνотικά φάσματα ενός ηχητικού κομματιού πριν και μετά την εφαρμογή της τεχνικής του pseudostereo. (α) Το αριστερό κανάλι ανεπεξεργαστο. (β) Το δεξί κανάλι ανεπεξεργαστο. (γ) Το αριστερό κανάλι αφού έχει υποστεί την επεξεργασία. (δ) Το δεξί κανάλι αφού έχει υποστεί την επεξεργασία.

Παράρτημα Γ

Σε αυτό το παράρτημα δίδεται δύο κώδικες του matlab που βοηθούν πολύ στην κατανόηση της λειτουργίας των limiter.

Κώδικας 'A':

```
% This program computes and plots of a sinusoid
% with amplitudes 1, 2 and 4. If desired the sinusoid can
be
% clipped to simulate the effect of a limiting device.
% The frequency is 1 Hz and the time duration is 10
seconds.
% The sample interval is 0.1 seconds. The code is not
efficient
% but is meant to illustrate MATLAB statements.

clear all % clear all variables from workspace
delt=0.01; % set sampling time interval
F0=1; % set frequency
t=[0:delt:10]'; % compute time samples 0, 0.01,
0.02,...,10
A=[1 2 4]'; % set amplitude
clip='yes'; % set option to clip

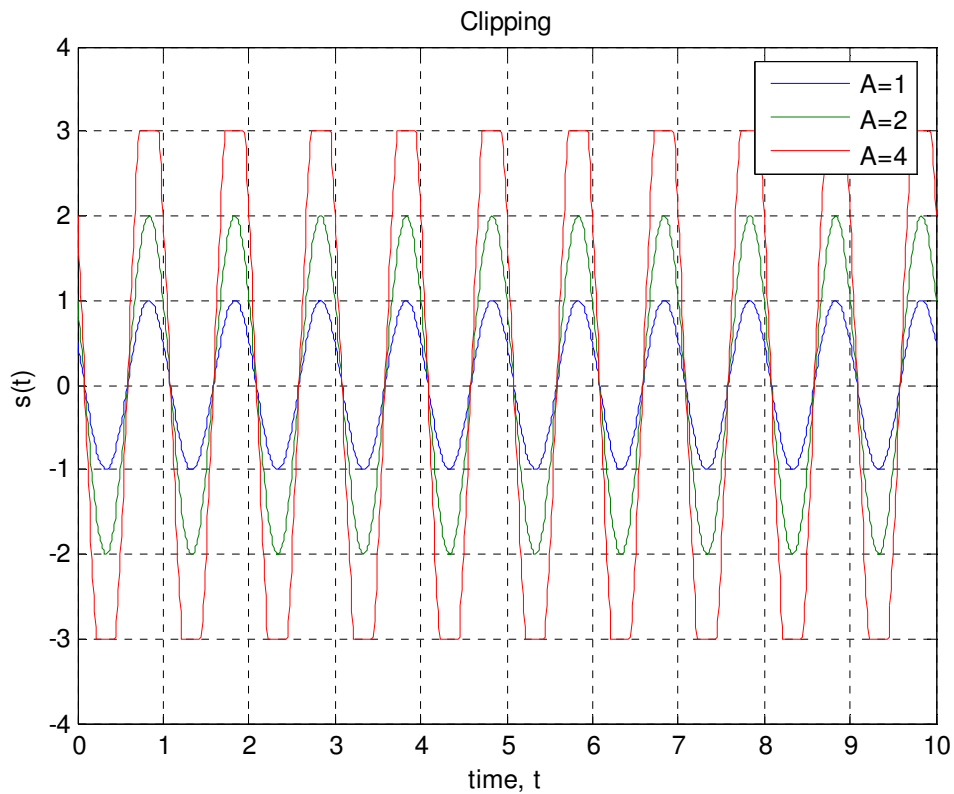
for i=1:length(A) % begin computation of sinusoid
samples
    s(:,i)=A(i)*cos(2*pi*F0*t+pi/3); % note that samples
for sinusoid
                                % are computed all at
once and
                                % stored as columns
in a matrix
    if clip=='yes' % determine if clipping desired
        for k=1:length(s(:,i)) % note that number of
samples given as
                                % dimension of column
using length command
                if s(k,i)>3 % check to see if sinusoid sample
exceeds 3
                    s(k,i)=3; % if yes then clip
elseif s(k,i)<-3 % check to see if sinusoid
is less
                    s(k,i)=-3; % than -3 if yes, then
clip
                end
            end
        end
    end
end
end
```

```

figure % open up a new figure window
plot(t,s(:,1),t,s(:,2),t,s(:,3)) % plot sinusoid
samples versus time...
...samples

grid % add grid to plot
xlabel('time, t') % label x-axis
ylabel('s(t)') % label y-axis
title('Clipping') % puts a title on the plot
axis([0 10 -4 4]) % set up axes using axis([xmin xmax
ymin ymax])
legend('A=1','A=2','A=4') % display a legend to
distinguish...
...different sinusoids

```



Σχ. Π.Γ.1: Σχηματική αναπαράσταση του φαινομένου του φαλιδισμού ενός ημιτονοειδούς σήματος το οποίο έχει ξεπεράσει ένα δεδομένο πλάτος.

Κώδικας Β:

```
% clear all
n=0:0.0001:1/10;
A=[2];
f1=20;
f2=40;
f3=63;
f4=125;
f5=250;
x1=2*cos(2*pi*f1*n+pi/3);
x2=2*cos(2*pi*f2*n+pi/4);
x3=2*cos(2*pi*f3*n+pi);
x4=2*cos(2*pi*f4*n+pi/6);
x5=2*cos(2*pi*f5*n+pi/16);
x=x1+x2+x3+x4+x5;

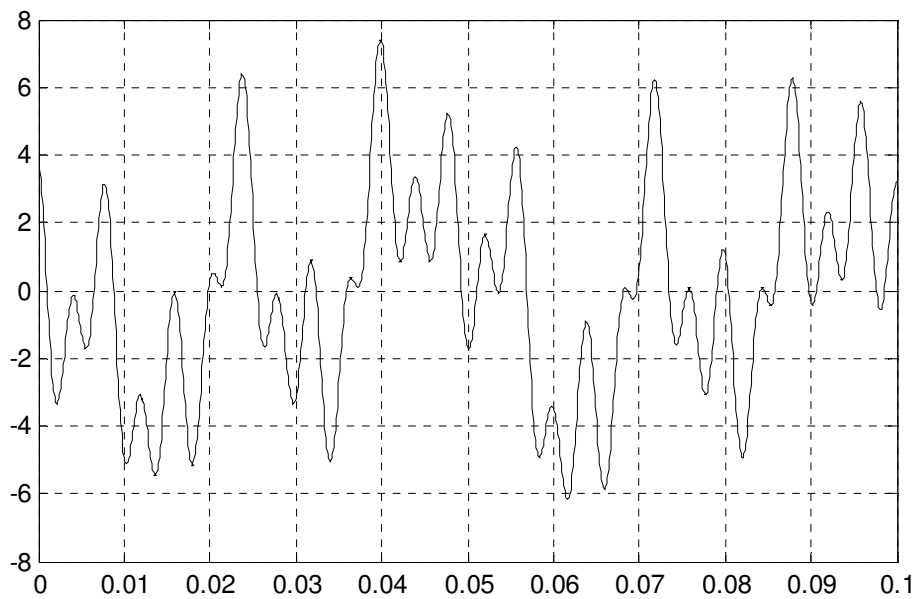
clip='yes'; % set option to clip

for i=1:length(A) % begin computation of sinusoid
samples
    s(:,i)=x; % note that samples for sinusoid
                % are computed all at once and
                % stored as columns in a matrix
    if clip=='yes' % determine if clipping desired
        for k=1:length(s(:,i)) % note that number of
samples given as
                                % dimension of column
using length command
                if s(k,i)>3 % check to see if sinusoid sample
exceeds 3
                    s(k,i)=3; % if yes then clip
                elseif s(k,i)<-3 % check to see if sinusoid
is less
                    s(k,i)=-3; % than -3 if yes, then
clip
                end
            end
        end
    end
end

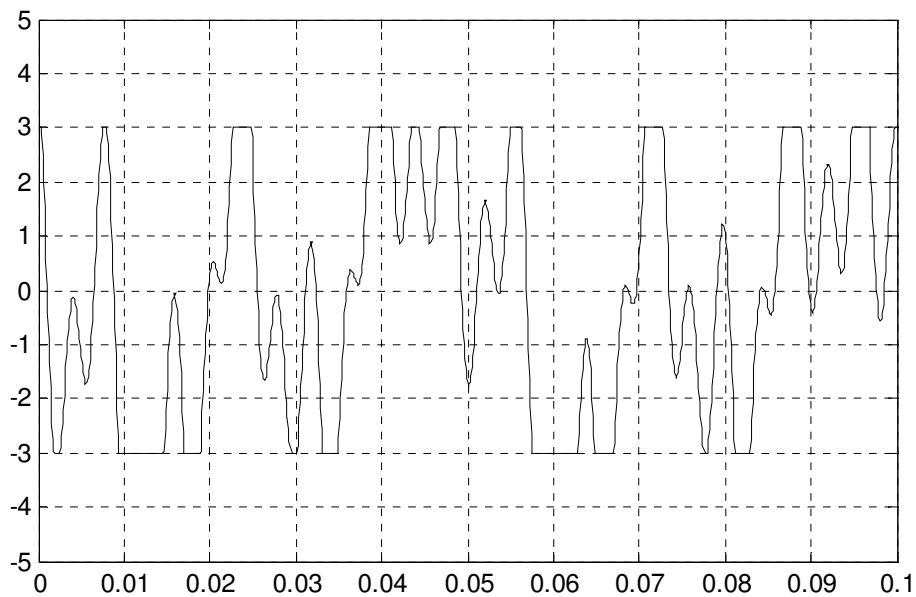
figure (1)
plot(n,x1,n,x2,n,x3,n,x4,n,x5)
grid on
figure (2)
plot(n,x)
grid on
figure (3)
plot(n,s(:,i))
grid on
```



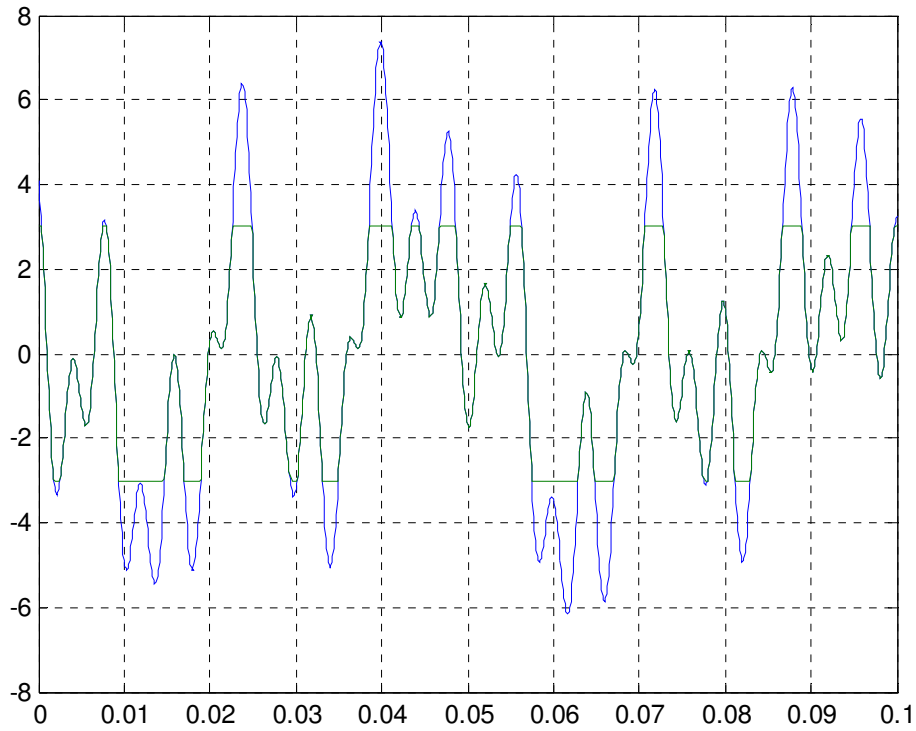
```
figure (4)
subplot (3,1,1)
plot (n,x)
grid on
subplot (3,1,2)
plot (n,s(:,i))
grid on
subplot (3,1,3)
plot (n,x,n,s(:,i))
grid on
figure (5)
plot (n,x,n,s(:,i))
grid on
```



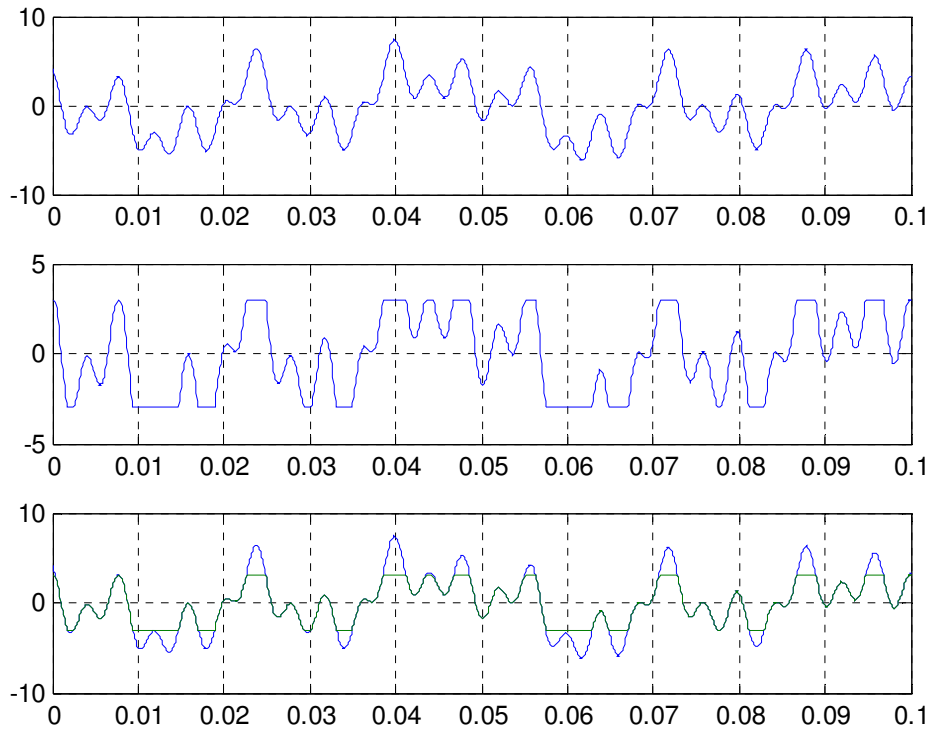
Σχ. Π.Γ.2: Ένα ημιτονοειδές σήμα το οποίο αποτελεί το άθροισμα των συχνοτήτων των 20, 40, 63, 125 και 250 Hz.



Σχ. Π.Γ.3: Το φαινόμενο του ψαλιδισμού, όμοιο με αυτό που θα προκαλούσε ένας limiter, στο σήμα του σχήματος Π.Γ.2



Σχ. Π.Γ.4: Το ψαλιδισμένο και το μη ψαλιδισμένο σήμα σε παράθεση.



Σχ. Π.Γ.5: Τα δύο σήματα σε μία συνολική αναλυτική γραφική παράθεση.

Παράρτημα 'Δ:

Η μέση ισχύς, υπολογιζόμενη σε ένα χρονικό διάστημα $t_2 - t_1$, δίδεται από τη σχέση:

$$\bar{P} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} P(t) dt \quad (\text{Π.Δ.1})$$

Για να παρασταθεί σωστά η μέση τιμή, η διάρκεια $t_2 - t_1$ πρέπει να είναι αρκετά μεγαλύτερη από τις χαρακτηριστικές χρονικές μεταβολές του σήματος. Εάν το σήμα είναι περιοδικό τότε το $t_2 - t_1$ πρέπει να είναι μεγάλο σε σύγκριση με την περίοδο ή διαφορετικά θα πρέπει να ισούται με έναν ακέραιο αριθμό περιόδων. Για ένα περιοδικό σήμα η καλύτερη επιλογή για να βρούμε τη μέση τιμή είναι να πάρουμε μόνο μία περίοδο. Η ενεργός τιμή του σήματος (RMS) ορίζεται σαν μία σταθερή αριθμητική τιμή, η οποία θα ισούται με τη μέση τιμή της ισχύος \bar{P} του σήματος. Με άλλα λόγια ορίζεται ως:

$$\frac{x^2(t)}{R} = \bar{P} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} P(t) dt \quad (\text{Π.Δ.2})$$

από όπου προκύπτει:

$$x_{RMS} = \sqrt{\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} P(t) dt} \quad (\text{Π.Δ.3})$$

Η εξίσωση (Π.Δ.3) δίνει τον ορισμό της ενεργού τιμής του x με λεπτομέρεια. Οι πράξεις γίνονται με αντίθετη σειρά από τον τρόπο με τον οποίο αναφέρονται στην ενεργό τιμή στα αγγλικά, δηλ. (RMS), από τα αρχικά Root Mean Square. Δηλαδή πρώτα τετραγωνίζουμε το σήμα, στη συνέχεια βρίσκουμε τη μέση τιμή ως προς το χρόνο και τέλος υπολογίζουμε την τετραγωνική του ρίζα.

Συχνά απαιτείται να εξάγουμε μέσες τιμές γινομένων από ποσότητες οι οποίες μεταβάλλονται αρμονικά και εκφράζονται με μιγαδική εκθετική μορφή (φάσορες). Αυτό μπορεί να γίνει παίρνοντας πρώτα τα πραγματικά τους μέρη και ολοκληρώνοντας τα σε μία περίοδο. Υπάρχει, παρόλα αυτά, ένας πολύ πιο εύκολος τρόπος μέσω του τύπου που ακολουθεί. Αν έχουμε δύο αρμονικές ποσότητες που παριστάνονται ως: $x(t) = \tilde{X} e^{-j\omega t}$ και $y(t) = \tilde{Y} e^{-j\omega t}$, τότε η μέση τιμή του γινομένου των ποσοτήτων αυτών δίδεται από τη σχέση:

$$\overline{x(t) y(t)} = \frac{1}{T} \int_0^T x(t) y(t) dt = \frac{1}{2} \text{Re} \{ \tilde{X} \tilde{Y}^* \} \quad (\text{Π.Δ.4})$$

Στην περίπτωση τώρα , όπως παραπάνω με τη μέση τιμή, όπου έχουμε μέση τιμή τετραγώνου, η σχέση (Π.Δ.4) εκφράζεται ως:

$$\overline{x(t)^2} = \frac{1}{T} \int_0^T x(t) x(t) dt = \frac{1}{2} \operatorname{Re} \{ \tilde{X} \tilde{X}^* \} = \frac{1}{2} |\tilde{X}|^2 \quad (\text{Π.Δ.5})$$

Συνεχίζοντας, ο συντελεστής κορυφής (crest factor) ορίζεται με το σύμβολο C , και είναι ο λόγος της ακρότατης τιμής ενός σήματος, προς την ενεργό τιμή του:

$$C = \frac{\operatorname{Max}(|x(t)|)}{x_{RMS}} \quad (\text{Π.Δ.6})$$

Η απόλυτη τιμή σημαίνει ότι μπορεί να χρησιμοποιηθεί είτε το θετικό μέγιστο, είτε το αρνητικό, ανάλογα με το ποιο απο τα δύο είναι το μεγαλύτερο. Σε κάθε περίπτωση, ο συντελεστής κορυφής είναι ένα θετικό νούμερο. Εάν το σήμα είναι τυχαίο, υπάρχει αβεβαιότητα για τις ακρότατες τιμές του και έτσι ο συντελεστής κορυφής δεν καθορίζεται εύκολα. Για ένα περιοδικό σήμα το $\operatorname{Max}(|x(t)|)$ είναι η μέγιστη απόλυτη τιμή του για τη διάρκεια μιας περιόδου.

Ο συντελεστής κορυφής παρουσιάζει ενδιαφέρον όταν ένα σήμα με μεγάλα μέγιστα ή ελάχιστα είναι κάτι που θέλουμε να αποφύγουμε κυρίως λόγω των ορίων του συστήματος που λαμβάνει ή που επεξεργάζεται ή και που αποθηκεύεται το σήμα ή ακόμη και για αισθητικούς λόγους (π.χ. το compression στο audio mastering).

Τέλος η peak to peak τιμή ενός σήματος ορίζεται ως το άθροισμα της τιμής του πλάτους ενός μέγιστου με αυτή με την τιμή του πλάτους ενός ελαχίστου, για περιοδικά σήματα και για διάρκεια μιας περιόδου. Είναι, και αυτή, πάντοτε θετικός αριθμός. Η σχέση που μας δίνει την peak to peak τιμή είναι ιδιαιτέρως απλή:

$$P_{oP} = |\operatorname{Max} x(t)| + |\operatorname{Min} x(t)| \quad (\text{Π.Δ.7})$$

Παράρτημα Έ:

Κώδικας Α:

```
fs=1000; %Συχνότητα δειγματοληψίας: 1000 Hz.
Ts=1/fs; %Βρίσκουμε την περίοδο δειγματοληψίας.
n=0:Ts:1; %Δημιουργούμε το χρονικό παράθυρο.
x1=cos(2*pi*3*n);
x2=cos(2*pi*7*n);
x3=cos(2*pi*13*n);
x=x1+x2+x3;
figure(1)

subplot(411)
plot(n,x1);
xlabel('Time')
ylabel('Amp')
subplot(412)
plot(n,x2);
xlabel('Time')
ylabel('Amp')
subplot(413)
plot(n,x3);
xlabel('Time')
ylabel('Amp')
subplot(414)
plot(n,x);
xlabel('Time')
ylabel('Amp')

for i=1:length(x)
    greconstructed(i)=sum(x.*sinc(fs*(n(i)-n)));
end

figure(2)
subplot(211)
plot(n,x);
xlabel('Time')
ylabel('Amp')
subplot(212)
plot(n,greconstructed);
xlabel('Time')
ylabel('Amp')
```

%Η εντολή αυτή μας μεταφέρει στο...
...κατάλληλο παράθυρο σχημάτων

Κώδικας Β:

```
fs=1000; %Συχνότητα δειγματοληψίας: 1000 Hz.
Ts=1/fs; %Βρίσκουμε την περίοδο δειγματοληψίας.
n=0:Ts:1; %Δημιουργούμε το χρονικό παράθυρο.

x1=cos(2*pi*3*n);
x2=cos(2*pi*7*n);
x3=cos(2*pi*13*n);
x=x1+x2+x3;
```

```

for i=1:length(x)
    greconstructed1(i)=sum(x.*sinc(fs*(n(i)-n)));
end

fs1=15; %Συχνότητα δειγματοληψίας 15 HZ.
Ts1=1/fs1; %Βρίσκουμε την περίοδο δειγματοληψίας.
n1=0:Ts1:2; %Δημιουργούμε το χρονικό παράθυρο.

x11=cos(2*pi*3*n1);
x12=cos(2*pi*7*n1);
x13=cos(2*pi*13*n1);
x1=x11+x12+x13;

for i=1:length(x)
    greconstructed2(i)=sum(x1.*sinc(fs1*(n(i)-n1)));
end

subplot(311)
plot(n,x);
xlabel('Time')
ylabel('Amp')
subplot(312)
plot(n,greconstructed1);
xlabel('Time')
ylabel('Amp')
subplot(313)
plot(n,greconstructed2);
xlabel('Time')
ylabel('Amp')

```

Κώδικας Γ:

```

fs=1000; %Συχνότητα δειγματοληψίας: 1000 Hz.
Ts=1/fs; %Βρίσκουμε την περίοδο δειγματοληψίας.
n=0:Ts:1; %Δημιουργούμε το χρονικό παράθυρο.

x1=cos(2*pi*3*n);
x2=cos(2*pi*7*n);
x3=cos(2*pi*13*n);
x=x1+x2+x3;

for i=1:length(x)
    greconstructed1(i)=sum(x.*sinc(fs*(n(i)-n)));
end

fs1=10; %Συχνότητα δειγματοληψίας 15 HZ.
Ts1=1/fs1; %Βρίσκουμε την περίοδο δειγματοληψίας.
n1=0:Ts1:2; %Δημιουργούμε το χρονικό παράθυρο.

x11=cos(2*pi*3*n1);
x12=cos(2*pi*7*n1);
x13=cos(2*pi*13*n1);
x1=x11+x12+x13;

for i=1:length(x)
    greconstructed2(i)=sum(x1.*sinc(fs1*(n(i)-n1)));
end

```

```

subplot(311)
plot(n,x);
xlabel('Time')
ylabel('Amp')
subplot(312)
plot(n,greconstructed1);
xlabel('Time')
ylabel('Amp')
subplot(313)
plot(n,greconstructed2);
xlabel('Time')
ylabel('Amp')

```

Κώδικας Δ:

```

N=1000; %Αριθμός samples στο σήμα.
fs=2000; %Συχνότητα δειγματοληψίας.
nyquist=fs/2; %Συχνότητα Nyquist.
t=[1:N]/fs; %Time axis of length N at sampling...
           ...frequency fs.

y1=cos(2*pi*300*t);
y2=cos(2*pi*700*t);
y3=cos(2*pi*1400*t);
y=y1+y2+y3;

X=fft(y,N); %Convert to frequency domain by FFT.
nfreq=floor(N/2)+1; %Number of points in frequency domain.
db=20*log10(abs(X(1:nfreq))); %Convert to dB.

frequency=linspace(0,nyquist,nfreq);

plot(frequency,db)
axis([0,1000,-300,100])
xlabel('Frequency, (Hz)')
ylabel('Magnitude, (dB)')

```


Παράρτημα 'ΣΤ:

Όταν κάνουμε κβαντισμό του σήματός μας, αντιστοιχίζουμε την τιμή που αυτό έχει μια συγκεκριμένη χρονική στιγμή, σε ένα από τα επίπεδα κβαντισμού. Η κβαντοποίηση του σήματος, όμως, εισάγει θόρυβο σε αυτό ο οποίος καλείται θόρυβος κβαντοποίησης (Quantization noise). Τα σφάλματα που προκύπτουν από την κβαντοποίηση στην ουσία είναι η διαφορά του αρχικού μας σήματος με αυτό που προέκυψε έπειτα από την κβαντοποίηση.

Γνωρίζουμε, επίσης, πως ο θόρυβος κβαντοποίησης δεν είναι τυχαίος, αλλά εξαρτάται από τη συχνότητα του σήματος μας και από τη συχνότητα δειγματοληψίας. Όταν η συχνότητα Nyquist (f_s) είναι πολλαπλάσιο της συχνότητας του σήματος, τότε προκαλείται αρμονική παραμόρφωση με συνιστώσες οι οποίες είναι ακέραια πολλαπλάσια της συχνότητας του σήματος. Όπου η συχνότητα Nyquist ή αλλιώς συχνότητα δειγματοληψίας δίδεται από τη σχέση:

$$f_s = 2 \cdot f_m \quad (\text{Π.ΣΤ.1})$$

όπου f_s : η συχνότητα δειγματοληψίας.

και f_m : η υψηλότερη συχνότητα που περιέχεται στο σήμα μας.

Τέλος, απότοκο το οποίο εκπορεύεται από όλα όσα έχουν αναφερθεί μέχρι τώρα, είναι η εμφάνιση συχνοτήτων ακόμη και πέρα από το όριο Nyquist, οπότε θα εμφανίζεται και το φαινόμενο του aliasing. Οι συχνότητες που προστίθενται με αυτό τον τρόπο στο σήμα μας παράγουν έναν «περίεργο» ήχο ο οποίος στην αρχή των τεχνικών ήχου καλείται «bird singing» ή «birdies». Το aliasing προσθέτει συχνότητες στο σήμα μας σύμφωνα με τη σχέση:

$$f_a = \pm N \cdot f_s \pm f_j \quad (\text{Π.ΣΤ.2})$$

όπου $N = \pm 1, \pm 2, \pm 3, \dots$, κ.ο.κ.

f_a : οι νέες συχνότητες που προστίθενται στο σήμα.

f_s : η συχνότητα δειγματοληψίας.

και f_j : η συχνότητα που έχει τιμή μεγαλύτερη από το μισό της συχνότητας δειγματοληψίας.

Για την αντιμετώπιση των παραπάνω φαινομένων γίνεται η πρόσθεση ενός τυχαίου θορύβου (dither) στο σήμα μας και έτσι ο θόρυβος κβαντοποίησης που έχει προκύψει

με τη μορφή αρμονικής παραμόρφωσης θα ξαναγίνει τυχαίος. Επίσης, με αυτό τον τρόπο, θα απαλειφθεί το φαινόμενο του aliasing.

Τελειώνοντας και για την καλύτερη κατανόηση όλων των παραπάνω, σχεδιάσαμε στο Matlab έναν κώδικα στον οποίο εφαρμόζονται όλα όσα έχουν αναφερθεί. Ο κώδικας, καθώς και τα διαγράμματα (με περαιτέρω επεξηγήσεις σε αυτά) που προκύπτουν δίδονται παρακάτω (σελ. 235 και πέρα).

Κώδικας του Matlab αναφορικά με το dither:

```
%*****Matlab code for dither*****
% clear all

N=16384; %Number of samples in signal.
fs=11025; %Sampling Frequency.
nyquist=fs/2; %Nyquist Frequency.
t=[1:N]/fs; %Time axis of length N at sampling...
           ...frequency fs.

t=t';

%*****InputSignal(1)*****
f=400; %Input frequency in Hz.
x=sin(2.*pi.*f.*t); %Input signal is a sinusoid.
X=fft(x,N);

nfreq=floor(N/2)+1; %Number of points in frequency domain.
db=20*log10(abs(X(1:nfreq))); %Convert to dB.
frequency=linspace(0,nyquist,nfreq); %Frequency values used for...
           ...plotting below.

%*****InputSignal(2)*****
f2=400; %Input frequency in Hz.
x2=sin(2.*pi.*f2.*t); %Input signal is a sinusoid.

%*****InputSignal(3)*****
f3=408; %Input frequency in Hz.
x3=sin(2.*pi.*f3.*t); %Input signal is a sinusoid.

%*****InputSignal(4)*****
f4=393.75; %Input frequency in Hz.
x4=sin(2.*pi.*f4.*t); %Input signal is a sinusoid.

%*****Dither*****
d=(rand(N,1)-0.5)/10.0; %Dither between -0.5 & +0.5
x5=x4+d; %Add dither to input signal.

%*****Quantization(1)*****
y=round(x2*10)/10; %Uncomment this line to quantize.
X2=fft(y,N);

nfreq=floor(N/2)+1; %Number of points in frequency domain.
db2=20*log10(abs(X2(1:nfreq))); %Convert to dB.
frequency2=linspace(0,nyquist,nfreq); %Frequency values used for
           ...plotting below.

%*****Quantization(2)*****
y2=round(x2*3)/3; %Uncomment this line to quantize.
X3=fft(y2,N);

nfreq=floor(N/2)+1; %Number of points in frequency domain.
db3=20*log10(abs(X3(1:nfreq))); %Convert to dB.
frequency3=linspace(0,nyquist,nfreq); %Frequency values usedfor...
           ...plotting below.
```

```

%*****Quantization(3)*****
y3=round(x3*10)/10;          %Uncomment this line to quantize.
X4=fft(y3,N);

nfreq=floor(N/2)+1;          %Number of points in frequency domain.
db4=20*log10(abs(X4(1:nfreq))); %Convert to dB.
frequency4=linspace(0,nyquist,nfreq);%Frequency values used for...
                                     ...plotting below.

%*****Quantization(4)*****
y4=round(x4*10)/10;          %Uncomment this line to quantize.
X5=fft(y4,N);

nfreq=floor(N/2)+1;          %Number of points in frequency domain.
db5=20*log10(abs(X5(1:nfreq))); %Convert to dB.
frequency5=linspace(0,nyquist,nfreq);%Frequency values used for...
                                     ...plotting below.

%*****Quantization(5)*****
y5=round(x5*10)/10;          %Uncomment this line to quantize.
X6=fft(y5,N);

nfreq=floor(N/2)+1;          %Number of points in frequency domain.
db6=20*log10(abs(X6(1:nfreq))); %Convert to dB.
frequency6=linspace(0,nyquist,nfreq); %Frequency values used
                                     %for plotting below.

%*****Plots*****
figure(1)
plot(frequency,db)          %Plot a graph of the spectrum (400 Hz).
xlabel('Frequency, (Hz)')
ylabel('Amplitude, (dB)')
title('Pure tone at 400Hz')

figure(2)
plot(frequency2,db2)        %Plot a graph of the spectrum (21).
xlabel('Frequency, (Hz)')
ylabel('Amplitude, (dB)')
title('Truncation - 21 Q. Steps (Pure tone at 400 Hz)')

figure(3)
stairs(t,y)                 %Quantization levels (21) steps).
axis([0,0.004,-2,2])
xlabel('Time, (msec)')
ylabel('Level, (mV)')
title('21 Quantization steps')

figure(4)
plot(frequency3,db3)        %Plot a graph of the spectrum (7).
xlabel('Frequency, (Hz)')
ylabel('Amplitude, (dB)')
title('Truncation - 7 Q. Steps (Pure tone at 400 Hz)')

figure(5)
stairs(t,y2)                %Quantization levels (7) steps).
axis([0,0.004,-2,2])
xlabel('Time, (msec)')
ylabel('Level, (mV)')

```

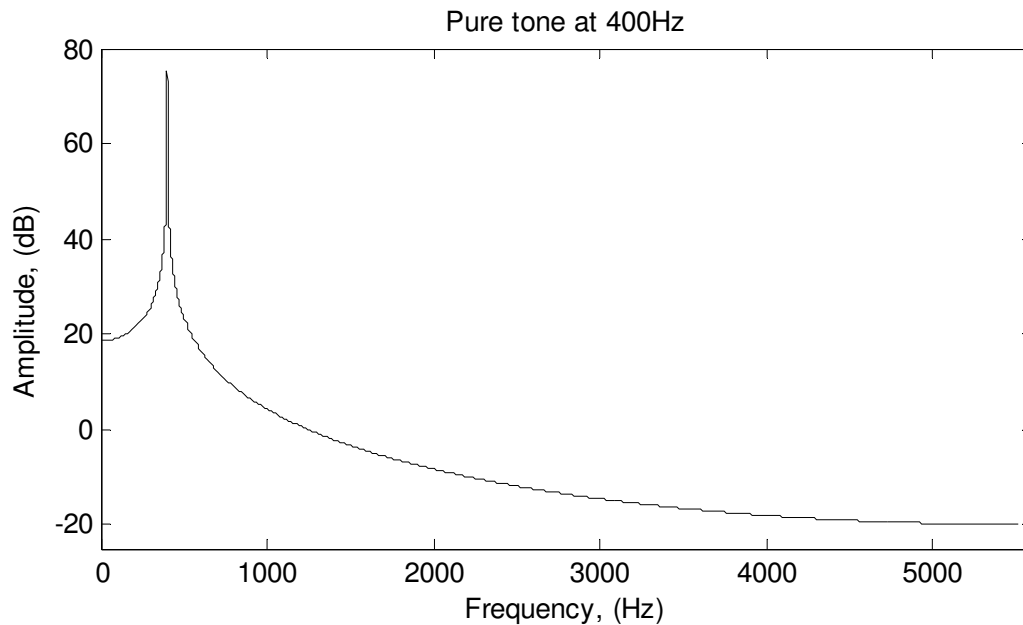
```
title('7 Quantization steps')
```

```
figure(6)
plot(frequency4,db4)      %Plot a graph of the spectrum (408 Hz).
xlabel('Frequency, (Hz)')
ylabel('Amplitude, (dB)')
title('Pure tone at 408 Hz')
```

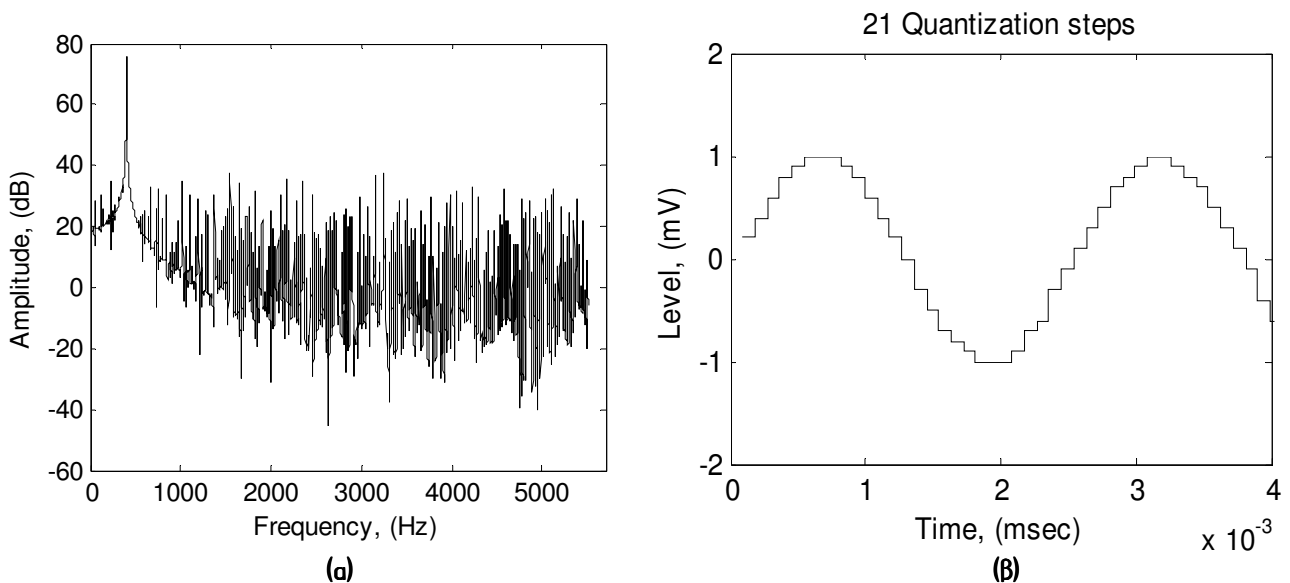
```
figure(7)
plot(frequency5,db5)      %Plot a graph of the spectrum (393.75 Hz).
xlabel('Frequency, (Hz)')
ylabel('Amplitude, (dB)')
title('Aliasing effect')
```

```
figure(8)
plot(frequency6,db6)      %Plot a graph of the spectrum (393.75 Hz).
xlabel('Frequency, (Hz)')      %Dithered
ylabel('Amplitude, (dB)')
title('Dithered tone (400 Hz)')
```

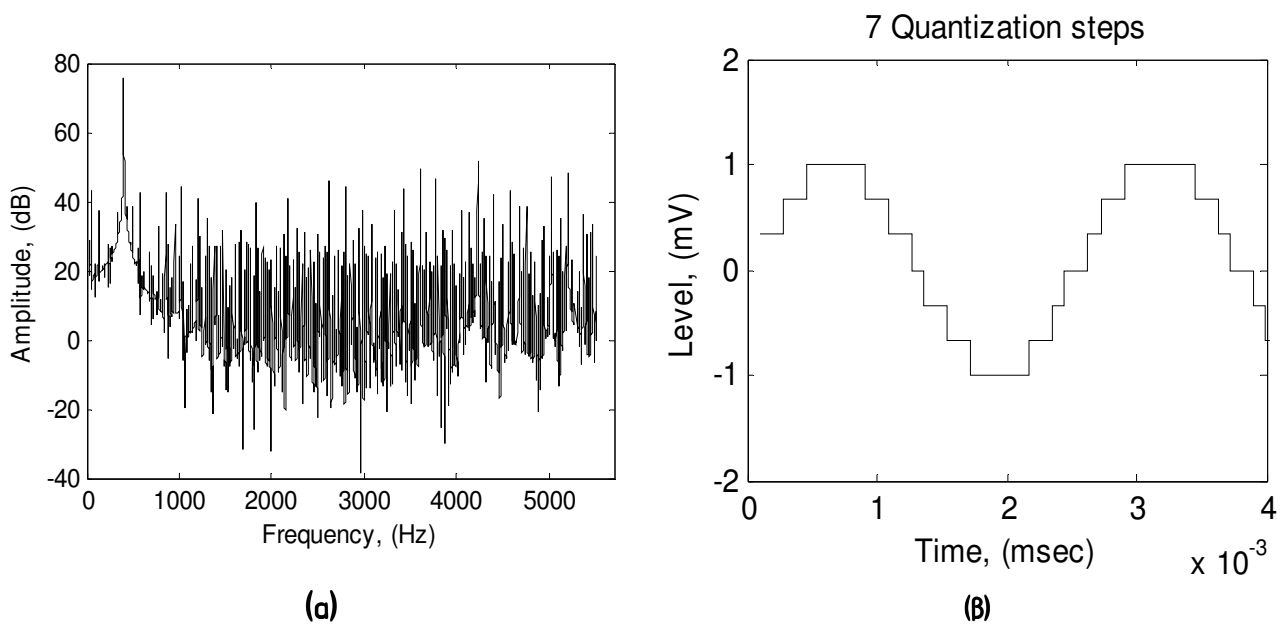
Διαγράμματα κι επεξηγήσεις που προκύπτουν απ' τον κώδικα:



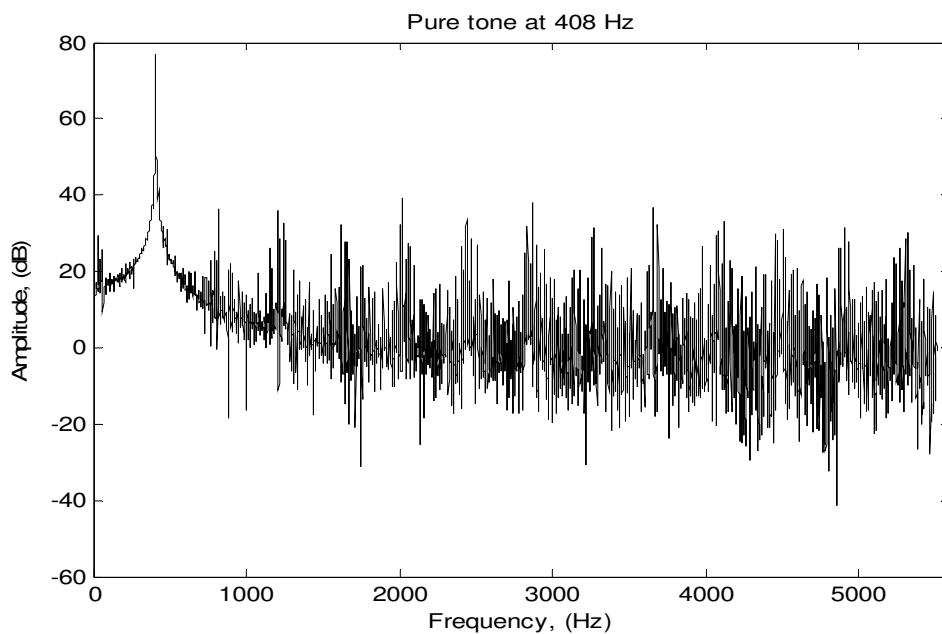
Σχ. Π.ΣΤ.1: Το φάσμα του καθαρού τόνου των 400 Hz που προκύπτει αρχικά απ' τον κώδικα.



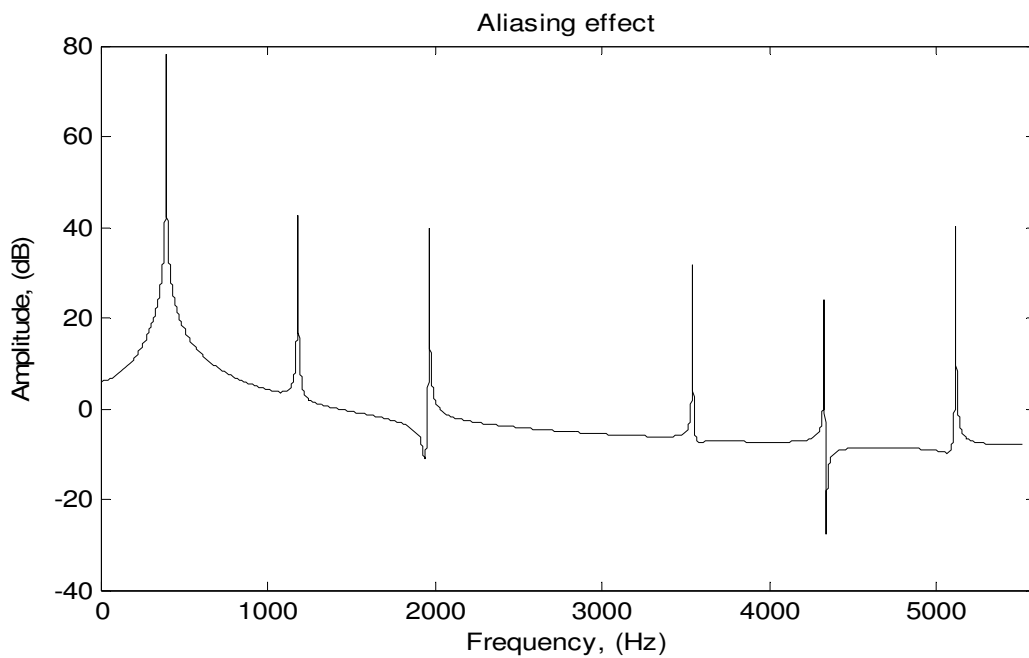
Σχ. Π.ΣΤ.2: (α) Το φάσμα του καθαρού τόνου των 400 Hz του σχήματος 1 έπειτα απο την κβαντοποίηση του απο 21 στάθμες κβαντισμού. Φαίνεται ξεκάθαρα ο θόρυβος που έχει προστεθεί στο σήμα. (β) Το ίδιο σήμα στο πεδίο του χρόνου για μία και μισή περίοδο έπειτα απο την κβαντοποίηση του. Φαίνονται ξεκάθαρα οι 21 στάθμες κβαντισμού που έχουν χρησιμοποιηθεί για την αναπαράσταση του.



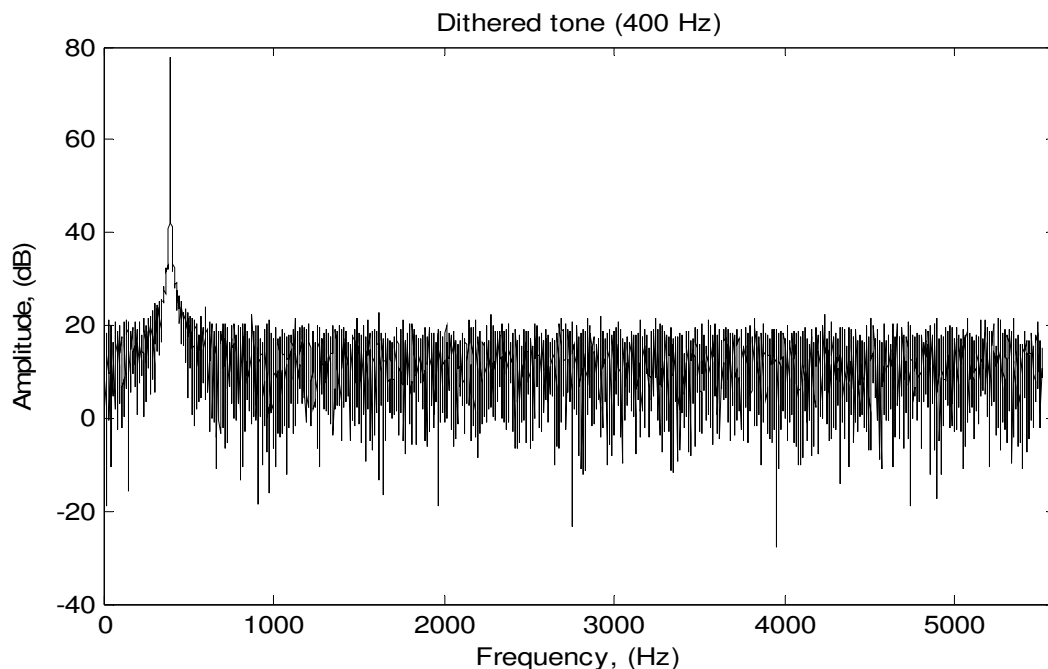
Σχ. Π.ΣΤ.3: (α) Το φάσμα του καθαρού τόνου των 400 Hz του σχήματος 1 έπειτα από την κβαντοποίηση του από 7 αυτή τη φορά στάθμες κβαντισμού. Φαίνεται ξεκάθαρα ο θόρυβος που έχει προστεθεί στο σήμα είναι κατά ένα αρκετά μεγάλο ποσοστό περισσότερος. (β) Το σήμα στο πεδίο του χρόνου για μία και μισή περίοδο. Φαίνονται ξεκάθαρα οι 7 στάθμες κβαντισμού, αυτή τη φορά, που έχουν χρησιμοποιηθεί για την αναπαράστασή του.



Σχ. Π.ΣΤ.4: Αλλάζοντας τη συχνότητα του σήματος εισόδου στον κώδικα από τα 400 Hz στα 408 Hz παρατηρούμε στο φάσμα που προκύπτει ότι ο θόρυβος κβαντοποίησης δεν τόσο τυχαίος όσο ήταν για την περίπτωση του σήματος εισόδου των 400 Hz.



Σχ. 5: Παρουσίαση του φαινομένου του aliasing. Όταν οι συνθήκες είναι ευνοϊκές (συχνότητα Nyquist, συχνотικό περιεχόμενο σήματος, φίλτρο αποκοπής υψηλών συχνοτήτων του converter) εμφανίζεται το εν λόγω φαινόμενο. Για τη δημιουργία του, στο συγκεκριμένο κώδικα, χρησιμοποιήθηκε ως σήμα εισόδου ένας καθαρός τόνος των 393,75 Hz. Παρατηρούμε ότι πλέον εμφανίζονται στο σήμα μας συχνοότητες που προηγουμένως δεν υπήρχαν σε αυτό.



Σχ. 6: Κβαντοποίηση και στη συνέχεια εφαρμογή flat dither στο αρχικό μας σήμα (καθαρός τόνος των 400 Hz). Παρατηρούμε ότι πλέον, έπειτα από την εφαρμογή του dither, ο θόρυβος στο σήμα μας απέκτησε ξανά τυχαία μορφή.

Παράρτημα Ζ:

Dither Internal NT₁ NS₁ 8bit

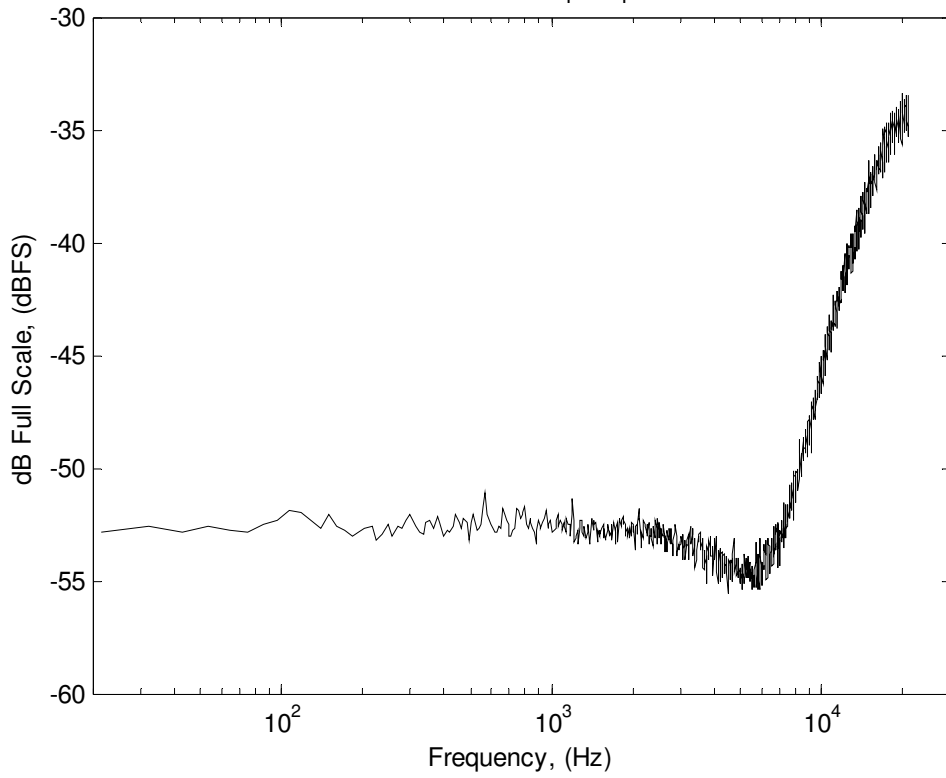


Fig. 1: Dither, Internal, Noise type 1, Noise Shape 1, 8 Bits.

Dither Internal NT₁ NS₁ 16bit

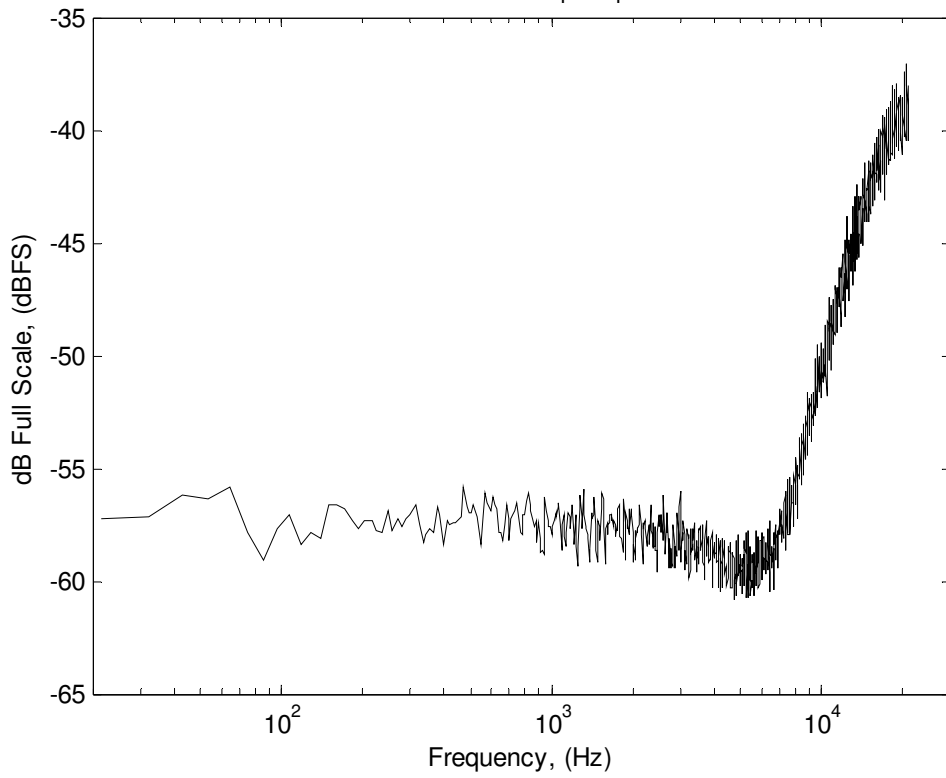


Fig. 2: Dither, Internal, Noise type 1, Noise Shape 1, 16 Bits.

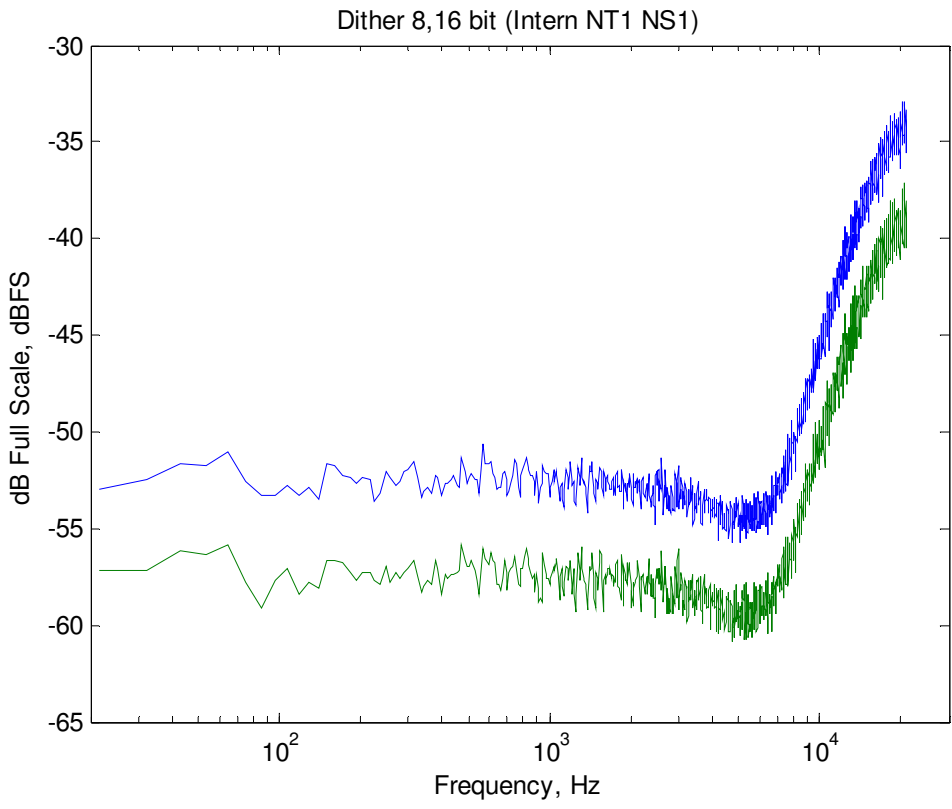


Fig. 3: Dither, Internal, Noise type 1, Noise Shape 1, 8 & 16 Bits.

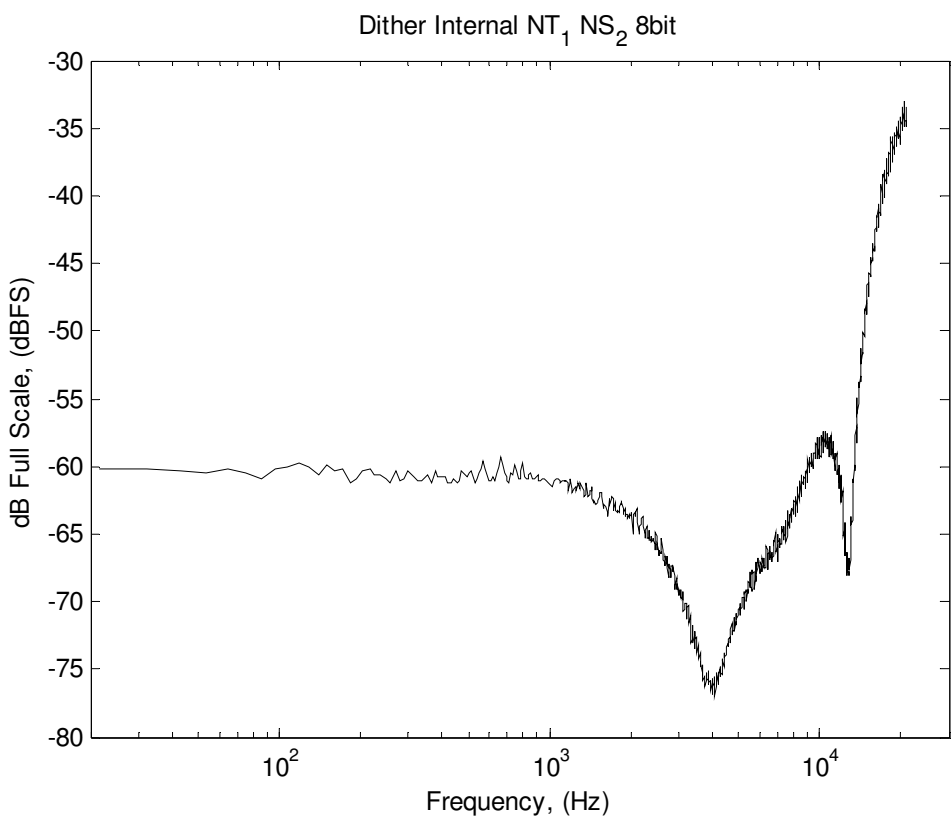


Fig. 4: Dither, Internal, Noise type 1, Noise Shape 2, 8 Bits.

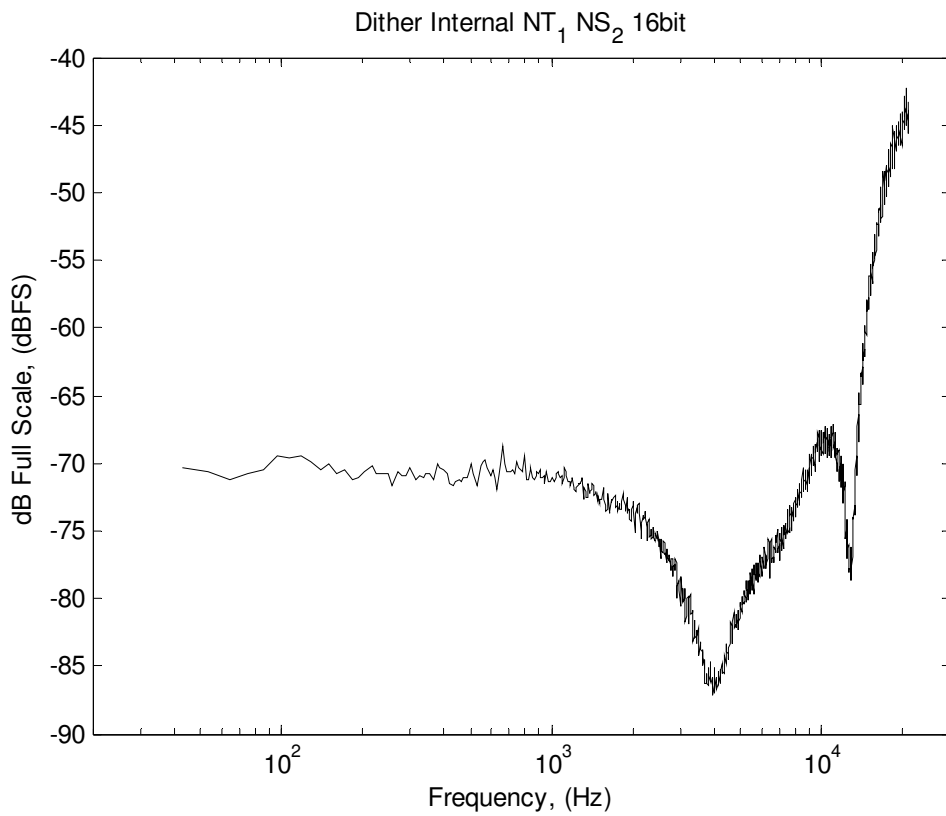


Fig. 5: Dither, Internal, Noise type 1, Noise Shape 2, 16 Bits.

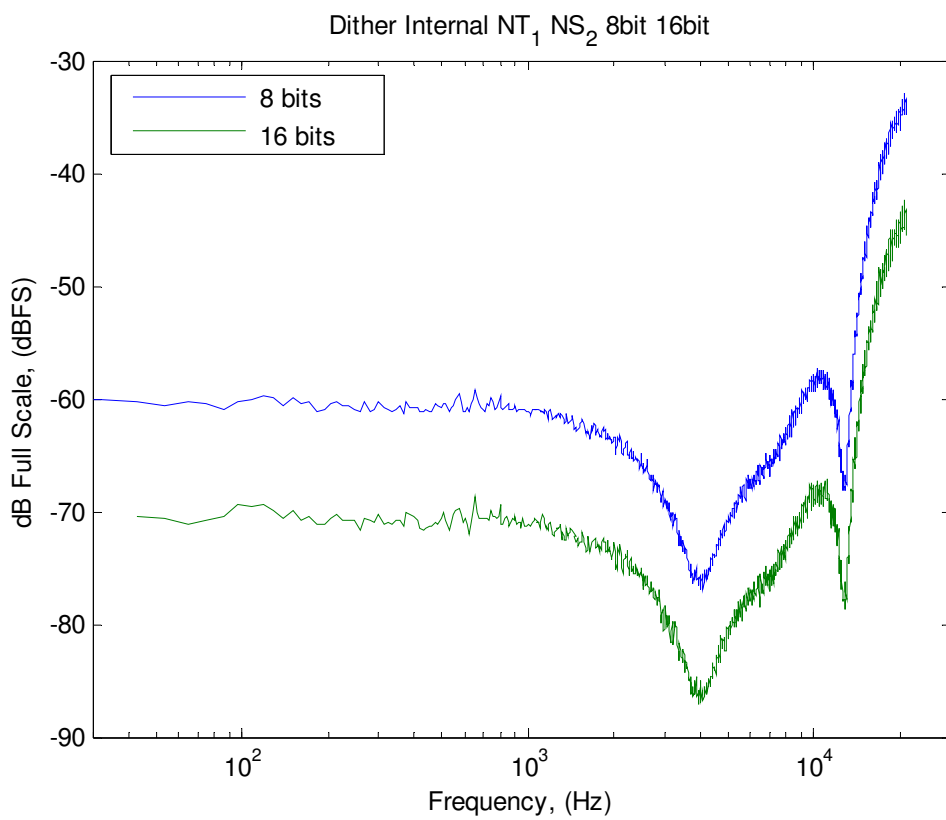


Fig. 6: Dither, Internal, Noise type 1, Noise Shape 2, 8 & 16 Bits.

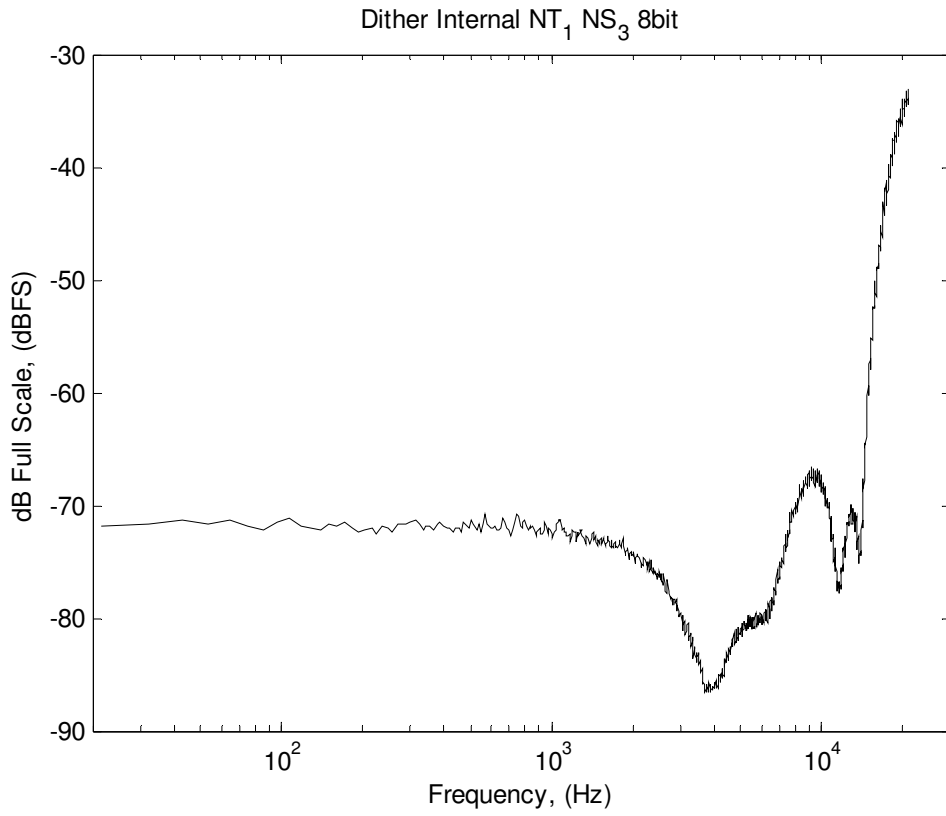


Fig. 7: Dither, Internal, Noise type 1, Noise Shape 3, 8 Bits.

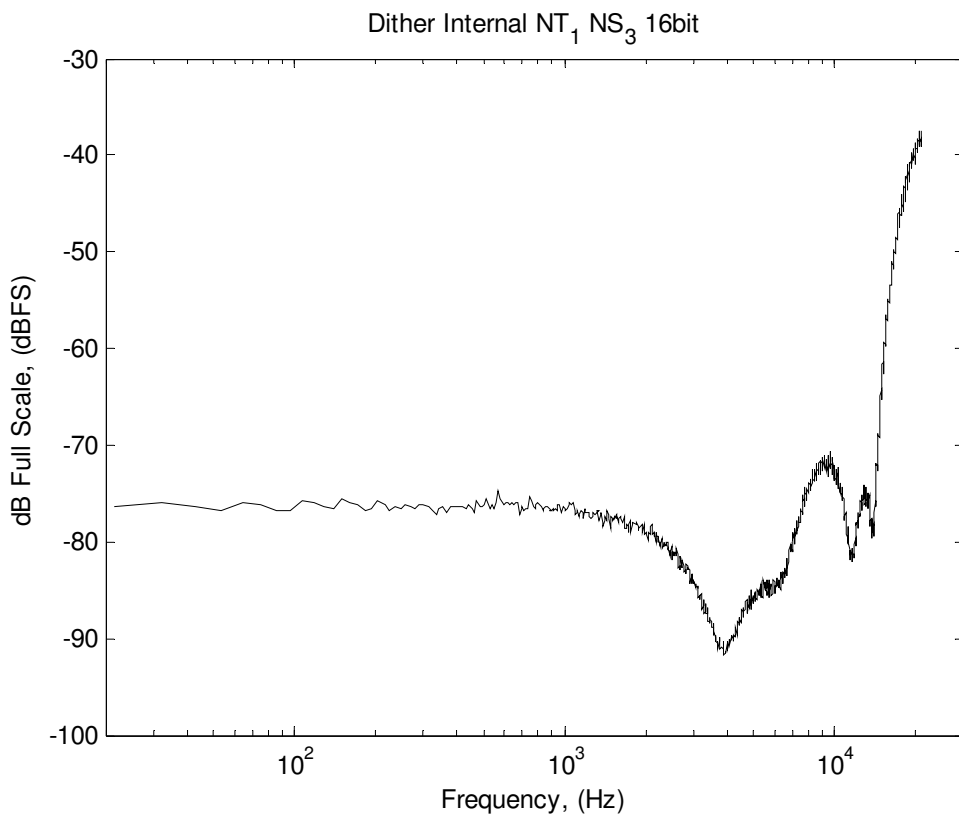


Fig. 8: Dither, Internal, Noise type 1, Noise Shape 3, 16 Bits.

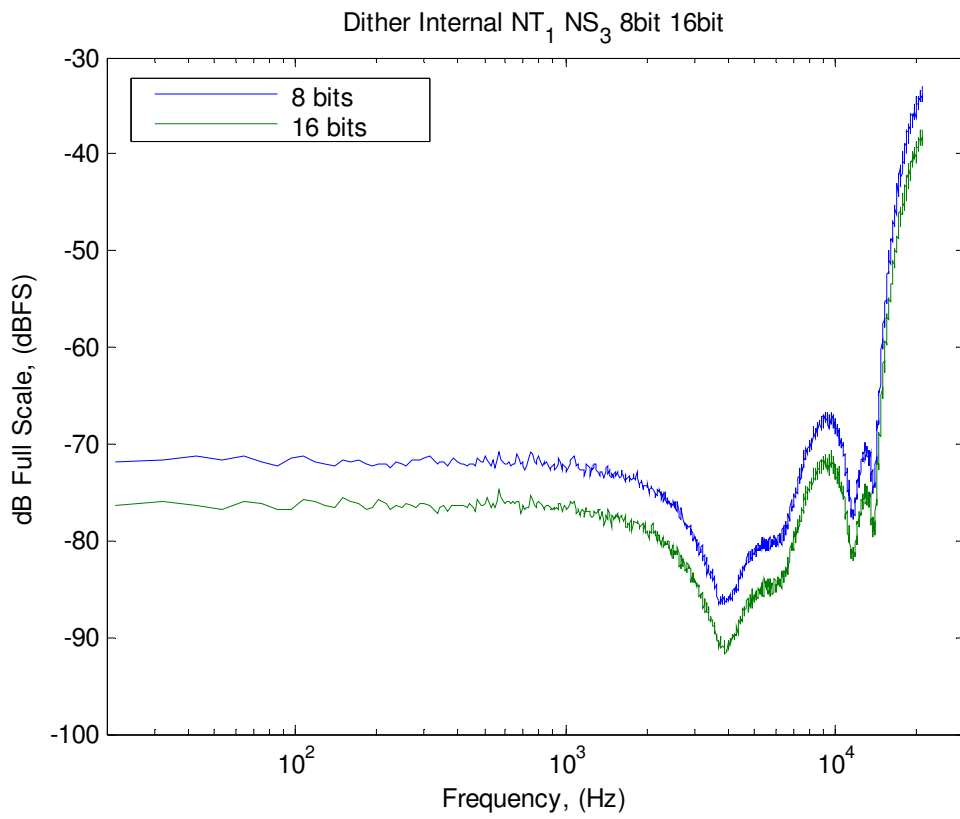


Fig. 9: Dither, Internal, Noise type 1, Noise Shape 3, 8 & 16 Bits.

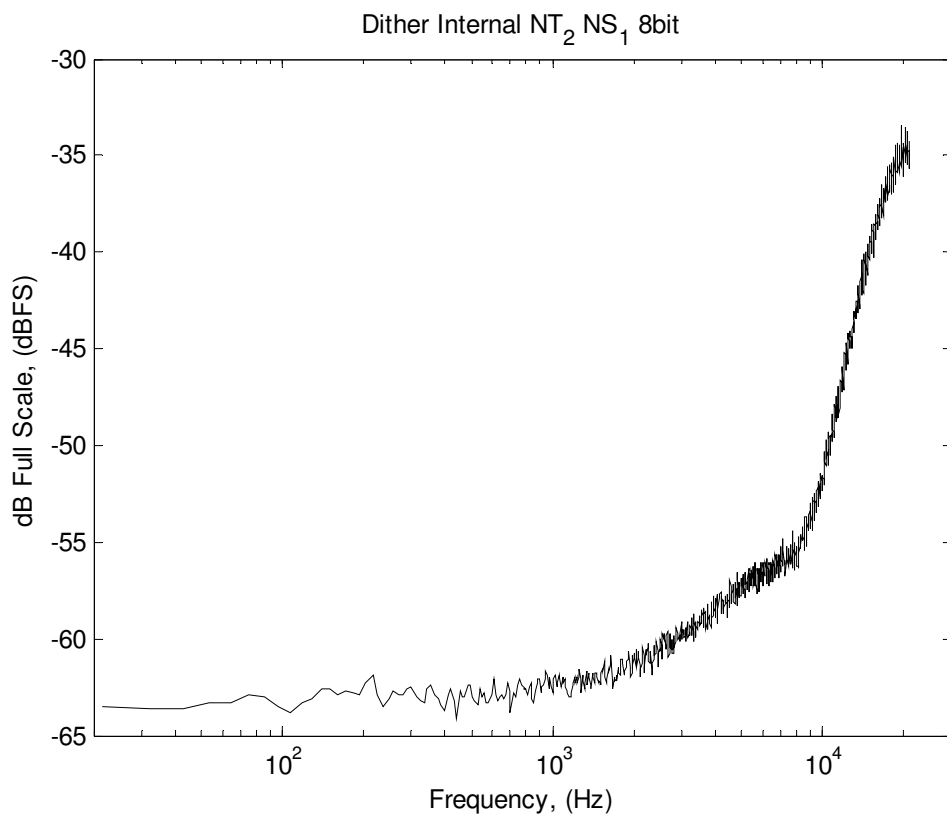


Fig. 10: Dither, Internal, Noise type 2, Noise Shape 1, 8 Bits.

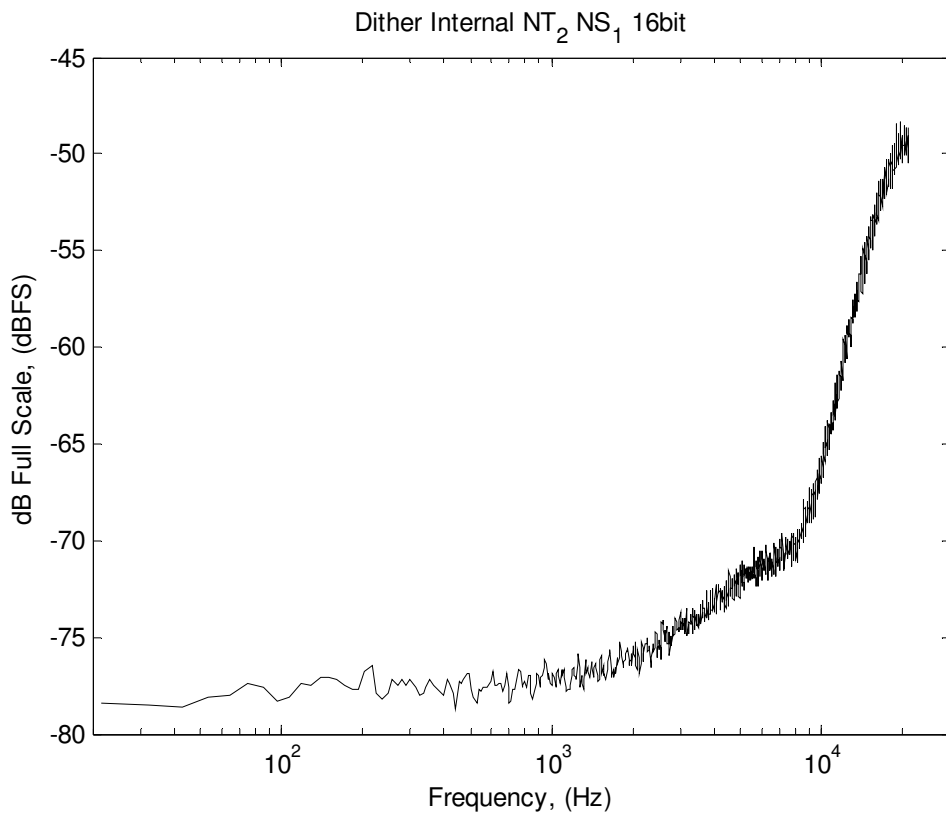


Fig. 11: Dither, Internal, Noise type 2, Noise Shape 1, 16 Bits.

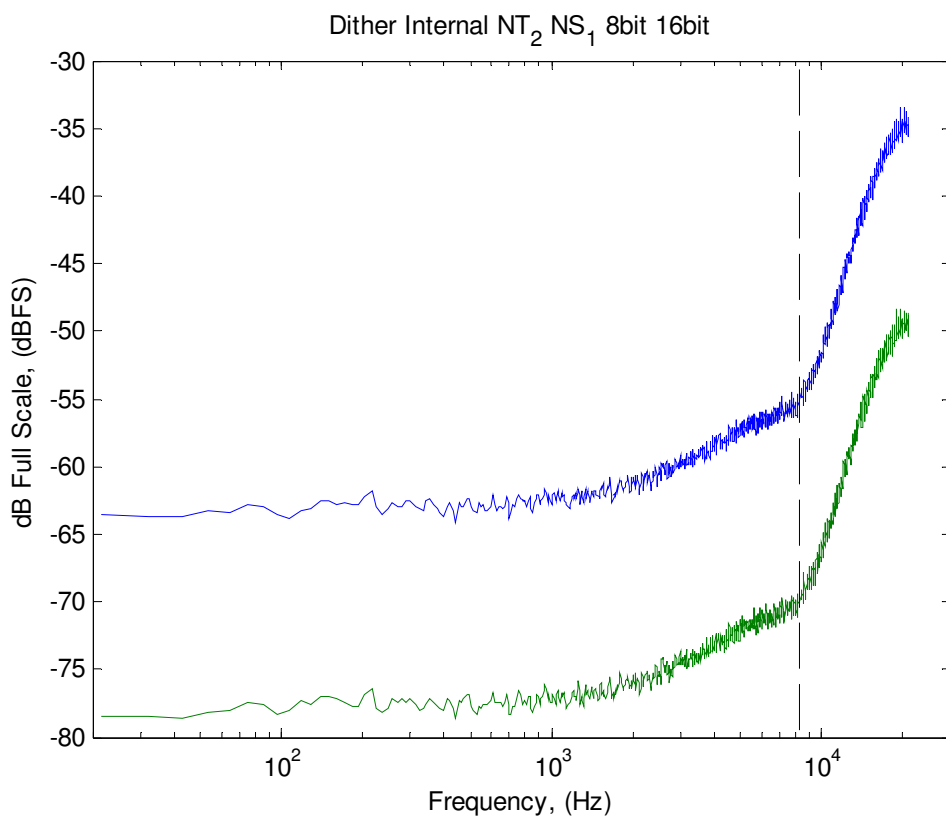


Fig. 12: Dither, Internal, Noise type 2, Noise Shape 1, 8 & 16 Bits.

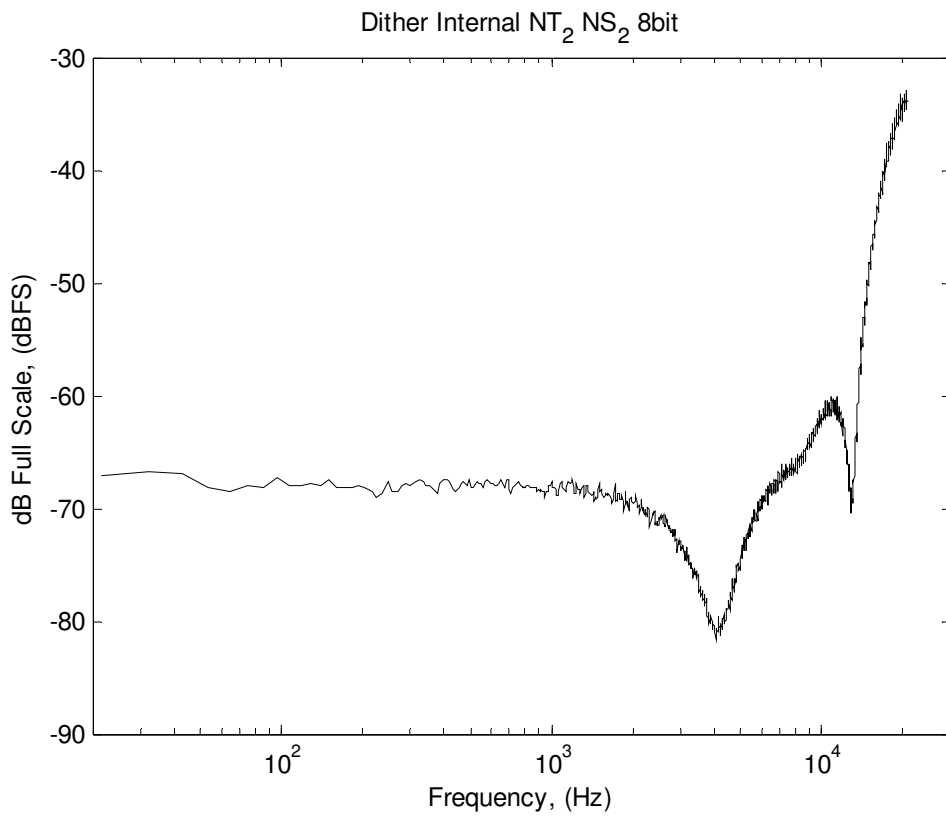


Fig. 13: Dither, Internal, Noise type 2, Noise Shape 2, 8 Bits.

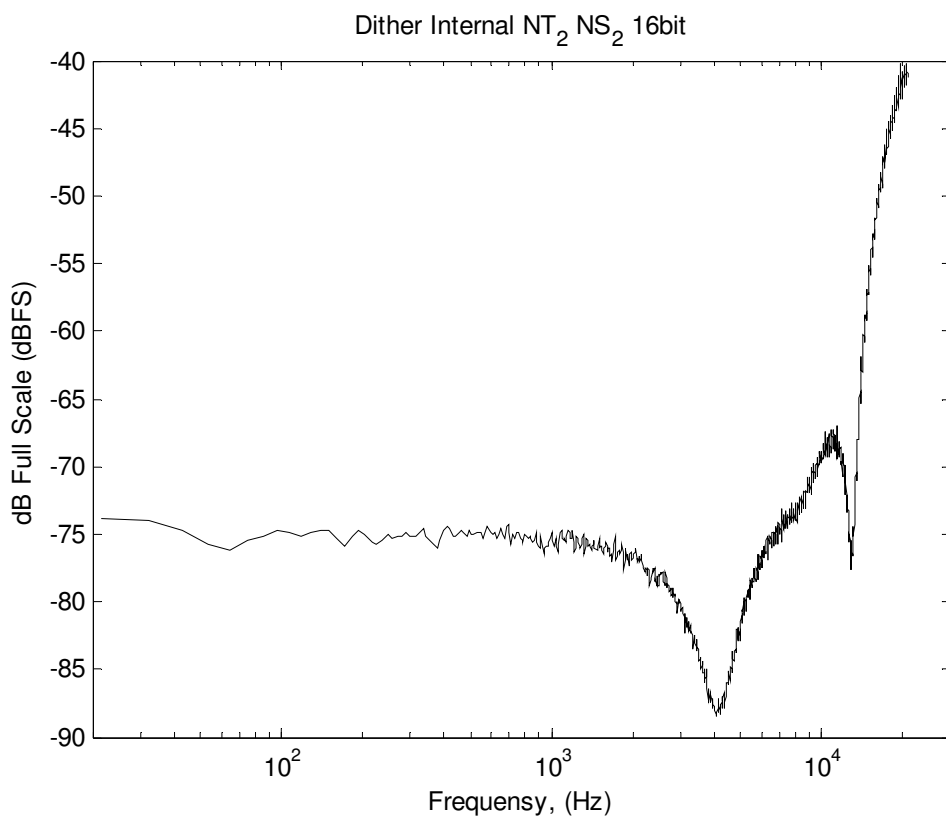


Fig. 14: Dither, Internal, Noise type 2, Noise Shape 2, 16 Bits.

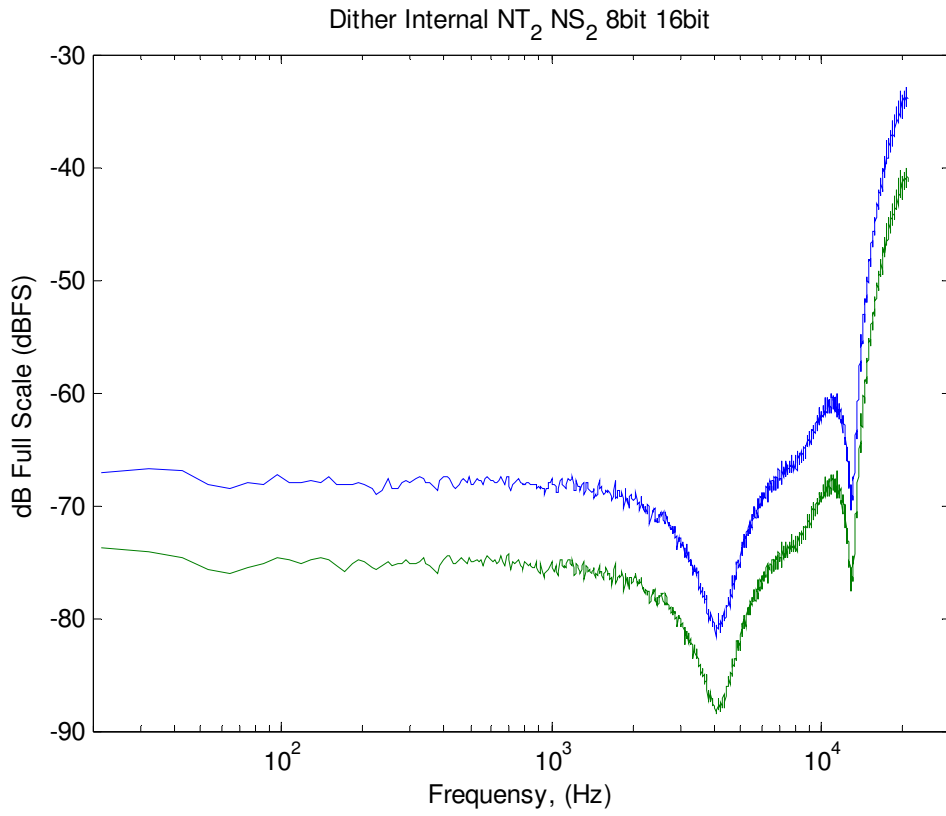


Fig. 15: Dither, Internal, Noise type 2, Noise Shape 2, 8 & 16 Bits.

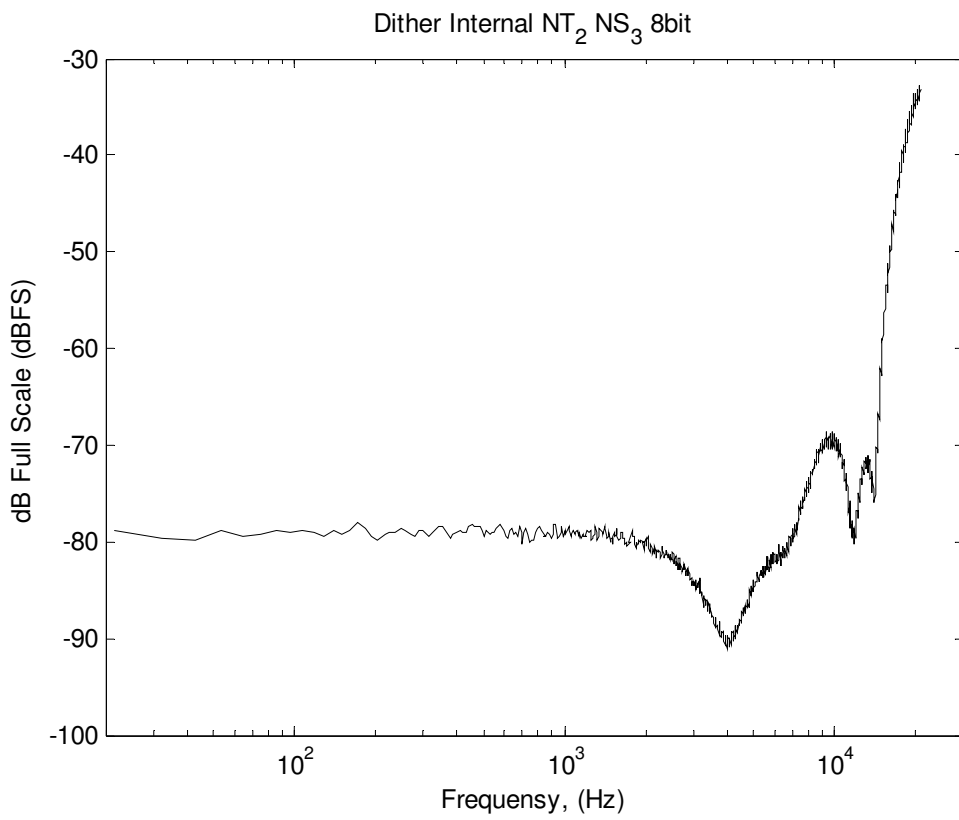


Fig. 16: Dither, Internal, Noise type 2, Noise Shape 3, 8 Bits.

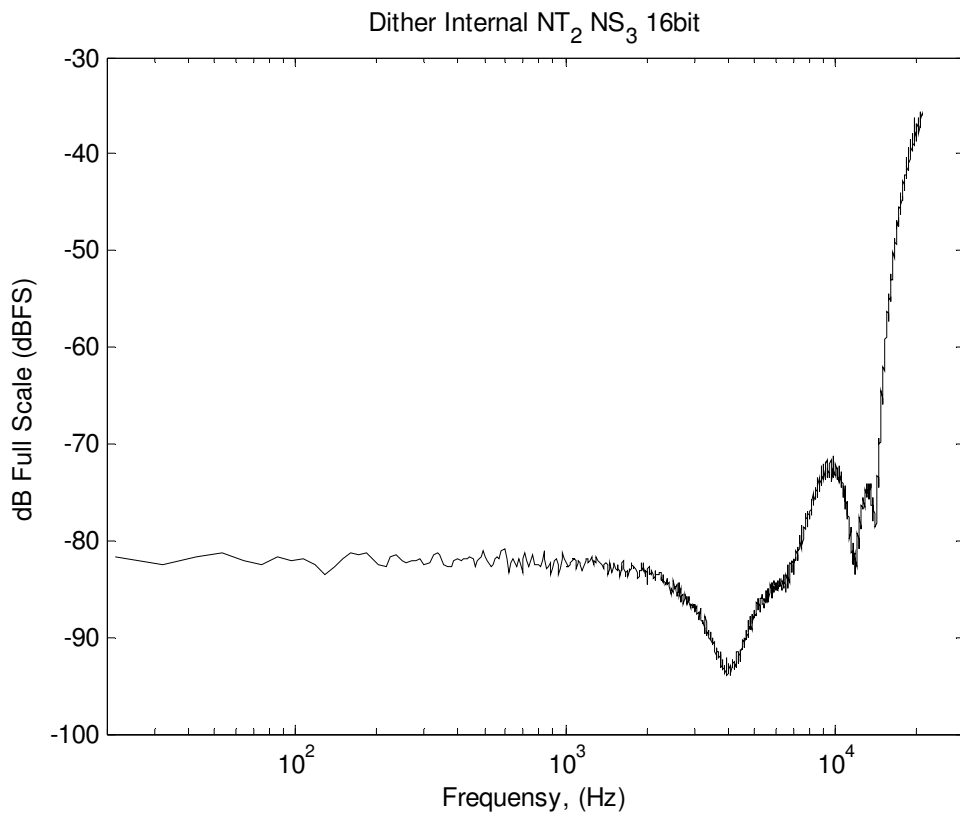


Fig. 17: Dither, Internal, Noise type 2, Noise Shape 3, 16 Bits.

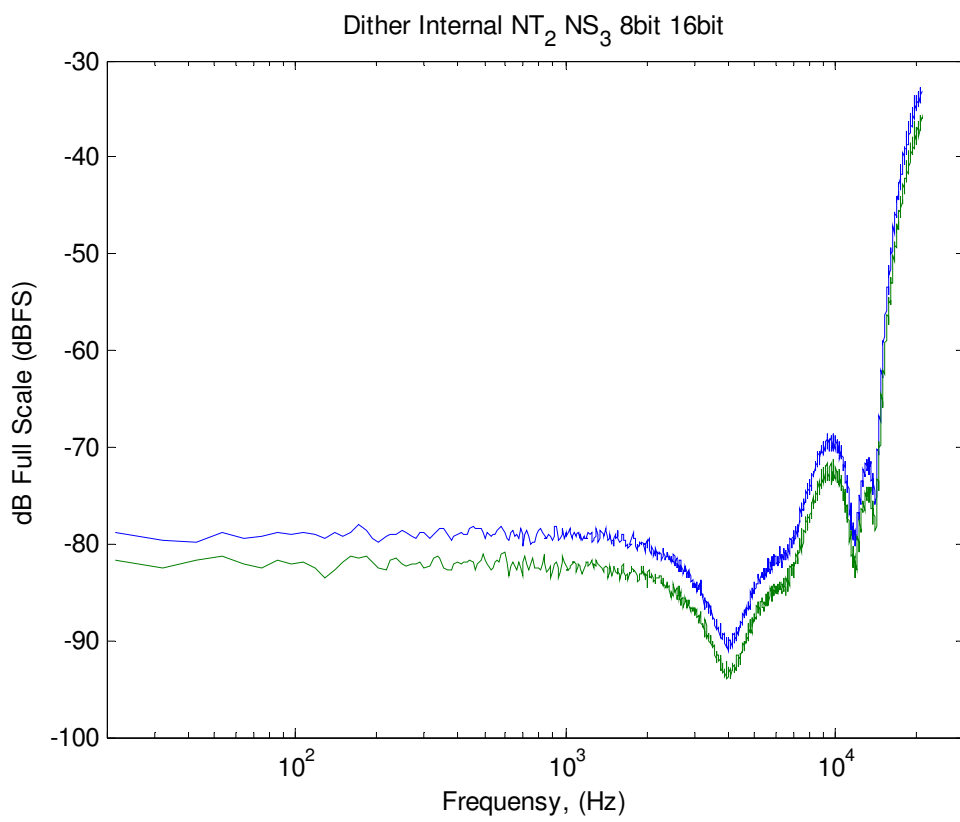


Fig. 18: Dither, Internal, Noise type 2, Noise Shape 3, 8 & 16 Bits.

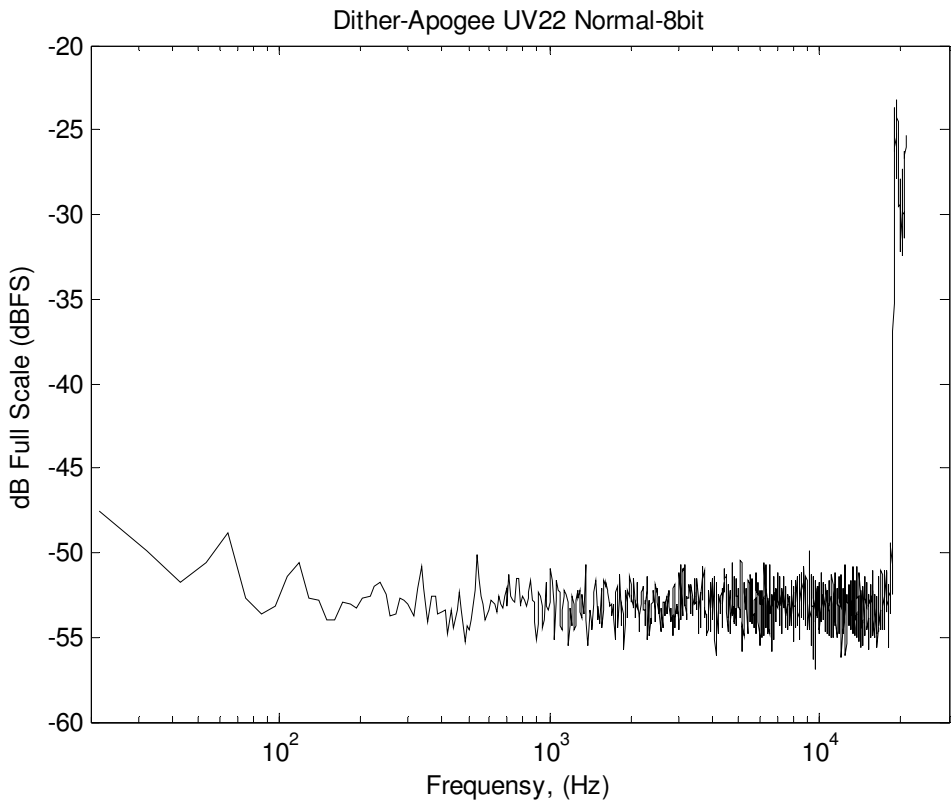


Fig. 19: Dither, Apogee UV22, Normal, 8 Bits.

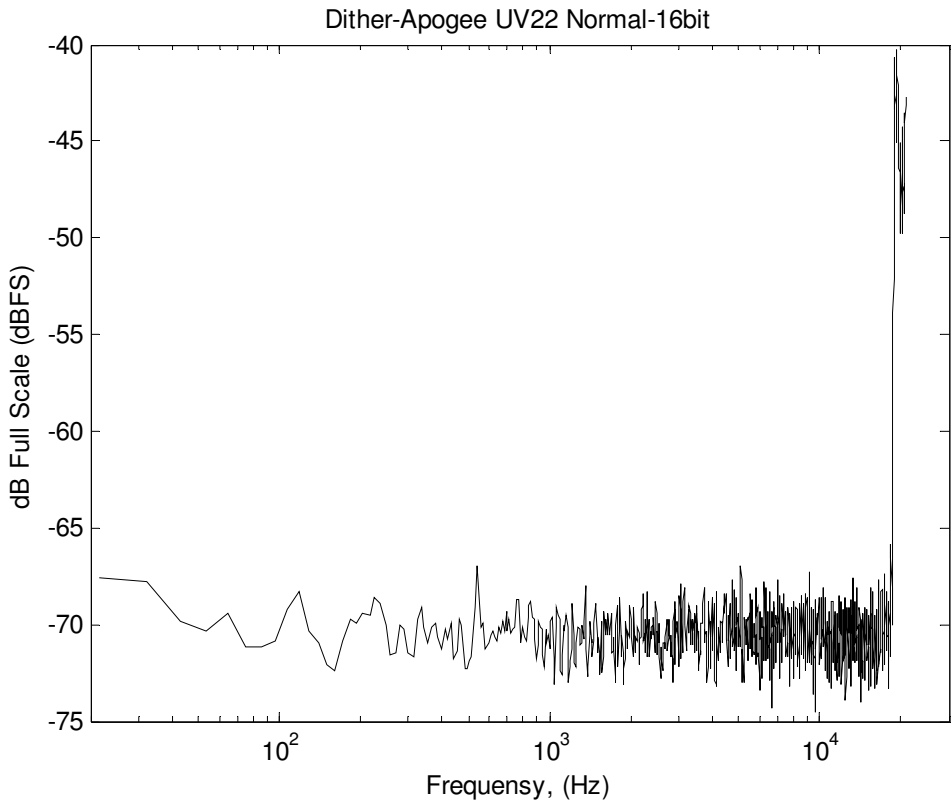


Fig. 20: Dither, Apogee UV22, Normal, 16 Bits.

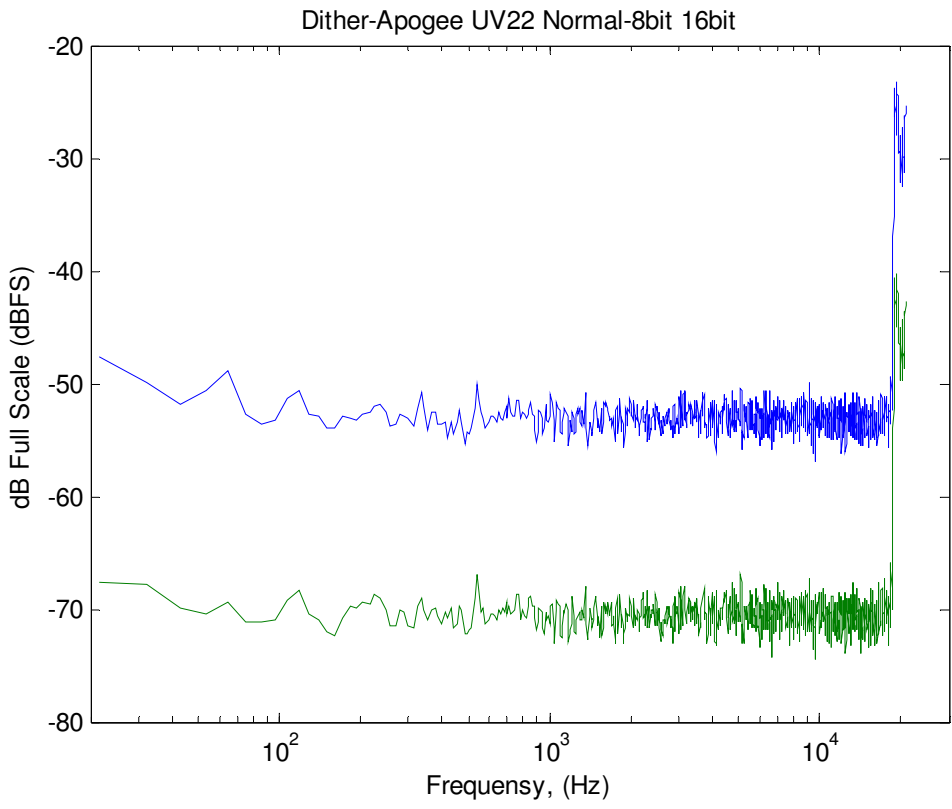


Fig. 21: Dither, Apogee UV22, Normal, 8 & 16 Bits.

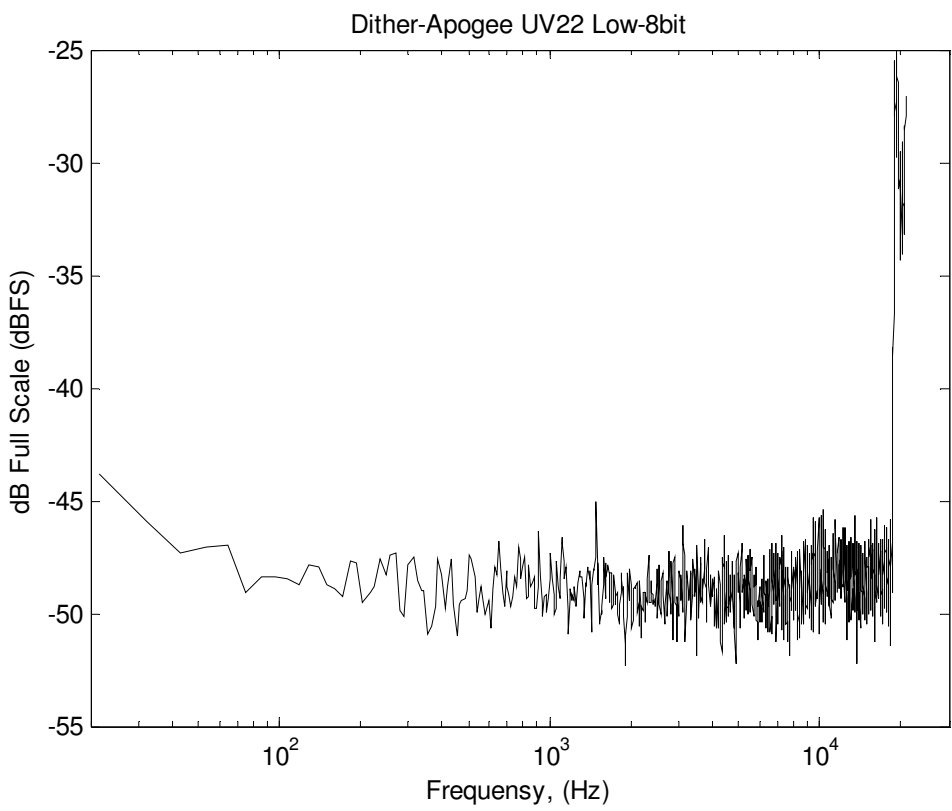


Fig. 22: Dither, Apogee UV22, Low, 8 Bits.

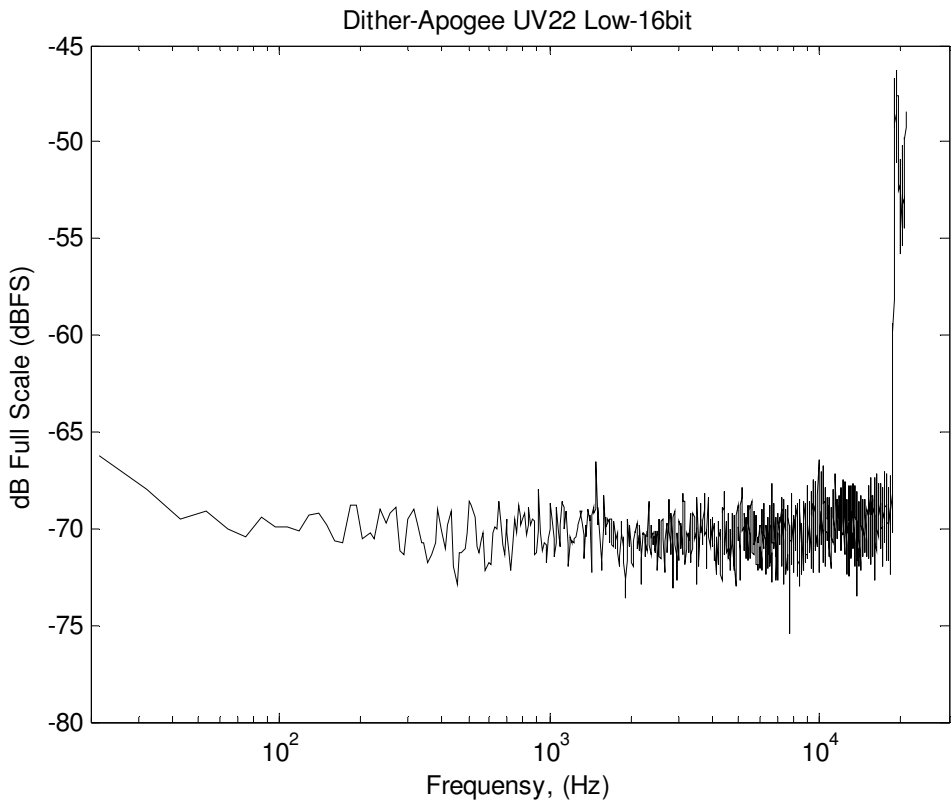


Fig. 23: Dither, Apogee UV22, Low, 16 Bits.

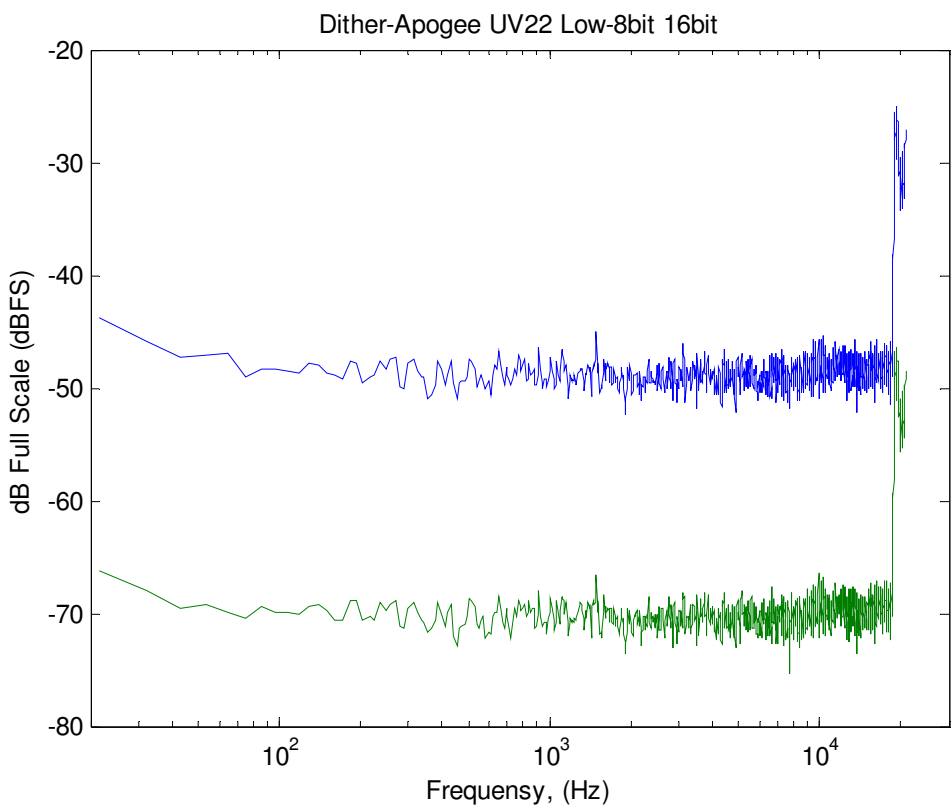


Fig. 24: Dither, Apogee UV22, Low, 8 & 16 Bits.

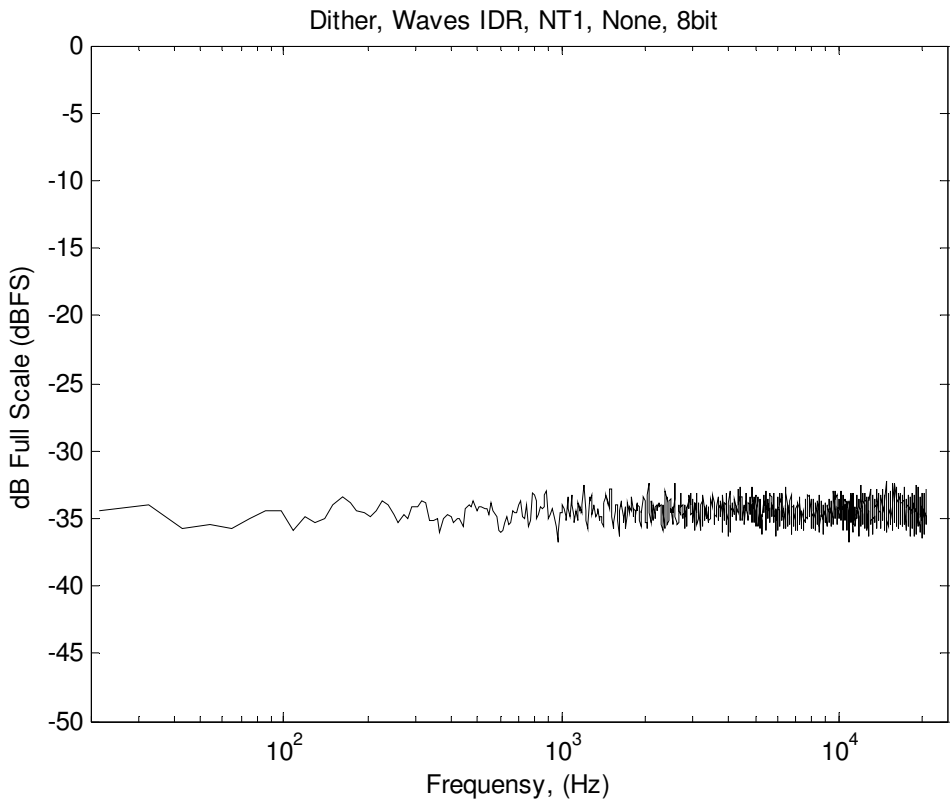


Fig. 25: Dither, Waves IDR, Noise Type 1, None, 8 Bits.

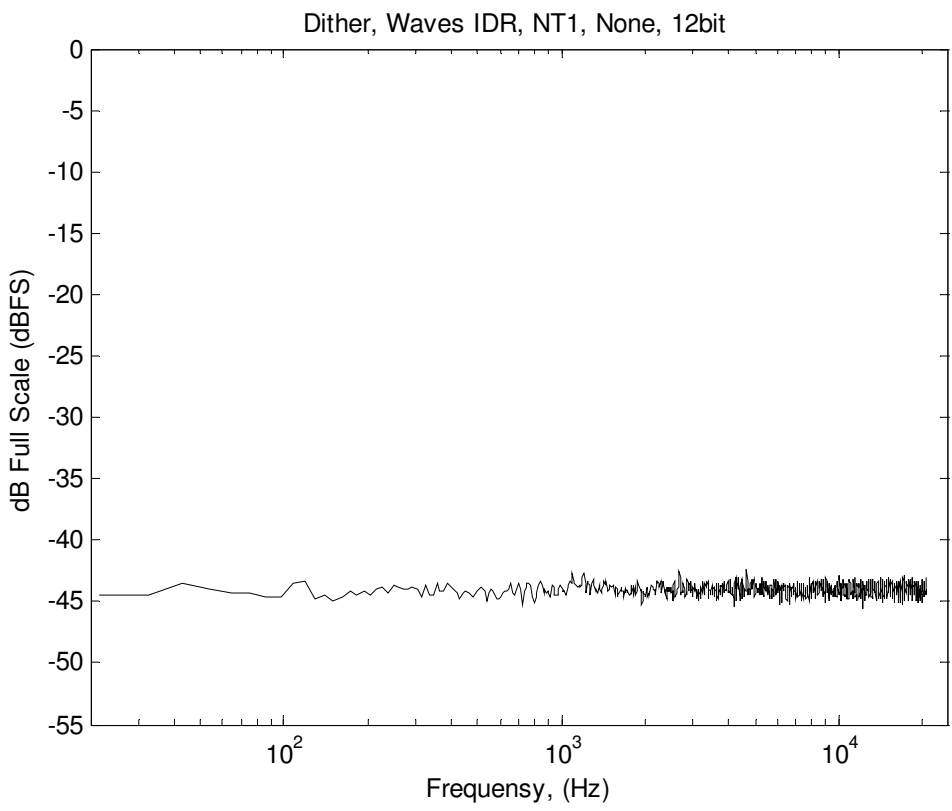


Fig. 26: Dither, Waves IDR, Noise Type 1, None, 12 Bits.

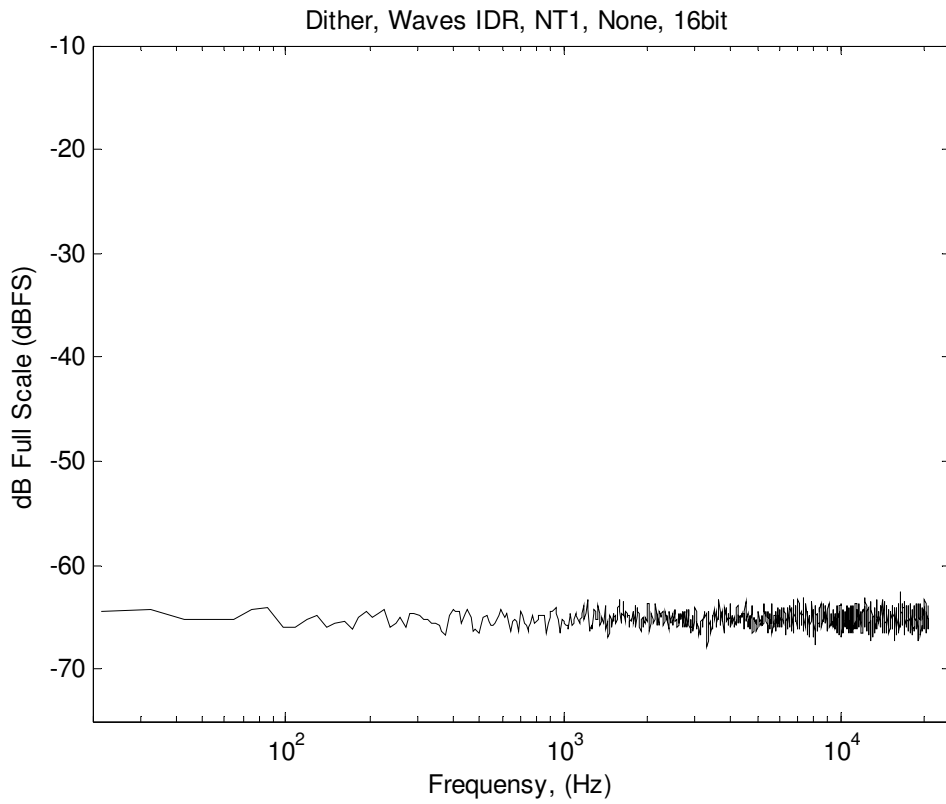


Fig. 27: Dither, Waves IDR, Noise Type 1, None, 16 Bits.

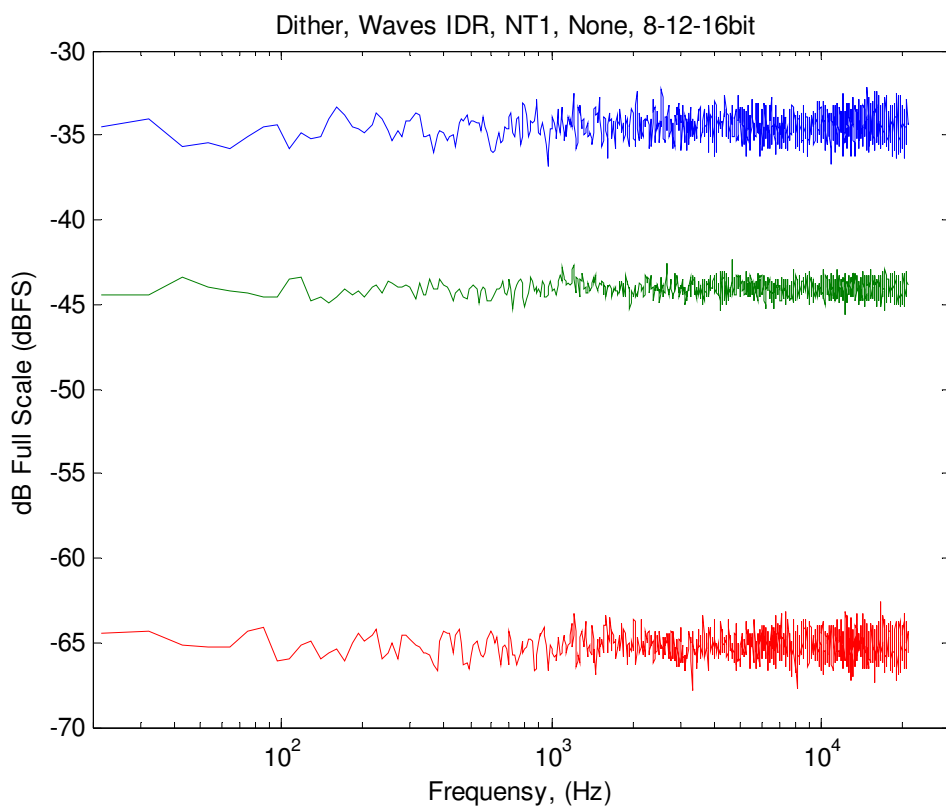


Fig. 28: Dither, Waves IDR, Noise Type 1, None, 8, 12 & 16 Bits.

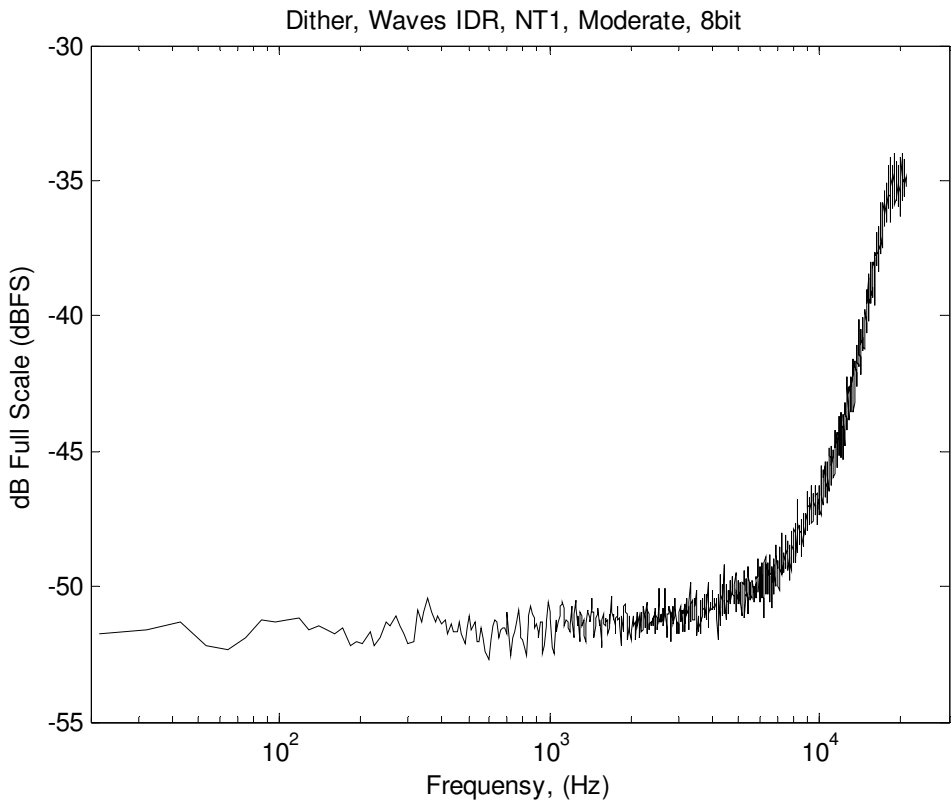


Fig. 29: Dither, Waves IDR, Noise Type 1, Moderate, 8 Bits.

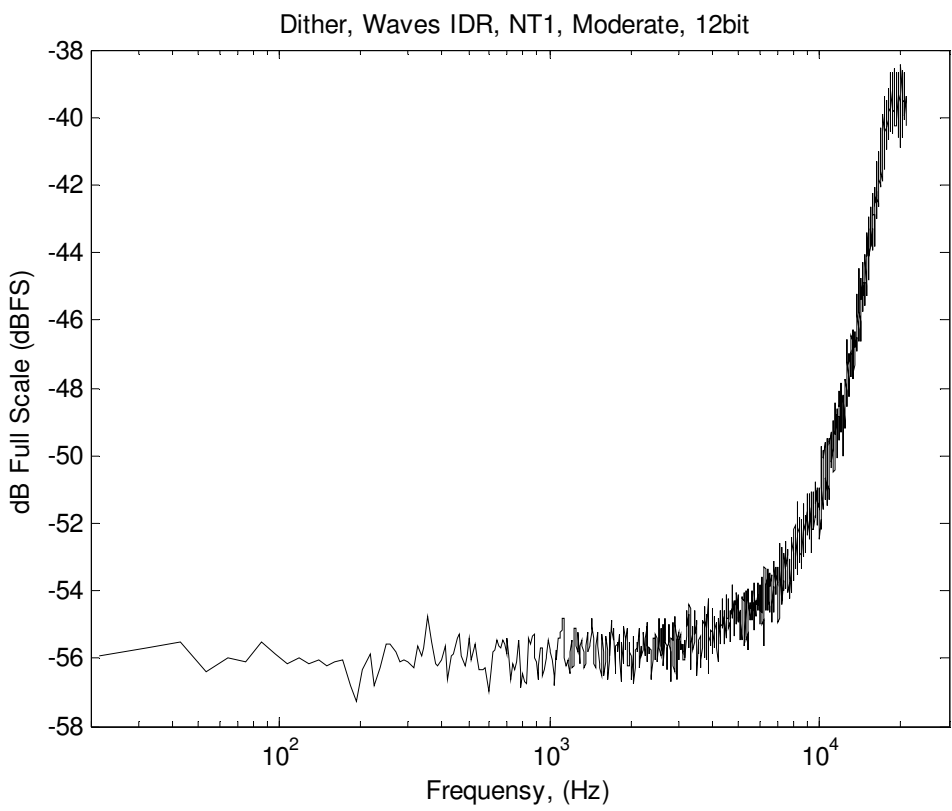


Fig. 30: Dither, Waves IDR, Noise Type 1, Moderate, 12 Bits.

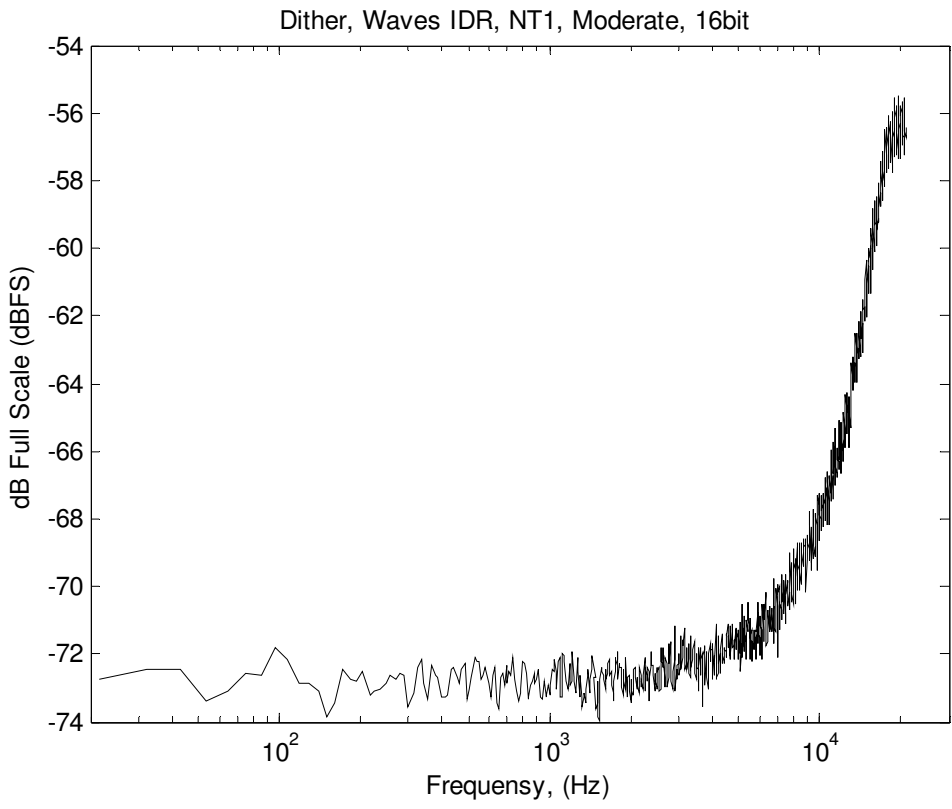


Fig. 31: Dither, Waves IDR, Noise Type 1, Moderate, 16 Bits.

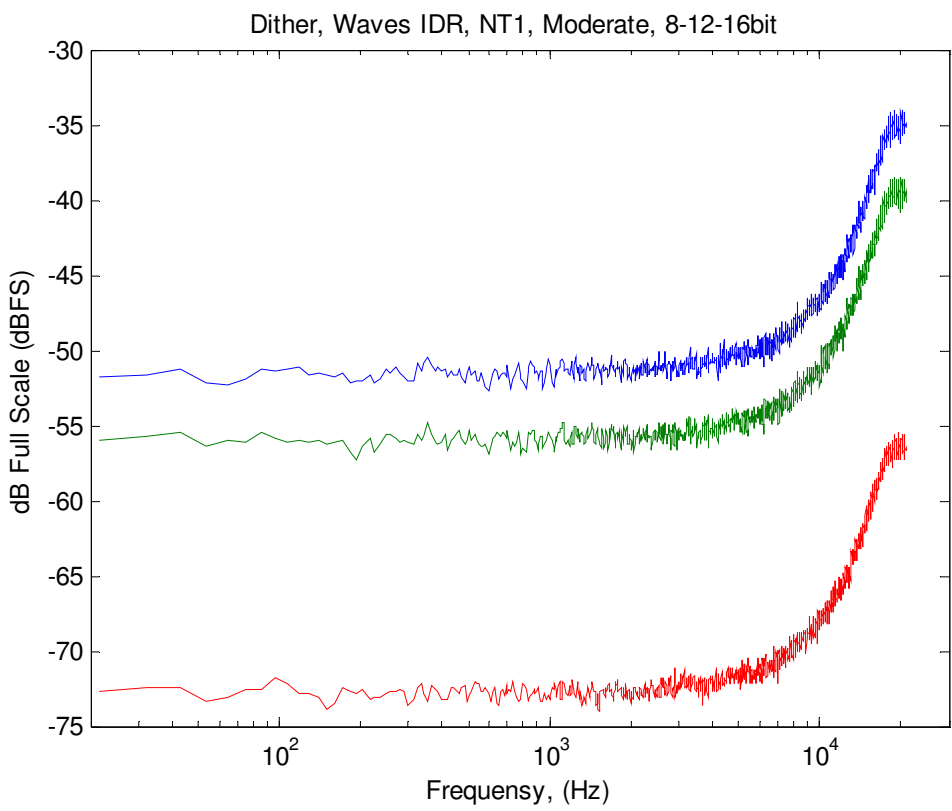


Fig. 32: Dither, Waves IDR, Noise Type 1, Moderate, 8, 12 & 16 Bits.

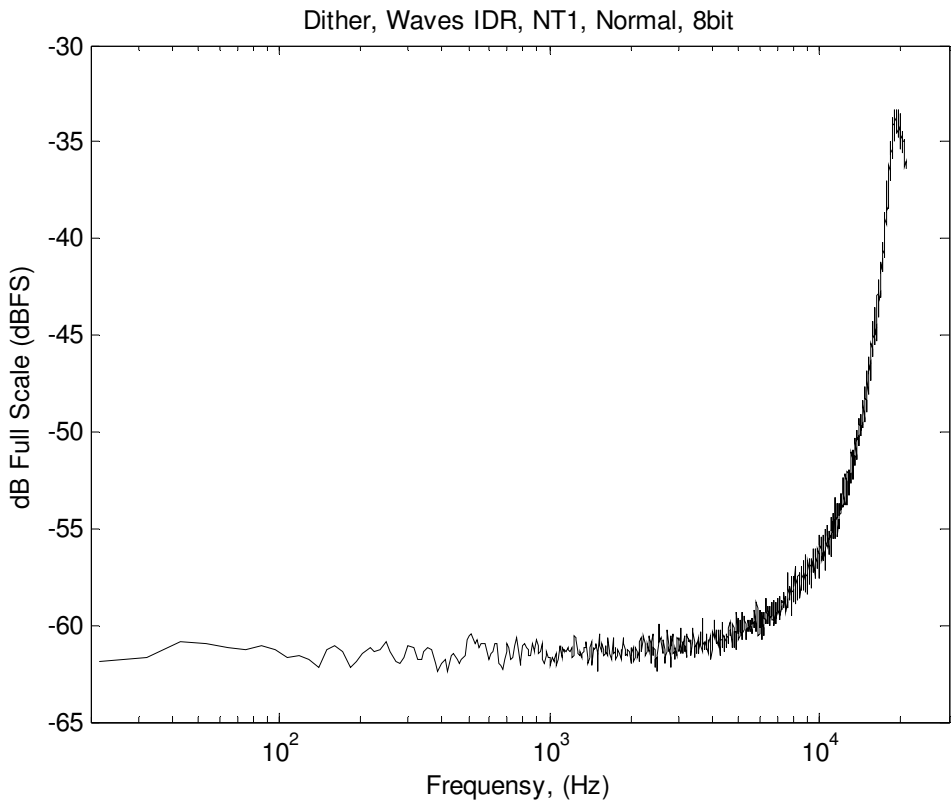


Fig. 33: Dither, Waves IDR, Noise Type 1, Normal, 8 Bits.

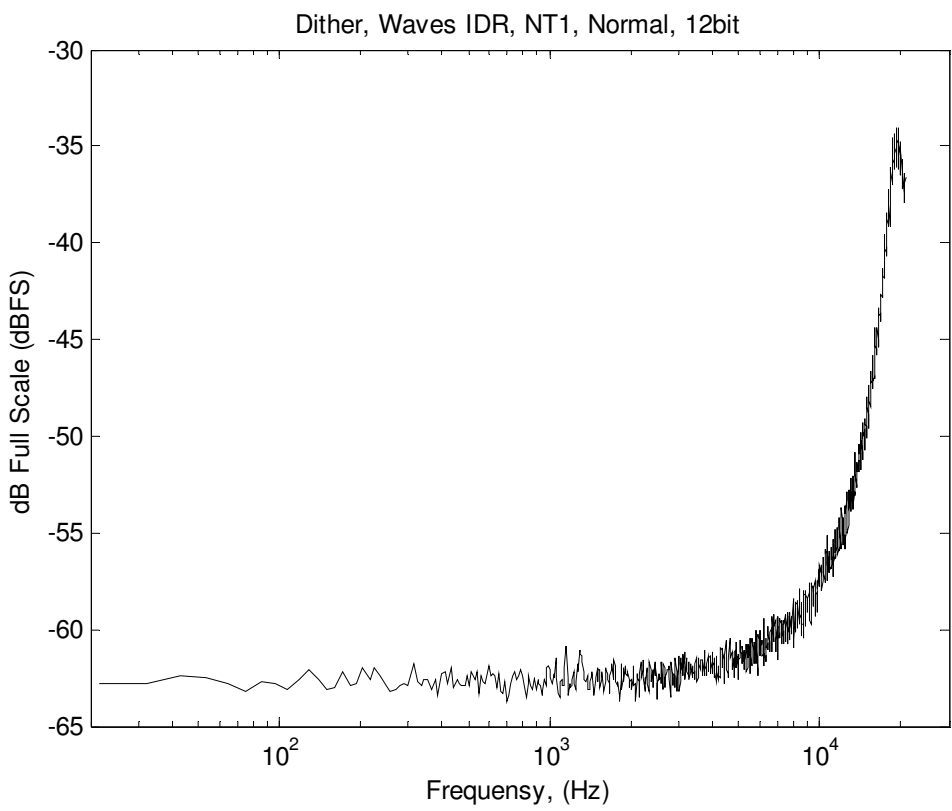


Fig. 34: Dither, Waves IDR, Noise Type 1, Normal, 12 Bits.

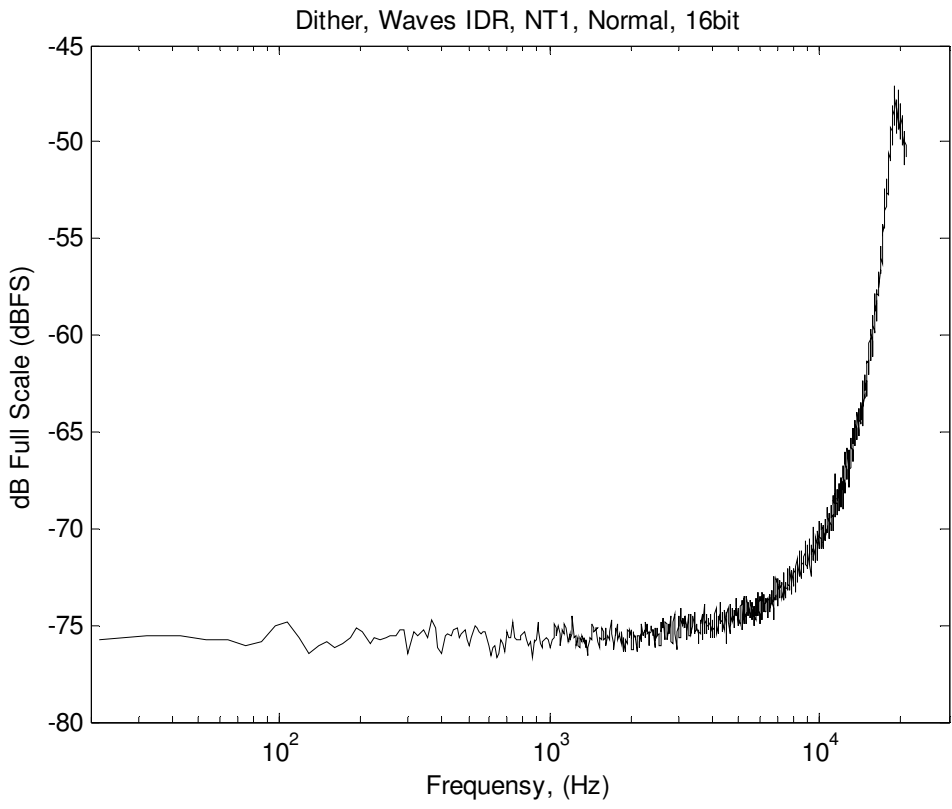


Fig. 35: Dither, Waves IDR, Noise Type 1, Normal, 16 Bits.

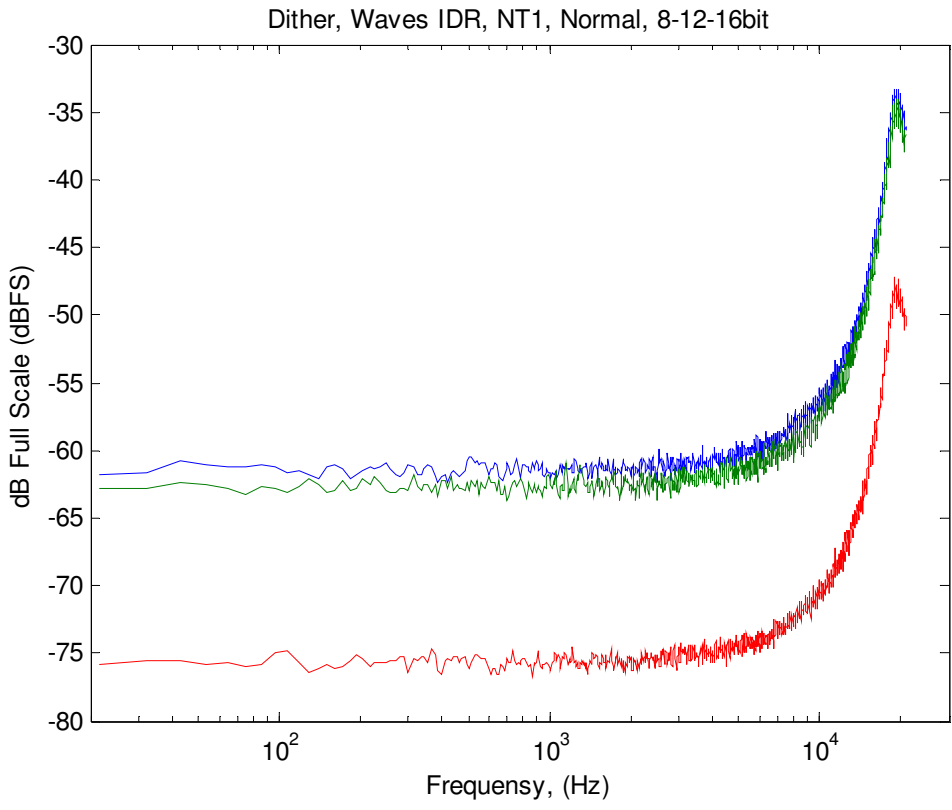


Fig. 36: Dither, Waves IDR, Noise Type 1, Normal, 8, 12 & 16 Bits.

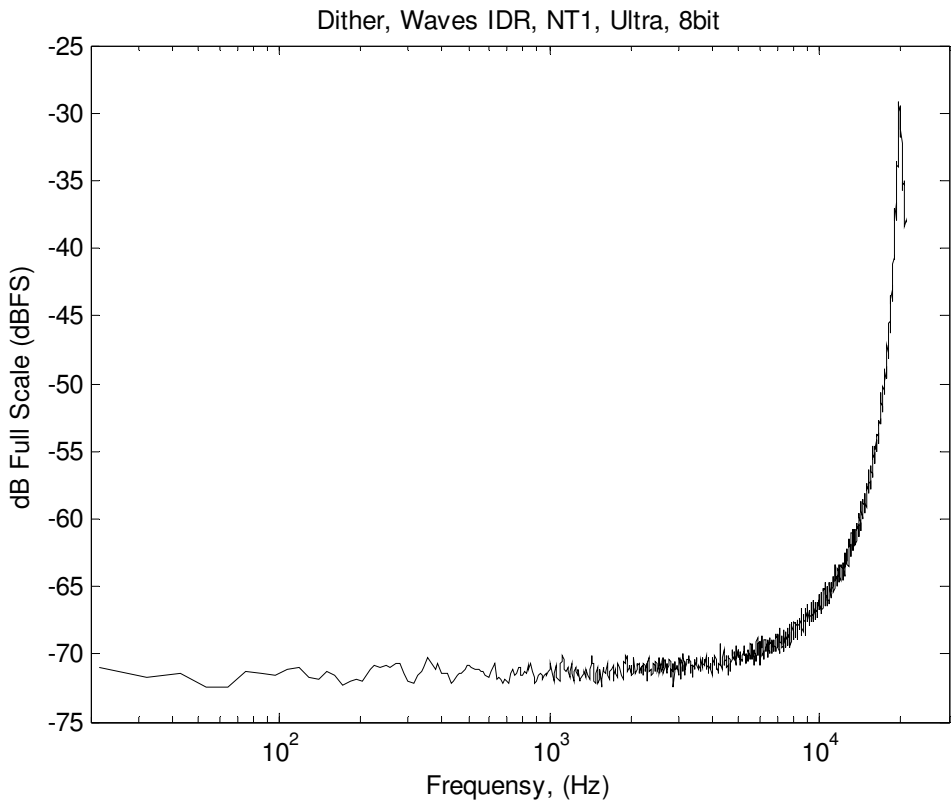


Fig. 37: Dither, Waves IDR, Noise Type 1, Ultra, 8 Bits.

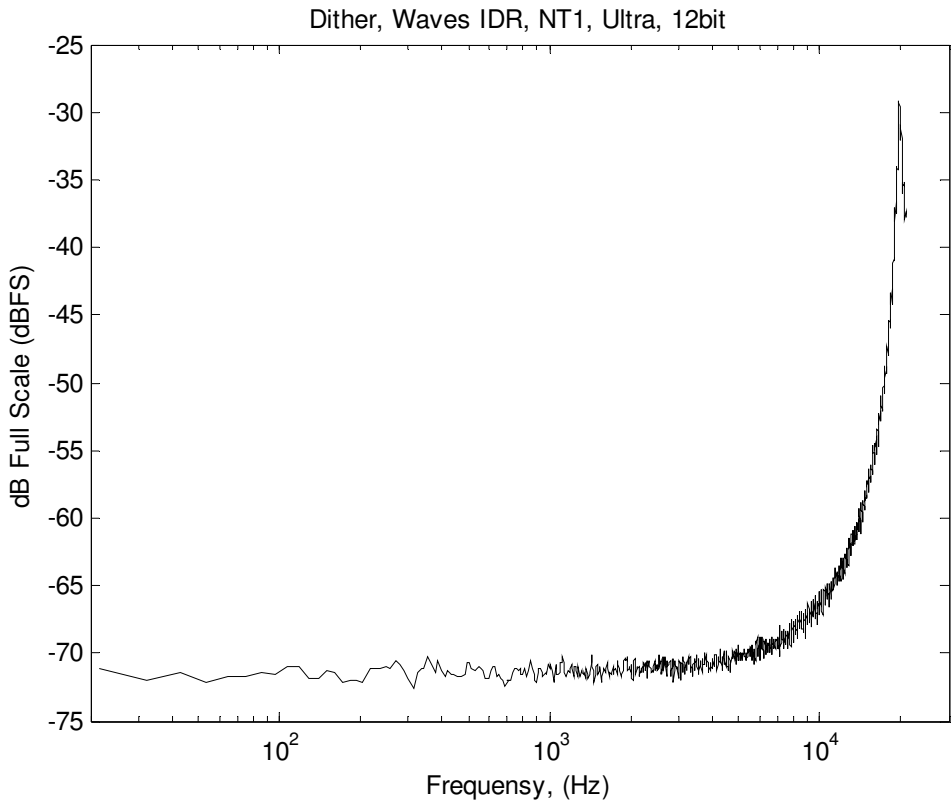


Fig. 38: Dither, Waves IDR, Noise Type 1, Ultra, 12 Bits.

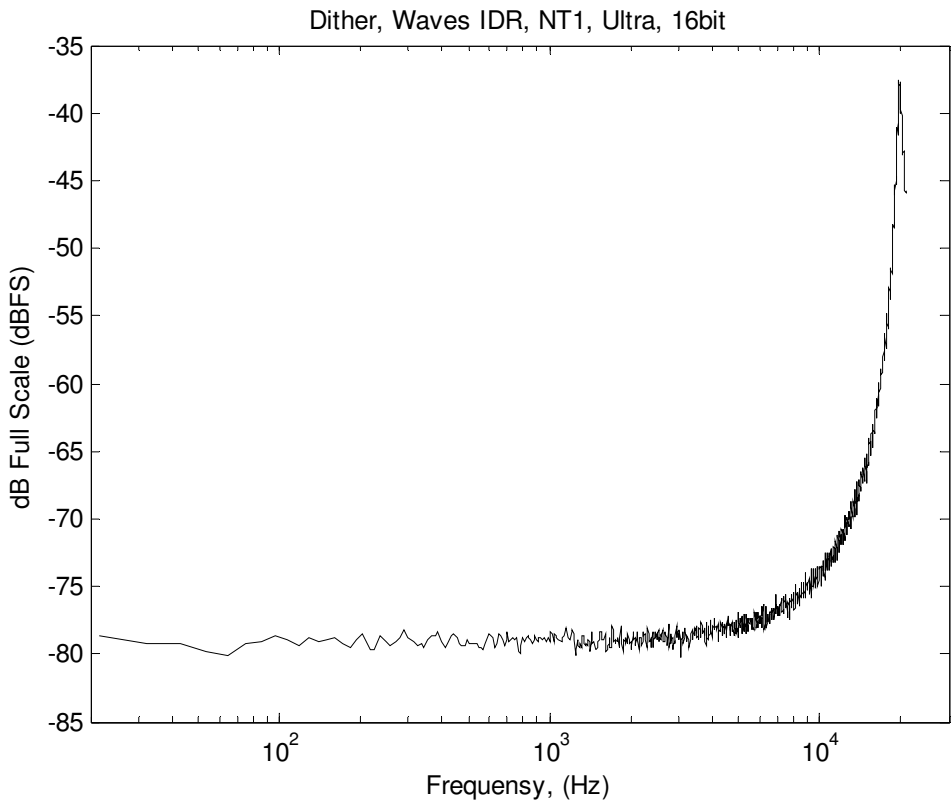


Fig. 39: Dither, Waves IDR, Noise Type 1, Ultra, 16 Bits.

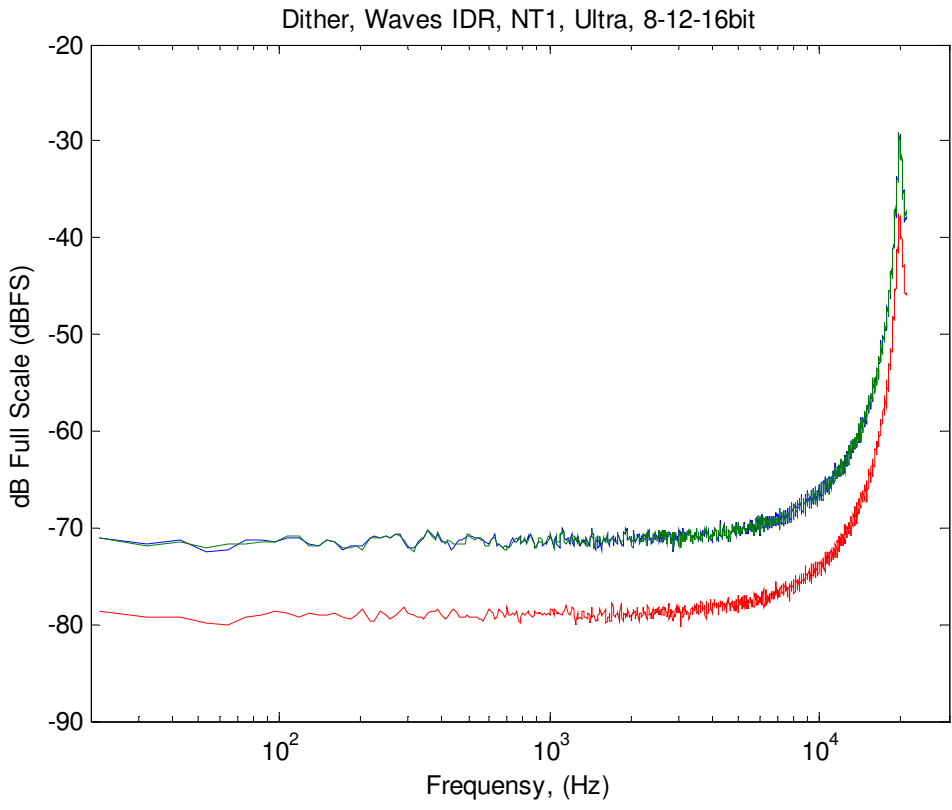


Fig. 40: Dither, Waves IDR, Noise Type 1, Ultra, 8, 12 & 16 Bits.

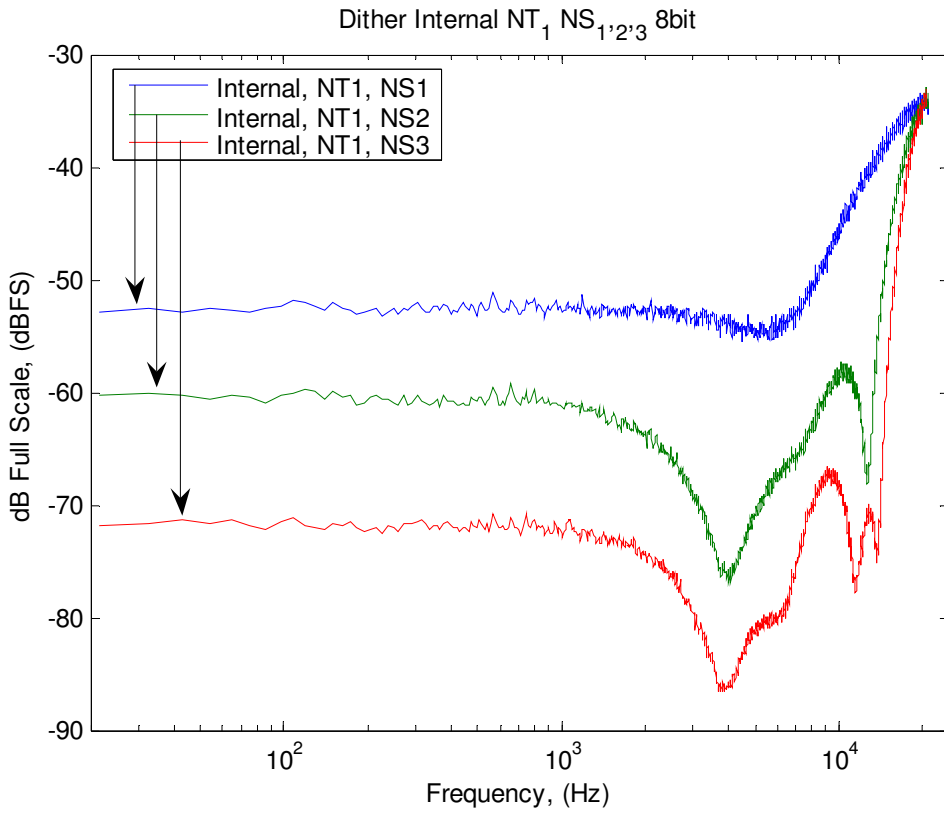


Fig. 41: Dither, Internal, Noise Type 1, Noise Shape 1, 2 & 3 8 Bits.

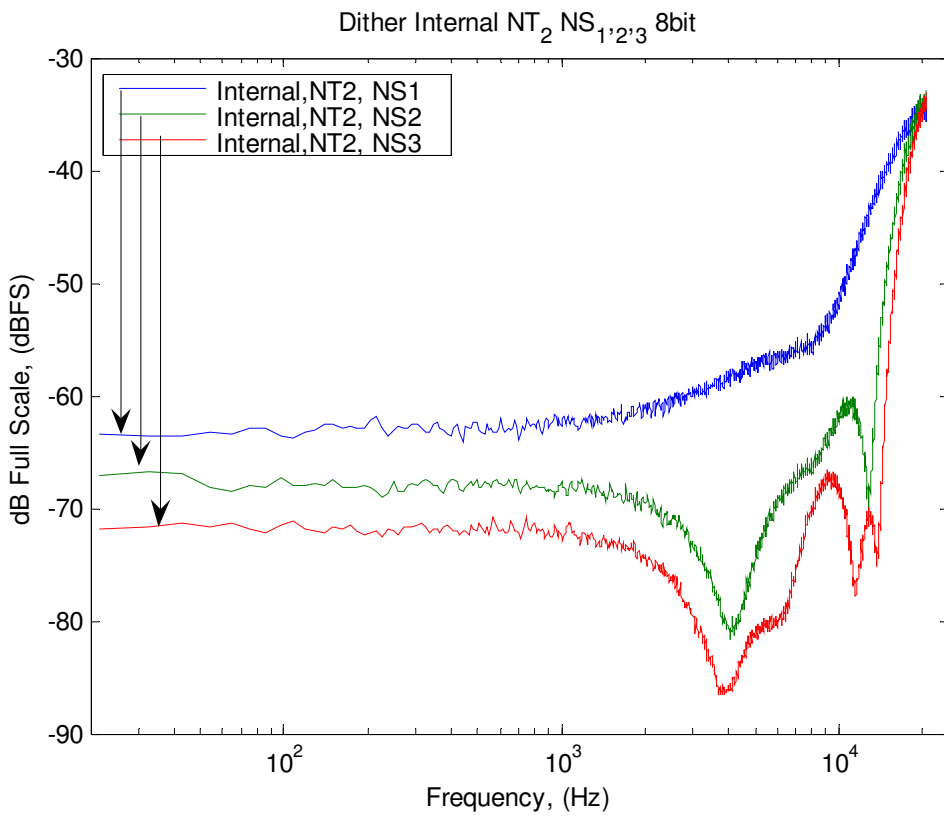


Fig. 42: Dither, Internal, Noise Type 2, Noise Shape 1, 2 & 3 8 Bits.

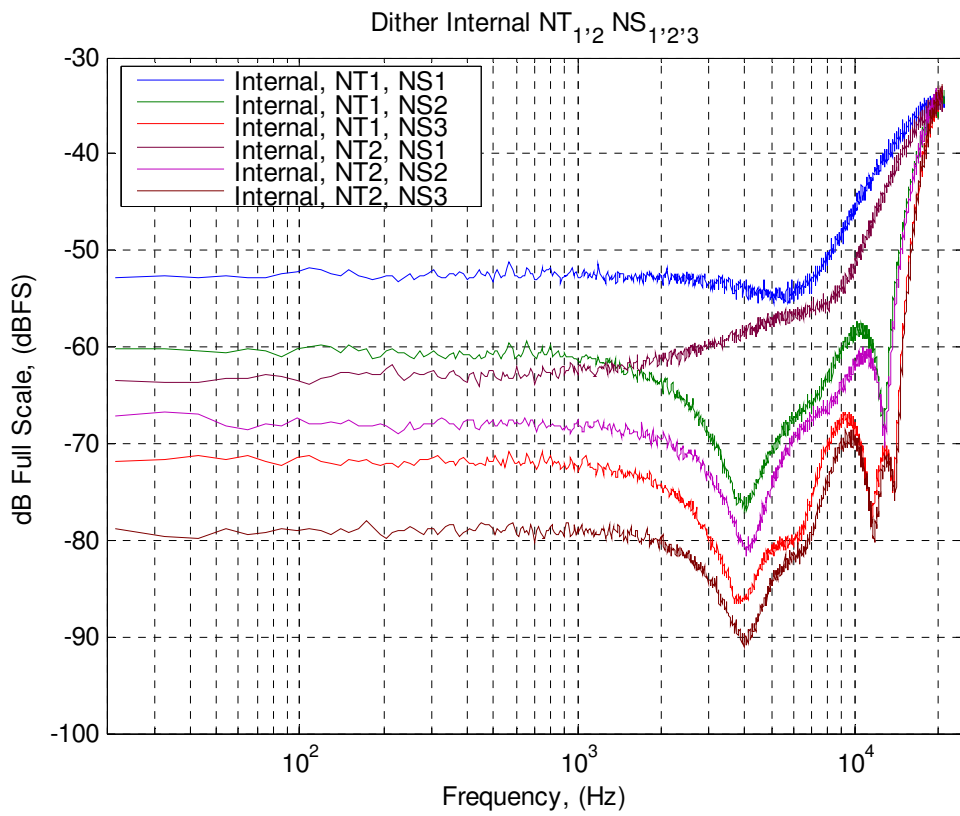


Fig. 43: Dither, Internal, Noise Type 1 & 2, Noise Shape 1, 2 & 3 8 Bits.

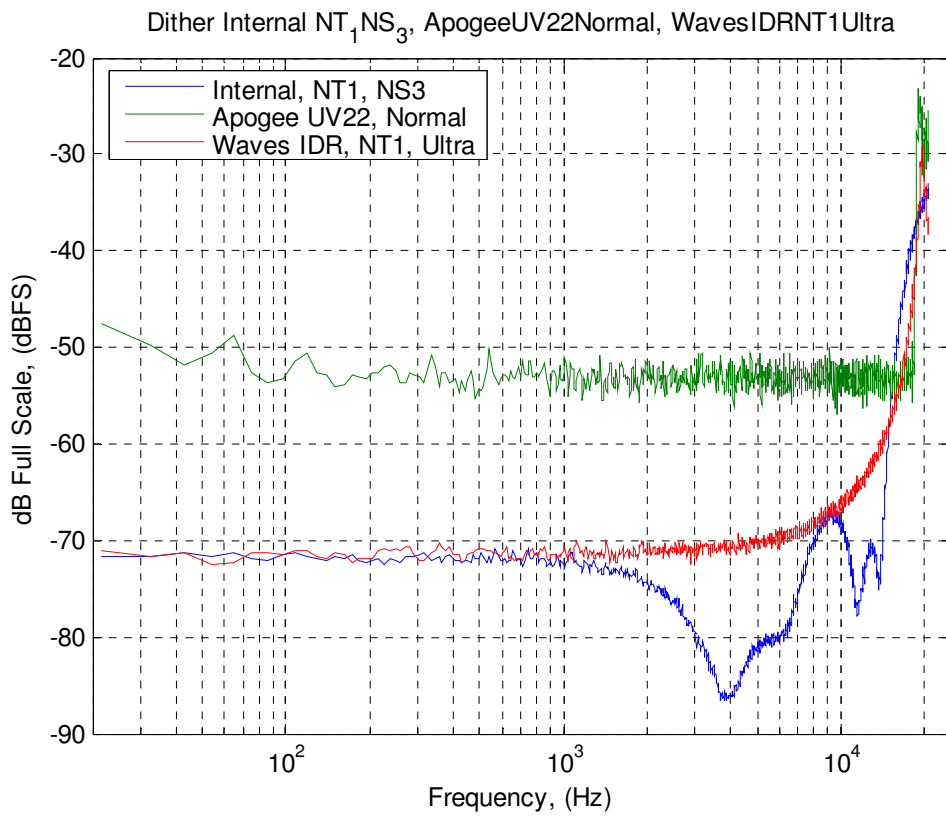


Fig. 44: Dither /Internal, Noise Type 1, Noise Shape 3/Apogee UV22, Normal/ Waves IDR, NT1, Ultra/ 8 Bits.

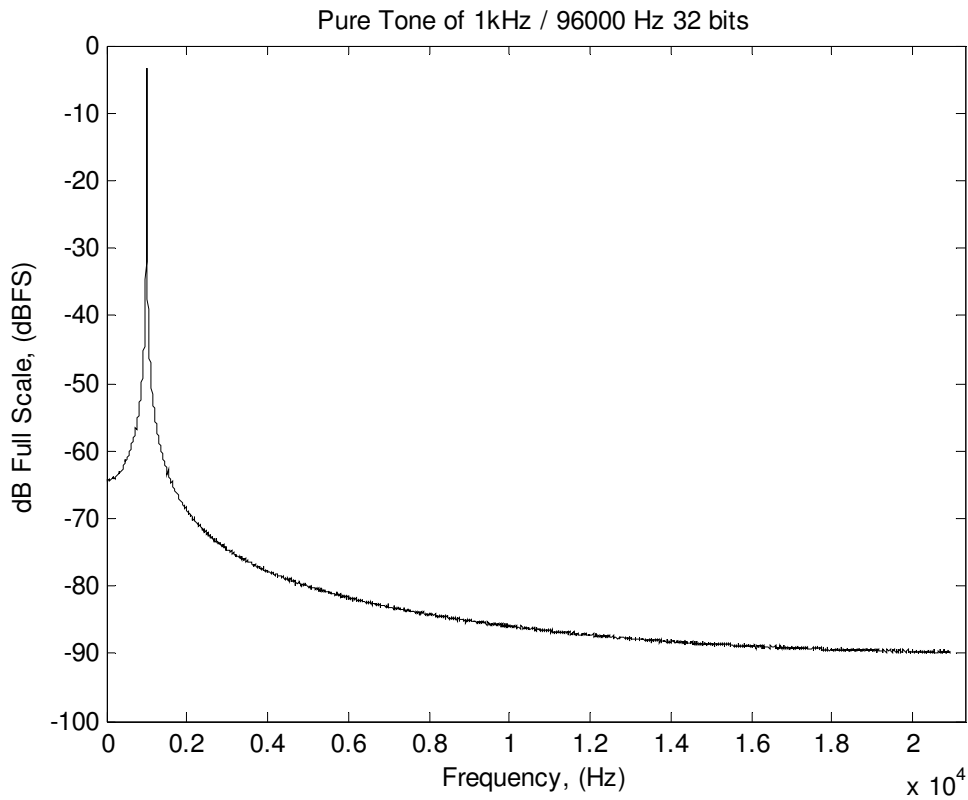


Fig. 45: Καθαρός τόνος του 1 kHz στα 96000 Hz και στα 32 bit.

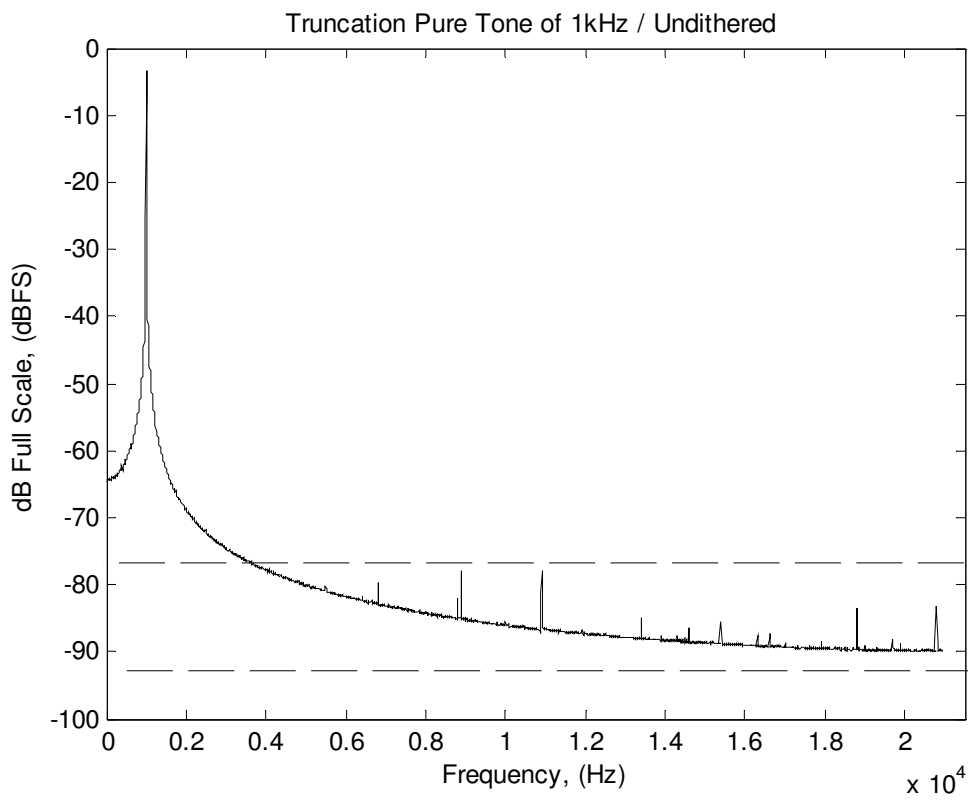


Fig. 46: “κατέβασμα” του τόνου απο τα 96000 Hz και τα 32 bit στα 44100 και τα 16 bit, φαίνονται καθαρά τα σφάλματα που προκύπτουν απο τη στρωγγυλοποίηση που λαμβάνει χώρα κατά τη διάρκεια της κβαντοποίησης.

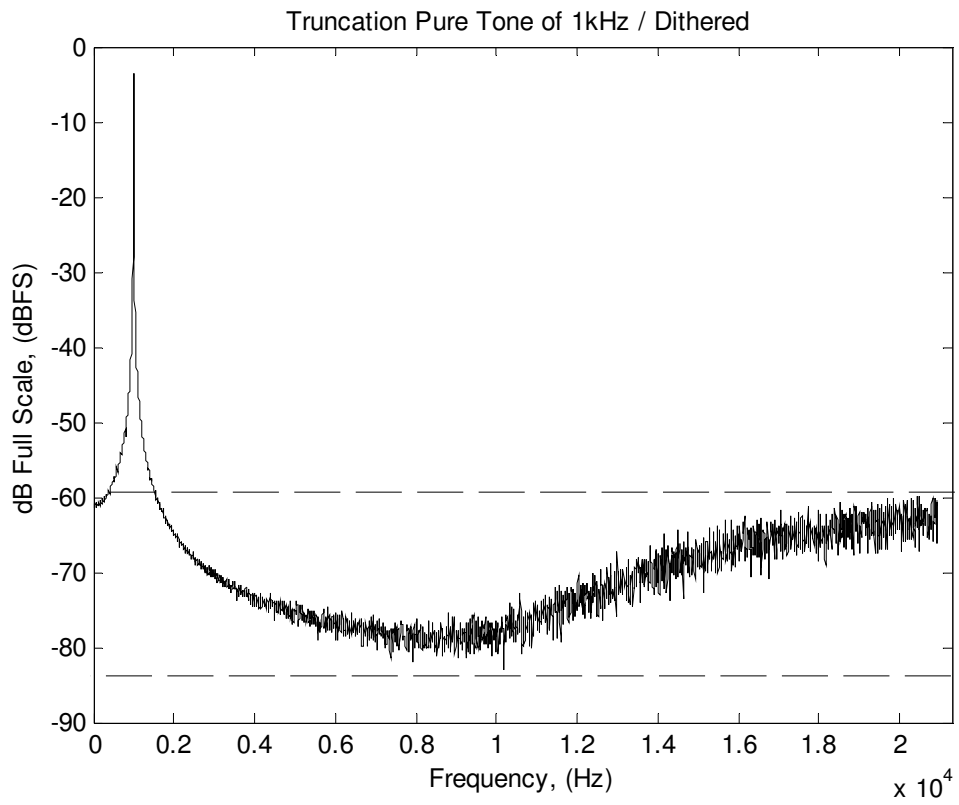


Fig. 47: Απαλοιφή των σφαλμάτων και αύξηση του πλάτους της ενέργειας του καθαρού τόνου με τη χρήση dither.

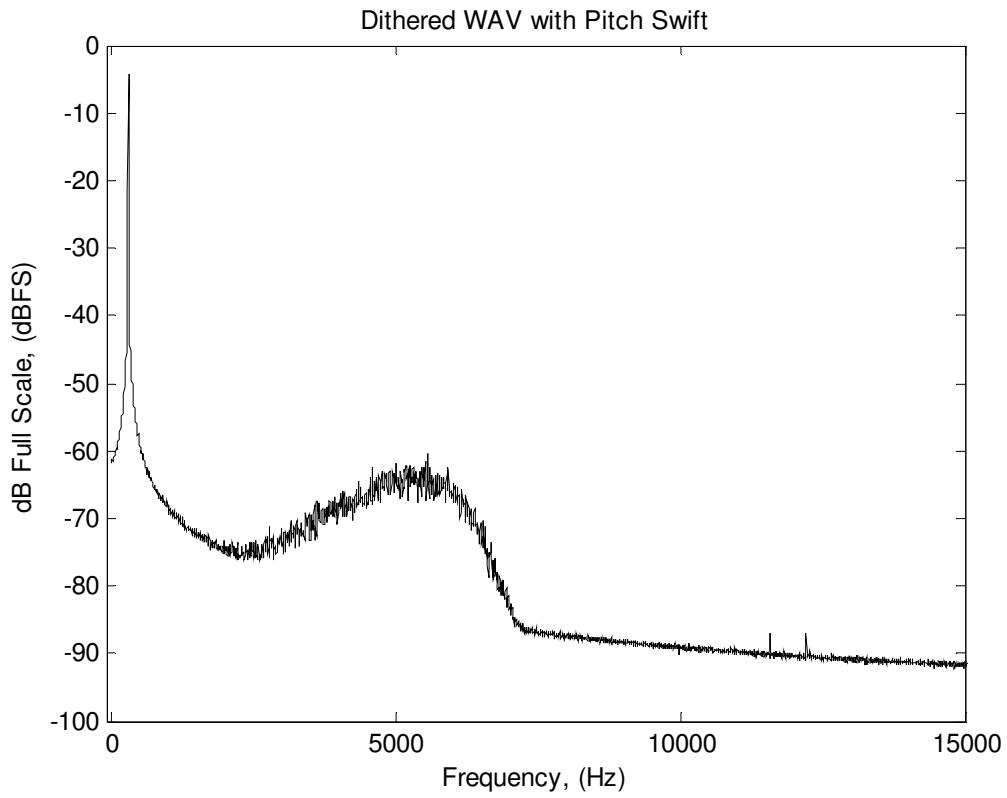


Fig. 48: Ο τόνος του 1 kHz που παρουσιάστηκε στο Σχ. 47 αφού έχει υποστεί μετατόπιση του τονικού του ύψους κατά 20 ημίτονα προς τα κάτω. Παρατηρούμε τη μετατόπιση προς τα κάτω όλου του φάσματος καθώς και του dither.

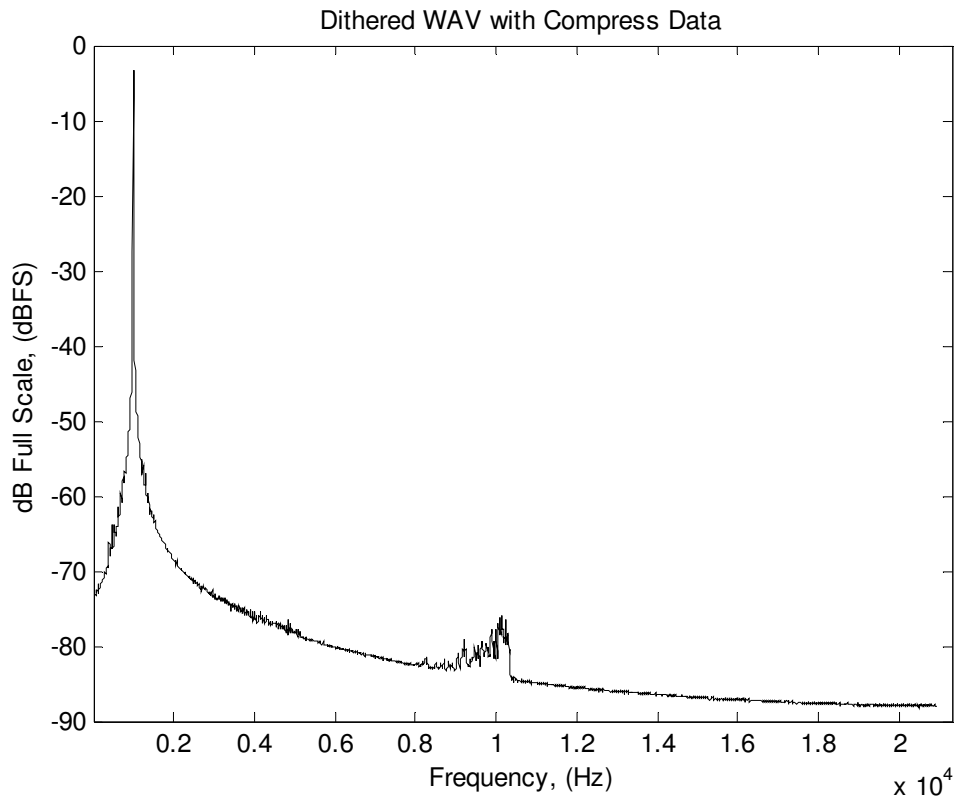


Fig. 49: Ο τόνος του 1 kHz που παρουσιάστηκε στο Σχ. 47 αφού έχει υποστεί συμπίεση των δεδομένων του (mp3_128 kbps) αυτή τη φορά. Παρατηρούμε αφενός πως όλα τα πλεονεκτήματα του dither έχουν εξαφανιστεί από τη συμπίεση των δεδομένων και αφετέρου – και πολύ σημαντικό – πως έχουν εμφανιστεί αλλοιώσεις υπό τη μορφή παραμόρφωσης στο σήμα μας κυρίως γύρω από την περιοχή των 10 kHz. Παρατηρούμε ακόμα τη μείωση του πλάτους της ενέργειας ανά συχνότητα στις υψηλές συχνότητες όπως ήταν αναμενόμενο.