

Ανώτατο Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης
Σχολή Εφαρμοσμένων Επιστήμων
Τμήμα Μηχανικών Μουσικής Τεχνολογίας και Ακουστικής

Τίτλος Πτυχιακής Εργασίας

*«Σύγκριση Audio Engines-Digital Audio Workstations (D.A.W)
με χρήση listening tests» (μέθοδος ABX double blind test)*

Σπουδαστής:

Συμεωνίδης Ιωάννης (Α.Μ 1167)

Επιβλέπων Καθηγητής:

κ. Παπαδάκης Νικόλαος

Περιεχόμενα

Ευχαριστίες.....	5
Πρόλογος.....	6
ΚΕΦΑΛΑΙΟ 1 ^ο : ΑΝΘΡΩΠΙΝΟ ΣΥΣΤΗΜΑ ΑΚΟΗΣ ΚΑΙ ΨΥΧΟΑΚΟΥΣΤΙΚΗ.....	7
1.1 Ανθρώπινο Ακουστικό Σύστημα	7
1.1.1 Εξωτερικό Αυτί.....	7
1.1.2 Μέσο Αυτί.....	8
1.1.3 Εσωτερικό Αυτί	8
1.2 Ψυχοακουστική.....	9
1.3 Critical Bands	9
1.4 Κατώφλι Ακουστότητας	12
1.5 Masking	13
1.5.1 Simultaneous Masking.....	13
1.5.2 Temporal Masking	14
ΚΕΦΑΛΑΙΟ 2 ^ο : ΒΑΣΙΚΕΣ-ΕΙΣΑΓΩΓΙΚΕΣ ΕΝΝΟΙΕΣ ΓΙΑ ΤΑ DAW (Digital Audio Workstations) ΣΥΣΤΗΜΑΤΑ.....	16
2.1 Digital audio workstation (DAW).....	16
2.1.1 Βασικές έννοιες και ανάλυση λειτουργίας ενός DAW.....	17
2.2 Audio Engine ενός DAW.....	18
2.3 Ανάλυση του όρου fader και pan pot.....	18
2.3.1 Faders.....	18
2.3.2 Pan Law.....	19
2.4 Σύγκριση του pan rule/law για κάθε DAW που θα χρησιμοποιήσουμε στην πειραματική διαδικασία.....	20
2.4.1 Logic pro 9.....	20
2.4.2 Pro tools HD 9.....	21
2.4.3 Cubase 5.....	22
2.5 Plug-ins και η σχέση τους με τα DAW συστήματα.....	22
2.5.1 DAW (fixed and floating point μηχανή ήχου).....	24
2.6 Αρχικά θεωρητικά συμπεράσματα για την πειραματική διαδικασία.....	24
ΚΕΦΑΛΑΙΟ 3 ^ο : ΑΝΑΛΥΣΗ ΤΗΣ ΛΕΙΤΟΥΡΓΙΑΣ ΚΑΙ ΤΗΣ ΔΟΜΗΣ ΤΗΣ ΜΗΧΑΝΗΣ ΗΧΟΥ (AUDIO ENGINE) ΕΝΟΣ DAW.....	25
3.1 Ψηφιοποίηση του σήματος.....	25
3.1.1 Δειγματοληψία.....	26
3.1.2 Κβαντισμός (Quantization).....	28
3.1.3 Κωδικοποίηση.....	30
3.2 Fixed και floating point συστήματα διαφορές και ανάλυση των δύο αυτών λειτουργιών.....	31
3.2.1 Fixed point-floating point dsp's.....	32
3.3 Μηχανή ήχου (audio engine) ενός DAW.....	35

3.4 Shifting (μετακίνηση).....	36
3.4.1 Η διαδικασία του shifting αναλυτικά.....	36
3.5 Απαντήσεις σε ερωτήματα που έχουν να κάνουν με τα audio engines.....	37
Κεφάλαιο 4 ^ο : ΜΕΘΟΔΟΣ ABX DOUBLE BLIND TEST.....	39
4.1 Σύντομο ιστορικό abx μεθόδου.....	40
4.1.1 Τι είναι ένα τυφλό πείραμα (blind experiment).....	41
4.1.2 Double-blind trial.....	41
4.2 Ανάλυση της Μεθόδου.....	42
4.2.1 Η Λογική της Μεθόδου.....	42
4.3 Δομή ABX double blind test.....	43
4.3.1 Null hypothesis.....	43
4.3.2 Στατιστικά σημαντικά αποτελέσματα (statistical significance results).....	44
4.3.3 p-value	45
4.3.3.1 Binomial table (πίνακας/κατανομή).....	46
4.3.4 Ανάλυση του binomial πίνακα.....	48
4.4 Επιλογή αντικειμένων.....	49
4.5 Επιλογή κριτηρίων.....	50
4.6 Ενστάσεις και αντιπαραθέσεις που δημιουργούνται κατά την πραγματοποίηση των blind test.....	52
Κεφάλαιο 5 ^ο : ΣΧΕΔΙΑΣΜΟΣ ΤΗΣ ΠΕΙΡΑΜΑΤΙΚΗΣ ΔΙΑΔΙΚΑΣΙΑΣ – ΜΕΘΟΔΟΛΟΓΙΑ..	54
5.1 Ερευνητικά ερωτήματα.....	54
5.2 Σχεδιασμός πειραματικής διαδικασίας.....	55
5.3 Επιλογή των δειγμάτων.....	55
5.3.1 Αντικειμενικό κομμάτι και δείγματα (Null sound test).....	55
5.3.2 Υποκειμενικό κομμάτι και δείγματα (listening tests).....	56
5.4 Τα ζευγάρια των δειγμάτων που δημιουργούνται και η χρήση τους με το foobar2000..	57
5.4.1 Η λειτουργία του λογισμικού foobar2000 (abx comparator).....	58
5.5 Επιλογή Υποκειμένων.....	59
5.6 Εξοπλισμός και λογισμικό.....	61
5.7 Πειραματική διαδικασία.....	61
Κεφάλαιο 6 ^ο : ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΕΡΕΥΝΑΣ-ΣΥΓΚΡΙΣΕΙΣ.....	63
6.1 Αντικειμενικό μέρος πειραματικής διαδικασίας (Null tests).....	63
6.1.1 Πρώτο ζευγάρι: Cubase-Logic (χωρίς plug-ins).....	64
6.1.2 Δεύτερο ζευγάρι: Cubase-Pro tools (χωρίς plug-ins).....	65
6.1.3 Τρίτο ζευγάρι: Logic-Pro tools (χωρίς plug-ins).....	66
6.1.4 Τέταρτο ζευγάρι: Cubase-Logic (με plug-ins).....	67
6.1.5 Πέμπτο ζευγάρι: Cubase-Pro tools (με plug-ins).....	68
6.1.6 Έκτο ζευγάρι: Logic-Pro tools (με plug-ins).....	69
6.1.7 Null test για ημίτονο στα 440Hz (export σε Logic, Pro tools, Cubase).....	70

6.1.8 Global analysis κενών export των τριών DAW στο wavelab	72
6.2 Υποκειμενικό μέρος πειραματικής διαδικασίας (listening tests με την μέθοδο abx double blind)	74
6.2.1 Διατύπωση της μηδενικής υπόθεσης (null hypothesis).....	75
6.2.2 Πρώτο ζευγάρι: Cubase-Logic (χωρίς plug-ins).....	75
6.2.3 Δεύτερο ζευγάρι: Cubase-Pro tools (χωρίς plug-ins).....	76
6.2.4 Τρίτο ζευγάρι: Logic-Pro tools (χωρίς plug-ins).....	77
6.2.5 Τέταρτο ζευγάρι: Cubase-Logic (με plug-ins).....	78
6.2.6 Πέμπτο ζευγάρι: Cubase-Pro tools (με plug-ins).....	79
6.2.7 Έκτο ζευγάρι: Logic-Pro tools (με plug-ins).....	80
Κεφάλαιο 7 ^ο : ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΙ ΤΗΣ ΕΡΕΥΝΑΣ.....	82
7.1 Συμπεράσματα.....	82
7.1.1 Αντικειμενικό κομμάτι.....	82
7.1.2 Υποκειμενικό κομμάτι (ακροάσεις).....	83
7.2 Περιορισμοί της έρευνας.....	84
7.3 Γενικό συμπέρασμα υποκειμενικό πτυχιακής εργασίας.....	84
7.4 Περαιτέρω προτάσεις της έρευνας.....	85
Βιβλιογραφία.....	86
Ελληνική.....	86
Ξενόγλωσση.....	86
ΠΑΡΑΡΤΗΜΑ Α.....	88
ΠΑΡΑΡΤΗΜΑ Β.....	90

Ευχαριστίες

Θα ήθελα να ευχαριστήσω όλους τους ανθρώπους που βοήθησαν προκειμένου να ολοκληρωθεί η παρούσα πτυχιακή εργασία. Πρώτο απ' όλους, ευχαριστώ τον επιβλέποντα καθηγητή μου κ. Παπαδάκη Νίκο, για τη βοήθεια και συμπαράσταση που έδειξε καθ' όλη τη διάρκεια της εργασίας. Επίσης, ευχαριστώ θερμά τον καθηγητή κ. Νίκο Κεφαλογιάννη για την βοήθεια του σε κομμάτια της πτυχιακής εργασίας και τους κ. Νικόλα Βαλσαμάκη και κ. Μίνω Φιτσανάκη για την παραχώρηση του εργαστήριου «Mastering», προκειμένου να διεξαχθούν τα listening tests. Τέλος, ευχαριστώ θερμά όλα τα άτομα τα οποία προσφερθήκαν εθελοντικά να συμπεριληφθούν στο δείγμα της έρευνας.

Πρόλογος

Εδώ και πολλά χρόνια και με την εξέλιξη των υπολογιστών και των προγραμμάτων στον τομέα της μουσικής τεχνολογίας πολλοί υποστηρίζουν ότι υπάρχουν ακουστές διαφορές στον ήχο που παράγει κάθε Digital Audio Workstations (DAW). Το αντικείμενο μελέτης της παρούσας πτυχιακής εργασίας είναι να εξεταστεί αν ισχύει αυτό, δηλαδή κατά πόσο παρατηρούνται διαφορές αρχικά θεωρητικά (τρόπος κατασκευής) στα audio engines του κάθε DAW και στην συνέχεια με χρήση ακουστικών δειγμάτων να δούμε σε πρακτικό επίπεδο αν μπορούν οι ακροατές να καταλάβουν διαφορές στον ήχο. Σκοπός μας δεν είναι να προτείνουμε ποιο είναι το καλύτερο DAW αλλά να ελέγξουμε αν ισχύει αυτός ο ισχυρισμός.

Η εργασία χωρίζεται σε δύο μέρη: το θεωρητικό και το ερευνητικό.

Στο θεωρητικό κομμάτι, αναφέρεται η λειτουργία του ακουστικού μας συστήματος και ορισμένα ψυχοακουστικά φαινόμενα (κεφάλαιο 1^ο), αναλύονται ορισμένες βασικές έννοιες για τα DAW (Digital Audio Workstations) συστήματα και από τί αποτελούνται (κεφάλαιο 2ο), επίσης αναλύεται η δομή και η λειτουργία της μηχανής ήχου (κεφάλαιο 4^ο). Στη συνέχεια, αναλύεται η μέθοδος ABX double blind test (Κεφάλαιο 4^ο) σύμφωνα με την οποία πραγματοποιήθηκαν τα listening tests.

Στο πειραματικό μέρος, περιγράφεται η μεθοδολογία της έρευνας (κεφάλαιο 5ο), τα αποτελέσματα που προέκυψαν (κεφάλαιο 6ο), καθώς και τα συμπεράσματα και οι περιορισμοί της (κεφάλαιο 7ο).

ΚΕΦΑΛΑΙΟ 1^ο: ΑΝΘΡΩΠΙΝΟ ΣΥΣΤΗΜΑ ΑΚΟΗΣ ΚΑΙ ΨΥΧΟΑΚΟΥΣΤΙΚΗ

Εισαγωγή 1^ο Κεφαλαίου

Στο κεφάλαιο αυτό γίνεται αναφορά στις βασικές λειτουργίες του ακουστικού μας συστήματος και ανάλυση των ψυχοακουστικών φαινομένων τα οποία είναι το masking και τα critical bands, τα οποία έχουν σχέση με την διπλωματική καθώς και το κατώφλι ακουστότητας.

Σκοπός του κεφαλαίου αυτού είναι η εισαγωγή σε βασικές έννοιες και λειτουργίες της ανθρώπινης ακοής.

1.1 Ανθρώπινο Ακουστικό Σύστημα

Η ανθρώπινη ακοή είναι η πιο ανεπτυγμένη μας αίσθηση και τα όργανα της έχουν αξιοθαύμαστα χαρακτηριστικά. Για παράδειγμα, το αυτί μας έχει πολύ μεγάλο δυναμικό εύρος μιας και το κατώφλι του πόνου για την ένταση των ήχων είναι στα 120dB_{SPL}. Όσον αφορά τη συχνότητα, το αυτί μας αντιλαμβάνεται ένα εύρος συχνοτήτων το οποίο μπορεί να φτάσει από τα 20Hz έως τα 20kHz.

Το ανθρώπινο ακουστικό σύστημα είναι ένας περίπλοκος μηχανισμός με συγκεκριμένα χαρακτηριστικά και ιδιότητες. Το σύστημα αυτό αντιλαμβάνεται τους ήχους του περιβάλλοντός μας, απορρίπτει ήχους που δε μας ενδιαφέρουν, προσαρμόζεται σε πολύ δυνατούς ήχους και έχει μηχανισμό αυτοπροστασίας για την αποφυγή βλάβης. Στο σημείο αυτό θα αναφερθεί η λειτουργία του ακουστικού μας συστήματος περιληπτικά μιας και είναι απαραίτητο να γνωρίζουμε τις βασικές λειτουργίες προκειμένου να προχωρήσουμε σε ανάλυση της πτυχιακής εργασίας.

Τα τμήματα που συνθέτουν το σύστημα ακοής μας είναι τρία: το εξωτερικό αυτί, το μέσο αυτί και το εσωτερικό αυτί. Καθένα από αυτά αποτελείται από διάφορα μικρά τμήματα τα οποία επιτελούν συγκεκριμένες διεργασίες.

1.1.1 Εξωτερικό Αυτί

Το εξωτερικό αυτί αποτελείται από το πτερύγιο, το ακουστικό κανάλι και το τύμπανο. Το πτερύγιο είναι υπεύθυνο για τη συλλογή των ήχων από το περιβάλλον και συμβάλλει σημαντικά στον εντοπισμό της θέσης μιας ηχητικής πηγής. Λόγω της θέσης και του σχήματός του διαφοροποιεί σημαντικά την ηχητική πληροφορία που εισέρχεται σε αυτό.

Το ακουστικό κανάλι έχει μήκος 3 εκατοστά και έχει την ικανότητα να αυξάνει την ηχηρότητα των συλλεγμένων ήχων. Είναι ουσιαστικά ένας φυσικός ενισχυτής με ικανότητα ενίσχυσης κοντά στα 10 με 15dB_{SPL}. Είναι επίσης υπεύθυνο για την προστασία του μέσου αυτιού από εξωτερικούς κινδύνους (εξωτερικά σώματα, αλλαγές στη θερμοκρασία κλπ).

1.1.2 Μέσο Αυτί

Το μέσο αυτί αποτελείται από το τύμπανο, τα οστά και τους μύες του μέσου αυτιού, την κοιλότητα και την ευσταχιακή σάλπιγγα (Παπαδάκης, 2009, 11).

Το τύμπανο είναι μια μεμβράνη η οποία έχει την ιδιότητα να αυξομειώνει την καμπυλότητα και τη σκληράδα της ανάλογα με την ένταση του ήχου που δέχεται. Η ικανότητα αυτή ονομάζεται *aural reflex* και προστατεύει το τύμπανο από την παραμόρφωση σε περίπτωση απότομης αύξησης της έντασης του ήχου.

Τα οστά του μέσου αυτιού είναι τρία: η σφύρα, ο άκμονας και ο αναβολέας. Ενώνονται μεταξύ τους με δύο μύες. Η σφύρα είναι ενωμένη και με το τύμπανο και ο αναβολέας με το οβάλ παράθυρο. Τα οστά αυτά είναι υπεύθυνα για τη μεταφορά της ενέργειας από την ταλάντωση του τυμπάνου στο οβάλ παράθυρο και από εκεί στον κοχλία. Οι μύες που βρίσκονται ενδιάμεσα επιτελούν διάφορες λειτουργίες όπως ο έλεγχος της έντασης για πολύ δυνατούς ήχους χαμηλής συχνότητας, ελάττωση της αντίληψης των ήχων που παράγει το ανθρώπινο σώμα κ.α.

Η ευσταχιακή σάλπιγγα συνδέει το μέσο αυτί με το άνω μέρος του λαιμού και η βασική της λειτουργία είναι η εξίσωση της στατικής πίεσης του αέρα στο μέσο αυτί με την εξωτερική ατμοσφαιρική πίεση προκειμένου να λειτουργεί σωστά το τύμπανο. Μια ακόμα λειτουργία της είναι η απαγωγή υγρών σε περίπτωση μόλυνσης του αυτιού.

1.1.3 Εσωτερικό Αυτί

Το εσωτερικό αυτί αποτελείται από διάφορα μέρη. Οι δυνατότητες του εσωτερικού αυτιού είναι πολλές και εκτείνονται από την αντίληψη των σωματικών κινήσεων μέχρι και την ικανότητα ανάλυσης κυματικών μετακινήσεων των ηχητικών πιέσεων.

Ο μεμβρανώδης λαβύρινθος περιλαμβάνει δύο όργανα: το αιθουσαίο και τον κοχλία. Το αιθουσαίο είναι υπεύθυνο για την αντίληψη της στάσης του σώματός στους τρεις άξονες του χώρου. Ο κοχλίας είναι ίσως το σημαντικότερο όργανο της ακοής μας και αποτελείται από πολλά μικρότερα τμήματα με συγκεκριμένες λειτουργίες. Τα σημαντικότερα από αυτά είναι η βασική μεμβράνη, τα τριχοειδή κύτταρα και το όργανο του Corti.

Το ηχητικό σήμα ταξιδεύει από το τύμπανο, μέσω των προαναφερθέντων οργάνων και καταλήγει στο ακουστικό νεύρο από όπου και στέλνεται στον εγκέφαλο. Η αντίληψη της συχνότητας του ήχου προέρχεται από τη συχνότητα στην οποία συντονίζεται κάθε φορά η βασική μεμβράνη. Καθώς τα κύματα «ταξιδεύουν» κατά μήκος της μεμβράνης προκαλούν ταλάντωση αυτής, της οποίας το πλάτος εξαρτάται από την ένταση του και την συχνότητα του ερεθίσματος. Συγκεκριμένα, οι περιοχές της βασικής μεμβράνης κοντά στην βάση του κοχλία παρουσιάζουν μέγιστη διέγερση στις υψηλές συχνότητες, οι περιοχές κοντά στην κορυφή στις χαμηλές, ενώ οι μεσαίες περιοχές στις μεσαίες συχνότητες. Έτσι ένα ερέθισμα 10KHz θα διεγείρει κυρίως την βασική μεμβράνη κοντά στην βάση του κοχλία, των 2KHz στη μεσότητα του κοχλία, και των 200Hz στην κορυφή του κοχλία. Αυτή η κατανομή των συχνοτήτων πάνω στην βασική μεμβράνη είναι λογαριθμική. Χάριν σ' αυτές τις μηχανικές ιδιότητες της βασικής μεμβράνης, επιτυγχάνεται η διάκριση των συχνοτήτων.

1.2 Ψυχοακουστική

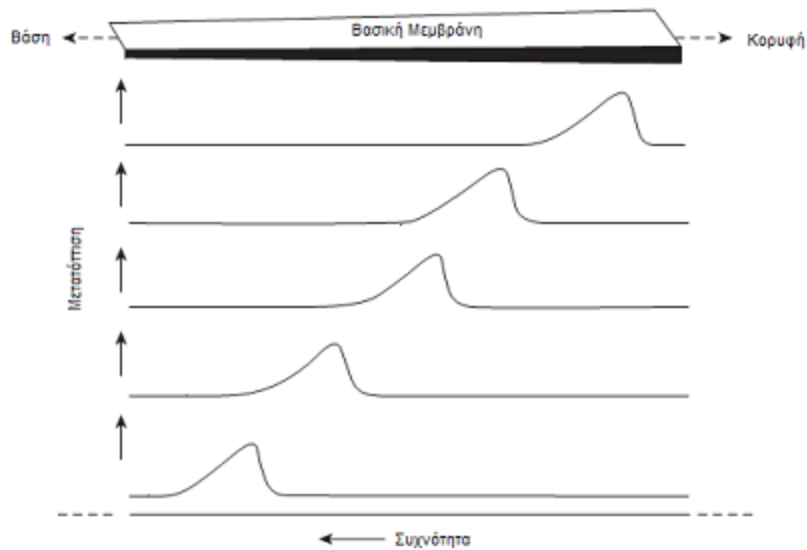
Η Ψυχοακουστική είναι ένας όρος ο οποίος περιλαμβάνει τη μελέτη κατασκευής και λειτουργίας του αυτιού, τις διαδρομές κίνησης του ήχου, την αντίληψή του καθώς και τις μεταξύ τους σχέσεις. Σαν ορισμό, λοιπόν, θα μπορούσαμε να πούμε πως η ψυχοακουστική είναι ο τομέας της επιστήμης ο οποίος εξετάζει τη σύνδεση ανάμεσα στις φυσικές ιδιότητες των ακουστικών ερεθισμάτων και τις φυσιολογικές και ψυχολογικές αντιδράσεις που αυτά προκαλούν στον άνθρωπο. Μέσω αυτής εξάγεται το συμπέρασμα πως η σημασία που δίνεται στον ήχο από τους ανθρώπους προέρχεται αποκλειστικά και μόνο από την αντίληψη που έχουν οι άνθρωποι γι' αυτόν. Διαφορετικά, ο ήχος θα ήταν μια αδιάφορη και ασαφής ακαδημαϊκή έννοια.

Ο τομέας της ψυχοακουστικής δίνει απαντήσεις σε ερωτήματα που εγείρονται σχετικά με τον τρόπο που ακούμε, αποκαλύπτοντας τη μη γραμμικότητα του αυτιού μας και τα φαινόμενα που συνεπάγονται της ιδιότητας αυτής.

1.3 Critical Bands

Το αυτί μας δέχεται ανά πάσα στιγμή πολλά ταυτόχρονα ακουστικά ερεθίσματα από διαφορετικές ηχητικές πηγές. Πόσο εύκολα όμως μπορεί να ξεχωρίσει δύο συχνότητες, ειδικά όταν αυτές είναι κοντινές; Και ποια είναι τα όρια μέσα στα οποία αντιλαμβανόμαστε τη διαφορά αυτή;

Κάθε φορά που ένα ακουστικό ερέθισμα φτάνει στο αυτί μας, η βασική μεμβράνη διεγείρεται σε συγκεκριμένα σημεία ανάλογα με τη συχνότητα του ακουστικού ερεθίσματος, όπως φαίνεται και στην Εικόνα 1.1.



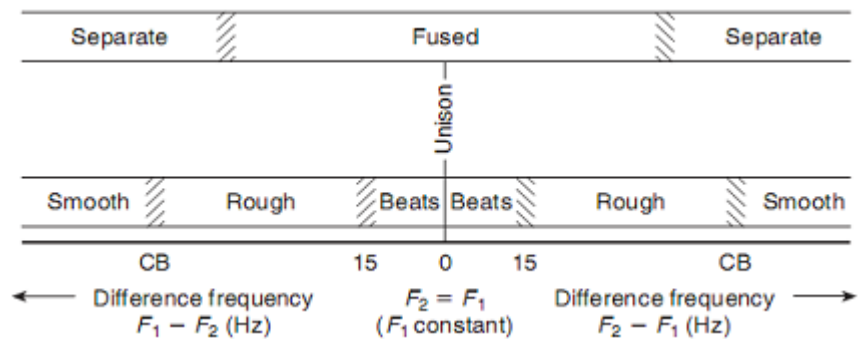
Εικόνα 1.1: Διέγερση της βασικής μεμβράνης ανάλογα με τη συχνότητα¹

Η διάκριση της συχνότητας ανάμεσα σε δύο ακουστικά ερεθίσματα με κοντινές συχνότητες εξαρτάται από τη θέση στην οποία βρίσκεται η μετατόπιση της βασικής μεμβράνης για κάθε ένα από τα ερεθίσματα. Εάν οι μετατοπίσεις είναι σε διαφορετικά σημεία, τότε η διάκριση είναι εφικτή. Επί παραδείγματι, υποθέτουμε πως έχουμε δύο ημιτονοειδή σήματα με ίδιο πλάτος και συχνότητες f_1 και f_2 που συνηχούν. Η f_1 παραμένει σταθερή ενώ η f_2 μεταβάλλεται σταδιακά, ξεκινώντας από την τιμή της f_1 .

Όταν η f_2 έχει την ίδια τιμή με την f_1 τότε ακούγεται μια και μόνο συχνότητα. Όταν η f_2 αρχίζει να μεταβάλλεται προς τα πάνω ή προς τα κάτω, για ένα συγκεκριμένο εύρος τιμών, δημιουργείται διακρότημα. Το διακρότημα μας δίνει την αίσθηση της ύπαρξης μιας συχνότητας αλλά κάπως θολής και μπερδεμένης. Όσο μεταβάλλεται περισσότερο η f_2 και περνάει τα όρια του διακροτήματος, το αυτί μας αντιλαμβάνεται δύο συχνότητες με την αίσθηση όμως της σκληράδας στον ήχο τους. Καθώς η διαφορά των f_1 και f_2 μεγαλώνει περισσότερο, η αίσθηση της σκληράδας εξαφανίζεται και το αυτί μας αντιλαμβάνεται πλέον δυο καθαρές διαφορετικές συχνότητες.

Για την πλειοψηφία των ανθρώπων, διακροτήματα σχηματίζονται όταν η διαφορά των δύο τόνων είναι μικρότερη από 12.5Hz και η αίσθηση της σκληράδας εμφανίζεται όταν η διαφορά των δύο συχνοτήτων ξεπεράσει τα 15Hz. Όσο περισσότερο για συγκεκριμένο εύρος συχνοτήτων αυξάνεται η συχνοτική διαφορά, η αίσθηση της σκληράδας εξαφανίζεται και ακούγονται άνετα δύο τόνοι διαφορετικής συχνότητας, όπως φαίνεται στην Εικόνα 1.2.

¹ Howard, D. M. & Angus J. (2006). Acoustics and Psychoacoustics, 3rd edition Focal Press 73.



Εικόνα 1.2: Critical bands²

Το σημείο όπου η σκληράδα και η άνετη ακρόαση των ήχων διαχωρίζεται δεν έχει σταθερή τιμή αλλά διαφέρει από άνθρωπο σε άνθρωπο, ενώ ανάλογα με τη συχνότητα αυξομειώνεται και το εύρος της σκληρής περιοχής το οποίο υπολογίζεται σύμφωνα με την παρακάτω σχέση:

$$ERB = \{24.7 * [(4.37 * f_c) + 1]\} \quad (1.1)$$

όπου f_c είναι η κεντρική συχνότητα και ERB (Equivalent Rectangular Bandwidth) το εύρος σε Hz.

Με την πάροδο της αίσθησης της σκληράδας σηματοδοτείται η έναρξη της επόμενης κρίσιμης περιοχής. Το ακουστικό μας φάσμα χωρίζεται σε 24 περιοχές για συχνότητες έως και 15kHz, ενώ υπάρχει και μια τελευταία, 25η περιοχή, η οποία εκτείνεται από τα 15kHz ως τα 20kHz. Στον Πίνακα 1.1 φαίνεται ένα παράδειγμα διαχωρισμού του ακουστικού μας φάσματος σε κρίσιμες ζώνες.

² Παπαδάκης Νίκος 2009. Πανεπιστημιακές σημειώσεις για το μάθημα «εισαγωγή στην Ψυχοακουστική ΑΤΕΙ Κρήτης». 57.

Critical band number	Center frequency(Hz)	Critical band(Hz)	Lower cutoff frequency(Hz)	Upper cutoff frequency(Hz)
1	50	-	-	100
2	150	100	100	200
3	250	100	200	300
4	350	100	300	400
5	450	110	400	510
6	570	120	510	630
7	700	140	630	770
8	840	150	770	920
9	1000	160	920	1080
10	1170	190	1080	1270
11	1370	210	1270	1480
12	1600	240	1480	1720
13	1850	280	1720	2000
14	2150	320	2000	2320
15	2500	380	2320	2700
16	2900	450	2700	3150
17	3400	550	3150	3700
18	4000	700	3700	4400
19	4800	900	4400	5300
20	5800	1100	5300	6400
21	7000	1300	6400	7700
22	8500	1800	7700	9500
23	10500	2500	9500	12000
24	13500	3500	12000	15500
25	18775	6550	15500	22050

Πίνακας 1.1: Διαχωρισμός ακουστικού φάσματος σε κρίσιμες ζώνες³

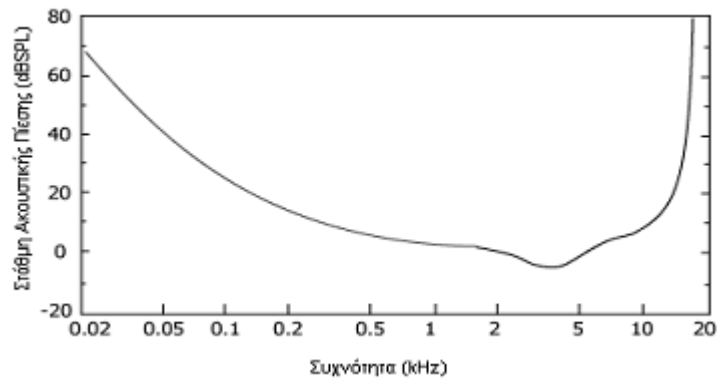
1.4 Κατώφλι Ακουστότητας

Το κατώφλι ακουστότητας κατατάσσεται στην κατηγορία των ορίων ακουστότητας του ανθρώπινου ακουστικού συστήματος και σχετίζεται με την ένταση του ήχου. Συγκεκριμένα, το κατώφλι ακουστότητας ενός ήχου ορίζεται ως η ελάχιστη αντιληπτή ένταση του ήχου αυτού, χωρίς την παρουσία άλλων ήχων (Παπαδάκης, 2009, 42).

³ Pohlmann, K. C. (2011). Principles of digital audio. 6th edition. 342

Στο κατώφλι ακουστότητας αντιστοιχούν τα 0dB SPL (20μPa), στη συχνότητα των 4kHz. Οποιοσδήποτε ήχος κάτω από το όριο αυτό δε γίνεται αντιληπτός. Είναι σημαντικό να σημειωθεί ότι το κατώφλι ακουστότητας δεν είναι σταθερό αλλά μεταβάλλεται ανάλογα με τη συχνότητα του ήχου, όπως φαίνεται στην Εικόνα 1.3.

Εύκολα γίνεται αντιληπτό ότι το αυτί μας αντιλαμβάνεται πιο εύκολα τις συχνότητες που βρίσκονται στην περιοχή από 2 έως 5kHz ενώ απαιτούνται πολύ μεγαλύτερες εντάσεις για τις πολύ χαμηλές και πολύ υψηλές συχνοτικές περιοχές.



Εικόνα 1.3: Κατώφλι ακουστότητας συναρτήσει της συχνότητας

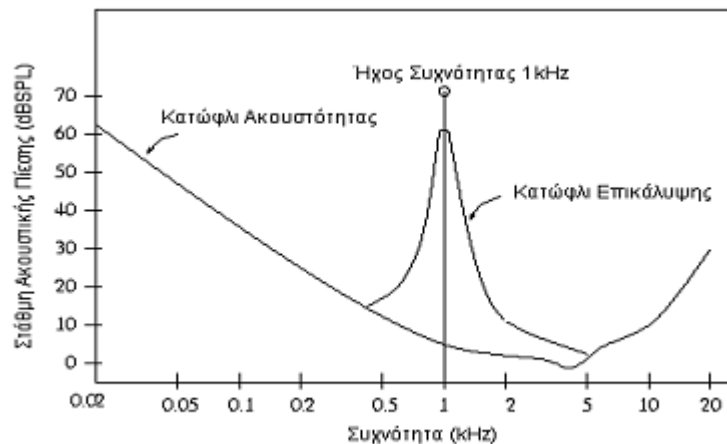
1.5 Masking

Όπως αναφέρθηκε νωρίτερα, μια ηχητική πληροφορία της οποίας η ένταση δεν ξεπερνά το κατώφλι της ακουστότητας δε γίνεται αντιληπτή από το σύστημά μας. Παρ'όλα αυτά, υπάρχουν περιπτώσεις που μεν ο ήχος έχει αρκετή ένταση αλλά και πάλι δε γίνεται αντιληπτός. Αυτό μπορεί να συμβεί αν ο ήχος αυτός επικαλυφθεί από κάποιον άλλο. Η επικάλυψη αυτή ονομάζεται masking και λαμβάνει χώρα υπό ορισμένες προϋποθέσεις, που αφορούν το χρόνο και την τονικότητα των ήχων.

1.5.1 Simultaneous Masking

Η επικάλυψη αυτή λαμβάνει χώρα όταν ένας ήχος καθίσταται μη ακουστός λόγω κάποιου άλλου ήχου της ίδιας διάρκειας που συνηχεί με τον πρώτο. Ο ήχος που επικαλύπτεται ονομάζεται maskee και ο ήχος που επικαλύπτει ονομάζεται masker. Οι λόγοι για τους οποίους προκύπτει το φαινόμενο αυτό σχετίζονται με τη συχνότητα και το πλάτος των masker και maskee και συνδέεται άμεσα με τη λειτουργία της βασικής μας μεμβράνης. Σε γενικές γραμμές, ένας ήχος με πλάτος A και συχνότητα f επικαλύπτει έναν ήχο πλάτους $B < A$ και παραπλήσιας στην f συχνότητας. Έχει παρατηρηθεί ότι ήχοι με ψηλές συχνότητες επικαλύπτονται πολύ πιο εύκολα από ήχους με χαμηλές συχνότητες.

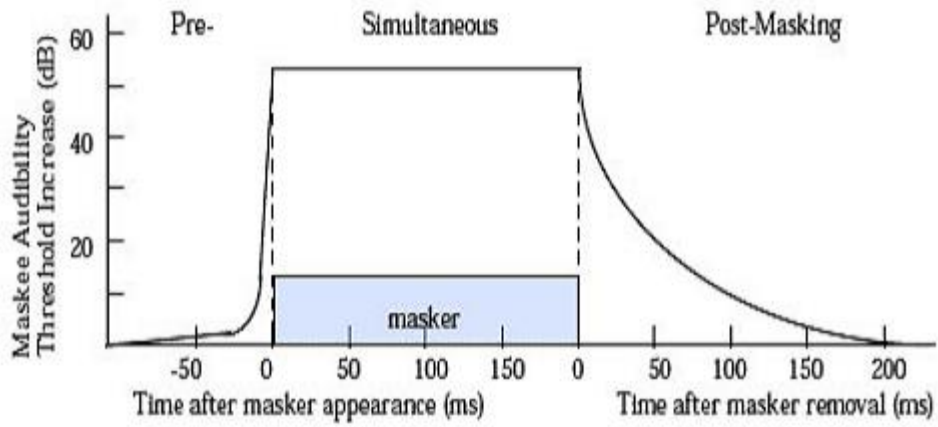
Στην Εικόνα 1.4 φαίνεται ένα παράδειγμα της δυνατότητας επικάλυψης ενός τόνου με συχνότητα 1kHz. Το κατώφλι ακουστότητας δείχνει τις απαιτούμενες εντάσεις για κάθε συχνότητα προκειμένου ο ήχος να γίνει αντιληπτός όταν αυτός δε συνηχεί με άλλους ήχους. Το κατώφλι επικάλυψης δείχνει τις εντάσεις που χρειάζονται οι ήχοι για να γίνουν αντιληπτοί όταν συνηχούν με έναν τόνο συχνότητας 1kHz. Από το παραπάνω παράδειγμα γίνεται αντιληπτό ότι οι παραπλήσιες στο masker συχνότητες χρειάζονται αρκετά μεγαλύτερη ένταση για να ακουστούν σε σχέση με το κατώφλι ακουστότητας.



Εικόνα 1.4: Περιοχή επικάλυψης για ήχο συχνότητας 1kHz

1.5.2 Temporal Masking

Η επικάλυψη αυτή προκύπτει όταν οι ήχοι ακούγονται διαδοχικά σε μικρή χρονική απόσταση. Χωρίζεται σε δύο υποκατηγορίες: το backward masking στο οποίο ο masker ακολουθεί το maskee και το forward masking στο οποίο ο masker προηγείται του maskee. Όσο κοντινότερη είναι η χρονική απόσταση των δύο τόνων, τόσο πιο αποτελεσματική είναι η επικάλυψη. Το forward masking λαμβάνει χώρα όταν η διαφορά μεταξύ των ηχητικών ερεθισμάτων είναι έως 200ms και εξαρτάται από την ένταση και τη διάρκεια του masker ενώ το backward masking συμβαίνει για διαφορά έως 50ms (Εικόνα 1.5).



Εικόνα 1.5: Temporal Masking⁴

⁴ Jobert S. Jacaba. (2001). Audio compression using modified discrete cosine transform: The MP3 coding standard. Undergraduate research paper. 20.

ΚΕΦΑΛΑΙΟ 2ο: ΒΑΣΙΚΕΣ-ΕΙΣΑΓΩΓΙΚΕΣ ΕΝΝΟΙΕΣ ΓΙΑ ΤΑ DAW (Digital Audio Workstations) ΣΥΣΤΗΜΑΤΑ

Εισαγωγή 2^ο Κεφαλαίου

Στο κεφάλαιο αυτό γίνεται αναφορά σε βασικές έννοιες σχετιζόμενες με τα DAW συστήματα, την φιλοσοφία τους, την δομή τους αλλά και τα χαρακτηριστικά τους.

Σκοπός αυτού του κεφαλαίου είναι η κατανόηση των χρήσιμων αυτών εννοιών οι οποίες κρίνονται απαραίτητες για να καταλάβουμε πως λειτουργούν αλλά και θεωρία που θα μας βοηθήσει στο πειραματικό μέρος.

2.1 Digital audio workstation (DAW)

Με τα αρχικά DAW θα αναφερόμαστε από εδώ και πέρα για συντομία στον όρο (Digital Audio Workstation). Είναι ένας όρος που χρησιμοποιήθηκε για πρώτη φορά στη δεκαετία του '80 για να περιγράψει τις "tapeless" μηχανές εγγραφής/ δειγματοληψίας, όπως το Fairlight και synclavier. Σήμερα, η λέξη DAW χρησιμοποιείται για να περιγράψει προγράμματα ήχου και MIDI λογισμικού, όπως Cubase, Logic Pro, Pro tools, Reaper κτλ. Τα DAW είναι ηλεκτρονικά συστήματα σχεδιασμένα αποκλειστικά ή κατά κύριο λόγο για την καταγραφή-επεξεργασία και αναπαραγωγή ψηφιακού ήχου. Ήταν αρχικά tapeless συστήματα βασιζόμενα σε μικροεπεξεργαστή, όπως το synclavier, είναι λογισμικά (software) που εγκαθίστανται και τρέχουν στον υπολογιστή χρησιμοποιώντας κάρτες ήχου ή audio converters.

Ένα ολοκληρωμένο DAW αποτελείται από μια κονσόλα μίξης, επιφάνειες ελέγχου (control surface), μετατροπέα ήχου (audio converter), και μια συσκευή αποθήκευσης δεδομένων (σκληρός δίσκος). Ένα DAW που τρέχει σε υπολογιστή έχει τέσσερις βασικές συνιστώσες: έναν υπολογιστή, μία κάρτα ήχου ή audio converters (Analog to digital - Digital to Analog), ένα ψηφιακό λογισμικό επεξεργασίας ήχου, και τουλάχιστον μία συσκευή εισόδου για την προσθήκη ή τροποποίηση των μουσικών δεδομένων. Αυτό θα μπορούσε να είναι τόσο απλό όσο ένα ποντίκι, και τόσο περίπλοκο όσο ένα MIDI keyboard ή μία κονσόλα με αυτοματοποιημένα fader για την μίξη των διαφόρων καναλιών. Ο υπολογιστής ενεργεί ως διαχειριστής/συντονιστής (host) για την κάρτα ήχου και το λογισμικό και παρέχει την επεξεργαστική ισχύ για την επεξεργασία ήχου. Η εξωτερική κάρτα ήχου ή ο audio converter μετατρέπει τα αναλογικά σήματα ήχου σε ψηφιακή μορφή όταν πρόκειται για εγγραφή μουσικών δεδομένων και για την αναπαραγωγή μετατρέπει ψηφιακά δεδομένα που εξάγονται από το πρόγραμμα σε αναλογικό ήχο στα ηχεία. Μπορεί επίσης να βοηθήσει στην περαιτέρω επεξεργασία του ήχου πχ. όταν χρησιμοποιούμε plug-ins. Το λογισμικό ελέγχει όλες τις σχετικές συνιστώσες των περιφερειακών συσκευών (hardware) και παρέχει ένα περιβάλλον εργασίας για να καταστεί δυνατή η καταγραφή, επεξεργασία και αναπαραγωγή των ηχητικών δεδομένων. Τα περισσότερα DAW που τρέχουν σε υπολογιστή μας παρέχουν εκτεταμένη MIDI εγγραφή, επεξεργασία και δυνατότητα αναπαραγωγής, τέλος μερικά έχουν ακόμη και δυνατότητα επεξεργασίας βίντεο.

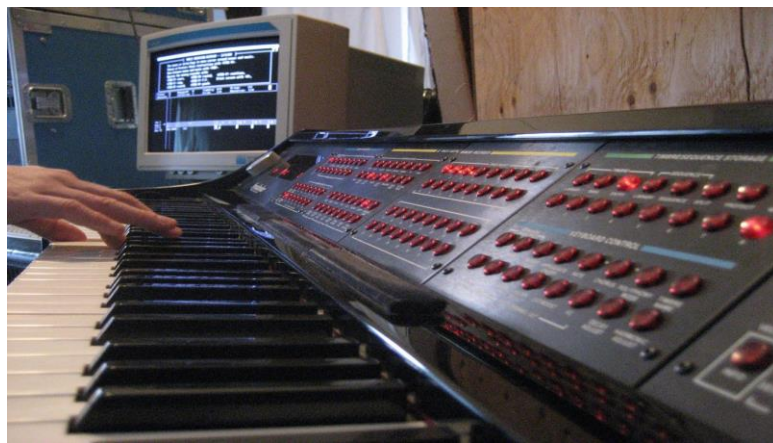
2.1.1 Βασικές έννοιες και ανάλυση λειτουργίας ενός DAW

Σε γενικές γραμμές η φιλοσοφία τους βασίζεται στην αρχική πολυκάναλη αναλογική ηχογράφηση σε κασέτα ή μπομπίνα, κάτι που είναι γνώριμο και καθιστά ευκολότερο για τους μηχανικούς ήχου και τους μουσικούς που είναι ήδη εξοικειωμένοι με τη χρήση μαγνητοφώνων να εξοικειωθούν και με το νέο σύστημα. Ως εκ τούτου, τα DAW που βασίζονται σε υπολογιστή τείνουν να έχουν μια τυπική διάταξη η οποία περιλαμβάνει στοιχεία ελέγχου μεταφοράς (transport controls όπως play, rewind, record, κλπ.), διαχείριση του κάθε καναλιού, οθόνη μίκτη (mix panel) και τέλος οθόνη επεξεργασίας των κυματομορφών (arrange panel). Σε κάθε κανάλι των DAW, υπάρχει ξεχωριστά επιλογή είτε να επιλέγουμε μονοφωνικό ή στερεοφωνικό σήμα στο κανάλι μας.

Έχοντας, λοιπόν, κάποιον εξωτερικό εξοπλισμό (hardware), σε συνδυασμό με τον υπολογιστή και κάποιο DAW της προτίμησής μας όπως Logic, Pro Tools, Reaper, Cubase ή κάποιο άλλο, μπορούμε να επεξεργαστούμε ξεχωριστά κάθε κανάλι του κομματιού και να το μετατρέψουμε σε ένα αρεστό και ευχάριστα ακουστό αποτέλεσμα.

Βασικό στοιχείο της λειτουργίας των DAW είναι η δυνατότητα που προσφέρουν για ομαδοποίηση των δεδομένων, έτσι ώστε να είναι δυνατή η επιλεκτική επεξεργασία και δρομολόγησή τους. Το “εργαλείο” για αυτή τη δουλειά είναι το track ή αλλιώς κανάλι. Ένα track είναι ένας νοητός χώρος στον οποίο ομαδοποιούνται οι πληροφορίες, έτσι ώστε να μπορούμε να τις επεξεργαστούμε με διαφορετικό τρόπο και να τις δρομολογήσουμε σε διαφορετικό αποδέκτη (έξοδο ήχου) από κάποιες άλλες. Μπορούμε να επέμβουμε στην ένταση, το ηχόχρωμα, το τονικό ύψος, τη θέση στη στερεοφωνική εικόνα, αλλά και σε άλλα χαρακτηριστικά του περιεχομένου του track, χωρίς να επηρεάσουμε τα χαρακτηριστικά των άλλων tracks.

Γι’ αυτή τη δουλειά έχουμε τα plug-ins. Αφού έχουμε επεξεργαστεί κάθε κανάλι (track) και του έχουμε δώσει το επιθυμητό αποτέλεσμα, δημιουργούμε ένα στερεοφωνικό αρχείο (και όχι πολυφωνικό), σε μορφή .WAV ή .AIFF ή MP3 που προορίζεται για τον mastering engineer, ο οποίος θα προχωρήσει στο mastering.



Εικόνα 2.1 Το synclavier μία από τις πρώτες μηχανές εγγραφής στην δεκαετία του '80.

2.2 Audio Engine ενός DAW

Αρχικά, μέσα στα συστήματα DAW και ειδικότερα στα audio engines, πρώτον και κύριων το πιο σημαντικό που πρέπει να καταλάβουμε είναι από πού αρχίζει και προκύπτει αυτή η διαφορά που μπορεί παρουσιάζουν στον ήχο;

Ένα DAW σύστημα έχει αυστηρά μηδενική επιρροή στα αρχεία ήχου που ηχογραφεί. Κατά την εγγραφή του ήχου, ένα DAW παίρνει τις πληροφορίες από μηδενικά και ένα (0010...) που προέρχονται από τη κάρτα ήχου ή από τους audio converters (DAC) και τα βάζει σε ένα αρχείο ήχου. Έτσι, 00101 πρόκειται να είναι 00101 αν χρησιμοποιούμε Cubase, pro tools, Logic ή ότι άλλο θέλουμε.

Εκεί που διαφέρουν τα DAW είναι στην άθροιση. Όταν ξεκινάμε την μίξη διαφορετικών αρχείων ήχου σε ένα στερεοφωνικό ήχο (ή αρχείο), τότε το DAW έχει να κάνει υπολογισμούς χρησιμοποιώντας μια μηχανή ήχου (audio engine). Διαφορετικές μηχανές ήχου χρησιμοποιούν διαφορετικούς τρόπους για τον υπολογισμό της άθροισης και αποφέρουν διαφορετικά αποτελέσματα οι οποίες είναι ο λόγος που ίσως μπορεί να παίρνουμε και διαφορετικά αποτελέσματα στον ήχο μας από DAW σε DAW.

2.3 Ανάλυση του όρου fader και pan pot

Πρίν όμως αναλύσουμε την δομή και την λειτουργία της μηχανής ήχου πρέπει να λάβουμε υπ' όψιν μας πως όλες οι άλλες επιλογές και ρυθμίσεις πρέπει να είναι κοινές σε όλα τα DAW που θέλουμε να εξετάσουμε. Δύο σημαντικοί παράγοντες που επηρεάζουν το αποτέλεσμα σημαντικά και πως αυτό το ακούμε εμείς είναι τα faders του mix panel ενός DAW αλλά και τα pan pots, μιας και ξέρουμε ότι εκτός από τα faders ένταση σε ένα κανάλι μπορούμε να δώσουμε και με σωστή χρήση των pan pots. Στην συνέχεια θα αναλύσουμε αυτές τις δύο υποκατηγορίες.

2.3.1 Faders

Αρχίζοντας με τα faders μπορούμε να πούμε ότι είναι οποιαδήποτε συσκευή που χρησιμοποιείται για fading, και ειδικότερα όταν πρόκειται για ένα κουμπί ή για ένα κουμπί που ανεβοκατεβαίνει κατά μήκος ενός track. Ένα fader μπορεί να είναι είτε αναλογικό, να ελέγχει δηλαδή άμεσα την αντίσταση ή την αντίσταση στην πηγή (π.χ. ένα ποτενσιόμετρο) ή ψηφιακό, να ελέγχει αριθμητικά έναν ψηφιακό επεξεργαστή σήματος (DSP). Τα ψηφιακά fader αναφέρονται επίσης ως εικονικά faders (virtual faders), δεδομένου ότι μπορούμε να τα δούμε στην οθόνη ενός DAW. Σύγχρονοι high end ψηφιακοί μίκτες διαθέτουν συχνά motorized faders, faders δηλαδή με συνδεδεμένους πιεζοηλεκτρικούς ενεργοποιητές, που μεταβαίνουν αυτόματα σε μία στάθμη που έχουμε ορίσει. Στην συνέχεια είναι πολύ σημαντικό να αναφέρουμε ότι όλα αυτά τα ψηφιακά συστήματα χρησιμοποιούν την dBFs κλίμακα (Decibels relative to full scale) μία κλίμακα που μετρά ντεσιμπέλ επίπεδα έντασης σε ψηφιακά συστήματα. Τα 0 dB_{FS} αποδίδονται στο μέγιστο δυνατό ψηφιακό επίπεδο έντασης (level). Για παράδειγμα, ένα σήμα

που φτάνει στο 50% του μέγιστου επιπέδου του σε κάθε σημείο θα φτάσει -6 dB_{FS} σε εκείνο το σημείο, 6 dB κάτω από την πλήρη κλίμακα. Με βάση την σύντομη αυτή θεωρία μπορούμε εύκολα να καταλάβουμε ότι δεδομένου ότι είναι όλα βαθμονομημένα σε ντεσιμπέλ και πιο συγκεκριμένα σε dBFS δεν παίζει κάποιο σημαντικό ρόλο στην διαφορά του ηχητικού αποτελέσματος από DAW σε DAW μιας και όλα χρησιμοποιούν την ίδια παραπάνω βαθμονόμηση. Αυτό που χρειάζεται προσοχή είναι σε όλη την πειραματική διαδικασία όλα τα faders των κάθε DAW που εξετάζουμε να έχουν την ίδια στάθμη μεταξύ τους ή τουλάχιστον το ίδιο αποτέλεσμα έντασης μιας και μπορεί να προκύψουν διαφορές έντασης κατά την άθροιση των καναλιών (export).

2.3.2 pan law

Pan law, ή pan rule, είναι μια αρχή στο χώρο της μουσικής παραγωγής που ορίζει ότι κάθε σήμα ίσου πλάτους και φάσης που παίζεται σε δύο κανάλια ενός στερεοφωνικού συστήματος θα αυξηθεί σε ένταση μέχρι 6.02 dB_{SPL} , εφόσον υπάρχει τέλεια απόκριση από το σύστημα μεγαφώνων και τέλεια ακουστική στο δωμάτιο. Συχνά, η ακουστική αθροίσεως ενός δωματίου και του συστήματος είναι πολύ κατώτερη από το ιδανικό έτσι η συγκεκριμένη σχετική ένταση θα αλλάξει από -3 dB έως 0 dB καθώς το μονοφωνικό σήμα μετακινηθεί (panning) από το κέντρο προς τέρμα αριστερά ή δεξιά, και αυτό ακούγεται φυσικό. Η ιδέα είναι ότι όταν κάποιος κατευθύνει τα σήματα αριστερά και δεξιά με το pan pot, η παρατηρούμενη ένταση θα παραμείνει ίδια, ανεξάρτητα από την θέση που βρισκόμαστε (regardless of latitude).

Ωστόσο, τόσο η κατεύθυνση της εξασθένησης σε όλο το panoramic sweep και το ποσό κατά το οποίο το σήμα είναι εξασθενημένο ποικίλλουν ανάλογα με το pan rule. Για παράδειγμα, οι ψηφιακές κονσόλες Yamaha χρησιμοποιούν ένα τυπικό (συμβιβασμός) 3 dB pan rule, όμως το σήμα είναι σε μέγιστη ένταση στο κέντρο και γίνεται σταδιακά πιο δυνατό (μέχρι $+3 \text{ dB}$), όταν πανάρετε προς τα δεξιά ή προς τα αριστερά.

Το pan rule -3 dB συνήθως εφαρμόζεται συμβιβαστικά για να συμβαδίζει με τις ακουστικές δυνατότητες άθροισης των περισσότερων control rooms. Ωστόσο, ο κατασκευαστής της κονσόλας SSL χρησιμοποιεί pan rule 4.5 dB , επειδή πιστεύει ότι οι ακριβείς τους κονσόλες θα πρέπει κανονικά να χρησιμοποιηθούν σε control rooms που έχουν δεχτεί ακουστική μελέτη για να είναι πιο κοντά σε αυτή την ιδέα.

Πολλές κονσόλες που έχουν μόνο ένα pan rule λειτουργούν ώστε όταν ένα σήμα πανάρετε τέρμα αριστερά ή δεξιά να βρίσκεται σε μέγιστη ένταση και να γίνεται προοδευτικά χαμηλότερο σε ένταση όταν το panning κατευθύνεται προς το κέντρο.

Σύμφωνα με τον mastering engineer Glenn Meadows, τα δωμάτια Kinoshita - Hidley στο Masterfonics (Nashville, Tennessee), παράγουν περίπου $5,9 \text{ dB}$ ακουστικής άθροισης (acoustic summing) όταν και τα δύο ηχεία παρουσιάζουν το ίδιο σήμα σε φάση.

2.4 Σύγκριση του pan rule/law για κάθε DAW που θα χρησιμοποιήσουμε στην πειραματική διαδικασία.

Έτσι όπως διαφέρουν οι αναλογικές κονσόλες αλλά και οι ψηφιακές στο pan law έτσι και τα διάφορα DAW που θέλουμε να συγκρίνουμε τα audio engine, σημαντικό ρόλο στην ποιότητα του ήχου αλλά και στην εμφάνιση διαφορών ανάμεσα τους παίζει και το pan law. Για να συγκρίνουμε χωρίς αυτός ο κανόνας να επηρεάσει το αποτέλεσμα της άθροισης (export) θα πρέπει να φροντίσουμε πριν αρχίσει η πειραματική διαδικασία όλα τα DAW που θα συγκρίνουμε να έχουν το ίδιο pan law. Παρακάτω θα παραθέσουμε τι επιλογές από pan law έχει ξεχωριστά το κάθε DAW που θα συγκρίνουμε. Για την συγκεκριμένη πειραματική διαδικασία θα χρησιμοποιήσουμε για να βγάλουμε αξιόπιστα αποτελέσματα 3 DAW τα Pro Tools HD 9, το logic pro 9, και το cubase 5. Στην συνέχεια παραθέτουμε τα pan law του κάθε DAW.

2.4.1 Logic pro 9

Από τη έκδοση 7.1 και μετά το logic μας δίνει την δυνατότητα να επιλέξουμε ποιο pan law θέλουμε να χρησιμοποιήσουμε, ποιο δηλαδή μας βολεύει καλύτερα στις ανάγκες μας. Έτσι έχουμε την δυνατότητα να επιλέξουμε μία από τις τρεις επιλογές που έχει. Έχουμε 0dB, -3dB, and -3dB Compensated.

- **Οι -3dB και -3dB Compensated αλγόριθμοι** στην πραγματικότητα μειώνουν την ένταση (level) ενός σήματος όταν είναι στο κέντρο (ή αυξάνουν την ένταση όταν το panning είναι αριστερά ή δεξιά), έτσι ώστε το κέντρο δεν ακούγεται πιο δυνατά.
- **Ο 0dB αλγόριθμος** από την άλλη δεν αλλάζει την ένταση του σήματος όταν το panning είναι αριστερά ή δεξιά, που σημαίνει ότι όταν ένα σήμα βρίσκεται αυστηρά (dead) στο κέντρο στην πραγματικότητα είναι πιο δυνατό σε ένταση από ότι θα ήταν τέρμα αριστερά ή δεξιά.

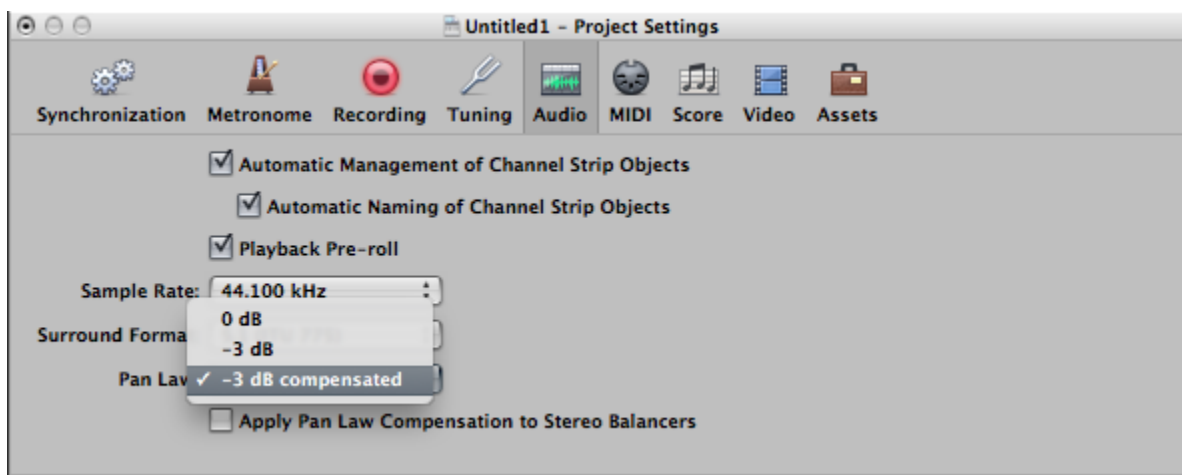
Οι δύο αυτές ρυθμίσεις (-3dB και -3dB Compensated) είναι ουσιαστικά ίδιες, εκτός από μία αύξηση της έντασης 3dB:

-3dB ρύθμιση: 0dB Αριστερά, Κέντρο-3dB, δεξιά 0dB

-3dB Compensated ρύθμιση: Αριστερά +3 dB, Κέντρο 0dB, Δεξιά +3 dB

Με λίγα λόγια δηλαδή:

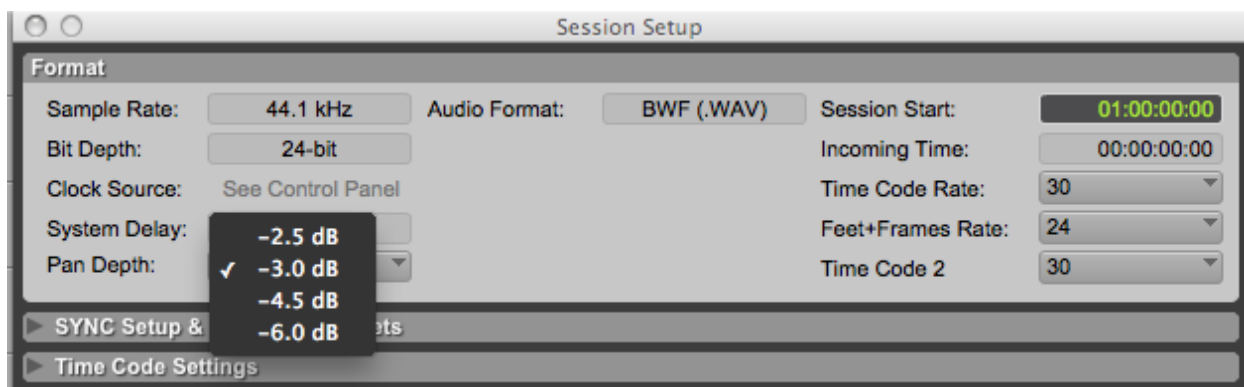
- -3dB: Μειώνει ένα σήμα κατά 3dB όταν το panning είναι στο κέντρο.
- -3dB Compensated: Αυξάνει την ένταση ενός σήματος κατά 3dB όταν το panning είναι πλήρως αριστερά ή δεξιά.
- 0dB: Πιο δυνατά στο κέντρο παρά τέρμα δεξιά ή αριστερά.



Εικόνα 2.2 Ρυθμίσεις για το pan law στο Logic Pro 9.

2.4.2 Pro tools HD 9

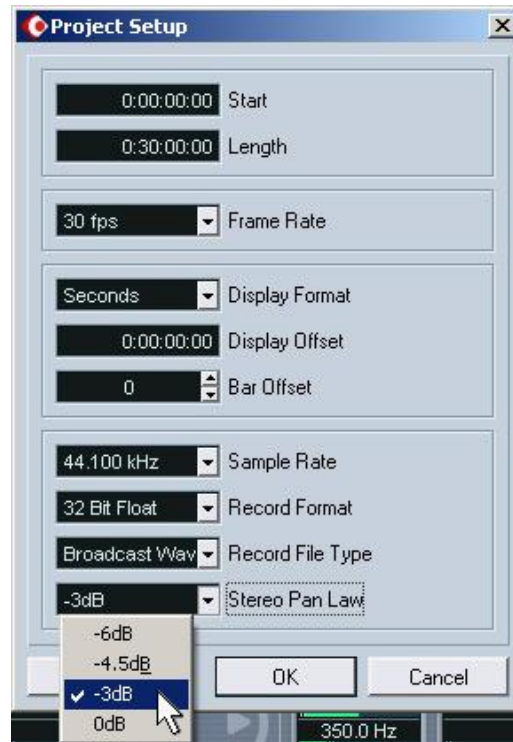
Στα pro tools ο όρος pan law αναφέρεται ως pan depth και μπορούμε να το βρούμε στο session setup. Σε σχέση με το Logic pro τα Pro tools έχουν 4 αλγορίθμους για το pan depth και όπως βλέπουμε μόνο ο ένας είναι κοινός και στα δύο που είναι τα -3dB. Έτσι για τα pro tools έχουμε τους εξής αλγορίθμους: **-2.5dB**, **-3.0dB**, **-4.5dB** και **-6.0dB**. Από τα παραπάνω παρατηρούμε ότι τα pro tools μας δίνουν την δυνατότητα να έχουμε μέχρι και **6dB** πτώση έντασης του σήματος όταν αυτό βρίσκεται στο κέντρο.



Εικόνα 2.3 Ρυθμίσεις για το pan law στα Pro Tools 9 HD.

2.4.3 Cubase 5

Στο cubase ο όρος pan law αναφέρεται ως stereo pan law και μας δίνει τις εξής επιλογές που σχεδόν είναι ίδιες με αυτές των pro tools. Έτσι οι αλγόριθμοι που χρησιμοποιούνται είναι: **0dB, -3dB, -4.5dB και -6dB.**



Εικόνα 2.4 Ρυθμίσεις για το pan law στο Cubase.

2.5 Plug-ins και η σχέση τους με τα DAW συστήματα

Τα plug-ins είναι προγράμματα λογισμικού που λειτουργούν μέσα στον κεντρικό λογισμικό του προγράμματος και επεκτείνουν την δυνατότητές του. Στις περισσότερες περιπτώσεις, τα plug-ins είναι δύο κατηγοριών:

- 1^η κατηγορία: επεξεργαστές σήματος (processors) δηλαδή επεξεργάζονται το σήμα (πχ. Compressors, limiters, equalizers):
- 2^η κατηγορία: εφέ, όπως (reverb, delay, chorus, pitch shift)

Υπάρχουν αμέτρητα plugins για τα σύγχρονα DAW λογισμικά, το καθένα έρχεται με την δικιά του μοναδική λειτουργία, διευρύνοντας έτσι τη συνολική ποικιλία των ήχων και των χειρισμών που μπορούμε να έχουμε. Το καθένα έχει τη δική του μορφή χειρισμού των ηχητικών κυμάτων, του τόνου και τον άλλον χαρακτηριστικών για να μετατρέψει έναν απλό ήχο σε κάτι διαφορετικό.

Υπάρχει και η επιλογή όσον αναφορά τις δυνατότητες και την ποιότητα στο θέμα επεξεργασίας μέσω plug-ins η ύπαρξη ολοκληρωμένων κυκλωμάτων DSP (Digital Signal Processing) τα οποία αναλαμβάνουν την ψηφιακή επεξεργασία του σήματος αποσπώντας την διαδικασία (και την επιβάρυνση που αυτή δημιουργεί) από την κεντρική μονάδα επεξεργασίας του Η/Υ (CPU). Πρόσθετα περιφερειακά που συμπληρώνουν τις δυνατότητες και την ευχρηστία τέτοιων συστημάτων είναι τα MIDI interfaces, καθώς η διαχείριση MIDI πληροφορίας παρέμεινε και βελτιώθηκε στην εξέλιξη των DAW συστημάτων.

Στην συγκεκριμένη πτυχιακή τα DAW συστήματα που μας ενδιαφέρουν και ασχολούμαστε στο πειραματικό μέρος είναι τα Pro tools, το Logic και το Cubase και παρακάτω θα αναλύσουμε τι τύπου DSP προγράμματα εμπεριέχει το κάθε πρόγραμμα.

Pro tools HD 9: Διαθέτει το σύστημα Pro tools HD Accel της εταιρείας Avid όπου τόσο το software όσο και το hardware μέρος του συστήματος κατασκευάζεται από την εταιρεία με υποστήριξη τόσο σε συμβατό hardware όσο και σε software (plug-ins) υψηλών προδιαγραφών, από άλλες εταιρείες. Το περιβάλλον εργασίας είναι αρκετά εύχρηστο και ευέλικτο τόσο σε θέματα δρομολόγησης σημάτων όσο και επεξεργασίας audio, πολυκάναλα. Το σύστημα υποστηρίζει το πρωτόκολλο επεξεργασίας RTAS (Real Time Audiosuite) και TDM (Time Domain Multiplex) που προσφέρει επεξεργασία σήματος σε πραγματικό χρόνο (Real time) με χρήση DSP κυκλωμάτων.

Logic Pro 9: Διαθέτει το λογισμικό Logic Audio της εταιρείας Emagic (που αγοράστηκε από την Apple και μετονομάστηκε σε Logic Studio). Διαθέτει επεξεργασία audio και ενσωματωμένων plug-ins υψηλών προδιαγραφών και ευελιξία στο επίπεδο των δρομολογήσεων. Υποστηρίζει το πρωτόκολλο AU (Audio Unit) και το TDM της Avid. Σε επίπεδο hardware υποστηρίζει τις συσκευές της Avid καθώς και άλλων εταιριών όπως είναι η Apogee.

Cubase 5: Το λογισμικό Cubase της εταιρείας Steinberg διαθέτει αρκετά ευέλικτες λειτουργίες διαχείρισης στην επιφάνεια edit και πλήθος λογισμικού plug-ins τόσο για επεξεργασία σήματος όπως και παράγωγης ήχου. Υποστηρίζει το πρωτόκολλο επεξεργασίας VST (Virtual Studio Technology). Σε επίπεδο hardware υποστηρίζει μεγάλο αριθμό από audio interfaces. (Κεφαλογιάννης Νίκος, 2007, 13)

Για το υποκειμενικό κομμάτι του πειραματικού μέρους (ακροάσεις) θα χρησιμοποιήσουμε τα plug-ins από την εταιρεία Waves στα exports αυτής της κατηγορίας. Αυτό γίνεται διότι σκοπός μας δεν είναι να χρησιμοποιήσουμε τα plug-ins που έχει το κάθε το DAW που είναι διαφορετικά μεταξύ τους. Για να είναι αντικειμενικά τα exports με plug-ins θα χρησιμοποιήσουμε τα *ίδια ακριβώς* plug-ins της εταιρείας Waves με το *ίδιο ακριβώς preset* και παραμέτρους και στα τρία DAW (logic, pro tools, Cubase). Έτσι πχ το κανάλι της φωνής θα έχει το ίδιο reverb με το ίδιο preset και στα τρία DAW. Με την ίδια λογική θα χρησιμοποιηθούν και στα υπόλοιπα κανάλια.

2.5.1 DAW (fixed and floating point μηχανή ήχου)

Pro tools HD 9: Υποστηρίζουν 16bit, 24bit, 32bit και 64bit float, ήχου σε ρυθμούς δειγματοληψίας έως 192 kHz, και μπορεί να χειριστούν WAV, AIFF, AIFC, MP3, WMA, και SDII αρχεία ήχου και αρχεία βίντεο QuickTime. Η μηχανή ήχου παραδοσιακά είναι 48bit fixed point arithmetic, αλλά floating point χρησιμοποιείται επίσης σε ορισμένες περιπτώσεις, όπως με τα Pro Tools HD Native. Τα pro tools HD X χρησιμοποιούν αθροιστή 64bit floating point.

Logic pro 9: Η μηχανή ήχου του Logic pro 9 είναι 32bit floating point.

Cubase 5: Η μηχανή ήχου που χρησιμοποιεί το Cubase 5 είναι 32bit floating point.

Για fixed και floating point ψηφιακή επεξεργασία θα αναφερθούμε στην συνέχεια της πτυχιακής στο Κεφάλαιο 3.

2.6 Αρχικά θεωρητικά συμπεράσματα για την πειραματική διαδικασία

Με βάση τα παραπάνω και για να υπάρχει μία κοινή βάση στο συγκεκριμένο πείραμα θα χρησιμοποιήσουμε όπως αναφέραμε τον ίδιο αλγόριθμο ran law και για τα 3 DAW. Οι εταιρείες βλέπουμε από την ανάλυση των ran law χρησιμοποιούν διαφορετική default τιμή. Έτσι εμείς θα φροντίσουμε στο πείραμα μας να έχουμε την ίδια τιμή διότι δεν θέλουμε οι διαφορετικές default ρυθμίσεις του κάθε DAW να παίζουν ρόλο στο ηχητικό αποτέλεσμα μιας και εξετάζουμε αν τελικά προκύπτουν διαφορές από DAW σε DAW εξ' αιτίας της κατασκευής τους.

Απ' όσα έχουμε παραθέσει και παραπάνω ο αλγόριθμος που μας βολεύει και είναι κοινός σε όλα τα DAW είναι ο -3dB και αυτόν θα χρησιμοποιήσουμε στην συνέχεια για την πειραματική διαδικασία.

Τέλος είδαμε ότι κάθε DAW χρησιμοποιεί τα δικά του plug-ins και την δικιά του DSP επεξεργασία για την διαχείριση του σήματος. Στην κατηγορία δειγμάτων που έχουν χρησιμοποιηθεί plug-ins θα χρησιμοποιήσουμε τα ίδια plug-ins της εταιρείας Waves και στα 3 DAW που εξετάζουμε με το ίδιο preset για να μην έχουμε έτσι διαφορές στον ηχητικό αποτέλεσμα που προήλθαν από τα διαφορετικά plug-ins.

ΚΕΦΑΛΑΙΟ 3^ο: ΑΝΑΛΥΣΗ ΤΗΣ ΛΕΙΤΟΥΡΓΙΑΣ ΚΑΙ ΤΗΣ ΔΟΜΗΣ ΤΗΣ ΜΗΧΑΝΗΣ ΗΧΟΥ (AUDIO ENGINE) ΕΝΟΣ DAW

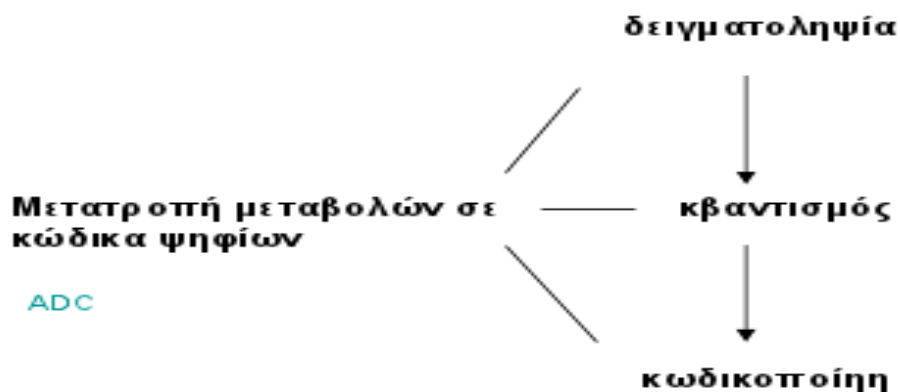
Εισαγωγή 3^ο Κεφαλαίου

Στο κεφάλαιο αυτό γίνεται αναφορά στα στάδια που περνά το αναλογικό σήμα για να μετατραπεί σε ψηφιακό καθώς και ανάλυση των δύο βασικών κατηγοριών της ψηφιακής επεξεργασίας του σήματος (fixed και floating point). Τέλος αναλύεται η δομή των audio engines των DAW.

3.1 Ψηφιοποίηση του σήματος

Για να εμβαθύνουμε και να καταλάβουμε που δημιουργούνται οι πραγματικές διαφορές ανάμεσα στα DAW πρέπει να αναλύσουμε κάποια σημαντικά κομμάτια που αποτελούν και την δομή των audio engines τους. Αρχικά θα αναλύσουμε τους όρους που προκύπτουν από την ψηφιοποίηση του σήματος αλλά και από τα τεχνικά χαρακτηριστικά που αποτελείται ένα DAW.

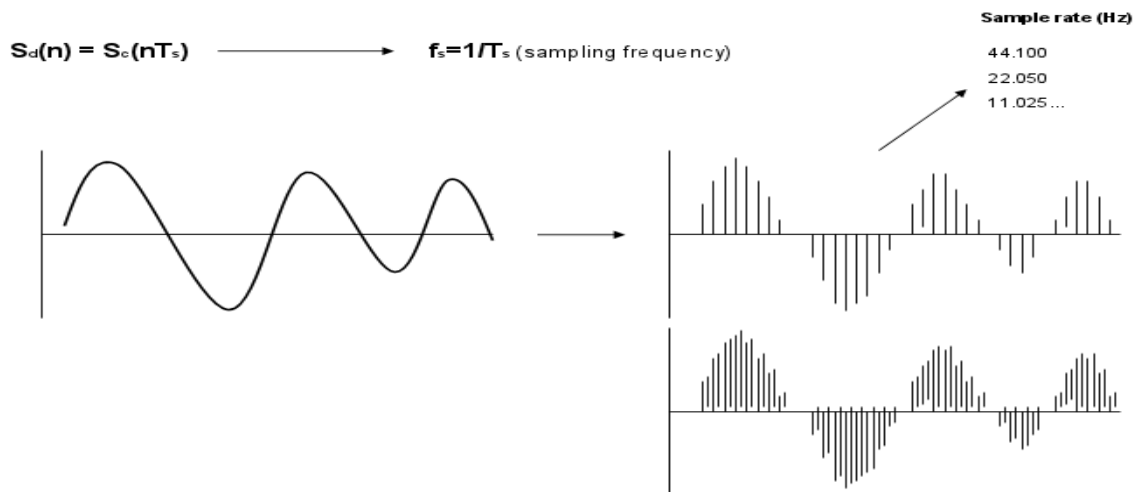
Το αναλογικό σήμα περνά μέσα από πολλά στάδια μέχρι να καταλήξει σε ψηφιακή μορφή. Αρχικά όλα ξεκινούν κατά την είσοδο του αναλογικού σήματος στον κονβέρτορα A/D (analog to digital). Όπως βλέπουμε και από την παρακάτω εικόνα τρία είναι τα στάδια που ακολουθεί το σήμα. Είναι η δειγματοληψία, η κβαντοποίηση ή κβαντισμός και τέλος η κωδικοποίηση του σήματος. Στην συνέχεια θα αναλύσουμε ένα-ένα αυτά τα στάδια για να παρατηρήσουμε και να καταλάβουμε τι αλλαγές και μεταβολές επιφέρουν στο σήμα.



Εικόνα 3.1 Τα στάδια που περνάει το αναλογικό σήμα για να καταλήξει σε ψηφιακή μορφή.⁵

⁵ Η εικόνα προέρχεται από το μάθημα «κβαντισμός και κωδικοποίηση» του τμήματος Μουσικής. Τεχνολογίας κ' οργάνων.

3.1.1 Δειγματοληψία: Είναι η μετατροπή ενός συνεχούς σήματος $S_c(t)$ σε σήμα διακριτού χρόνου $S_d(n)$. Το δείγμα συνήθως αναφέρεται σε μια τιμή του συνεχούς σήματος σε μια συγκεκριμένη χρονική στιγμή. Ο δειγματολήπτης είναι ένα σύστημα το οποίο εξάγει δείγματα (σε ίσα χρονικά διαστήματα) από ένα συνεχές σήμα. Ένας θεωρητικός ιδανικός δειγματολήπτης δημιουργεί δείγματα τα οποία αντιστοιχούν στην στιγμιαία τιμή του συνεχές σήματος στα επιθυμητά χρονικά σημεία.



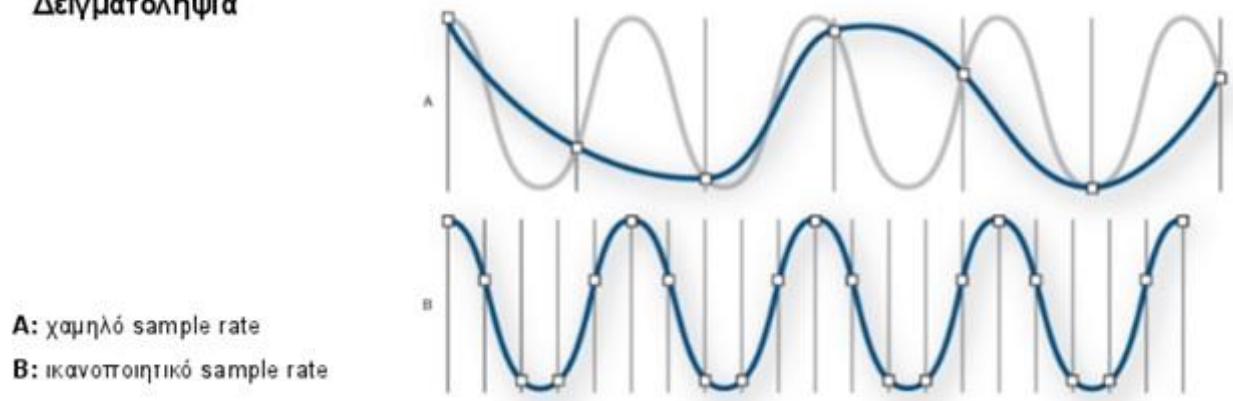
Εικόνα 3.2 Μετατροπή του συνεχούς αναλογικού σήματος σε σήματα διακριτού χρόνου⁶.

Η συχνότητα δειγματοληψίας ή **ρυθμός δειγματοληψίας** μετριέται στην μονάδα μέτρησης **Χέρτζ (H_z)** και μας δείχνει πόσα δείγματα έχουν ληφθεί από τον δειγματολήπτη σε διάρκεια ενός δευτερολέπτου. Εάν η συχνότητα δειγματοληψίας f_s (ο συμβολισμός f_s προκύπτει από τα αρχικά sampling frequency - συχνότητα δειγματοληψίας) είναι 1000Hz, σημαίνει ότι ο δειγματολήπτης δημιουργεί 1000 δείγματα σε κάθε δευτερόλεπτο του σήματος. Η χρονική απόσταση των δειγμάτων υπολογίζεται από τον τύπο $f_s = 1/T$ (το T είναι η **περίοδος δειγματοληψίας**) και για συχνότητα δειγματοληψίας 1000Hz έχουμε:

$$\boxed{T = 1/f_s = 1/1000 = 0.001s} \quad (3.1)$$

⁶ Η εικόνα προέρχεται από το μάθημα «κβαντισμός και κωδικοποίηση» του τμήματος Μουσικής. Τεχνολογίας κ' οργάνων.

Δειγματοληψία



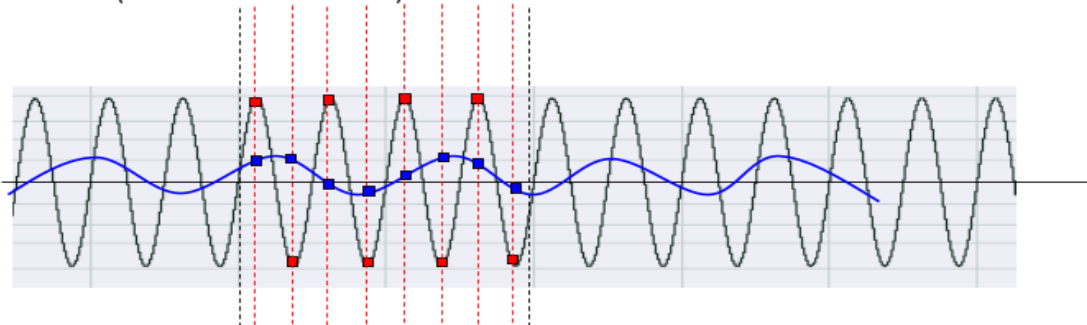
Εικόνα 3.3 Διαφορά χαμηλής και ικανοποιητικής συχνότητας δειγματοληψίας⁷.

Στην συνέχεια το **Θεώρημα Nyquist** μας εξηγεί ότι η συχνότητα δειγματοληψίας f_s θα πρέπει να είναι υπερδιπλάσια της μέγιστης συχνότητας f_{\max} που μπορεί να εμπεριέχεται στο ψηφιοποιημένο σήμα δηλαδή $f_s > 2f_{\max}$. Αυτό λαμβάνεται σοβαρά υπόψη μιας και η διαδικασία της δειγματοληψίας δημιουργεί αντίγραφα συχνοτήτων εκατέρωθεν της συχνότητας δειγματοληψίας τα αποκαλούμενα images (είδωλα) και ενώ το άνω είδωλο θα εμφανιστεί ακριβές αντίγραφο της αρχικής συχνότητας με μια μετατόπιση το κάτω είδωλο θα εμφανιστεί ανεστραμμένο κατά 180° . Το ίδιο θα συμβεί και για όλες τις συχνότητες που περιέχει το αρχικό σήμα με κίνδυνο τα κάτω είδωλα των υψηλότερων συχνοτήτων που εμπεριέχονται στο εύρος του αρχικού σήματος να επικαλύπτονται με τις αρχικές συχνότητες δημιουργώντας ανεπιθύμητα αποτελέσματα που ονομάζονται **aliasing**. Γι αυτό τον λόγο έχουν δημιουργηθεί τα **anti-aliasing** φίλτρα.

⁷ Η εικόνα προέρχεται από το μάθημα «κβαντισμός και κωδικοποίηση» του τμήματος Μουσικής. Τεχνολογίας κ' οργάνων.

$$F_s \geq 2F_{max}$$

CD Quality: $F_s=2F_n$ (22050Hz: cutoff frequency) → 44100Hz
 Τηλέφωνο: $F_s > 2F_n$ (3300Hz >> >>) → 8000Hz



Εικόνα 3.4 Διαφορά στην συχνότητα αποκοπής για ποιότητα cd και ομιλία τηλεφώνου⁸.

Aliasing: $f_s < 2f_n$ → Αλληλοκάλυψη συχνοτήτων, παραμόρφωση



Anti-alias filter: αφαίρεση συχνοτήτων μεγαλύτερων από $f_s/2$

Εικόνα 3.5 Anti-alias filter⁹

3.1.2 Κβαντισμός (Quantization): Είναι η διαδικασία της προσέγγισης ενός αναλογικού (συνεχούς) δείγματος με ένα πεπερασμένο αριθμό bit.

Πιο αναλυτικά ξέρουμε ότι αυτό που ουσιαστικά ψηφιοποιείται είναι η ένταση του σήματος σε διαδοχικές χρονικές στιγμές που ορίζονται από την συχνότητα δειγματοληψίας όπως είδαμε και παραπάνω. Η διαδικασία αυτή προϋποθέτει μια κβαντοποίηση, δηλαδή την αντιστοιχία της πραγματικής έντασης στην πλησιέστερη κβαντική στάθμη ο αριθμός των οποίων άρα η ανάλυση του συστήματος ορίζεται από τον αριθμό των bits που χρησιμοποιεί το σύστημα για την διαδικασία αυτή. Σε ένα σύστημα 4 bit για παράδειγμα, ο αριθμός των διαφορετικών τιμών άρα των κβαντικών σταθμών που μπορούν να χρησιμοποιηθούν για να απεικονίσουν τις διακυμάνσεις της έντασης του αναλογικού σήματος είναι ελάχιστος (16).

Το αποτέλεσμα της ψηφιοποίησης που δημιουργεί ένα τέτοιο σύστημα είναι μεγάλης παραμόρφωσης κβαντισμού (Quantization distortion) και αναφέρεται στο γεγονός ότι μεγάλο δυναμικό εύρος του αναλογικού σήματος αντιστοιχεί στην ίδια κβαντική στάθμη άρα θα πάρει την ίδια ψηφιακή τιμή. Αυτό το λάθος κβαντισμού, όταν το σήμα μεταβάλλεται μεταξύ πολλών

^{8,9} Οι εικόνες προέρχονται από το μάθημα «κβαντισμός και κωδικοποίηση» του τμήματος Μουσικής Τεχνολογίας κ' οργάνων.

κβαντικών σταθμών εμφανίζεται τυχαία άρα κατά την αναπαραγωγή εμφανίζεται ως λευκός θόρυβος. Η αντιμετώπιση αυτού του φαινομένου γίνεται με την αναβάθμιση της ανάλυσης του συστήματος όπου με την προσθήκη ενός επιπλέον bit επιτυγχάνεται ο διπλασιασμός των κβαντικών σταθμών άρα ο υποδιπλασιασμός της απόστασης μεταξύ δυο εντάσεων (δηλ. της πιθανότητας κβαντικού λάθους). Ο διπλασιασμός ή υποδιπλασιασμός στην κλίμακα των dB αντιστοιχεί σε τιμή των 6dB. Επομένως ο υποδιπλασιασμός των κβαντικών λαθών αυξάνει το δυναμικό εύρος του ψηφιακού συστήματος κατά 6dB διπλασιάζοντας τις κβαντικές στάθμες.

$$\text{Quantizing Levels} = 2^{(\# \text{ of bits})} \quad (3.2)$$

$$\text{Dynamic Range} = (\# \text{ of bits}) \times 6\text{dB} \quad (3.3)$$

Ένα άλλο πρόβλημα που παρουσιάζεται με την αναγωγή των μεταβολών της έντασης σε συγκεκριμένες κβαντικές στάθμες είναι ότι χάνεται η γραμμικότητα της αλλαγής έντασης καθώς το σήμα παραμένει σε μια κβαντική στάθμη έως ότου η διάφορα πλησιάσει στην επομένη κβαντική στάθμη οπότε παίρνει αυτή την τιμή. Έτσι για να δημιουργήσουμε το χαρακτηριστικό της γραμμικής μεταβολής προσθέτουμε ένα **ποσοστό λευκού θορύβου (dither noise)**, ίσο με μια κβαντική στάθμη το οποίο προστιθέμενο στην πραγματική ένταση του σήματος ουσιαστικά συμπληρώνει την διάφορα από την επομένη κβαντική στάθμη.

Υπάρχουν πολλές μορφές dither noise η πιο συνηθισμένη είναι η **TDDF (Triangular distribution dither function)** και επίσης υπάρχει και το **noise-shaping** με κατανομή του θορύβου dither σε μικρότερο ποσοστό στις μεσαίες συχνοτικές περιοχές και περισσότερο στα άκρα του ακουστικού φάσματος όπου η ακοή έχει λιγότερη ευαισθησία.

Τελειώνοντας όσον αναφορά το quantizing οφείλουμε να πούμε ότι η διαδικασία αναγωγής των μεταβολών έντασης σε κβαντικές στάθμες δημιούργησε την ανάγκη μιας νέας κλίμακας μέτρησης σε dB που ονομάζεται dBFS (decibel Full Scale) με βαθμονόμηση από $-\infty$ έως 0 dB. Αυτή την καινούργια κλίμακα dB την χρησιμοποιούν όλα τα DAW που θέλουμε να εξετάσουμε. Η λογική μιας τέτοιας κλίμακας βασίζεται στο γεγονός ότι το δυναμικό εύρος κάθε ψηφιακής συσκευής προκαθορίζεται από την ανάλυση δειγματοληψίας (bitrate) πχ. 144dB στα 24bit τα οποία ανάγονται στις προκαθορισμένες κβαντικές στάθμες. Στις πρακτικές ηχοληψίας και ηχογράφησης η ρύθμιση της στάθμης ενός σήματος απαιτεί τον καθορισμό ενός nominal level (μια τιμή έντασης γενικώς αποδεκτή από όλες τις κατασκευάστριες εταιρείες μεταξύ της όποιας και της μέγιστης δυνατής). Η αναγωγή λοιπόν του nominal level +4dBu στην κλίμακα dBFS είναι καθορισμένη από την EBU (European Broadcasting Union) στα -14dBFS.

3.1.3 Κωδικοποίηση: Τελευταία άλλα όχι ασήμαντη διαδικασία στην μετατροπή του αναλογικού σήματος σε ψηφιακό είναι η διαδικασία της κωδικοποίησης. Κωδικοποίηση είναι η αντιστοίχιση των κωδικών που προκύπτουν από τον κβαντισμού στα δείγματα δηλαδή η δημιουργία μιας ψηφιακής σειράς από 0 και 1 (πχ 00110, 01110 κτλ).

Υπάρχουν πολλοί μέθοδοι κωδικοποίησης που χρησιμοποιούμε. Παρακάτω θα αναφέρουμε μερικές και πως λειτουργούν.

PCM (Pulse Code Modulation): Παλμοκωδική κωδικοποίηση. Ξεχωριστή αποθήκευση δειγμάτων, γραμμική διαμόρφωση.

Mu-Law / A-Law PCM: Λογαριθμική κωδικοποίηση. Αντιστοίχιση περισσότερων σταθμών στις χαμηλές συχνότητες

DPCM (Differential Pulse Code Modulation): Διαφορική παλμοκωδική κωδικοποίηση. Κωδικοποίηση δείγματος σε σχέση με τα προηγούμενα (πλήρης αποθήκευση 1^{ου} δείγματος, αποθήκευση διαφορών για τα επόμενα δείγματα).

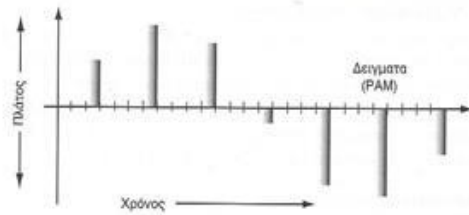
Διαμόρφωση Δέλτα (Delta Modulation): 1bità 1= μεγαλύτερο εύρος, 0=μικρότερο εύρος

ADPCM (Adaptive Differential Pulse Code Modulation): Προσαρμοζόμενη διαφορική παλμοκωδική κωδικοποίηση. Πρόβλεψη τιμών à υπολογισμός διαφοράς μεταξύ πραγματικών - προβλεπόμενων τιμών.

LPC (Linear Predictive Coding): Γραμμική προβλεπόμενη κωδικοποίηση. Σύγκριση χαρακτηριστικών του σήματος με ένα αποθηκευμένο αναλυτικό μοντέλο μετάδοσης των χαρακτηριστικών που συμπίπτουν με τα χαρακτηριστικά του μοντέλου. (εφαρμογή σε τηλεπικοινωνίες, συμπίεση ομιλίας)

Συνοψίζοντας μπορούμε να πούμε ότι η μετατροπή του αναλογικού σήματος σε ψηφιακό είναι μια πολύπλοκη διαδικασία που γίνεται σε πολλά στάδια. Παρακάτω θα παραθέσουμε ένα πίνακα που περιέχει επιγραμματικά όλα τα παραπάνω βήματα που αναλύσαμε.

ΨΗΦΙΟΠΟΙΗΣΗ ΗΧΟΥ – ΣΥΝΟΠΤΙΚΑ:



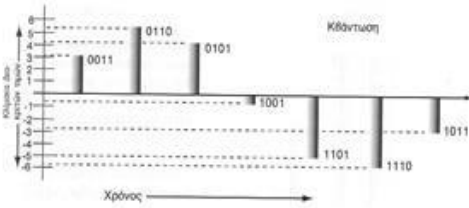
1. Δειγματοληψία

$$S_d(n) = S_c(nT_s)$$

Παραγωγή δειγμάτων (samples), διαμόρφωση παλμού (PAM=Pulse Amplitude Modulation)

$$f_s = 1/T_s \text{ (sampling frequency – sample rate)}$$

$$f_s = 2f_n \text{ (συχνότητα Nyquist)}$$



2 Κβαντισμός

$$S(n) = Q(S_d(n)), \quad \tilde{S}(n) = S_d(n) + e(n)$$

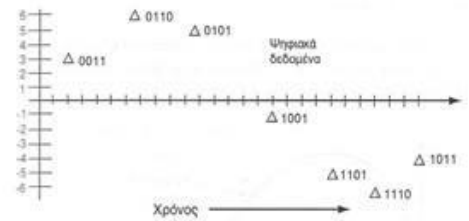
Μετατροπή συνεχών τιμών πλάτους σε διακριτές

N = τάξη κβαντιστή / μέγεθος δείγματος

2^N = επίπεδο κβαντισμού / αριθμός διαθέσιμων bits

Βήμα κβαντισμού (Δ)

Σφάλμα κβαντισμού



3. Κωδικοποίηση

Αντιστοίχιση κωδικών κβαντισμού στα δείγματα, δημιουργία σειρών ψηφίων / ψηφιακής πληροφορίας (δυναμικός κώδικας)

Αποκωδικοποίηση (decoding) / Διαδικασία παρεμβολής (interpolation)

Bit Rate (ρυθμός μετάδοσης δεδομένων) = $F \times M$ (sampling freq x sample size)

Εικόνα 3.6 Συνοπτικός πίνακας σταδίων ψηφιοποίησης ενός αναλογικού σήματος¹⁰

3.2 Fixed και floating point συστήματα διαφορές και ανάλυση των δύο αυτών λειτουργιών

Η ψηφιακή επεξεργασία σήματος μπορεί να χωριστεί σε δύο κατηγορίες fixed point (σταθερού σημείου) και floating point. Οι χαρακτηρισμοί αυτοί αναφέρονται στην μορφή που χρησιμοποιείται για την αποθήκευση και το χειρισμό των αριθμητικών αναπαραστάσεων των δεδομένων.

Οι **fixed point DSPs** έχουν σχεδιαστεί για να εκπροσωπούν και να χειρίζονται ακέραιους θετικούς και αρνητικούς με την χρήση τουλάχιστον 16bit, αποδίδοντας έτσι έως 65.536 μοτίβα bit (216). Οι **Floating-point DSPs** έχουν σχεδιαστεί να χειρίζονται ρητούς αριθμούς με την χρήση τουλάχιστον 32bit με έναν τρόπο παρόμοιο με επιστημονική σημειογραφία, όπου ένας αριθμός αναπαρίσταται με ένα δεκαδικό και έναν εκθέτη (π.χ. $A \times 2^B$, όπου «A» είναι το δεκαδικό και «B» είναι ο εκθέτης), αποδίδοντας έως και 4294967296 πιθανά μοτίβα bit (232). Με floating point αναπαράσταση, η τοποθέτηση της υποδιαστολής μπορεί να αλλάζει (float), σε σχέση με τα σημαντικά ψηφία του αριθμού.

¹⁰ Η εικόνα προέρχεται από το μάθημα «κβαντισμός και κωδικοποίηση» του τμήματος Μουσικής. Τεχνολογίας κ' οργάνων.

3.2.1 Fixed point-floating point dsp's

Ο όρος fixed point (σταθερό σημείο) αναφέρεται στον αντίστοιχο τρόπο με τον οποίο οι αριθμοί αντιπροσωπεύονται, με ένα σταθερό αριθμό ψηφίων μετά, και μερικές φορές και πριν την υποδιαστολή. Με floating point αναπαράσταση, η τοποθέτηση της υποδιαστολής μπορεί να αλλάζει (float), σε σχέση με τα σημαντικά ψηφία του αριθμού. Για παράδειγμα, μια αναπαράσταση σταθερής υποδιαστολής (fixed point) με μια παρόμοια σύμβαση τοποθέτησης της υποδιαστολής (with a uniform decimal point placement convention) μπορεί να αντιπροσωπεύει τους αριθμούς 123.45, 1234.56, 12345.67 κτλ, ενώ μία αναπαράσταση floating point θα μπορούσε να αναπαριστά τους αριθμούς 1.234567, 123456.7, 0.00001234567, 1234567000000000 κτλ. Το αποτέλεσμα είναι πως τα floating point συστήματα μπορούν να υποστηρίξουν ένα πολύ ευρύτερο φάσμα τιμών από ότι τα fixed point συστήματα, έτσι τα floating point έχουν την ικανότητα να αναπαριστούν και να επεξεργάζονται πολύ μικρούς αριθμούς άλλα και πολύ μεγάλους αριθμούς.

Στην λειτουργία τους στα fixed point συστήματα, τα κενά μεταξύ γειτονικών αριθμών είναι πάντα ίσα μεταξύ τους, ενώ σε floating-point συστήματα, τα κενά μεταξύ γειτονικών αριθμών δεν είναι ομοιόμορφα κατανομημένα - η απόσταση μεταξύ δύο αριθμών είναι περίπου δέκα εκατομμύρια φορές μικρότερη από την αξία των αριθμών (ANSI / IEEE Std . 754 standard format), με μεγάλες αποκλίσεις μεταξύ μεγάλων αριθμών και μικρά κενά μεταξύ των μικρών αριθμών.

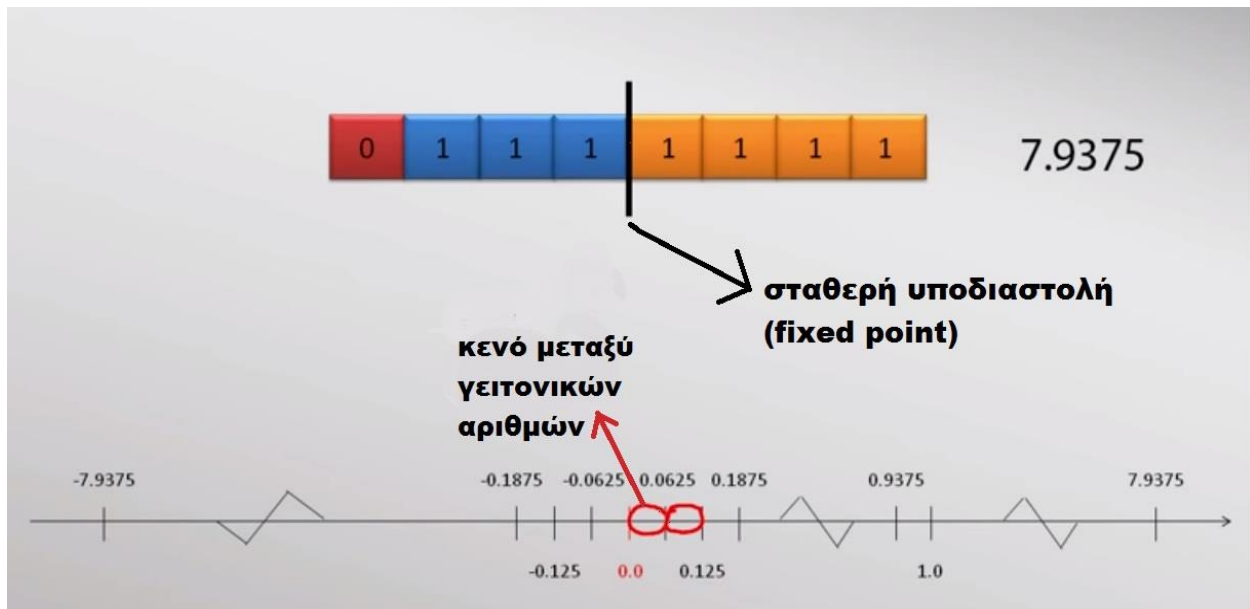
Αναλυτικότερα τα floating point συστήματα χρησιμοποιούν μια αρχιτεκτονική PCM τροποποιημένη για να δεχτεί μια τιμή κλίμακας (scaling value). Πρόκειται για μια προσαρμοστική προσέγγιση, με ανομοιόμορφη κβάντιση. Στα πραγματικά floating point συστήματα, ο συντελεστής κλίμακας εφαρμόζεται ακαριαία από δείγμα σε δείγμα. Σε άλλες περιπτώσεις, όπως στα block floating-point συστήματα, ο συντελεστής κλίμακας εφαρμόζεται σε ένα σχετικά μεγάλο όγκο δεδομένων. Αντί να σχηματίζουν μια γραμμική ψηφιακή λέξη δεδομένων, ένα floating point σύστημα χρησιμοποιεί μία ανομοιόμορφη κβάντιση για να δημιουργήσει μια λέξη η οποία χωρίζεται σε δύο μέρη: το δεκαδικό (τιμή δεδομένων) και τον εκθέτη (συντελεστής κλίμακας). Το δεκαδικό αντιπροσωπεύει την τιμή της κυματομορφής και το κλιμακούμενο πλάτος, το μέγεθος του βήματος κβαντισμού εκπροσωπείται από τον εκθέτη. Ειδικότερα, ο εκθέτης ενεργεί ως ένας βαθμονόμος που ποικίλλει/εναλλάσσει το κέρδος του σήματος στο μετατροπέα PCM A / \bar{v} . Ρυθμίζοντας το κέρδος του σήματος, ο μετατροπέας A/D χρησιμοποιείται πιο αποτελεσματικά. Τα σήματα χαμηλής στάθμης ενισχύονται και τα σήματα υψηλής στάθμης εξασθενούν. Ειδικότερα, το επίπεδο ενός σήματος έχει οριστεί στο υψηλότερο δυνατό επίπεδο που δεν υπερβαίνει το εύρος του μετατροπέα. Αυτό εναλλάσσει αποτελεσματικά το μέγεθος του βήματος κβαντισμού σύμφωνα με το πλάτος του σήματος και βελτιώνει την ακρίβεια της κωδικοποίησης σήματος χαμηλής στάθμης, κατάσταση όπου το σφάλμα κβαντισμού (quantizing error) είναι σχετικά πιο προβληματικό. Μετά την μετατροπή D/A το κέρδος και πάλι ρυθμίζεται ώστε να αντιστοιχεί στην αρχική του τιμή.

Ένα floating point σύστημα χρησιμοποιεί ένα short-word A/D μετατροπέα για να επιτευχθεί μια μέτρια δυναμική περιοχή, ή ένα longer word μετατροπέα για ένα μεγαλύτερο δυναμικό εύρος. Για παράδειγμα, ένα σύστημα floating point χρησιμοποιώντας ένα 16bit μετατροπέα A/D και ένα εκθέτη 3bit προσαρμοσμένο σε ένα εύρος 42dB σε 6dB βήματα θα δώσει 138dB δυναμικό εύρος (96dB + 42dB). Αυτός ο τύπος συστήματος θα ήταν χρήσιμο για την κωδικοποίηση ιδιαίτερα ακραίων σημάτων. Επιπλέον, αυτό το σύστημα απαιτεί μόνο μια διάσπαση λέξης των 19bit, αλλά το ισοδύναμο σταθερό γραμμικό σύστημα PCM θα απαιτούσε μια γραμμική λέξη 23bit. Επιπλέον όταν τα στάδια κέρδους (gain stages) τοποθετούνται κατά διαστήματα 6dB, οι κωδικοποιημένες λέξεις μπορούν εύκολα να μετατραπούν σε ένα ενιαίο κώδικα για την επεξεργασία ή την αποθήκευση χωρίς υπολογισμό. Το δεκαδικό μέρος υποβάλλεται σε μια μετατόπιση λειτουργίας σύμφωνα με την τιμή του εκθέτη.

Μολονότι το δυναμικό εύρος ενός floating point συστήματος είναι μεγάλο, η φύση της δυναμικής περιοχής του διαφέρει από εκείνη ενός σταθερού γραμμικού συστήματος είναι εγγενώς λιγότερη από ότι η αναλογία του S/N (signal to noise ratio). Αυτό συμβαίνει επειδή το δυναμικό εύρος μετρά την αναλογία μεταξύ του μέγιστου σήματος και του θορύβου, όταν δεν υπάρχει σήμα (no signal is present). Με την αναλογία S/N, από την άλλη πλευρά, ο θόρυβος μετράται όταν υπάρχει σήμα. Σε ένα σταθερό γραμμικό σύστημα το δυναμικό εύρος είναι περίπου ίσο με την αναλογία S/N, όταν υπάρχει σήμα. Ωστόσο σε ένα floating point σύστημα η αναλογία S/N κατά προσέγγιση καθορίζεται από την ανάλυση του σταθερού A/D μετατροπέα (περίπου 6n) η οποία είναι ανεξάρτητη από την μεγαλύτερη δυναμική περιοχή. Αλλαγές στο σήμα υπαγορεύουν αλλαγές στη δομή κέρδους το οποίο επηρεάζει το σχετικό πλάτος του σφάλματος κβαντισμού.

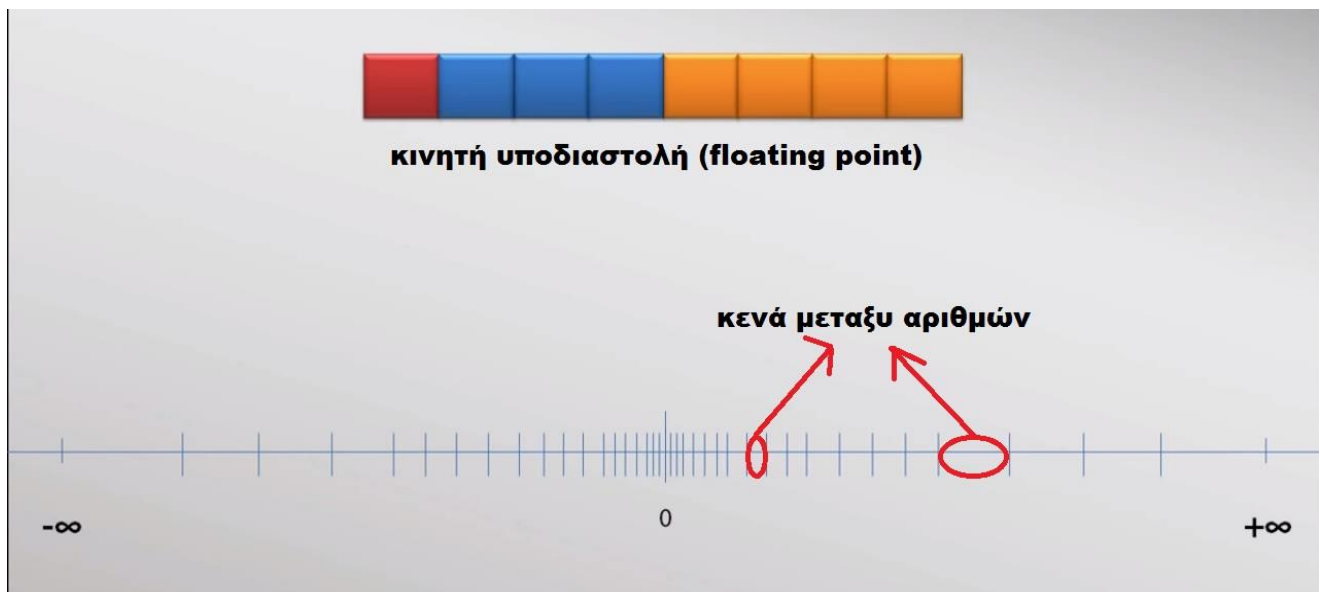
Οι αλλαγές στη δομή του κέρδους μπορεί να επηρεάσουν την ακουστικότητα του σφάλματος. Στιγμιαία αλλαγή από δείγμα σε δείγμα τείνει να οξύνει το πρόβλημα. Αντ' αυτού η αλλαγή του κέρδους πρέπει να γίνεται με αλγόριθμους που ακολουθούν την διακύμανση στο πλάτος του σήματος, με βάση τον τύπο του προγράμματος που πρόκειται να κωδικοποιηθεί. Για παράδειγμα, οι συλλαβικοί αλγόριθμοι¹¹ προσαρμόζονται στην ταχύτητα με την οποία οι συλλαβές ποικίλουν και αλλάζουν στην ομιλία. Οι μειώσεις κέρδους είναι στιγμιαίες, αλλά οι αυξήσεις κέρδους καθυστερήσουν. Εν πάση περιπτώσει, το κέρδος πρέπει να είναι ενεργοποιημένο για να αποφευχθεί οποιαδήποτε υπερφόρτωση του μετατροπέα A/D.

¹¹ Οι συλλαβικοί αλγόριθμοι χρησιμοποιούνται στην αναγνώριση ομιλίας και στην αυτόματη φωνητική μεταγραφή από τον προφορικό λόγο. Χρησιμοποιούν ακριβή κατάτμηση του εισερχόμενου σήματος ομιλίας σε τμήματα μεγέθους συλλαβής. Ένας αρκετά κοινός και απλός αλγόριθμος είναι να χρησιμοποιούμε την ένταση (intensity) από το αρχικό σήμα και την ένταση από το ένα ή περισσότερα αρχικά σήματα που έχουν επεξεργαστεί με band-pass φίλτρο. Στην συνέχεια αυτά συγκρίνονται χρησιμοποιώντας διαφορετικά κριτήρια για να καθοριστούν τα συλλαβικά όρια στο σήμα της ομιλίας.



Εικόνα 3.7 Αναπαράσταση 8bit fixed point επεξεργασίας.

Στην εικόνα 3.7 βλέπουμε μία αναπαράσταση για 8bit fixed point επεξεργασία. Αναλυτικότερα βλέπουμε την υποδιαστολή που είναι σταθερή και χωρίζει το ακέραιο και το δεκαδικό μέρος του αριθμού. Στον άξονα βλέπουμε τις τιμές των κβαντικών σταθμών που αντιστοιχούν και παρατηρούμε ότι κάθε τιμή απέχει την ίδια απόσταση μεταξύ τους (είναι ισόποσα μοιρασμένες) από το μέγιστο θετικό μέχρι τον μέγιστο αρνητικό αριθμό. Ένα αρνητικό της fixed point επεξεργασίας που παρατηρούμε από την θεωρία και την εικόνα 3.7 είναι ότι χρειαζόμαστε περισσότερα bits για να επεξεργαστούμε μεγαλύτερους αριθμούς σε σχέση με τα floating point συστήματα.



Εικόνα 3.8 Αναπαράσταση 8bit floating point επεξεργασίας.

Στην εικόνα 3.8 βλέπουμε μία αναπαράσταση για 8bit fixed point επεξεργασία. Αναλυτικότερα βλέπουμε ότι η υποδιαστολή είναι κινητή δεν είναι σταθερή και μετακινείται ανάλογα με τα σημαντικά ψηφία των αριθμών που επεξεργάζεται. Λόγο του πλεονεκτήματος ότι η υποδιαστολή κινείται η floating point επεξεργασία μας δίνει την δυνατότητα εκεί που χρειαζόμαστε μεγαλύτερη ανάλυση δηλαδή κοντά στο μηδέν η απόσταση των τιμών των κβαντικών σταθμών που αντιστοιχούν να είναι μικρότερη (μικρότερα κενά) προσφέροντας έτσι μεγαλύτερη λεπτομέρεια στις χαμηλές στάθμες και όσο η τιμή των κβαντικών σταθμών μεγαλώνει τα κενά γίνονται μεγαλύτερα για αντιστοιχούν σε αυτά και μεγαλύτεροι αριθμοί. Αυτό μπορούμε να το παρατηρήσουμε από τα κενά του άξονα της εικόνας 3.8.

3.3 Μηχανή ήχου (audio engine) ενός DAW

Όσο αναφορά το θεωρητικό κομμάτι για την μηχανή ήχου των διαφόρων DAW τα πράγματα δεν είναι και τόσο ξεκάθαρα μιας και πίσω από αυτόν τον ορό “κρύβεται” το λογισμικό της κάθε εταιρείας που κάνει όλο το σύστημα να δουλεύει και όπως είναι λογικό καμιά εταιρεία δεν αποκαλύπτει τον τρόπο με τον οποίο είναι κατασκευασμένος ο κώδικας, πάρα μόνο να διαφημίζει ότι είναι ο καλύτερος στην αγορά. Η μόνη χρήσιμη αναφορά που βρέθηκε είναι ένα white paper της Avid που περιγράφει θέματα γύρω από το D.A.E (Digital Audio Engine) και το mixer της, όμως τα θέματα που μελετά είναι γενικώς σχετικά και σε γενικές γραμμές κοινά σε όλα τα DAW που ξέρουμε δίνοντας μας έτσι την δυνατότητα να πάρουμε μια ιδέα γύρω από το αλγοριθμικό κομμάτι των DAW. Στην συνέχεια θα παραθέσουμε κάποιες χρήσιμες πληροφορίες που πρέπει να γνωρίζουμε πριν μπούμε στο πειραματικό μέρος.

Αναλυτικότερα, θα ξεκινήσουμε με το αναλογικό σήμα που φθάνει στο A/D μετατροπέα για τους σκοπούς αυτής της συζήτησης, θα υποθέσουμε ότι εργαζόμαστε με 24bit στα 48KHz. Το αναλογικό σήμα εισέρχεται στο σύστημα ως δείγματα ήχου 24bit τα οποία έχουν ένα θεωρητικό δυναμικό εύρος 144dB (θυμηθείτε 1bit = 6.02dB δυναμικό εύρος). Αυτό, φυσικά είναι περιορισμένο με βάση την απόδοση του μετατροπέα. Αξίζει να σημειωθεί ότι όταν συγκρίνουμε προδιαγραφές μετατροπέων, το δυναμικό εύρος ορίζεται ως η διαφορά μεταξύ των peak power levels του δυνατότερου ηχητικού σήματος μείον το self noise του συγκεκριμένου μηχανήματος. Οι πολύ καλής απόδοσης μετατροπείς προσφέρουν περίπου 120dB δυναμικό εύρος. Στη συνέχεια, το 24bit σήμα τοποθετείται σε 24bit TDM bus και μεταφέρεται στους DSPs για την plug-in επεξεργασία. Το TDM επιτρέπει η περίοδο δείγματος να κομματιάζεται σε 512 " timeslots (χρονοθύρες) ". Σε επίπεδο DSP , πολλά από τα σημερινά plug- ins επεξεργασίας ήχου χρησιμοποιούν «διπλής ακρίβειας» μαθηματικά. Διπλή ακρίβεια σημαίνει ακριβώς αυτό αντί για 24bits, τα αποτελέσματα των υπολογισμών να μεταφέρονται και να αποθηκεύονται ως 48bits. Αυτό σημαίνει ότι περισσότερα από το δεκαδικά μετά την υποδιαστολή που αντιπροσωπεύει ο ήχος διατηρούνται κατά τη διάρκεια της ψηφιακής επεξεργασία σήματος, η οποία, πέρα από πολλαπλά στάδια επεξεργασίας, μπορεί να μεταφραστεί σε υψηλής ποιότητας ήχο.

Υπάρχει ένα στάδιο dithering στα περισσότερες plug-ins διπλής ακρίβειας (double precision) και ένα τελικό στάδιο dithering μετά το master output του μίκτη άθροισης. Το dither

noise είναι ένας θόρυβος που προστίθεται στο σήμα για να αποσυσχετίσει τον θόρυβο δαπέδου (noise floor) από το αρχικό σήμα. Το dither noise προστίθεται στο σήμα 48bit, πριν μετατραπεί σε 24bit και τοποθετηθεί στο TDM bus για να σταλεί έπειτα στους DACs.

3.4 Shifting (μετακίνηση)

Προχωρώντας θα αναλύσουμε τον ορό sifting (μετακίνηση) που συναντάμε μιας και είναι μια διεργασία που γίνεται στο audio engine των DAW. Ας ρίξουμε μια πιο προσεκτική ματιά στο πώς ο μίκτης χρησιμοποιεί την μετακίνηση των bit (bit shifting) για την βελτιστοποίηση του headroom και του “footroom” διατηρώντας παράλληλα τη μέγιστη δυνατή διαφάνεια. Η μετακίνηση των bit περιγράφει την κίνηση μιας δυαδικής λέξης σε ένα καταχωρητή (register), και χρησιμοποιείται για την αύξηση ή την μείωση του κέρδους (gain). Μπορεί να είναι μια ρητή εντολή (αριθμητική μετακίνηση αριστερά ή δεξιά) ή μπορεί να επιτευχθεί με πολλαπλασιασμό επί τον συντελεστή με την δύναμη του 2. Μία αρνητική δύναμη του 2, προκαλεί «δεξιά μετακίνηση» (δηλαδή, μείωση της στάθμης) και μια θετική δύναμη προκαλεί μια «αριστερή στροφή» (δηλαδή, να αποκτήσουν αύξηση). Μετακίνηση με πολλαπλασιασμό μετακινεί τον αριθμό με το μέγεθος της δύναμης με την οποία υψώνεται (π.χ. 2 στην 4η δύναμη παράγει μια αριστερή μετατόπιση από 4 bits). Στο μίξερ των Pro Tools, τα σήματα εισόδου πολλαπλασιάζονται με συντελεστές που αντιπροσωπεύουν gain και το κέρδος (gain), καθώς εισέρχονται στον μίξερ, τα οποία παράγουν λέξεις 48bit από τις αρχικές λέξεις που ήταν 24bit (ένας αριθμός 24bit πολλαπλασιάζομενος με έναν αριθμό 24bit παράγει έναν αριθμό 48bit) . Αυτή η αρχική πράξη του πολλαπλασιασμού τοποθετεί την αρχική λέξη των 24bit στη μέση περίπου του καταχωρητή (register) των 56bit με κορυφές και μηδενικά που συνθέτουν το υπόλοιπο των 48bits. Μετά την αρχική πράξη του πολλαπλασιασμού, τα σήματα εισόδου αθροίζονται και μετακινούνται προς τα δεξιά και πάλι. Στη συνέχεια, το τελικό αποτέλεσμα του μίξερ - η έξοδος στέλνεται στο DAC ή σε plug- ins στο Master Fader, το dithered αποτέλεσμα των 24bit μετακινείται αριστερά με ακριβώς το ίδιο αντίστροφο ποσό με τη συσσωρευμένη δεξιά μετακίνηση, επιστρέφοντας έτσι την αρχική στάθμη σήματος στους DACs.

3.4.1 Η διαδικασία του shifting αναλυτικά

Στην συνέχεια παραθέτουμε την παραπάνω διαδικασία με μία εικόνα και θα την αναλύσουμε βήμα-βήμα.

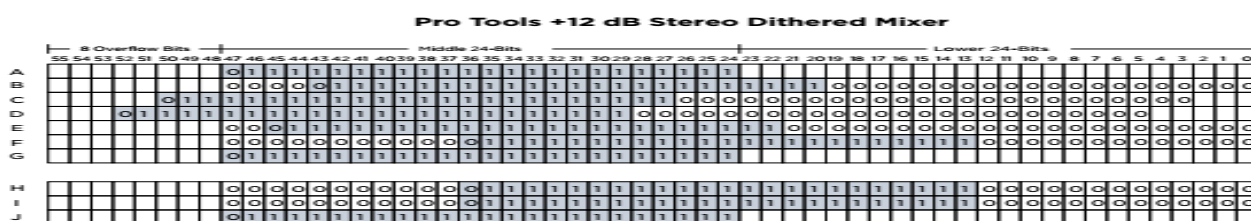


Figure 3: Pro Tools +12 Stereo mixer from the 56-bit register perspective

Εικόνα 3.9 Ο +12dB stereo mixer των pro tools από την πλευρά του 56bit καταχωρητή¹²

¹² Η εικόνα προέρχεται από το white paper της avid «The pro tools 48bit mixer».

- **Η Σειρά Α** απεικονίζει ένα ενιαίο πλήρη κωδικό εισόδου 24bit πριν εισέλθει στο μίξερ του plug-in.
- **Η Σειρά Β** δείχνει την ίδια είσοδο καθώς αυτή τοποθετείται στον mixer. Έχει εφαρμοστεί μοναδιαίο κέρδος (unity gain) και pan left. Είναι το ίδιο σήμα σε μία λέξη 48bit "μετατοπισμένο" 4 δυαδικά ψηφία προς τα δεξιά.
- **Οι σειρές C και D** δείχνουν 128 ολοκληρωμένου κώδικα κανάλια αθροιζόμενα μαζί και +12 στην συνέχεια, καταναλώνοντας κάποια από τα bits του head-room.
- **Η Σειρά Ε** δείχνει την έξοδο του μίκτη, καθώς όλα τα κανάλια αθροίζονται σε ένα ζεύγος εξόδων. Η λέξη των 48bit μετακινείται δεξιά κατά 7bit για βγει έξω από την περιοχή υπερχείλισης (overflow area) και να δώσει +12 στάδιο του κέρδους (gain stage).
- **Η Σειρά F** απεικονίζει τη θέση της λέξης των 48bit στην έξοδο του mixer όταν το Master Fader έχει οριστεί στα -42dB. Το αποτέλεσμα είναι να οδηγήσει την έξοδο του mixer στο να μην κλιπάρει και να την επαναφέρει στα 0,0dB.
- **Η σειρά G** δείχνει την έξοδο 24bit του mixer από τη γραμμή F όπως φαίνεται στο DAC. Η λέξη των 48bit περνά από το στάδιο του dither πριν μειωθεί στα 24bit, διατηρώντας έτσι χαμηλά επίπεδα λεπτομέρειας. Το αποτέλεσμα έχει μετακινηθεί προς τα αριστερά κατά 11bit για να αναπαραστήσει το αρχικό σήμα εισόδου με τα 24bit του να παραμένουν άθικτα.
- **Η σειρά Η** απεικονίζει τη χρήση του "footroom". Αυτό είναι μια λέξη 48bit που αντιπροσωπεύει ένα hard panned channel Fader που έχει μειωθεί στα -42.1dB. Και πάλι, η αρχική λέξη των 24-bit παραμένει άθικτη και μετατοπισμένη προς τα δεξιά. Παρατηρήστε ότι εξακολουθούν να υπάρχουν 8bit του footroom στα αριστερά. Τα 15bit κάτω από το μοναδιαίο κέρδος χρησιμοποιούνται για τη διατήρηση χαμηλής έντασης λεπτομέρεια (low level detail) κάτω στα -90dB. Η αρχική λέξη των 24bit δεν θα υποστεί καμία απώλεια μέχρι το fader του καναλιού κινηθεί κάτω από τα -90dB.
- **Η σειρά Ι** δείχνει την έξοδο του mixer με 128 εισόδους πλήρους κώδικα αθροισμένες στα -42.1dB και το Master Fader να έχει οριστεί στα 0dB. Το αποτέλεσμα εξόδου του mixer είναι σχεδόν πανομοιότυπο με το αρχικό κανάλι στο μοναδιαίο κέρδος (unity gain), καθώς και τα 128 κανάλια αθροίζονται σε ένα με το Master Fader να ορίζεται στα -42.1dB - αυτές οι τρεις λειτουργίες έχουν ως αποτέλεσμα σχεδόν ακριβώς την ίδια έξοδο.
- **Τέλος η σειρά J** δείχνει την έξοδο των 24bit, όπως αυτή εμφανίζεται στον DAC. Όπως και στη σειρά F, το αποτέλεσμα έχει μετατοπιστεί προς τα αριστερά κατά 11bit, πράγμα που επαναφέρει το αρχικό σήμα.

3.5 Απαντήσεις σε ερωτήματα που έχουν να κάνουν με τα audio engines

Στην συνέχεια θα δώσουμε κάποιες απαντήσεις σε θέματα που υπάρχουν μέσα στο συγκεκριμένο white paper της εταιρείας avid. Η εταιρεία αναφέρει συγκεκριμένα ότι:

«Μπορεί η διαύγεια και η λεπτομέρεια να βελτιωθούν αθροίζοντας εξωτερικά;»

Μια άλλη παρανόηση είναι ότι υπάρχει σαφήνεια και διατήρηση της λεπτομέρειας που παρέχεται από την εξωτερική άθροιση δηλαδή μιζάροντας σε ομάδες (mixing to stems) και χρησιμοποιώντας ένα αναλογικό μίκτη για να τους ανασυνδυάσεις. Είναι πολύ κοινό σε αυτή τη δουλειά η εφαρμογή άλλων ηχητικών μεθόδων από hardware effects, όπως αναλογική συμπίεση και EQ που μπορεί να είναι υποκειμενικά ευχάριστο. Αλλά αυτά πρέπει να αναγνωρίζονται ως στοιχεία που είναι συνειδητά προστίθενται στη μίξη όχι για τι διόρθωση εξ' αιτίας μια ελλιπής ψηφιακής κονσόλας. Σκεφτείτε ότι κάθε σήμα που έχει καταγραφεί, έχει υποστεί panning, έχει αθροιστεί, ή έχει υποβληθεί σε επεξεργασία (δηλαδή, όλα τα σήματα) γίνεται από το mixer Pro Tools. Όλες οι υπομίξεις (submixes) καθώς και επιμέρους εξόδους των καναλιών γίνονται από το μίξερ, είτε αθροίζονται εσωτερικά είτε στέλνονται κατ' ευθείαν σε εξωτερικές εξόδους. Όπως και με κάθε υψηλής ποιότητας ψηφιακό σύστημα μίξης, ο mixer των Pro Tools έχει σχεδιαστεί για να παρέχει ένα αποτέλεσμα που είναι ηχητικά διάφανο αυτός είναι ο λόγος που παρέχεται αυτό το επίπεδο της ακρίβειας.

«Χαμηλώνοντας το master fader διακινδυνεύουμε την ανάλυση;»

Αυτό φέρνει στην επιφάνεια ένα άλλο σημαντικό σημείο - ενώ μπορούμε να αθροίσουμε εκατοντάδες κανάλια στον μίξερ Pro Tools χωρίς να κλιπάρουμε την είσοδο του mix bus, μπορούμε εύκολα να κλιπάρουμε την έξοδο του (όταν θα πάμε σε ένα 24bit DAC ή 24bit ψηφιακή έξοδο, ή σήματα Submix σε Aux Input). Γι' αυτό είναι τόσο σημαντικό και χρήσιμο να χρησιμοποιούμε master faders - που δεν χρησιμοποιούν DSP, παρέχει μέτρηση που μας επιτρέπει να γνωρίζουμε αν κλιπάρουμε στην πορεία της μίξης μας, και μας επιτρέπει να τριμάρουμε την ένταση εξόδου της υπομίξης ή της τελικής μίξης χωρίς να χάσουμε καμία ανάλυση (επειδή η ανάλυση διατηρείται από τον mixer των 48bit). Επειδή τα master Faders διατηρούν την ανάλυση κατά το κλιπάρισμα, δεν χρειάζεται να χαμηλώσουμε όλα τα faders που συμβάλλουν στη μίξη για να ληφθεί το βέλτιστο επίπεδο εξόδου. Αυτό είναι ακριβώς σχεδιασμένο όπως μια καλή αναλογική κονσόλα - απλώς θα χαμηλώσουμε το master fader για να αποφύγουμε το κλιπάρισμα του κυκλώματος εξόδου στην κονσόλα, ενώ το headroom του mix bus αφήνει τα input faders να μείνουν στο sweet spot τους.

Κεφάλαιο 4^ο: ΜΕΘΟΔΟΣ ABX DOUBLE BLIND TEST

Εισαγωγή 4^ο Κεφαλαίου

Στο κεφάλαιο αυτό γίνεται η ανάλυση της μεθόδου έρευνας ABX listening test η οποία χρησιμοποιείται στο πρακτικό μέρος (υποκειμενικό κομμάτι) για την αναγνώριση των DAW που θα χρησιμοποιήσουμε. Αναλύεται η λογική που διέπει τη μέθοδο, η δομή της, τα κριτήρια επιλογής καθώς επίσης και ορισμένα προβλήματα που παρουσιάζονται κατά τη χρήση της. Σκοπός του κεφαλαίου αυτού είναι η βαθιά και πλήρης κατανόηση της μεθοδολογίας της με σκοπό την ορθή και έγκυρη χρήση της κατά το πρακτικό μέρος.

Μιλώντας για DAW πάντα αναφερόμαστε, εξ ορισμού, στην κυρίως "μηχανή" (audio engine) του προγράμματος που το κάνει να μπορεί να επεξεργάζεται audio ψηφιακά αρχεία. Επίσης δεδομένου, ότι είναι δυνατόν θεωρητικά κάποιος να χρησιμοποιεί ένα DAW απλά και μόνο για να ηχογραφεί όργανα, φωνές κτλ. δηλαδή να χρησιμοποιεί εξωτερικές πηγές ήχου και όχι εικονικά όργανα και effects, είναι εμφανές ότι λέγοντας DAW εννοούμε ακριβώς τις δυνατότητες του DAW για sequencing και summing αυτών των ηχογραφήσεων (το αποτέλεσμα της σύμπτυξης πολλών ηχογραφήσεων σε μία), δηλαδή την "τοποθέτηση" τους στο χρόνο και mixing αυτών, δηλαδή το μίξάρισμα τους προκειμένου να προκύψει ένα τελικό audio αρχείο που συμπεριλαμβάνει μέρος (ή και το σύνολο) αυτών των ηχογραφήσεων.

Έτσι λοιπόν, και στα DAW δεν γίνεται να κρίνουμε τη "μηχανή" τους κάνοντας εγγραφή με διαφορετική κάρτα ήχου κάθε φορά, ή χρησιμοποιώντας διαφορετικό plug in instrument ή effect στο ένα από το άλλο. Θα πρέπει λοιπόν οι όποιες δοκιμές γίνουν στα DAW να είναι σίγουρο ότι διατηρούν όλες τις άλλες παραμέτρους ίδιες που μπορεί να άλλαζαν το αποτέλεσμα. Υπάρχουν λοιπόν πολλοί παράμετροι, που σε μια δοκιμή δύο DAW καθαρά για την μηχανή ήχου τους θα πρέπει να είναι ίσες. Είθισται στη περίπτωση των DAW οι δοκιμές να γίνονται με το ίδιο προηχογραφημένο υλικό σαν πρώτη πηγή ώστε να αποφύγουμε τις διαφορές που θα προέκυπταν από διαφορετικές πρώτες πηγές.

Όταν αναλύουμε ανεπαίσθητες διαφορές ήχου θα πρέπει να βρούμε έναν μηχανισμό σύγκρισης που να μην εξαρτάται από τις αισθήσεις μας και τις λάθος υποθέσεις του εγκεφάλου μας, αφού αυτά μπορούν να παρασυρθούν ή να ξεγελαστούν είτε λόγω της αδυναμίας των αισθήσεων είτε λόγω αδυναμίας του εγκεφάλου να είναι 100% αντικειμενικός (διότι ήδη από πριν π.χ. θεωρούμε ότι υπάρχει διαφορά, το λεγόμενο bias, προκατάληψη) και πάντα ασυνείδητα.

Η έρευνα για θέματα που άπτονται της υποκειμενικότητας και της κρίσης του κάθε ανθρώπου είναι μια δύσκολη διαδικασία. Υπάρχουν πολλοί παράγοντες οι οποίοι μπορούν να επηρεάσουν και να αλλοιώσουν τα αποτελέσματα, καθιστώντας τα άκυρα και μη ρεαλιστικά. Για να θεωρείται μια μεθοδολογία έρευνας σωστή και έγκυρη, είτε πρόκειται για μαθησιακή μέθοδο, είτε για ψυχολογική, είτε για μέθοδο αντίληψης πρέπει να πλήρη ορισμένα αντικειμενικά κριτήρια, τα βασικότερα εκ των οποίων είναι τα παρακάτω:

- **Αντικειμενικότητα:** Από τη μέθοδο θα πρέπει να εξάγονται αποτελέσματα τα οποία να είναι επαληθεύσιμα και ανεξάρτητα από την ιδιοσυγκρασία και την κρίση του ερευνητή.
- **Αξιοπιστία:** Σε περίπτωση επανάληψης ακριβώς της ίδιας διαδικασίας κάτω από τις ίδιες συνθήκες θα πρέπει να εξάγονται τα ίδια συμπεράσματα, με μικρά περιθώρια αποκλίσεων.
- **Εγκυρότητα:** Τα δεδομένα που θα προκύψουν πρέπει να είναι παραπλήσια με τα δεδομένα που θα εξάγονταν για το ίδιο θέμα άλλα με τη χρήση διαφορετικής μεθοδολογίας.
- **Ευαισθησία:** Η μέθοδος θα πρέπει να αποδίδει σωστά τις διαφοροποιήσεις ανάλογα με το θέμα το οποίο μελετάται.
- **Συγκρισιμότητα:** Η μέθοδος πρέπει να είναι εφαρμόσιμη σε μεγάλο εύρος θεμάτων και σε πολλά θεματικά πεδία και να δίνεται η δυνατότητα σύγκρισης των αποτελεσμάτων ανάμεσα σε διαφορετικές ομάδες ατόμων και σε μεμονωμένα άτομα.
- **Χρηστικότητα:** Η μέθοδος πρέπει να αποδίδει πληροφορίες σχετικές με τα θεωρητικά η πρακτικά ζητήματα τα οποία εξετάζει με τρόπο αποδοτικό. Για παράδειγμα, δεν θα πρέπει η μέθοδος να είναι κατασκευασμένη με τέτοιο τρόπο ώστε να κουράζει τα υποκείμενα ούτε και να καθοδηγεί με τον οποιοδήποτε τρόπο τη σκέψη τους.

Η τήρηση αυτών των προϋποθέσεων είναι απαραίτητη για τη σωστή ερευνητική μεθοδολογία και για την ορθή εξαγωγή των αποτελεσμάτων. Η επιλογή της μεθόδου πρέπει πάντα να ακολουθεί αυτά τα κριτήρια για να θεωρηθεί έγκυρη.

4.1 Σύντομο ιστορικό abx μεθόδου

Πολλά πειράματα έχουν αποδείξει ότι οι ηχητικές διαφορές που οι ακροατές μπορούν να ακούσουν μεταξύ των πηγών ήχου είναι μερικές φορές προϊόν της φαντασίας τους. Αυτές οι αυταπάτες μπορεί να είναι ισχυρές, να παραμένουν με το πέρασμα των χρόνων, να συμμερίζονται από πολλούς ακροατές και σχετίζονται καθαρά με τη γνώση και τα χαρακτηριστικά της πλατφόρμας και της πηγής ήχου που τα αναπαράγει.

Ιστορικά τώρα το 1977 ο Arnold B. Krueger και ο Bern Muller, και οι δύο μέλη της Michigan Woofers and Tweeters Marching Society (SMWTMS), εφηύραν το ABX Double Blind comparator system, προκειμένου να διευθετήσουν μια συζήτηση εάν υπάρχουν ακουστές διαφορές μεταξύ ενός καλά κατασκευασμένου ενισχυτή σε σχέση με έναν μέτριο. Στις 7 Μαΐου 1977 το SMWTMS οργάνωσε τα πρώτα τρία double blind listening test χρησιμοποιώντας τον ABX συγκριτή του Krueger και του Muller. Κατόπιν της συνεδρίασης μία εταιρεία που θα κατασκεύαζε και θα πουλούσε τους ABX συγκριτές πήρε το όνομά της ABX Corporation. Αργότερα ο David Clark, μέλος του Audio Engineering Society (AES), θα συνέχιζε την βελτίωση, την προώθηση και το marketing του ABX συγκριτή. Στις 22 Σεπ. 1999 ο Krueger ξεκίνησε μια ιστοσελίδα πλήρως αφιερωμένη στην εκπαίδευση σχετικά με τις δοκιμές που περιλαμβάνουν ABX listening tests με εφαρμογές λογισμικού για download.

4.1.1 Τι είναι ένα τυφλό πείραμα (blind experiment)

Ένα τυφλό πείραμα είναι ένα πείραμα στο οποίο οι πληροφορίες σχετικά με το τεστ που μπορούν να μας οδηγήσουν σε αποκλίσεις των αποτελεσμάτων έχουν αποκρυφτεί από τον υπεύθυνο του πειράματος, τον ακροατή, ή και τους δύο μαζί μέχρις ότου να τελειώσει το πείραμα. Οι αποκλίσεις μπορεί να είναι εκούσιες ή ασυνείδητες. Αν και οι δύο ο εξεταστής και ο ακροατής δεν γνωρίζουν στοιχεία για το πείραμα τότε η δοκιμή ονομάζεται double-blind test μιας και κανένας από τους δύο δεν γνωρίζει στοιχεία και αποτελέσματα.

Ένα blind test συνήθως επιβάλλεται για τους ερευνητές, τους τεχνικούς, τους ακροατές, και τους χρηματοδότες. Το αντίθετο ενός blind trial είναι ένα open trial. Τα blind πειράματα είναι ένα σημαντικό εργαλείο της επιστημονικής μεθόδου, σε πολλούς τομείς της έρευνας, της ιατρικής, της ψυχολογίας των κοινωνικών επιστημών, των φυσικών επιστημών, όπως η φυσική και η βιολογία, επίσης εφαρμοσμένες επιστήμες, όπως η έρευνα αγοράς, και πολλά άλλα. Σε ορισμένες ειδικότητες, όπως στον έλεγχο των φαρμακευτικών ναρκωτικών, τα τυφλά πειράματα θεωρούνται απαραίτητα.

Στην συγκεκριμένη πειραματική διαδικασία η υποκατηγορία του blind experiment που θα χρησιμοποιήσουμε ονομάζεται **double-blind trial** και παρακάτω θα αναλύσουμε τί είναι μια **double-blind** δοκιμή.

4.1.2 Double-blind trial

Η double-blind δοκιμή περιγράφει ένα ιδιαίτερα αυστηρό τρόπο διεξαγωγής ενός πειράματος που σκοπό έχει να εξαλείψει τις υποκειμενικές και μη βάσιμες προκαταλήψεις που μεταφέρονται από άτομα και υπευθύνους ενός πειράματος. Συχνά, αλλά όχι αποκλειστικά εφαρμόζεται για πειράματα σε δοκιμές που ασχολούνται με την υποκειμενικότητα των ανθρώπων και χρησιμοποιήθηκε για πρώτη φορά από τον WHR Rivers στην έρευνα του "war neurosis», πιο γνωστή ως «shell shock». Στις περισσότερες περιπτώσεις, τα double blind πειράματα καταφέρνουν να πετύχουν σε ένα υψηλότερο επιστημονικό επίπεδο τους σκοπούς τους σε σχέση με ένα single blind ή ένα non-blind experiment.

Σε αυτά τα double-blind πειράματα, ούτε οι συμμετέχοντες ούτε οι ερευνητές γνωρίζουν ποιος συμμετέχον ανήκει στην ομάδα ελέγχου και ποιος στην ομάδα δοκιμής. Μόνο μετά από όλα τα στοιχεία που έχουν καταγραφεί (και σε ορισμένες περιπτώσεις, αναλυθεί) μπορούν οι ερευνητές να μάθουν ποιός ήταν ποιός. Η εκτέλεση ενός πειράματος με double blind τρόπο μπορεί να ελαττώσει σημαντικά το αίσθημα της προκατάληψης ή του φυσικού αισθήματος (π.χ. το φαινόμενο placebo, προκαταλήψεις του παρατηρητή-προκατάληψη πειραματιστή) να νοθεύσει τα αποτελέσματα (κάνοντας ερευνητές/συμμετέχοντες να συμπεριφερθούν διαφορετικά από την καθημερινή ζωή). Τυχαία ανάθεση των θεμάτων δοκιμής για τους ακροατές και την ομάδα ελέγχου είναι ένα κρίσιμο μέρος οποιουδήποτε double blind τεστ.

Οι double blinds μέθοδοι μπορούν να εφαρμοστούν σε οποιαδήποτε πειραματική κατάσταση στην οποία υπάρχει μια πιθανότητα ότι τα αποτελέσματα θα επηρεαστούν από συνειδητή/ασυνειδητή προκατάληψη εκ μέρους των ερευνητών των συμμετεχόντων, ή και των δύο. Για παράδειγμα, σε μελέτες σε ζώα, τόσο ο φροντιστής των ζώων όσο και ο αξιολογητής των αποτελεσμάτων πρέπει να μην γνωρίζουν στοιχεία αλλιώς ο φροντιστής μπορεί να αντιμετωπίσει τα θέματα με διαφορετικό τρόπο και να αλλοιώσει τα αποτελέσματα.

Η ανάπτυξη των ερευνών που παρουσιάζονται σε θέματα μέσω των υπολογιστών δείχνουν ότι προκατάληψη μπορεί εύκολα να δημιουργηθεί πάνω στη διαδικασία. Συστήματα ψηφοφορίας είναι επίσης παραδείγματα όπου η μεροληψία μπορεί εύκολα να δημιουργηθεί σε ένα φαινομενικά απλό σύστημα που βασίζεται σε μηχανή. Σε αναλογία με την ανθρώπινη έρευνα που περιγράφεται παραπάνω, το τμήμα του λογισμικού που παρέχει την αλληλεπίδραση με τον άνθρωπο παρουσιάζεται στο ακροατή ως blind ερευνητής, ενώ το τμήμα του λογισμικού που καθορίζει το X είναι ο τρίτος. Ένα παράδειγμα είναι το ABX τεστ, όπου το ανθρώπινο υποκείμενο έχει να προσδιορίσει ένα άγνωστο ερέθισμα X που μπορεί να είναι είτε A είτε B.

4.2 Ανάλυση της Μεθόδου

Η ανάλυση της μεθόδου θα γίνει με βάση τη λογική την οποία ακολουθεί, τη δομή που πρέπει να έχουν τα μετρητικά όργανα που χρησιμοποιούνται (ερωτηματολόγια-λογισμικά), την αξιολόγηση της καθώς και ορισμένα προβλήματα που μπορεί κανείς να συναντήσει κατά την εφαρμογή της.

4.2.1 Η Λογική της Μεθόδου

Ένα διπλό double blind Test (DBT) είναι μια ρύθμιση ακρόασης που επιτρέπει να επιβεβαιώσουμε ότι μια συγκεκριμένη ηχητική διαφορά προκλήθηκε πράγματι από τις πηγές ήχου, και όχι από τις εντυπώσεις και απόψεις του ακροατή.

Σε ένα ABX listening test ο ακροατής έχει πρόσβαση σε τρεις πηγές που ονομάζονται A, B, και X. A και B είναι οι αναφορές. Για παράδειγμα, A το αρχείο .wav και B το αρχείο .mp3. X είναι η άγνωστη πηγή. Μπορεί να είναι ή A ή B. Ο ακροατής πρέπει να μαντέψει ποιά από τις δύο αναφορές (A ή B) αναπαράγεται κάθε φορά στην X επανάληψη.

Αλλά αν ο ακροατής λέει ότι ο X είναι A, και το X είναι στην πραγματικότητα A. Τι αποδεικνύει αυτό; Τίποτα φυσικά. Διότι και ένα κέρμα να ρίξουμε τις μισές φορές συνήθως θα έρθει η πλευρά που έχουμε επιλέξει. Αυτό είναι ακριβώς τύχη, τίποτα περισσότερο όπως και στην ακρόαση. Αυτός είναι ο λόγος για τον οποίο μια στατιστική ανάλυση είναι αναγκαία. Αναφέροντας αυτά γίνεται γρήγορα αντιληπτό ότι η ακρόαση δεν μπορεί να γίνει μία φορά διότι το αποτέλεσμα θα είναι θέμα τύχης. Έτσι θέλουμε ένα ικανοποιητικό αριθμό επαναλήψεων για να μπορέσουμε να είμαστε σίγουροι ότι η πειραματική διαδικασία έγινε σωστά και αξιόπιστα και πληρή τις προϋποθέσεις που αναφέραμε παραπάνω.

4.3 Δομή ABX double blind test

Όταν σχεδιάζουμε να πραγματοποιήσουμε ένα ABX test πρέπει να φροντίσουμε να εξαλείψουμε δύο πράγματα από την πειραματική μας διαδικασία το ένα είναι η προδιάθεση και το άλλο είναι η τύχη. Με τη λέξη προδιάθεση εννοούμε ότι όταν πραγματοποιείται το τεστ ούτε ο ακροατής που είναι εκεί να επιλέξει τις απαντήσεις ούτε και ο υπεύθυνος του πειράματος θα πρέπει να έχουν στοιχεία για τον άγνωστο X που αναπαράγεται. Το να μην γνωρίζει ο ακροατής τα αποτελέσματα είναι λογικό μιας και αλλιώς δεν θα μπορούσε να γίνει το πείραμα, αλλά γνωρίζοντας ο υπεύθυνος του πειράματος παραπάνω στοιχεία μπορεί με την συμπεριφορά του και τις αντιδράσεις του να προδώσει τα αποτελέσματα στον εξεταζόμενο (Clever Hans)¹³. Έτσι το τεστ θα πρέπει να είναι double blind.

Στην συνέχεια θα πρέπει να εξαλείψουμε και το θέμα τύχης. Λέγοντας τύχη εννοούμε ότι έστω και αν δεν ξέρουμε τίποτα για το ποιο είναι το αρχείο X που αναπαράγεται (δηλαδή A ή B) έχουμε 50% πιθανότητα να απαντήσουμε σωστά έχοντας επιλέξει στην τύχη την απάντηση μας χωρίς να έχουμε ακούσει καν το αρχείο X. Είναι βασικά το ίδιο σαν να στρίβουμε ένα νόμισμα και να ρωτάμε κάποιον αν θα έρθει κορώνα ή γράμματα έχει δηλαδή 50% πιθανότητες να απαντήσει σωστά. Έστω όμως ότι έχουμε ένα άτομο που υποστηρίζει ότι μαντεύει σωστά πότε θα έρθει κορώνα ή γράμματα πρέπει να σκεφτούμε κάτω από ποιές προϋποθέσεις θα πιστέψουμε τους ισχυρισμούς του. Εάν μαντέψει σωστά 1/1 δεν είναι κάτι σημαντικό διότι είπαμε ότι είχε 50% πιθανότητες. 3/3 όμως μπορούμε να πούμε ότι έχει ταλέντο στο να το μαντεύει. Τι γίνεται όμως στα 4/5 εδώ εισέρχονται τα μαθηματικά και η στατιστική.

Για τέτοιου είδους πειράματα έχει δημιουργηθεί ο όρος confidence level και παρακάτω θα δοθεί ο πίνακας με τα στοιχεία που υποστηρίζουν ποιά αποτελέσματα είναι αποδεκτά. Αυτό που δηλώνει ο όρος confidence είναι ότι ο ακροατής έχει μια σιγουριά-αυτοπεποίθηση (confidence) σε μια τέτοια συνεδρία. Έτσι σε ένα ABX test ισχύει γενικά σαν πρακτική να υπάρχει 95% confidence level¹⁴ όπου ένας ακροατής είναι ικανός να προσδιορίσει ή να βρει τις διαφορές ανάμεσα στα αρχεία X, A και B άρα αυτό μας δείχνει ότι ήταν 95% σίγουρος ότι δεν βρήκε τις διαφορές επιλέγοντας τυχαία κάθε φορά την απάντηση. Παρακάτω δίνονται δύο πίνακες θα κάνουμε μία εκτενή ανάλυση όλης της στατιστικής ανάλυσης και πως ερμηνεύονται τα αποτελέσματα τέτοιων πειραμάτων αλλά και βασικών όρων της abx λογικής.

4.3.1 Null hypothesis

Στη στατιστική συμπερασματολογία των παρατηρούμενων στοιχείων ενός επιστημονικού πειράματος, η μηδενική υπόθεση (null hypothesis) αναφέρεται σε μια γενική δήλωση ή θέση ότι δεν υπάρχει καμία σχέση μεταξύ των δύο φαινομένων που μετρούνται. Απόρριψη ή ανασκευή αυτής της υπόθεσης – και οδηγούμαστε στο συμπέρασμα ότι υπάρχουν

¹³ Άλογο το οποίο ήταν ικανό να κάνει μαθηματικές πράξεις. Ύστερα αποδείχτηκε ότι παρατηρούσε τις κινήσεις των ακροατών για να πραγματοποιήσει τις πράξεις.

¹⁴ David Clark, 1982 High-Resolution Subjective Testing Using a Double-Blind Comparator, ABX Company, p.333

λόγοι να πιστεύουμε ότι υπάρχει μια σχέση μεταξύ δύο φαινομένων ή ότι πχ. μια πιθανή θεραπεία έχει ένα μετρήσιμο αποτέλεσμα - είναι ένα βασικό έργο στη σύγχρονη πρακτική της επιστήμης, και δίνει μια ακριβή έννοια με την οποία η αξίωση είναι δυνατόν να αποδειχθεί ψευδής.

Στο στατιστικό κομμάτι, η μηδενική υπόθεση (null hypothesis) συχνά συμβολίζεται H_0 και, γενικά, θεωρείται αληθινή μέχρι τα στοιχεία να αποδείξουν το αντίθετο¹⁵. Η έννοια της μηδενικής υπόθεσης χρησιμοποιείται διαφορετικά με δύο προσεγγίσεις στη στατιστική συμπερασματολογία. Με την προσέγγιση του Ronald Fisher, η μηδενική υπόθεση δυνητικά απορρίπτεται ή διαψεύδεται επί τη βάση των δεδομένων αλλά γενικά αυτή η προσέγγιση ποτέ δεν αποδέχθηκε ή αποδείχτηκε. Στην προσέγγιση των Jerzy Neyman και Egon Pearson, η μηδενική υπόθεση είναι σε αντίθεση με μια εναλλακτική υπόθεση, και αυτά διακρίνονται με βάση τα δεδομένα, με ορισμένα ποσοστά σφάλματος. Οι υποστηρικτές αυτών των δύο προσεγγίσεων επικρίνουν ο ένας τον άλλον, αν και σήμερα μια υβριδική προσέγγιση είναι ευρέως διαδεδομένη πρακτική και παρουσιάζεται σε εγχειρίδια. Αυτή η υβριδική προσέγγιση με τη σειρά της επικρίθηκε ως εσφαλμένη και ασυνάρτητη. Η στατιστική σημαντικότητα διαδραματίζει έναν κεντρικό ρόλο στο στατιστικό έλεγχο υποθέσεων, όπου χρησιμοποιείται για να καθορίσει εάν μια μηδενική υπόθεση μπορεί να απορριφθεί ή να διατηρηθεί.

4.3.2 Στατιστικά σημαντικά αποτελέσματα (statistical significance results)

Τα στατιστικά σημαντικά αποτελέσματα παίζουν έναν κεντρικό ρόλο στον στατιστικό έλεγχο των υποθέσεων, όπου χρησιμοποιούνται για να καθορίσουν εάν μια μηδενική υπόθεση (null hypothesis) πρέπει να απορριφθεί ή να διατηρηθεί. Η μηδενική υπόθεση είναι η δήλωση εξ' αρχής ότι δεν έγινε τίποτα ή άλλαξε κάτι.¹⁶ Για να απορριφθεί ως ψευδής μια μηδενική υπόθεση πρέπει, τα αποτελέσματα να αναγνωριστούν ως στατιστικά σημαντικά, δηλαδή απίθανο να έχουν συμβεί από τύχη και μόνο.

Για να προσδιορίσουμε αν το αποτέλεσμα είναι στατιστικά σημαντικό, θα πρέπει να υπολογίσουμε την p-value, η οποία είναι η πιθανότητα παρατήρησης μιας επίπτωσης, δεδομένου ότι η μηδενική υπόθεση είναι αληθινή.¹⁷ Η μηδενική υπόθεση απορρίπτεται όταν η p-value είναι μικρότερη από το significance ή το α level. Το α level είναι η πιθανότητα απόρριψης της μηδενικής υπόθεσης, ότι είναι αλήθεια και πιο συχνά ορίζεται σε 0.05 (5%). Εάν το α level είναι 0,05 τότε η υπό όρους πιθανότητα ενός σφάλματος, δεδομένου ότι η μηδενική υπόθεση είναι

¹⁵ Moore, David; McCabe, George (2003). Introduction to the Practice of Statistics (4 ed.). New York: W.H. Freeman and Co. p. 438

¹⁶ Meier, Kenneth J.; Brudney, Jeffrey L.; Bohte, John (2011). Applied Statistics for Public and Nonprofit Administration (3rd ed.). Boston, MA: Cengage Learning. pp. 189–209.

¹⁷ Devore, Jay L. (2011). Probability and Statistics for Engineering and the Sciences (8th ed.). Boston, MA: Cengage Learning. pp. 300–344.

αληθινή, είναι 5%.¹⁸ Στη συνέχεια, ένα στατιστικά σημαντικό αποτέλεσμα είναι όταν η παρατηρούμενη p-value είναι μικρότερη από 5%, το οποίο είναι επίσημα γραμμένο ως $p < 0.05$.¹⁹

Στην συνέχεια θα αναλύσουμε τον όρο p-value αλλά και τους στατιστικούς πίνακες με βάση τους οποίους θα αξιολογήσουμε τα αποτελέσματα των ακροάσεων

4.3.3 p-value

Η στατιστική είναι η επιστήμη των πιθανοτήτων. Δεν υπάρχουν απόλυτες δοκιμές ή απόλυτες αποδείξεις. Το αποτέλεσμα μιας στατιστικής ανάλυσης είναι επίσης μια πιθανότητα. Εάν το αποτέλεσμα είναι τυχαίο ή όχι αναφέρεται στατιστικά, καθώς και η πιθανότητα το αποτέλεσμα να είναι τυχαίο. Εάν ένα αποτέλεσμα είναι τυχαίο καθορίζεται από τη σύγκριση του αποτελέσματος με πειραματικά αποτελέσματα που έχουν μελετηθεί θεωρητικά. Αυτά τα σημεία αναφοράς ονομάζονται "κατανομές" (distributions). Για να ελέγξουμε αν ένα abx double blind αποτέλεσμα είναι τυχαίο, το αποτέλεσμα θα πρέπει να συγκριθεί με μια γνωστή κατανομή.

Πετώντας ένα νόμισμα επανειλημμένα είναι στατιστικά όμοιο με τα αποτελέσματα των ακροατών σε ένα abx double blind test. Το νόμισμα είναι τυχαίο η κατανομή του είναι γνωστή. Σε έναν πολύ μεγάλο αριθμό ρίψεων η κορώνα έρχεται τις μισές φορές. Αν οι ακροατές μόνο μάντευαν σε μια abx σύγκριση, το σκορ θα ήταν επίσης 50%. Εάν η βαθμολογία είναι μεγαλύτερη από 50%, μπορεί να είναι τυχαία ή όχι. Αυτό εξαρτάται από τον αριθμό των επαναλήψεων της δοκιμής. (David Carlstrom, abx statistics, abx company publication P9) Γενικά, η απόδειξη ότι τα αποτελέσματα δεν προήλθαν από τυχαίες εικασίες είναι ισχυρότερη όσο ο αριθμός των επαναλήψεων μεγαλώνει.

Εάν ρίξουμε ένα κέρμα δεκαέξι φορές, δεν έρχεται πάντα 8 φορές κορώνα. Πέντε έως έντεκα φορές να έρθει κορώνα μπορεί να συμβεί αρκετά συχνά σε 16 επαναλήψεις της ρήψης.

Η διαχωριστική γραμμή που μπορούμε να μιλάμε για στατιστικά αποδεκτά αποτελέσματα είναι μεταξύ έντεκα και δώδεκα επαναλήψεων, διότι η κατανομή της ρίψης ενός κέρματος έχει συγκριθεί με άλλες γνωστές κατανομές. Οι πιο κοντινές κατανομές είναι η Chi-Squared και η Binomial κατανομή. Το Chi-Squared test είναι εύκολο στη χρήση με έναν απλό υπολογισμό και έναν πίνακα από Chi-Squared τιμές και τις αντίστοιχες τιμές πιθανοτήτων, αλλά είναι κατά προσέγγιση για τα μικρά αποτελέσματα. Η Binomial Κατανομή είναι ακριβής αλλά οι πίνακες δεν είναι άμεσα διαθέσιμα για πολύ μεγάλους αριθμούς.

^{18, 19} Healy, Joseph F. (2009). The Essentials of statistics: A tool for Social Research (2nd ed.). Belmont, CA: Cengage learning. pp. 177-205.

4.3.3.1 Binomial table (πίνακας/κατανομή)

Για την υποκειμενική διαδικασία (ακροάσεις-listening test) του πειραματικού μέρους θα χρησιμοποιήσουμε τον Binomial πίνακα/κατανομή διότι πρώτον είναι πιο εύκολος και κατανοητός στην χρήση και δεύτερον το πρόγραμμα foobar2000 που θα πραγματοποιήσουμε τα abx test χρησιμοποιεί αυτού του τύπου τον πίνακα για την ανάλυση των στατιστικών αποτελεσμάτων.

Ο πίνακας είναι εύκολος στη χρήση. Αν, για παράδειγμα, το αποτέλεσμα του πειράματος ήταν 20/30 που έχουν αναγνωριστεί σωστά, το 20 βρίσκεται στην κορυφή της σελίδας και 30 στο πλάι (πίνακας 4.1). Ο αριθμός που συμπίπτει είναι η πιθανότητα αυτό το αποτέλεσμα να είχε προκύψει από την ρίψη ενός νομίσματος ή να είχε απαντηθεί στην τύχη. Για 20/30, η πιθανότητα που αναγράφεται στον πίνακα είναι 0.049. Αυτό σημαίνει ότι όταν έχουμε 100 τυχαίες ρήψεις νομισμάτων και οι 30 ρήψεις είναι άγνωστες δηλαδή να μην γνωρίζουμε το αποτέλεσμα, το αποτέλεσμα 20/30 ή μεγαλύτερο αναμένεται να προκύψει περίπου πέντε φορές και όχι περισσότερο. Έτσι το αποτέλεσμα 20/30 είναι ένα αρκετά σπάνιο γεγονός. Σε μια επιστημονική αναφορά αυτό θα πρέπει να δηλώνεται: ως 67% σωστό ($p = 0,049$).

Σε αυστηρά επιστημονικά πειράματα η τιμή του p επιλέγεται εκ των προτέρων ως στόχος. Η πιο κοινή τιμή p που χρησιμοποιείται σε τέτοιου είδους πειράματα είναι το 0,05. Οι χαμηλότερες τιμές όπως 0,9-0,15 μερικές φορές κατορθώνονται να επιτευχθούν εύκολα, αλλά υψηλότερες τιμές όπως 0,03-0,01 σχεδόν ποτέ δεν μπορούν να επιτευχθούν. (David Carlstrom, abx statistics, abx company publication P9)

Σε ένα abx double blind test ο στόχος είναι να διαψεύσουμε στατιστικά τη μηδενική υπόθεση (null hypothesis) για να επιβεβαιώσουμε την υπόθεση.

Ο αριθμός των σωστών απαντήσεων συγκρίνεται με τον αριθμό των επαναλήψεων που ορίσαμε για να δοθεί το αποτέλεσμα όπως στο παράδειγμα 20 σωστές απαντήσεις στις 30 προσπάθειες, τα οποία γράφουμε εν συντομία ως ένα κλάσμα ή ποσοστό: $20/30 = 67\%$. Το σκορ στη συνέχεια συγκρίνεται με τον πίνακα πιθανοτήτων (στο συγκεκριμένο πείραμα τον binomial πίνακα) από το οποίο προσδιορίζεται η πιθανότητα το αποτέλεσμα να βασίζεται σε τυχαίες απαντήσεις. Έτσι, το αποτέλεσμα δηλώνεται: $20/30 = 67\%$, ($p = 0,049$). Αυτό σημαίνει κυριολεκτικά ότι αποτέλεσμα (20/30) δεν είναι τυχαίο εκτός από το ποσοστό 0,049 (ή 4,9%) πιθανότητας να ήταν τυχαίο. Έτσι το αποτέλεσμα του πειράματος είναι ότι η μηδενική υπόθεση δεν ισχύει εκτός από ποσοστό 0,049 πιθανότητας να είναι αλήθεια, ή λέγοντας το διαφορετικά, ο ήχος A είναι διαφορετικός από B εκτός από μία πιθανότητα 4,9% να είναι ίδιος.

Ακόμα σημειώνουμε ότι, ότι σκορ και να επιτευχθεί, δεν μπορούμε να αποδείξουμε ότι $A=B$. Αυτό είναι ο ρόλος της abx σύγκρισης δεν μπορεί να χρησιμοποιηθεί για να αποδείξει δύο ήχοι ακούγονται το ίδιο. Η αντίληψη ότι η abx σύγκριση μπορεί να αποδείξει ότι δύο στοιχεία ακούγονται ίδια είναι μια κοινή παρανόηση σχετικά με την abx λογική.

Ένα abx test δεν μας λέει ποιος ήχος από τους δύο ακούγεται καλύτερα αλλά επιβεβαιώνει ότι η διάφορα των δύο ήχων ακούγεται ότι είναι αντιληπτή δηλαδή από τον ακροατή.

Τέλος ένα τυχαίο παράδειγμα αν το αποτέλεσμα ενός τεστ ήταν, $19/30 = 63\%$ ($p = 0,10$) το μόνο που θα μπορούσε να αναφερθεί είναι ότι ο ακροατής δεν κατάφερε να διαψεύσει την μηδενική υπόθεση (null hypothesis). Κανένα περαιτέρω συμπέρασμα δεν μπορεί να εξαχθεί σχετικά με την ομοιότητα ή τη διαφορά του ήχου A με τον ήχο B από αυτό το πείραμα. Ωστόσο ένα τόσο κοντινό αποτέλεσμα μπορεί να παρακινήσει τον δοκιμαστέο και τον ακροατή να δοκιμάσουν ξανά το πείραμα.

Ανακεφαλαιώνοντας για την συγκεκριμένη πειραματική διαδικασία θα χρησιμοποιήσουμε

- την *Binomial* κατανομή και τον αντίστοιχο Binomial πίνακα
- p-value 0.05
- Ο αριθμός των επαναλήψεων θα είναι 12 για κάθε ζευγάρι έτσι με βάση τον Binomial πίνακα και την p-value θα αξιολογήσουμε αν η Null hypothesis απορρίφθηκε. Αν η Null hypothesis κατατριφθεί για το συγκεκριμένο πείραμα σημαίνει ότι όντως ο συγκεκριμένος ακροατής κατάλαβε διαφορές ανάμεσα στο κάθε ζευγάρι δειγμάτων των DAW.
- Η μηδενική υπόθεση που θα χρησιμοποιηθεί για το συγκεκριμένο πείραμα είναι “οι μηχανές ήχου (*audio engines*) των συγκεκριμένων DAW που μελετάμε δεν παρουσιάζουν διαφορές στο ηχητικό αποτέλεσμα έτσι τα δείγματα που συγκρίνουμε δεν έχουν διαφορές”

4.3.4 Ανάλυση του binomial πίνακα

Probability of Experimental Result Being the Same as Random Guesses

	NUMBER OF TRIALS WITH CORRECT RESPONSES																			
TRIALS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.500																			
2	0.750	0.250																		
3	0.875	0.500	0.125																	
4	0.938	0.688	0.312	0.063																
5	0.969	0.813	0.500	0.188	0.031															
6	0.984	0.891	0.656	0.344	0.109	0.016														
7	0.992	0.938	0.773	0.500	0.227	0.062	0.008													
8	0.996	0.965	0.855	0.637	0.363	0.145	0.035	0.004												
9	0.998	0.980	0.910	0.746	0.500	0.254	0.090	0.020	0.002											
10	0.999	0.989	0.945	0.828	0.623	0.377	0.172	0.055	0.011	0.001										
11	1.000	0.994	0.967	0.887	0.726	0.500	0.274	0.113	0.033	0.006	0.000									
12	1.000	0.997	0.981	0.927	0.806	0.613	0.387	0.194	0.073	0.019	0.003	0.000								
13	1.000	0.998	0.989	0.954	0.867	0.709	0.500	0.291	0.133	0.046	0.011	0.002	0.000							
14	1.000	0.999	0.994	0.971	0.910	0.788	0.605	0.395	0.212	0.090	0.029	0.006	0.001	0.000						
15	1.000	1.000	0.996	0.982	0.941	0.849	0.696	0.500	0.304	0.151	0.059	0.018	0.004	0.000	0.000					
16	1.000	1.000	0.998	0.989	0.962	0.895	0.773	0.598	0.402	0.227	0.105	0.038	0.011	0.002	0.000	0.000				
17	1.000	1.000	0.999	0.994	0.975	0.928	0.834	0.685	0.500	0.315	0.166	0.072	0.025	0.006	0.001	0.000	0.000			
18	1.000	1.000	0.999	0.996	0.985	0.952	0.881	0.760	0.593	0.407	0.240	0.119	0.048	0.015	0.004	0.001	0.000	0.000		
19	1.000	1.000	1.000	0.998	0.990	0.968	0.916	0.820	0.676	0.500	0.324	0.180	0.084	0.032	0.010	0.002	0.000	0.000	0.000	
20	1.000	1.000	1.000	0.999	0.994	0.979	0.942	0.868	0.748	0.588	0.412	0.252	0.132	0.058	0.021	0.006	0.001	0.000	0.000	0.000

Πίνακας 4.1 Η Πιθανότητα των πειραματικών αποτελεσμάτων να είναι τα ίδια με τυχαίες επιλογές (με χρήση p-value 5%)²⁰

Number of trials	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Minimum number correct	9	9	10	10	11	12	12	13	13	14	15	15	16	16	17	18

Πίνακας 4.2 Τα αποτελέσματα που απαιτούνται για 95% confidence level.²¹

^{20, 21} Οι πίνακες 4.1 και 4.2 προέρχονται από rappers της ABX company (abx home) και αναλύουν τις στατιστικές παραμέτρους ενός abx ABX Double Blind Subjective Listening Tests.

Οι παραπάνω δύο πίνακες μας δείχνουν πρακτικά πόσες σωστές απαντήσεις ή ποιο ποσοστό είναι αποδεκτό και μπορούμε να μιλήσουμε για **statistically significant result** (στατιστικά σημαντικό αποτέλεσμα). Αναλυτικότερα στον πίνακα 4.1 με τα ποσοστά βλέπουμε ότι για 1/1 οι πιθανότητες να βρέθηκε στην τύχη το αποτέλεσμα είναι 50% πολύ μεγάλο ποσοστό. Συνεχίζοντας στις 5 επαναλήψεις βλέπουμε ότι ακόμα και να απαντήσουμε σωστά 4/5 το ποσοστό να βρέθηκε στην τύχη το αποτέλεσμα είναι μεγάλο 18.8%. Για μεγαλύτερες επαναλήψεις ο ακροατής αρχίζει να είναι πιο σίγουρος δηλαδή δεν μαντεύει απλά τις απαντήσεις έτσι πχ για 10 επαναλήψεις βλέπουμε ότι για 9/10 το ποσοστό να βρέθηκε στην τύχη το αποτέλεσμα είναι πολύ μικρό μόλις 1.1% που αυτό μας δίνει ένα confidence level της τάξης του 98.9% ότι ο ακροατής δεν μάντεψε στην τύχη.

Για ένα αριθμό συνολικών επαναλήψεων N , σημαντικά αποτελέσματα (significant results - confidence level 95%) δηλαδή ο αριθμός των σωστών απαντήσεων που πρέπει να δοθεί για να απορρίψουμε την μηδενική υπόθεση, μπορεί να υποστηριχθεί από τα αποτελέσματα του τύπου (4.1):

$$k = N/2 + N^{1/2} \quad (4.1)$$

Όπου k ο αριθμός των σωστών απαντήσεων που απαιτούνται.

N ο αριθμός των επαναλήψεων.

Από όλα τα παραπάνω βγάζουμε δύο πολύ σημαντικά συμπεράσματα:

- για αποτελέσματα μεγαλύτερα ή ίσα του **95% confidence level** ή αλλιώς p -value 5% πιθανότητα ο ακροατής να μάντεψε στην τύχη το αποτέλεσμα, μπορούμε να μιλάμε για **statistically significant result** (στατιστικά σημαντικό αποτέλεσμα). Με άλλα λόγια μπορούμε να πούμε ότι ο ακροατής μάντεψε σωστά εξ' αιτίας των ικανοτήτων του και όχι εξ' αιτίας τύχης.
- Τέλος βλέπουμε ότι για να είναι σωστό και επιστημονικά αποδεκτό ένα ABX test θα πρέπει οι επαναλήψεις να είναι από 10 και πάνω για να μπορούμε να μιλάμε για έγκυρα αποτελέσματα.

4.4 Επιλογή αντικειμένων

Στην παρούσα εργασία, το αντικείμενο έρευνας είναι η σύγκριση Audio Engines Digital Audio Workstations (D.A.W) με χρήση listening tests και θα είναι δύο βασικών κατηγοριών. Η πρώτη κατηγορία θα είναι signal tests sounds και συγκεκριμένα θα χρησιμοποιήσουμε ένα ημίτονο στα 440Hz αλλά και κενά exports των DAW που εξετάζουμε. Η δεύτερη κατηγορία θα είναι με χρήση μουσικών αρχείων.

Το πειραματικό μέρος θα χωριστεί σε δύο κατηγορίες. Το αντικειμενικό μέρος όπου με την βοήθεια προγραμμάτων θα δούμε τις αντικειμενικές διαφορές των μουσικών δειγμάτων αλλά και το υποκειμενικό μέρος όπου θα εκτελεστούν listening tests από κοινό. Χρήσιμα συμπεράσματα θα εξαχθούν από τα αποτελέσματα των listening tests. Η επιλογή των αρχείων ήχου έγινε με τρόπο που να διασφαλίζει τη σχέση τους με το αντικείμενο που εξετάζουμε αλλά και την οικειότητα των υποκειμένων με αυτά.

4.5 Επιλογή κριτηρίων

Ο σχεδιασμός του πειράματος είναι ο πιο κρίσιμος παράγοντας όταν ετοιμαζόμαστε να εκτελέσουμε μια δοκιμή ακρόασης. Μπορεί να είναι αρκετά απογοητευτικό να ξοδέψουμε το χρόνο και την προσπάθεια της διεξαγωγής ενός listening test, μόνο για να καταλάβουμε ότι δεν μπορούμε να βγάλουμε κάποιο λογικό συμπέρασμα. Σε μια τέτοια περίπτωση, θα πρέπει να διορθώσουμε το πρόβλημα και να ξεκινήσουμε από την αρχή. Θα πρέπει πρώτα να αναφερθεί ότι υπάρχουν ποσοτικές μέθοδοι για τη μέτρηση των διαφορών μεταξύ των πηγών ήχου. Εάν ένας αντικειμενικός υπολογισμός θα απαντήσει στο ερώτημα που έχει τεθεί, αυτό είναι συχνά λιγότερο δαπανηρό ή και χρονοβόρο από ότι μια δοκιμή ακρόασης. Ωστόσο, τα listening tests συχνά χρησιμοποιούνται για να μετρήσουμε τις αντιληπτές διαφορές των γνωστών αντικειμενικών διαφορών. Σημειώστε, επίσης, ότι δεν είναι απαραίτητο να εκτελέσουμε δοκιμές ακρόασης σε υλικό που δεν έχει αντικειμενικές διαφορές.

Το πρώτο βήμα σε οποιαδήποτε δοκιμή ακρόασης είναι σαφώς να καθορίσουμε το ερώτημα που προσπαθούμε να απαντήσουμε και μια σχετική υπόθεση που πρόκειται να ελεγχθεί. Είναι πολύ σημαντικό να είναι συγκεκριμένο όσο το δυνατόν, και να περιέχει όλα τα απαραίτητα στοιχεία. Συχνά, η ερώτηση θα έχει πολλαπλές συνιστώσες και θα πρέπει να χωριστεί σε μικρότερες ερωτήσεις που μπορεί η κάθε μία να απαντηθεί από μία συγκεκριμένη μεθοδολογία δοκιμών.

Για παράδειγμα, το ABX test μπορεί μόνο να μας πει εάν υπάρχει μια αισθητή διαφορά ανάμεσα στους δύο ήχους. Δεν μπορεί να μας πει ότι δεν υπάρχει διαφορά ανάμεσα στους δύο ήχους. Το επόμενο βήμα και ίσως ένα από τα πιο συχνά που παραβλέπεται, είναι να εντοπίσουμε εξωγενείς παράγοντες ή μεταβλητές που θα μπορούσαν να οδηγήσουν σε ένα εσφαλμένο συμπέρασμα. Η δοκιμή ακρόασης θα πρέπει να λάβει χώρα σε ένα κατάλληλο χώρο για τους συγκεκριμένους ήχους που μας ενδιαφέρει. Τα ηχεία θα πρέπει επίσης να είναι κατάλληλα για τη δοκιμή μας (στην συγκεκριμένη πειραματική διαδικασία θα χρησιμοποιηθούν ακουστικά). Επίσης πρέπει να αποφασιστεί και η κατάλληλη ένταση στην οποία θα παρουσιαστούν τα δείγματα. Η ITU έχει τυποποιήσει πολλές πτυχές για το κατάλληλο περιβάλλον ακρόασης, καθώς και τα πρότυπα όπως η ITU-R BS.1116 που είναι χρήσιμος ως οδηγός. Αν έχουμε πρόβλημα στον έλεγχο μιας μεταβλητής που δεν μας ενδιαφέρει να μελετήσουμε, καλή πρακτική είναι να την εισάγουμε να επαναλαμβάνεται τυχαία όσο το δυνατόν περισσότερο. (Προγράμματα υπολογιστών είναι αρκετά χρήσιμα εδώ, γιατί αυτό που αντιλαμβανόμαστε ως τυχαίο συχνά δεν είναι.) Είναι σημαντικό να κρατήσουμε όλες τις περιβαλλοντικές μεταβλητές

ίδιες για όλους τους ακροατές. Αν κάτι έχει τη δυνατότητα να αλλάξει (π.χ., διαμόρφωση δωματίου, ένταση, κλπ), μπορεί να είναι πολύ δύσκολο στο να ερμηνεύσουμε τα αποτελέσματα στην συνέχεια. Στην πραγματικότητα, οποιαδήποτε αντιληπτή διαφορά μπορεί να έχει προκληθεί από μια από αυτές τις μεταβλητές που δεν φροντίσαμε να ελέγξουμε.

Το ακουστικό υλικό (μουσική, ομιλία, κλπ) θα πρέπει να είναι αντιπροσωπευτικό του τυπικού υλικού για το σύστημα που αξιολογείται, αλλά θα πρέπει επίσης να είναι ουσιαστικό υλικό. Με άλλα λόγια, θα πρέπει να ωθήσει το σύστημα στα όριά του και να αναδείξει τυχόν δυνητικά ευδιάκριτες διαφορές. Εάν ένα φαινόμενο υψηλής συχνότητας δοκιμάζεται, ποιότητα ομιλίας τηλεφώνου μπορεί να μην είναι αποδεκτό ερέθισμα. Επίσης, πρέπει να σκεφτούμε πόσο και τι είδους ακουστικό υλικό είναι απαραίτητο για να αποδείξουν την υπόθεσή μας. Πολλά πειράματα επιλέγουν ποικίλες μορφές της μουσικής ως ερέθισμα, αλλά κάθε πείραμα θα πρέπει να έχει διαφορετικά ερεθίσματα με βάση το πειραματικό θέμα και ό, τι είναι αναγκαίο για να αποδείξουμε την υπόθεσή σας. Για πολλές μελέτες θα χρειαστεί οι ακροατές να είναι ειδικοί και πεπειραμένοι. Ειδικοί ακροατές είναι εκείνοι που έχουν επιλεγεί ειδικά για να τις ικανότητες ακοής τους- είναι συνήθως αρκετά ευαίσθητοι στις ιδιότητες του ήχου που μας ενδιαφέρει και μελετάμε. Το πείραμα θα πρέπει να είναι double blind αν είναι δυνατόν- ο χειριστής της δοκιμής δεν θα πρέπει να γνωρίζει τις απαντήσεις. Αυτό θα εξασφαλίσει ότι ο διαχειριστής δεν δίνει ανεπαίσθητες λάθος ενδείξεις στον ακροατή. Η σειρά παρουσίασης του ακουστικού υλικού θα πρέπει να είναι τυχαία. Είναι αρκετά πιθανό η σειρά παρουσίασης να επηρεάσει τις αποφάσεις του ακροατή, και κάθε άλλη μεταβλητή. Είναι γενικά χρήσιμο να ζητήσουμε από τους ακροατές να γράψουν τις εντυπώσεις και τις απόψεις τους μετά το πείραμα. Συχνά, αυτά τα σχόλια θα αποκαλύψει πτυχές που ο υπεύθυνος για το πείραμα δεν θα έπαιρνε ποτέ υπόψη. Ουσιαστικά αυτό θα μας βοηθήσει να προχωρήσουμε. Είναι καλή πρακτική να τρέχει πάντα μια μικρή πιλοτική δοκιμή πριν από την έναρξη του πραγματικού πειράματος. Αυτό θα επιτρέψει σε μας να συλλέξουμε μερικά σχόλια, να ελέγξουμε την εγκατάσταση και το σύστημα, την επικύρωση του πειραματικού σχεδιασμού, και να βεβαιωθούμε ότι η στατιστική ανάλυση των αποτελεσμάτων μπορεί να οδηγήσει σε χρήσιμα συμπεράσματα.

Οι ακροατές θα πρέπει να λαμβάνουν όλοι τις ίδιες οδηγίες (και πάλι για να αποφευχθεί η απροσδόκητη μεταβλητότητα). Ένα γραπτό φύλλο οδηγιών είναι μια καλή λύση, το ίδιο ισχύει και για αρχεία ήχου/βίντεο. Ο πειραματικός σχεδιασμός των εκτιμήσεων σε ένα off-line πείραμα (πχ ακούγοντας προηχογραφημένα αρχεία ήχου) είναι συχνά διαφορετικός σε σχέση με εκείνα που γίνονται σε πραγματικό χρόνο ή σε on-line πειράματα. Στην συνέχεια πρέπει να κατανοήσουμε τι επιπτώσεις έχει ο εξοπλισμός που χρησιμοποιούμε πάνω στα υπό εξέταση αρχεία ήχου, εξοπλισμός όπως τα καλώδια, gain stages (στάδια ενίσχυσης του σήματος, ηχεία, κλπ.). Για παράδειγμα, αν ένα πείραμα επιχειρεί να μελετήσει την ακουστή διαφορά στο δυναμικό εύρος δύο ψηφιακών format όπως 16bit και 24bit, πρέπει κανείς να εξετάσει τις επιδράσεις των συσκευών που χρησιμοποιούνται για τη διεξαγωγή της δοκιμής. Πρέπει να σκεφτούμε την διαδρομή που θα ακολουθήσει το σήμα, όπως αναλογικό κύκλωμα, δρομολόγηση, θωράκιση, την ποιότητα μετατροπέα, και ούτω καθεξής. Πειράματα σε πραγματικό χρόνο απαιτούν εξελιγμένους μηχανισμούς εναλλαγής από το ένα αντικείμενο στο

άλλο (switching), με σκοπό να αποκλειστεί η πιθανότητα ένα άτομο να ακούσει την εναλλαγή που θα προδώσει την απάντηση σε ένα τεστ, όπως είναι το ABX.

Πειράματα σε πραγματικό χρόνο μπορεί να είναι δύσκολο να σχεδιαστούν ως double blinds μελέτες λόγω της αυξανόμενης ανάγκης για ειδικά μηχανήματα που θα βοηθήσουν να δημιουργηθεί μια τέτοια δοκιμασία, ωστόσο ένα double blind test είναι σημαντικό και θα πρέπει να εξετάζεται με προσοχή. Τα πειράματα σε πραγματικό χρόνο και τα on-line πειράματα είναι ιδιαίτερα ευαίσθητα σε amplitude και delay matching. Μικρές παραλλαγές είτε σε πλάτος ή σε φάση μπορεί να μειώσουν και πάλι την εγκυρότητα του πειράματος. Προσπάθειες για να διορθώσουμε τα χαρακτηριστικά του ήχου συχνά απαιτούν πρόσθετο εξοπλισμό στην αλυσίδα του σήματος (signal chain). Σε αυτή την περίπτωση είναι σημαντικό να έχουμε διπλό εξοπλισμό σε όλες τις διαδρομές του σήματος. Δεν είναι αποδεκτό να τοποθετήσετε ένα στάδιο κέρδους (gain stage) στο ένα μονοπάτι και όχι στο άλλο για να ταιριάζουν τα gain μεταξύ τους. Στην πραγματικότητα, είναι σωστό να χωρίσει την διαφορά κέρδους (gain difference) και να εφαρμόσουμε κάποιο κέρδος ή εξασθένηση και στις δύο διαδρομές και όχι αλλάζοντας μόνο ένα μονοπάτι. Άλλες εκτιμήσεις είναι διαισθητικά προφανείς, αλλά μερικές φορές παραβλέπονται και αξίζει να αναφερθούν. Για παράδειγμα, είναι σημαντικό να εξεταστεί ο πειραματικός σκοπός, όταν επιλέγουμε ακροατήριο. Εάν η δοκιμή ακρόασης συνεπάγεται απόκριση υψηλής συχνότητας, τα άτομα με περιορισμένη ακοή υψηλής συχνότητας μπορεί να μην είναι κατάλληλά. Ο σωστός πειραματικός σχεδιασμός είναι υψίστης σημασίας και είναι το θεμέλιο για την πειστική στατιστική ανάλυση.

4.6 Ενστάσεις και αντιπαραθέσεις που δημιουργούνται κατά την πραγματοποίηση των blind test.

Υπάρχουν διάφορες αντιρρήσεις όσον αναφορά τα blind test. Αλλά σχεδόν κάθε μία έχει λυθεί και δικαιολογηθεί. Ορισμένες από αυτές είναι:

- **Το άγχος της ακρόασης:** Ορισμένα blind tests πραγματοποιούνται από τρίτους που διαχειρίζονται την δοκιμή με αποτέλεσμα οι ακροατές να αισθάνονται κάποια πίεση στην ακρόαση. Αλλά με τα σύγχρονα λογισμικά έχουν δημιουργηθεί ABX προγράμματα (foobar2000-abx comparator) και μπορεί ο κάθε ακροατής να κάνει την ακρόαση μόνος του ακόμα και στο σπίτι του υπό τις συνθήκες που τους βολεύουν.
- **Η διάρκεια της ακρόασης:** Λέγεται συχνά ότι δεν μπορείς να εκτιμήσεις τις διαφορές μεταξύ των δειγμάτων ακούγοντας μόνο ένα κομμάτι ή μόνο για ένα μικρό χρονικό διάστημα πριν από την αλλαγή στο επόμενο. Με ένα ABX πρόγραμμα μπορούμε να το χρησιμοποιήσουμε ακόμα και στο σπίτι για όσο

χρονικό διάστημα επιθυμούμε εμείς. Μπορούμε να ακούσουμε το παράδειγμα Α για μια εβδομάδα, και το παράδειγμα Β για άλλη μια εβδομάδα, και να δώσουμε την απάντηση μας για τον Χ με σιγουριά ότι επιλέγουμε το σωστό. Αυτό θα μπορούσε να συνεχιστεί για όσο χρονικό διάστημα θέλουμε.

- **Το ηχητικό σύστημα που χρησιμοποιήθηκε στην ακρόαση δεν ήταν αρκετά υψηλής ανάλυσης:** Blind tests έχουν διεξαχθεί και με πιο φθηνά συστήματα ακρόασης και τα αποτελέσματα ήταν παρόμοια με τα ακριβά συστήματα ακρόασης.
- **Μη έγκυρα στατιστικά:** Υπήρξαν κάποιες δοκιμές με οριακά αποτελέσματα που αφήνουν περιθώρια αμφισβήτησης ως προς την στατιστική τους σημασία. Αλλά υπήρξαν πολλά περισσότερα που είχαν έγκυρο στατιστικό αποτέλεσμα σε πολλές μελέτες. Τα μαθηματικά και η στατιστική που χρησιμοποιούνται είναι πολύ σαφή και μια καλή απόδειξη ερμηνείας των αποτελεσμάτων.
- **Ικανότητες ακρόασης:** Ορισμένοι ισχυρίζονται ότι τα blind tests αποτυγχάνουν επειδή οι ακροατές που εμπλέκονται σε αυτά δεν είναι επαρκώς καταρτισμένοι. Αλλά είναι επίσης γνωστό σε αυτού του είδους τα τεστ ότι δημιουργείται ακούσια, φιλτράρισμα και προκατάληψη στην ακοή μας και όσο ικανός και να είναι ένας ακροατής σε μεγάλο ποσοστό αυτό δεν μπορεί να το αποβάλει την ώρα της ακρόασης. Αυτό το φαινόμενο μπερδεύει ακόμα και μηχανικούς ήχου (sound-recording engineers).

Κεφάλαιο 5^ο: ΣΧΕΔΙΑΣΜΟΣ ΤΗΣ ΠΕΙΡΑΜΑΤΙΚΗΣ ΔΙΑΔΙΚΑΣΙΑΣ – ΜΕΘΟΔΟΛΟΓΙΑ

Εισαγωγή 5^{ου} Κεφαλαίου

Στο κεφάλαιο αυτό, αναλύεται ο τρόπος με τον οποίο σχεδιάστηκε και διεξήχθη η πειραματική διαδικασία. Στην συνέχεια αναφέρονται τα ερωτήματα στα οποία εστιάζεται το ενδιαφέρον της παρούσας εργασίας. Σκοπός αυτού του κεφαλαίου είναι η κατανόηση της λογικής με την οποία σχεδιάστηκε η πειραματική διαδικασία καθώς και τα ερωτήματα τα οποία καλείται να απαντήσει.

Για τη διεξαγωγή της πειραματικής διαδικασίας κατασκευάστηκαν δείγματα για τα listening tests σύμφωνα με την μέθοδο ABX double blind test η οποία αναλύθηκε στο προηγούμενο κεφάλαιο.

Τα tests αυτά ήταν με τη μορφή ακουστικών δειγμάτων (samples) που δημιουργήθηκαν μέσα από τα DAW (πιο συγκεκριμένα Pro tools, Logic και Cubase) που θέλουμε να συγκρίνουμε.

5.1 Ερευνητικά ερωτήματα

Τα ερωτήματα στα οποία καλείται να απαντήσει η συγκεκριμένη πτυχιακή αφορούν τελικά στο αν οι μηχανές ήχου (audio engines) των DAW (cubase, logic, pro tools) κατά το export (άθροιση των καναλιών σε ένα στέρεο αρχείο) παρουσιάζουν διαφορές στο ηχητικό αποτέλεσμα και εάν αυτό μπορεί να το αντιληφθεί ο άνθρωπος. Πολλοί υποστηρίζουν ότι έχουν διαφορές και το ένα DAW είναι καλύτερο από το άλλο, αλλά αυτό πρέπει να ελεγχτεί διότι στις περισσότερες των περιπτώσεων είναι αποτέλεσμα προτίμησης κάποιου συγκεκριμένου προγράμματος μουσικής παραγωγής. Με την συγκεκριμένη πειραματική διαδικασία θα δούμε αν ισχύει αυτός ο μύθος ή όχι.

Έτσι αναλυτικά θα ελέγξουμε αν ο άνθρωπος (υποκείμενο) μπορεί:

- να καταλάβει διαφορές δηλαδή να κάνει την σωστή επιλογή στην πειραματική διαδικασία, μιας και με την συγκεκριμένη μέθοδο (abx double blind test) όπως είδαμε και στο κεφάλαιο 4 χρησιμοποιούνται μαθηματικά και στατιστική για την ερμηνεία των αποτελεσμάτων επειδή έχουμε να κάνουμε με πιθανότητες και παράγοντες τύχης.

5.2 Σχεδιασμός πειραματικής διαδικασίας

Για αυτή την πτυχιακή εργασία η συγκεκριμένη πειραματική διαδικασία θα χωριστεί σε δύο σκέλη για να συμβαδίζει με τα αντίστοιχα παγκόσμια πρότυπα (πειράματα και έρευνες που έχουν πραγματοποιηθεί παρόμοιους τομείς όπως μουσική τεχνολογία, ηλεκτρονικά αλλά και στον τομέα της υγείας και φυσικής).²²

Έτσι η πειραματική διαδικασία θα έχει:

- Το *αντικειμενικό κομμάτι* όπου θα αναλυθούν με την βοήθεια προγραμμάτων συγκεκριμένα μουσικά δείγματα όπως ημίτονα, και κενά exports. Ακόμα με την χρήση null sound tests θα συγκριθούν τα ζευγάρια των export των listening tests και θα εξεταστούν με την βοήθεια προγραμμάτων που παρέχουν edit και ανάλυση της κυματομορφής στο χρόνο. Έτσι αντικειμενικά το κάθε export από το κάθε DAW θα δείξει που διαφέρει συγκεκριμένα από το άλλο.
- Το *υποκειμενικό κομμάτι* όπου έχουμε δημιουργήσει μουσικά δείγματα και με την βοήθεια του προγράμματος foobar2000 και της λειτουργίας abx comparator που έχει θα μας βοηθήσει να πραγματοποιήσουμε την διαδικασία και να εξάγουμε χρήσιμα συμπεράσματα.

5.3 Επιλογή των δειγμάτων

Η δημιουργία των σωστών δειγμάτων είναι από τα σημαντικότερα και πιο κρίσιμα κομμάτια της πειραματικής διαδικασίας διότι με τα σωστά δείγματα τα listening tests θα λειτουργήσουν σωστά οι ακροατές θα ακούν τα σωστά παραδείγματα για να δώσουν τις απαντήσεις τους και έτσι τα αποτελέσματα θα είναι επιστημονικά αποδεκτά.

5.3.1 Αντικειμενικό κομμάτι και δείγματα (Null sound test)

Όσο αναφορά στον ήχο τα null sound tests είναι χρήσιμα για να προσδιορίσουμε εάν δύο αρχεία ήχου είναι πανομοιότυπα. Αν είναι διαφορετικά τότε θα ακούμε μόνο την διαφορά τους. Για όσους δεν είναι εξοικειωμένοι με τον όρο αυτό, η συγκεκριμένη δοκιμή είναι για να καθορίσει εάν δύο ψηφιακές εγγραφές (αρχεία) είναι ταυτόσημες ή όχι. Για να έχει νόημα αυτή η σύγκριση, οποιαδήποτε διαφορά έχουν τα αρχεία πρέπει να αντιμετωπιστεί έτσι ώστε όταν αυτά αθροιστούν εκτός φάσης (out of phase) να αλληλοαναιρούνται. Αυτό απαιτεί τα αρχεία να περιέχουν ακριβώς το ίδιο περιεχόμενο ήχου, να έχουν την ίδια συχνότητα δειγματοληψίας και τέλος η έντασή τους να είναι ίδια.

²² Leonard, Brett; Buttner-Schnirer, Pdraig (2012). Subjective Differences in Digital Audio Workstation Math. McGill University. 3-4

Συνήθως, τα αρχεία είναι τοποθετημένα σε δύο ξεχωριστά κανάλια και τα δύο έχουν τον ίδιο χρόνο αναπαραγωγής είναι δηλαδή συγχρονισμένα. Η πολικότητα (polarity) ενός από των δύο καναλιών αναστρέφεται έτσι ώστε το ακουστικό περιεχόμενο από τα δύο κομμάτια να είναι «εκτός φάσεως» με το άλλο. Όπως είπαμε η ένταση και των δυο αρχείων θα πρέπει να είναι ίδια αλλιώς το ένα κανάλι με το άλλο μπορεί να μην ακυρωθούν και η διαδικασία μας να μην γίνει σωστά.

Η χρησιμότητα ενός null test είναι εξαιρετικά περιορισμένη στον προσδιορισμό της διαφοράς από το "ιδανικό", όπως για παράδειγμα όταν ένα αρχείο είναι περισσότερο bright από ένα άλλο, η διαφορά θα είναι πολύ πιο bright. Δεν είναι δυνατόν να καθορίσουμε ποια από τα δύο αρχεία είναι πιο bright απλά ακούγοντας τη διαφορά διότι το ηχητικό αποτέλεσμα είναι ένας συνδυασμός και των δύο. Αυτό ισχύει για όλες τις διαφορές, συμπεριλαμβανομένης και της παραμόρφωσης.

Στην συγκεκριμένη πειραματική διαδικασία θα χρησιμοποιήσουμε το null sound test στο αντικειμενικό μέρος του πειράματος για να δούμε αντικειμενικά και με την βοήθεια του προγράμματος *audacity* αν τελικά προκύπτουν διαφορές ανάμεσα στα export των τριών DAW. Θα το χρησιμοποιήσουμε και για export χωρίς plug-ins αλλά και για τα export με plug-ins. Όπως αναφέραμε και παραπάνω στο αντικειμενικό σκέλος του πειράματος θα χρησιμοποιηθούν τα ζευγάρια των exports των listening tests (τα ζευγάρια είναι 6, τα αναφέρουμε στην παράγραφο 5.4) αλλά και δείγματα-αναφορές. Αρχικά με την χρήση null test θα συγκρίνουμε τα ζευγάρια των export πχ. θα τοποθετήσουμε στο *audacity* το export από το *Cubase* μαζί με το export από το *Logic*. Στην συνέχεια θα αναστρέψουμε σε ένα από τα δύο σήματα την φάση και στην συνέχεια θα τα αθροίσουμε. Αυτό που θα προκύψει θα είναι η διαφορά που έχουν αυτά τα δύο σήματα. Ακόμα θα χρησιμοποιήσουμε ένα ημίτονο που το έχουμε κάνει export και στα τρία DAW που μας ενδιαφέρουν (*Logic*, *Pro tools*, *Cubase*) και θα πραγματοποιήσουμε την ίδια παραπάνω διαδικασία. Τέλος θα δημιουργήσουμε κενά export δηλαδή χωρίς να υπάρχουν κανάλια με σήματα στα DAW και με την βοήθεια του *wavelab* θα πραγματοποιήσουμε *global analysis* για να δούμε τι πληροφορίες έχει το κάθε δείγμα και εάν έχουν διαφορές.

5.3.2 Υποκειμενικό κομμάτι και δείγματα (listening tests)

Σε αυτό το σκέλος του πειράματος θα χρησιμοποιηθεί μία πολυκάναλη παραγωγή ενός αρκετά δημοφιλούς συγκροτήματος τους *Radiohead* και το κομμάτι που θα χρησιμοποιήσουμε ονομάζεται *Nude*.

Τα συγκεκριμένα tracks είναι:

- σε .wav μορφή για να έχουμε ικανοποιητική ποιότητα αναπαραγωγής.
- είναι ηχογραφημένα στα 44100Hz, 24bit.
- αποτελούνται από ντραμς, μπάσο, πλήκτρα, κιθάρες, φωνή δηλαδή ηλεκτρικά όργανα αλλά και ακουστικά κάτι που ακούμε σε χιλιάδες κομμάτια, έτσι το

δείγμα θα είναι αντιπροσωπευτικό και θα πληροί τα στάνταρ των συγκεκριμένων listening test.

- Επίσης το είδος του κομματιού συνδυάζει το χαλαρό παίξιμο (jazzy feel) με το αρκετά έντονο και δυναμικό παίξιμο και ερμηνεία (rock feel) κάτι που μας δίνει την δυνατότητα να έχουμε μεγάλο εύρος δυναμικών σε ένα κομμάτι. Έτσι θα δούμε αν υπάρχει διαφορά στην άθροιση των audio engines αν μπορούμε να ακούσουμε διαφορές δηλαδή ανάλογα με τις δυναμικές που έχει το κομμάτι.

Το πιο βασικό που πρέπει να τονίσουμε είναι τα export του συγκεκριμένου κομματιού που θα δημιουργηθούν θα είναι δυο βασικών κατηγοριών.

- *1^η κατηγορία (export χωρίς την προσθήκη plug-ins):* Δηλαδή τα πολυκάναλα θα γίνουν export χωρίς να έχουν επεξεργαστεί μέσα από plug-ins ή άλλα εφέ και οι στάθμες των faders-master fader θα βρίσκονται στα 0dB και τα pan pots όλα κέντρο. Να τονίσουμε ότι έχουμε επιλέξει το pan law να είναι κοινό και στα τρία DAW για να είναι η σύγκριση όσο το δυνατόν πιο αντικειμενική και σωστή. Έτσι η επιλογή του pan law βρίσκεται στα -3dB.
- *2^η κατηγορία (export με την προσθήκη 2-3 plug-ins σε κάθε κανάλι):* Αυτό γίνεται με σκοπό να ζορίσουμε τον αθροιστή της μηχανής ήχου κατά το export να δώσει όποιες διαφορές και χαρακτηριστικά έχει από την κατασκευή του. Με μια απλή άθροιση χωρίς plug-ins τα συγκεκριμένα προγράμματα λειτουργούν σαν να προσθέτουν τα σήματα, είναι πολύ δύσκολο δηλαδή να δείξουν διαφορές. Με την χρήση των plug-ins και την χρησιμοποίηση μεγαλύτερης υπολογιστικής ισχύς για την dsp επεξεργασία και πρόσθεση των σημάτων υπάρχει μεγάλη πιθανότητα να έχουν διαφορά στον τρόπο λειτουργίας το ένα DAW με το άλλο. Αυτός είναι και ο σκοπός που πραγματοποιούμε αυτά τα δύο κομμάτια ξεχωριστά. Να αναφέρουμε ότι τα plug-ins που χρησιμοποιήσαμε και στα 3 DAW είναι της ίδιας εταιρείας (Waves) και ότι έχουμε χρησιμοποιήσει τον ίδιο τύπο plug-in με το ίδιο preset και για τα 3 λογισμικά (Logic, Pro tools, Cubase).

5.4 Τα ζευγάρια των δειγμάτων που δημιουργούνται και η χρήση τους με το foobar2000.

Από την παραπάνω παράγραφο είδαμε ότι για τις ακροάσεις έχουμε δυο κατηγορίες δειγμάτων με plug-ins και χωρίς plug-ins. Όσο αναφορά την μεθοδολογία abx double blind test, όπως είπαμε ο ακροατής έχει πρόσβαση σε τρεις πηγές που ονομάζονται A, B, και X. A και B είναι οι αναφορές. Για παράδειγμα, A το αρχείο ήχου από το Logic και B το αρχείο ήχου από τα Pro Tools. X είναι η άγνωστη πηγή. Μπορεί να είναι ή A ή B. Ο ακροατής πρέπει να επιλέξει ποιά από τις δύο αναφορές (A ή B) αναπαράγεται κάθε φορά στην X επανάληψη.

Με βάση τα παραπάνω και για την σωστή διεξαγωγή της πειραματικής ακρόασης έχουμε δημιουργήσει:

- 6 export (δείγματα) της ίδιας πολυκάναλης παραγωγής (*Radiohead, Nude*) που προήλθαν αντίστοιχα από Cubase, Logic, Pro tools. Τα 3 χωρίς την προσθήκη plug-ins και τα άλλα τρία με την προσθήκη plug-ins όπως αναφέραμε και παραπάνω.
- Έτσι δημιουργήθηκαν 6 ζευγάρια δειγμάτων για να τα συγκρίνουμε μεταξύ τους. Αυτά είναι:
 - 1°. Cubase-Logic (χωρίς plug-ins)
 - 2°. Cubase-Pro Tools (χωρίς plug-ins)
 - 3°. Logic-Pro Tools (χωρίς plug-ins)

 - 4°. Cubase-Logic (με plug-ins)
 - 5°. Cubase-Pro Tools (με plug-ins)
 - 6°. Logic-Pro Tools (με plug-ins)

5.4.1 Η λειτουργία του λογισμικού foobar2000 (abx comparator)

Αρχικά να πούμε ότι ένας αποδεκτός αριθμός επαναλήψεων για το κάθε ζευγάρι είναι πάνω από 10, στις συγκεκριμένες ακρόασεις θα πραγματοποιήσουμε 12 επαναλήψεις.²³ Δηλαδή για κάθε ζευγάρι θα πρέπει να επιλέξουμε τον άγνωστο X 12 φορές για να είναι τα αποτελέσματα μας αξιόπιστα. Άρα έχουμε 6 ζευγάρια από 12 επαναλήψεις το καθένα.

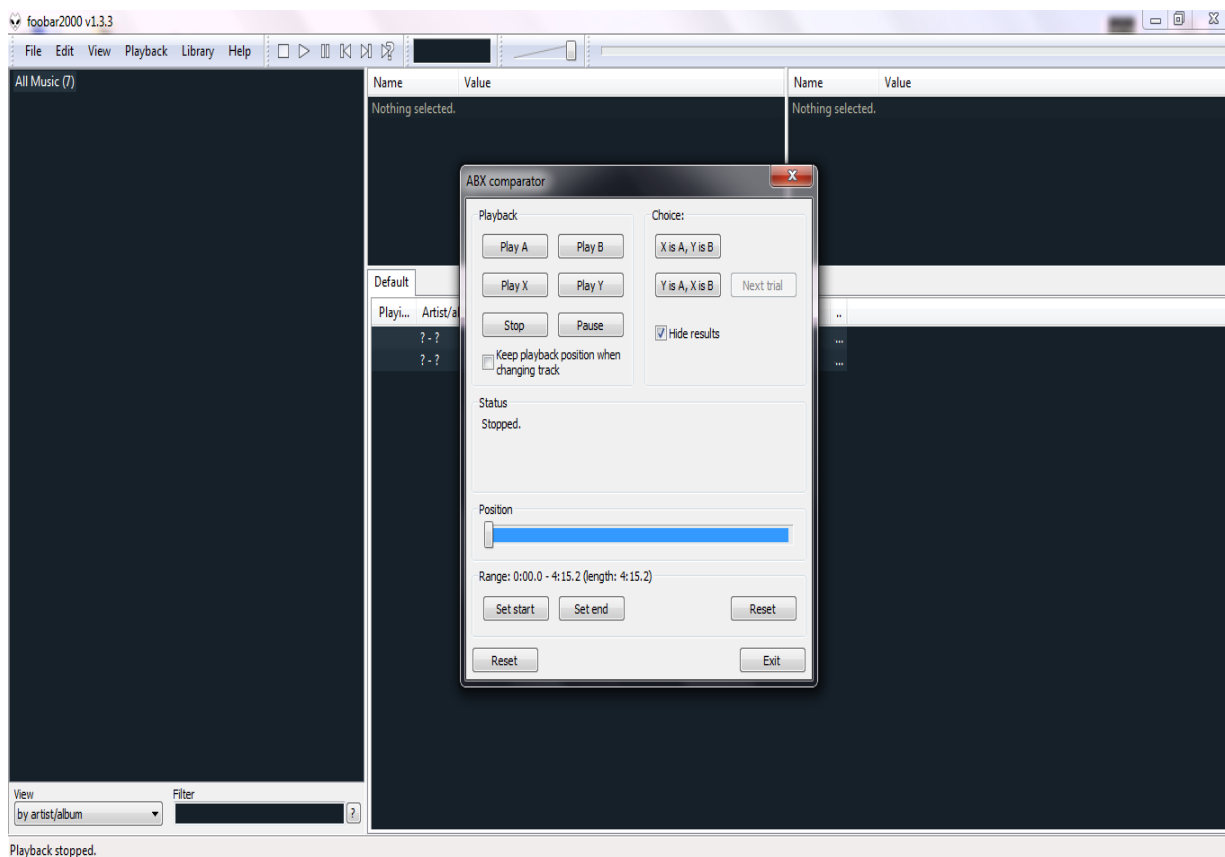
Για να πραγματοποιηθούν οι συγκεκριμένες ακρόασεις και για να βγάλουμε σωστά και έγκυρα στατιστικά αποτελέσματα θα χρησιμοποιήσουμε το λογισμικό foobar2000.

Η λειτουργία του προγράμματος είναι απλή, επιλέγουμε τα ζευγάρια μας, πχ. το 1° ζευγάρι A→Cubase και B→Logic και συνεχίζουμε στην ακρόαση. Τα A και B είναι αναφορές μας δεν αλλάζουν καθ' όλη την διάρκεια των επαναλήψεων. Αυτό που αλλάζει είναι από κάτω το X και το Y. Αυτά αλλάζουν θέσεις κατά την διάρκεια των επαναλήψεων και εμείς πρέπει να αντιστοιχίσουμε αν το A είναι το X ή το B είναι το Y ή το αντίθετο αν το B είναι το X και το A το Y. Δηλαδή το X και το Y σε κάθε επανάληψη παίρνουν τυχαία θέση που την καθορίζει το πρόγραμμα χωρίς να την γνωρίζει ούτε ο χειριστής του πειράματος ούτε ο ακροατής (double blind trial). Αυτές είναι και οι δυο απαντήσεις που έχουμε να δώσουμε και μία από τις δύο πρέπει να επιλέγουμε σε κάθε επανάληψη δηλαδή (**X is A, Y is B** ή **Y is A, X is B**). Έτσι

²³ Manuj Yadav, Densil Cabrera, Ralph Collins and William L. Martens (2011). Detection of headtracking in room acoustic simulations for one's own voice. University of Sydney, 3

επιλέγουμε μία από τις δυο απαντήσεις και στην συνέχεια επιλέγουμε next trial. Με αυτόν τον τρόπο έχει πραγματοποιηθεί μία επανάληψη. Να σημειώσουμε ότι τις αναφορές A και B μπορούμε να τις ακούμε καθ' όλη την διάρκεια των επαναλήψεων μαζί με τα X και Y για να έχουμε τις αναφορές και να συγκρίνουμε.

Μετά το τέλος του κάθε ζευγαριού δηλαδή μόλις τελειώσουμε και την 12 επανάληψη το foobar2000 μας παρέχει ένα .txt αρχείο με τον αριθμό των σωστών απαντήσεων που δώσαμε και το ποσοστό που εκφράζει την πιθανότητα να έχουμε δώσει στην τύχη τις απαντήσεις μας (p-value). Κάτω από ποσοστό 5% όπως αναλύσαμε στο κεφάλαιο 4 τα αποτελέσματα μας είναι στατιστικά αποδεκτά.



Εικόνα 5.1 Το foobar2000 player και η λειτουργία abx comparator. Η ανάλυση των παραμέτρων παραθέτονται στην παράγραφο 5.4.1

5.5 Επιλογή Υποκειμένων

Η επιλογή των υποκειμένων είναι μια διαδικασία ουσιώδης και πολύ σημαντική σε έρευνες τέτοιας φύσεως. Η διαδικασία αυτή γίνεται με βάση τρεις βασικούς παράγοντες επιλογής: τον αριθμό των υποκειμένων, την ηλικία και τη σχέση τους με το αντικείμενο.

Ο αριθμός των υποκειμένων που απαιτείται να λαμβάνουν μέρος σε έρευνες με τη χρήση της συγκεκριμένης μεθόδου είναι από 20 και άνω, σύμφωνα με την ITU-R (Recommendation

ITU-R BS. 111-2) αλλά και σχετικά papers που βρέθηκαν.²⁴ ²⁵ Έτσι, λοιπόν στην παρούσα έρευνα συμμετέχουν 20 άτομα με εύρος ηλικιών από 19-29 έτη.

Οι ηλικίες των υποκειμένων, όπως αναφερθήκαμε και παραπάνω είναι από 19 έως 29 ετών. Η επιλογή αυτή έγινε με βάση διάφορες μελέτες οι οποίες σχετίζονται με την απώλεια της ακοής λόγω ηλικίας και συγκεκριμένα συχνοτικών απωλειών όπως αυτή του Takeda.²⁶ Έτσι προτιμήθηκε η ηλικιακή ομάδα κάτω των 30 ετών για να μην υπάρχει κίνδυνος αλλοίωσης των αποτελεσμάτων λόγω προβλημάτων ακοής.

Ένας τρίτος παράγοντας που καθόρισε την επιλογή των υποκειμένων είναι η σχέση τους με το εξεταζόμενο αντικείμενο. Οι συγκεκριμένοι ακροατές προέρχονταν από το Τ.Ε.Ι Μηχανικών Μουσικής Τεχνολογίας και Ακουστικής έτσι με βάση το πρόγραμμα σπουδών αλλά και τις δικές τους ασχολίες με την μουσική παραγωγή και γενικά με την μουσική, έρχονται σε επαφή με τέτοιου είδους προγράμματα μουσικής παραγωγής (Logic, Cubase, pro tools) συχνά οπότε έχουν μια εμπειρία και μια εξοικείωση με αυτό το αντικείμενο. Επίσης σημαντικό είναι να προσθέσουμε ότι τα 20 άτομα που χρησιμοποιήσαμε χωρίστηκαν σε 2 ομάδες των 10. Αυτές οι ομάδες είχαν το χαρακτηριστικό ότι στην πρώτη ομάδα ήταν άτομα που ναι μεν ήταν από το χώρο του Τ.Ε.Ι όπως είπαμε και παραπάνω αλλά δεν έπαιζαν κάποιο μουσικό όργανο ή δεν ασχολούνταν με την μουσική. Στην δεύτερη ομάδα ήταν άτομα από το Τ.Ε.Ι που παίζουν κάποιο μουσικό όργανο ή ασχολούνταν με την μουσική. Αυτό έγινε για να διαπιστώσουμε αν ένα καλά δουλεμένο αντί ενός μουσικού ή το αντίθετο θα μπορούσε να διαπιστώσει περισσότερες διαφορές στην πειραματική διαδικασία.

Ανακεφαλαιώνοντας τα υποκείμενα των ακροάσεων που χρησιμοποιήσαμε ήταν 20 και είχαν και μουσικές και τεχνικές γνώσεις για τον ήχο και χωρίστηκαν σε δυο ομάδες:

- *A' ομάδα:* 10 άτομα, ενασχόληση με μουσική τεχνολογία-παραγωγή, όχι όμως μουσικοί
- *B' ομάδα:* 10 άτομα, ενασχόληση με μουσική τεχνολογία-παραγωγή και μουσικοί

Να αναφέρουμε τέλος ότι το δείγμα που χρησιμοποιήθηκε είναι μη πιθανοτικό και, επομένως όχι τυχαίο. Συνεπώς τα αποτελέσματα που παρατίθενται στο επόμενο κεφάλαιο είναι δύσκολο να γενικευτούν στον γενικό πληθυσμό. Το μη πιθανοτικό δείγμα επελέγη διότι κρίθηκε απαραίτητη η τεχνική και μουσική γνώση των υποκειμένων προκειμένου να κατανοήσουν την φύση της πειραματικής διαδικασίας.

²⁴ Recommendation sector of UTI (2014). Methods for the subjective assessment of small impairments in audio systems. 4

²⁵ Chung-Hsien Wu, Yi-Chin Huang, Chung-Han Lee, and Jun-Cheng Guo (2014). Synthesis of Spontaneous Speech With Syllable Contraction Using State-Based Context-Dependent Voice Transformation. IEEE/ACM. 593-594.

²⁶ Takeda, S et al (1992). Age variation in the upper limit of hearing. European journal of applied physiology. 65, 403-408

5.6 Εξοπλισμός και λογισμικό

Για την προετοιμασία της πειραματικής διαδικασίας χρησιμοποιήθηκαν:

- τα λογισμικά *Logic Pro 9 (Apple)*, *Pro tools HD 9 (Avid)* και *Cubase 5 (Steinberg)* όπου με την χρήση των παραπάνω πραγματοποιήσαμε τα export των δειγμάτων για το αντικειμενικό αλλά και το υποκειμενικό στάδιο της διαδικασίας. Πιο συγκεκριμένα και από τα τρία DAW κάναμε export την πολυκάναλη παραγωγή *Nude* των *Radiohead*, το ημίτονο που δημιουργήσαμε με την βοήθεια του λογισμικού *audacity* αλλά και τα τρία κενά export. Το sample rate και για την πολυκάναλη παραγωγή και τα κενά export ήταν στα 44100Hz και το bit rate στα 24bit ενώ για το ημίτονο ήταν στα 44100Hz και 16bit.
- Τα ακουστικά ημιανοιχτού τύπου (semi-open) *Akg K121 studio*, με συχνοτική απόκριση 18 με 22.500Hz και συνολική αρμονική παραμόρφωση $THD \approx 0.2\%$ για την ακρόαση των δειγμάτων.
- Την κάρτα ήχου *Tascam US-144 MkII* για την δημιουργία και αναπαραγωγή των δειγμάτων.

Για την ακρόαση των δειγμάτων και την αναγνώριση τους από τα υποκείμενα, χρησιμοποιήθηκαν:

- Τα ακουστικά ημιανοιχτού τύπου *Akg K121 studio*, που αναφέρθηκαν παραπάνω.
- Το λογισμικό *foobar2000* για την αναπαραγωγή των δειγμάτων. Το λογισμικό αυτό επελέγη διότι μας παρέχει λειτουργία *abx comparator* βγάζοντας στο τέλος το ποσοστό ανάλογα με τις απαντήσεις που δίνει ο ακροατής.

5.7 Πειραματική διαδικασία

Οι ακροάσεις έγιναν ατομικά (ένα υποκείμενο κάθε φορά) και διεξήχθησαν με τον ακόλουθο τρόπο, ο οποίος επαναλήφθηκε για κάθε υποκείμενο:

- Κάθε υποκείμενο εισέρχεται στην αίθουσα και καθόταν μπροστά στο laptop. Στην επιφάνεια εργασίας βρισκόταν ήδη ανοιχτό το foobar2000. Και στην λίστα αναπαραγωγής του βρίσκονταν τα 6 ζευγάρια δειγμάτων προς ακρόαση.
- Στην συνέχεια το υποκείμενο διάβαζε της οδηγίες χρήσης του foobar2000 αλλά και ένα κομμάτι θεωρίας σχετικά με το περιεχόμενο και τον σκοπό της πτυχιακής εργασίας αλλά και για την μεθοδολογία του abx double blind test τι είναι και από τι αποτελείται (στο τέλος της πτυχιακής εργασίας θα παραθέσουμε και αυτό το κομμάτι των οδηγιών). Αφού το υποκείμενο διάβαζε τις οδηγίες δίνονταν περεταίρω εξηγήσεις για τυχόν παρερμηνείες και προχωρούσαμε στο foobar2000

και στην λειτουργία του abx comparator. Σε αυτό το σημείο γινόταν επεξήγηση του προγράμματος και αναλύονταν όλοι οι παράμετροι για την σωστή λειτουργία από τον ακροατή.

- Στην συνέχεια προσαρμοζόταν η ένταση στα ακουστικά στην επιθυμητή στάθμη από το κάθε υποκείμενο και ξεκινούσε η πειραματική ακρόαση.
- Κατόπιν το υποκείμενο προχωρούσε στην ακρόαση και αναγνώριση των δειγμάτων. Οι επαναλήψεις για κάθε ζευγάρι δειγμάτων ήταν 12 σύμφωνα με τα πρότυπα και τις προδιαγραφές της συγκεκριμένης μεθοδολογίας. Ο μέσος όρος ολοκλήρωσης της πειραματικής διαδικασίας ήταν τα 90 λεπτά, με το γρηγορότερο υποκείμενο να την ολοκληρώνει σε 50 λεπτά και το πιο αργό σε 150 λεπτά.

Κεφάλαιο 6^ο: ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΕΡΕΥΝΑΣ- ΣΥΓΚΡΙΣΕΙΣ

Εισαγωγή 6^{ου} Κεφαλαίου

Στο κεφάλαιο αυτό, παρουσιάζονται τα αποτελέσματα της πειραματικής διαδικασίας με την βοήθεια εικόνων και πινάκων. Η πειραματική διαδικασία όπως είπαμε και στα προηγούμενα κεφάλαια χωρίστηκε σε αντικειμενικό μέρος (null tests) και σε υποκειμενικό μέρος (ακροάσεις) με την χρήση abx double blind listening test.

Στο πρώτο μέρος του κεφαλαίου θα παραθέσουμε τα αποτελέσματα του αντικειμενικού μέρους για κάθε ζευγάρι export αλλά και για το ημίτονο και τα κενά export με την βοήθεια εικόνων που παρουσιάζουν τις διαφορές όπως αυτές κατεγράφησαν μέσα από τα προγράμματα που τα αναλύσαμε (audacity, wavelab). Στην συνέχεια θα παραθέσουμε τα αποτελέσματα των ακροάσεων (υποκειμενικό μέρος) με την βοήθεια πινάκων για το ποσοστό των απαντήσεων που σημειώθηκε σε κάθε ζευγάρι export.

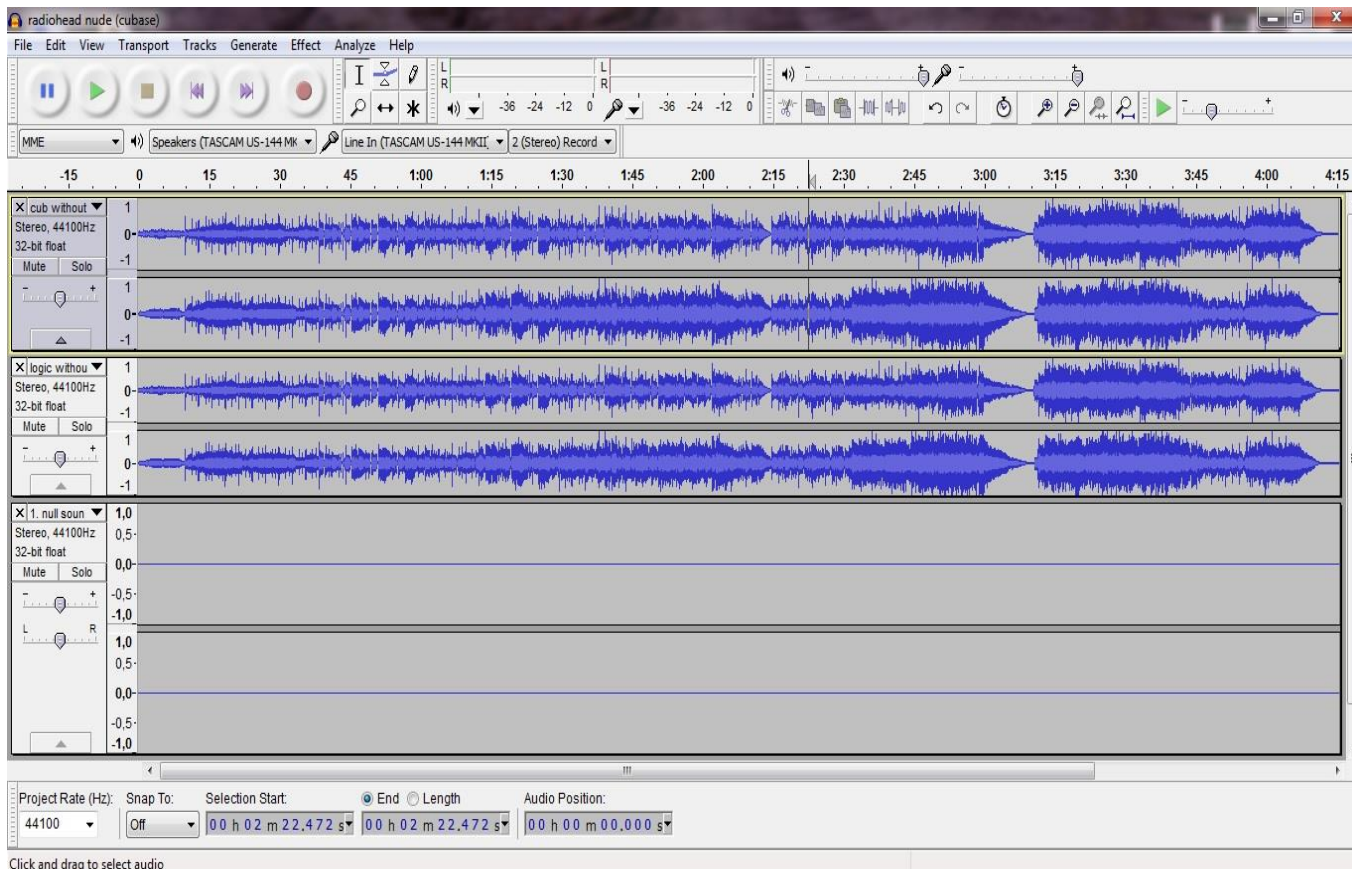
6.1 Αντικειμενικό μέρος πειραματικής διαδικασίας (Null tests)

Για κάθε ζευγάρι export που συγκρίναμε θα παραθέσουμε και την αντίστοιχη εικόνα αποτελεσμάτων. Να θυμίσουμε ότι τα ζευγάρια που έχουν δημιουργηθεί είναι 6 και είναι:

- 1^ο. Cubase-Logic (χωρίς plug-ins)
- 2^ο. Cubase-Pro Tools (χωρίς plug-ins)
- 3^ο. Logic-Pro Tools (χωρίς plug-ins)

- 4^ο. Cubase-Logic (με plug-ins)
- 5^ο. Cubase-Pro Tools (με plug-ins)
- 6^ο. Logic-Pro Tools (με plug-ins)

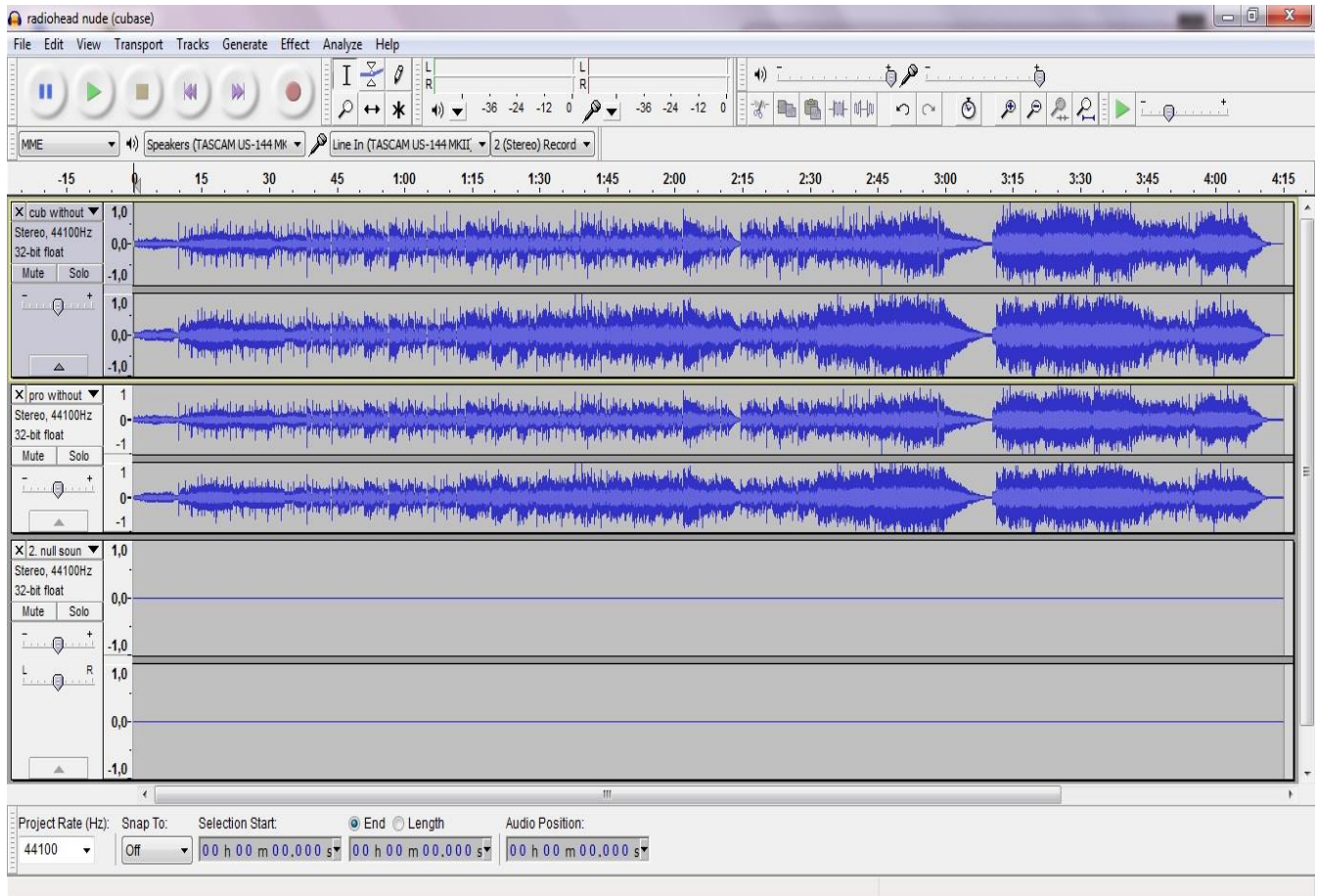
6.1.1 Πρώτο ζευγάρι: Cubase-Logic (χωρίς plug-ins)



Εικόνα 6.1 Null test για το ζευγάρι Cubase-Logic (χωρίς plug-ins).

Στην εικόνα 6.1 βλέπουμε τρία κανάλια στην περιοχή του arrange. Στο πρώτο κανάλι είναι το export από το Cubase στο δεύτερο κανάλι είναι το export από το Logic. Το τρίτο κανάλι είναι το αποτέλεσμα της αναστροφής του ενός από τα δύο πρώτα κανάλια (Cubase-Logic) και της πρόσθεσης τους. Αναλυτικότερα έχουμε πραγματοποιήσει αναστροφή φάσης σε ένα από τα δύο κανάλια στην συγκεκριμένη περίπτωση έχουμε κάνει αναστροφή φάσης στο κανάλι του Logic και στην συνέχεια τα έχουμε αθροίσει δηλαδή τα έχουμε κάνει export. Το αποτέλεσμα είναι όπως βλέπουμε το τρίτο κανάλι να έχει μηδενικό σήμα κάτι που σημαίνει ότι τα κανάλια Cubase και Logic έχουν αλληλοεξουδετερωθεί που αυτό με την σειρά του μας υποδεικνύει με μεγάλη σιγουριά ότι τα export από αυτά δυο προγράμματα (Cubase-Logic χωρίς plug-ins) είναι ίδια και δεν έχουν ηχητικές διαφορές.

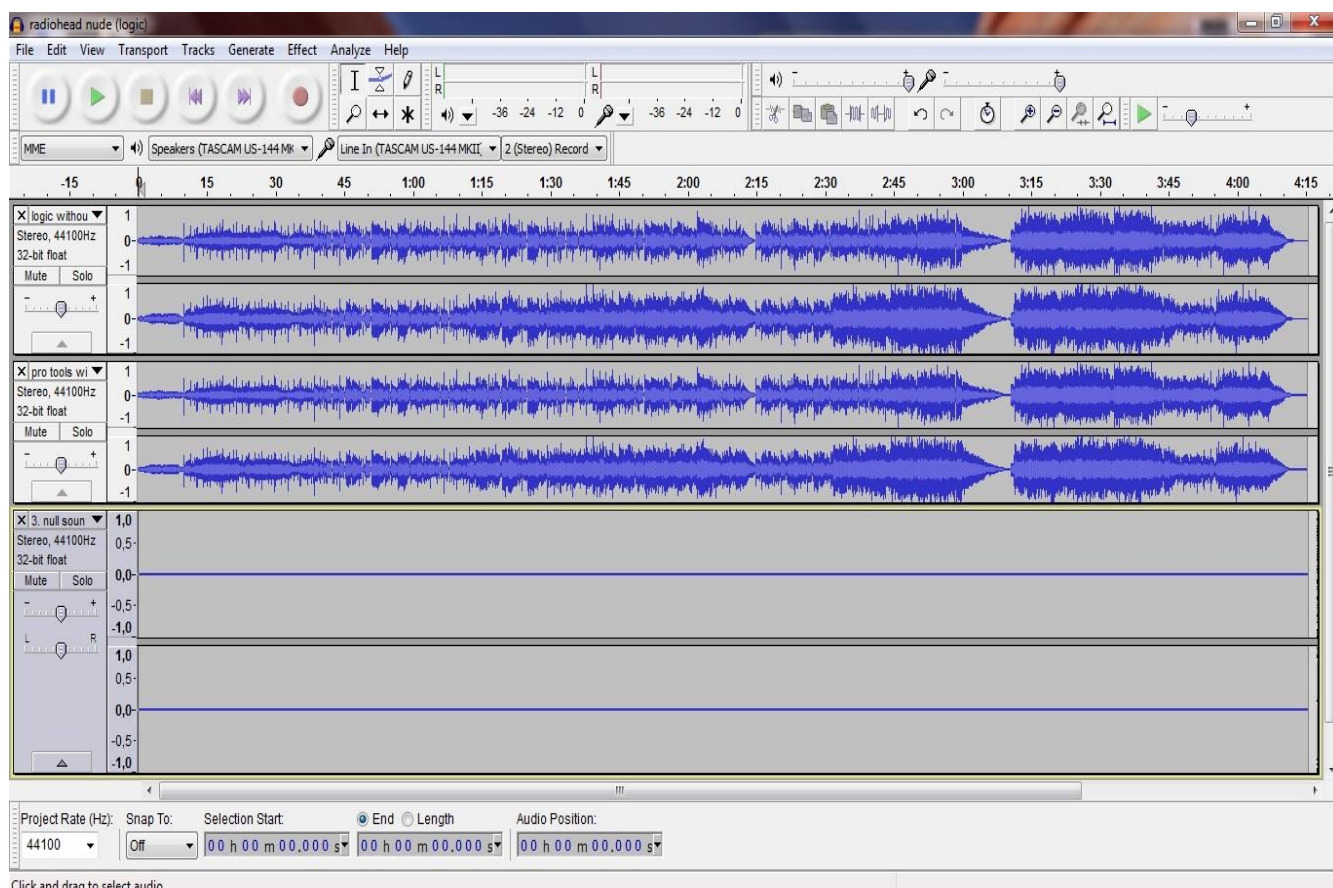
6.1.2 Δεύτερο ζευγάρι: Cubase-Pro tools (χωρίς plug-ins)



Εικόνα 6.2 Null test για το ζευγάρι Cubase-Pro tools (χωρίς plug-ins).

Στην εικόνα 6.2 έχουμε ακριβώς το ίδιο αποτέλεσμα με το πρώτο ζευγάρι. Βλέπουμε ότι με την ίδια ακριβώς διαδικασία το ζευγάρι Cubase-Pro tools (χωρίς plug-ins) αλληλοεξουδετερώθηκε άρα δεν παρουσιάζουν ηχητικές διαφορές στα export τους, είναι δηλαδή πανομοιότυπα.

6.1.3 Τρίτο ζευγάρι: Logic-Pro tools (χωρίς plug-ins)

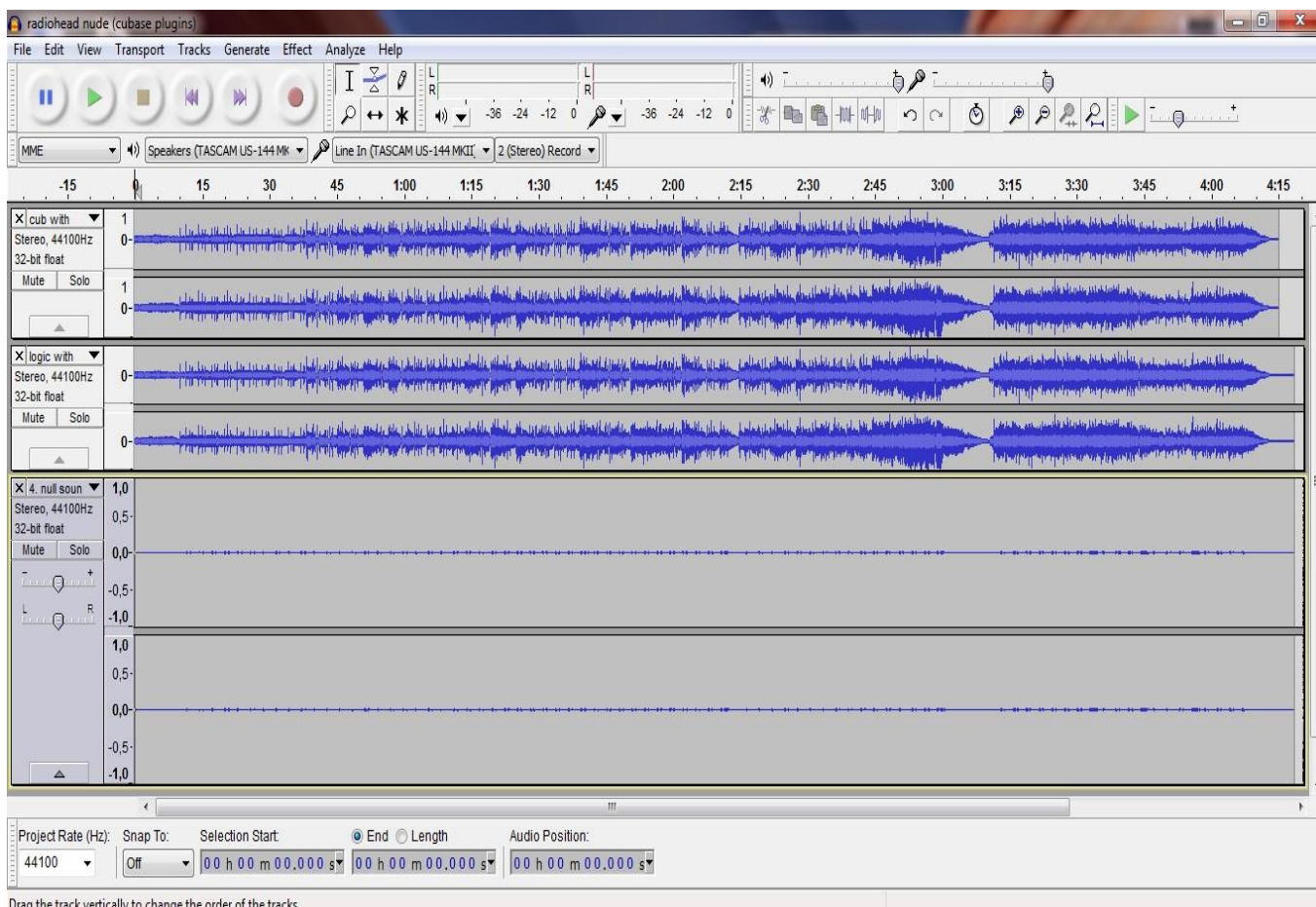


Εικόνα 6.3 Null test για το ζευγάρι Logic-Pro tools (χωρίς plug-ins).

Τα ίδια ακριβώς αποτελέσματα εξάγονται και για το τρίτο ζευγάρι χωρίς plug-ins Logic-Pro tools και τα δύο export είναι πανομοιότυπα δεν έχουν διαφορές στο ήχο με αποτέλεσμα το σήμα του τρίτου καναλιού να είναι μηδενικό.

Στην συνέχεια θα παρουσιάσουμε τα εναπομείναντα τρία ζευγάρια που στα exports έχουμε προσθέσει και plug-ins στο κάθε κανάλι. Να σημειώσουμε ότι έχουμε βάλει τα ίδια ακριβώς plug-ins της Waves με τα ίδια ακριβώς preset και στα 3 DAW και ότι ο σκοπός μας δεν ήταν να τα χρησιμοποιήσουμε για αισθητική παρέμβαση αλλά για να τονίσουμε τις διαφορές που έχουν τα DAW στην διαχείριση και την DSP επεξεργασία με την χρήση των plug-ins.

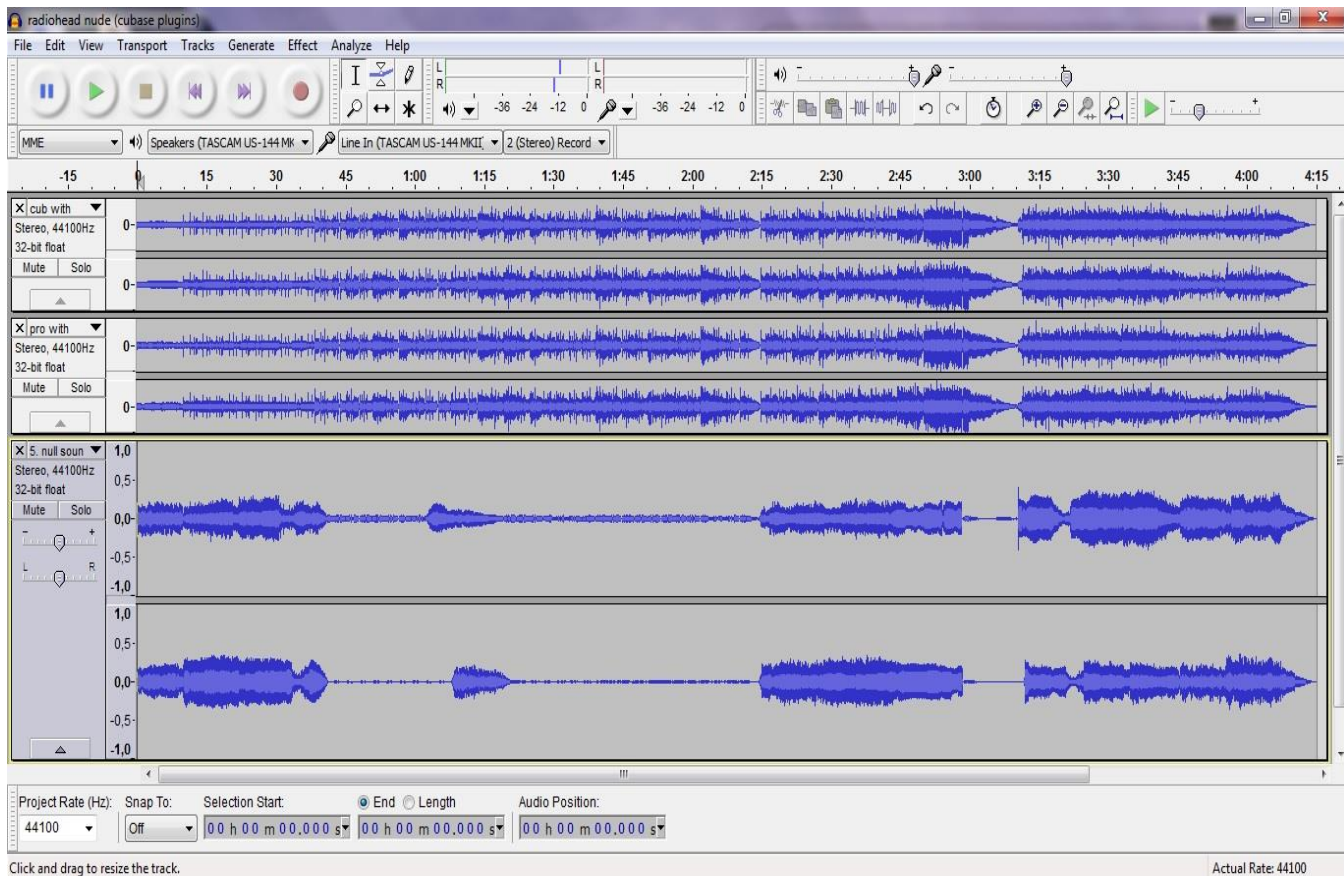
6.1.4 Τέταρτο ζευγάρι: Cubase-Logic (με plug-ins)



Εικόνα 6.4 Null test για το ζευγάρι Cubase-Logic (με plug-ins).

Με την χρήση plug-ins βλέπουμε ξεκάθαρα ότι αρχίζουν να προκύπτουν διαφορές στην άθροιση των DAW. Αυτό συμβαίνει διότι τα DAW έχουν διαφορετικό τρόπο το καθένα το πως δουλεύει και χρησιμοποιεί τα plug-ins (§ 2.5). Στο αποτέλεσμα του συγκεκριμένου null test μεταξύ Cubase-Logic (με plug-ins) παρουσιάζεται μια μικρή διαφορά που βλέπουμε στο 3^ο σήμα στην εικόνα. Αυτή η διαφορά είναι μικρή αλλά ακουστή όταν επιχειρούμε να ακούσουμε το συγκεκριμένο δείγμα. Σύμφωνα με όλα τα παραπάνω προκύπτουν διαφορές στο ηχητικό αποτέλεσμα στο συγκεκριμένο ζευγάρι.

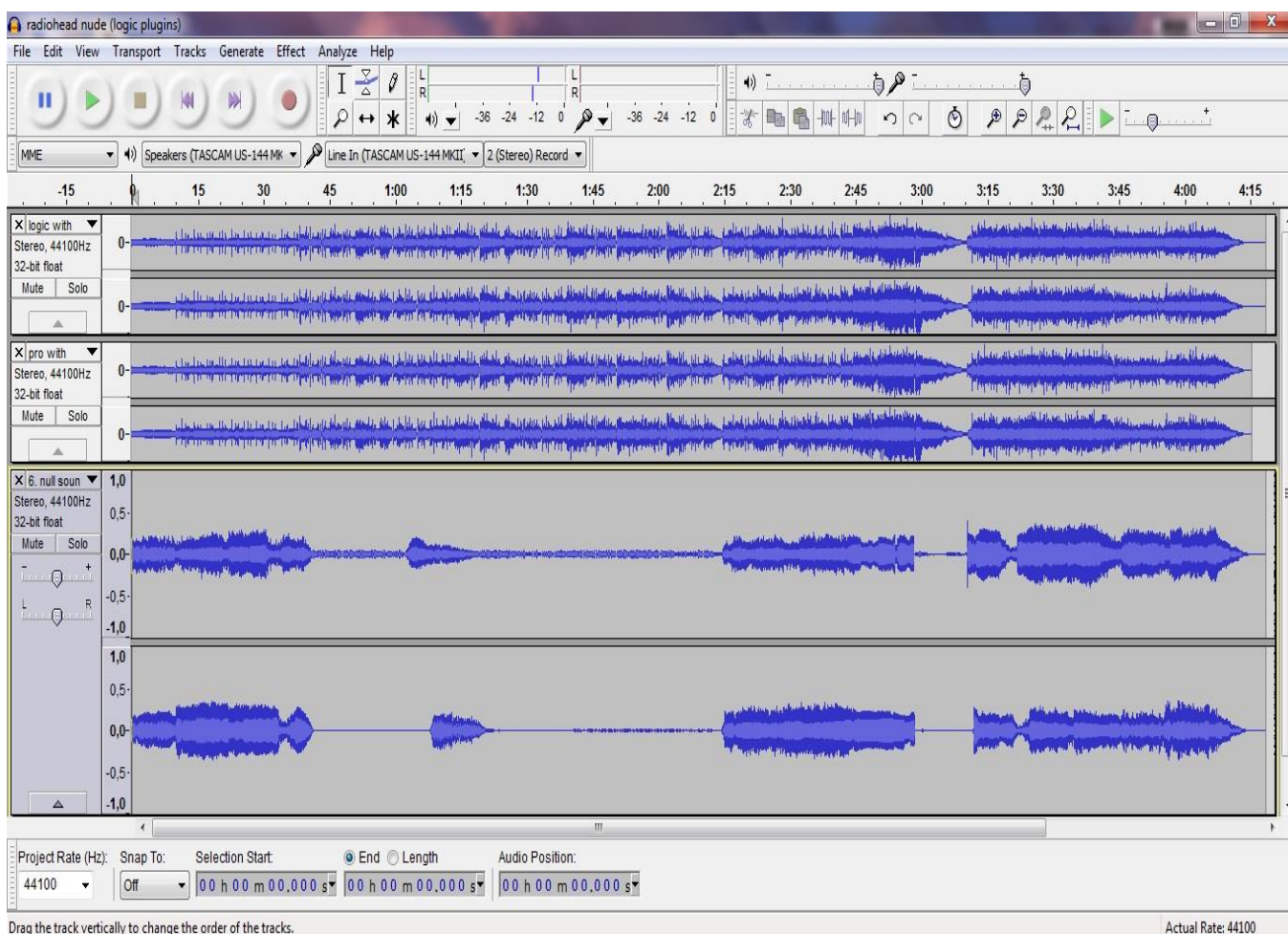
6.1.5 Πέμπτο ζευγάρι: Cubase-Pro tools (με plug-ins)



Εικόνα 6.5 Null test για το ζευγάρι Cubase-Pro tools (με plug-ins).

Στο πέμπτο ζευγάρι παρατηρούμε ότι οι διαφορές στο ηχητικό αποτέλεσμα είναι ακόμα μεγαλύτερες (βλέποντας την κυματομορφή του τρίτου σήματος βλέπουμε ότι είναι μεγαλύτερη σε σχέση με το προηγούμενο ζευγάρι). Αυτό μας οδηγεί στο συμπέρασμα ότι τα Pro tools έχουν μεγαλύτερη διαφορά στον τρόπο που χειρίζονται τα plug-ins σε σχέση με το Logic και το Cubase.

6.1.6 Έκτο ζευγάρι: Logic-Pro tools (με plug-ins)

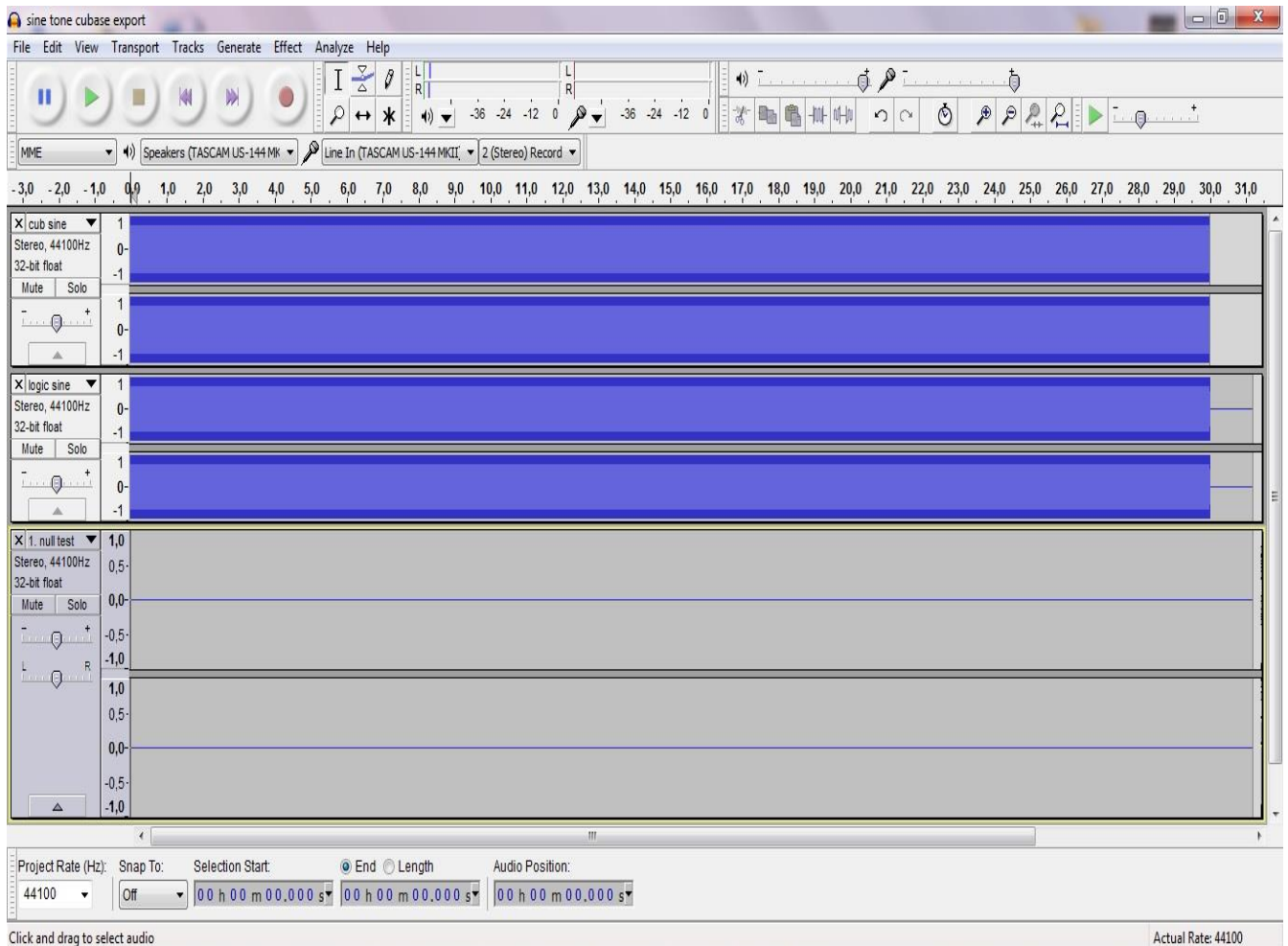


Εικόνα 6.6 Null test για το ζευγάρι Logic-Pro tools (με plug-ins).

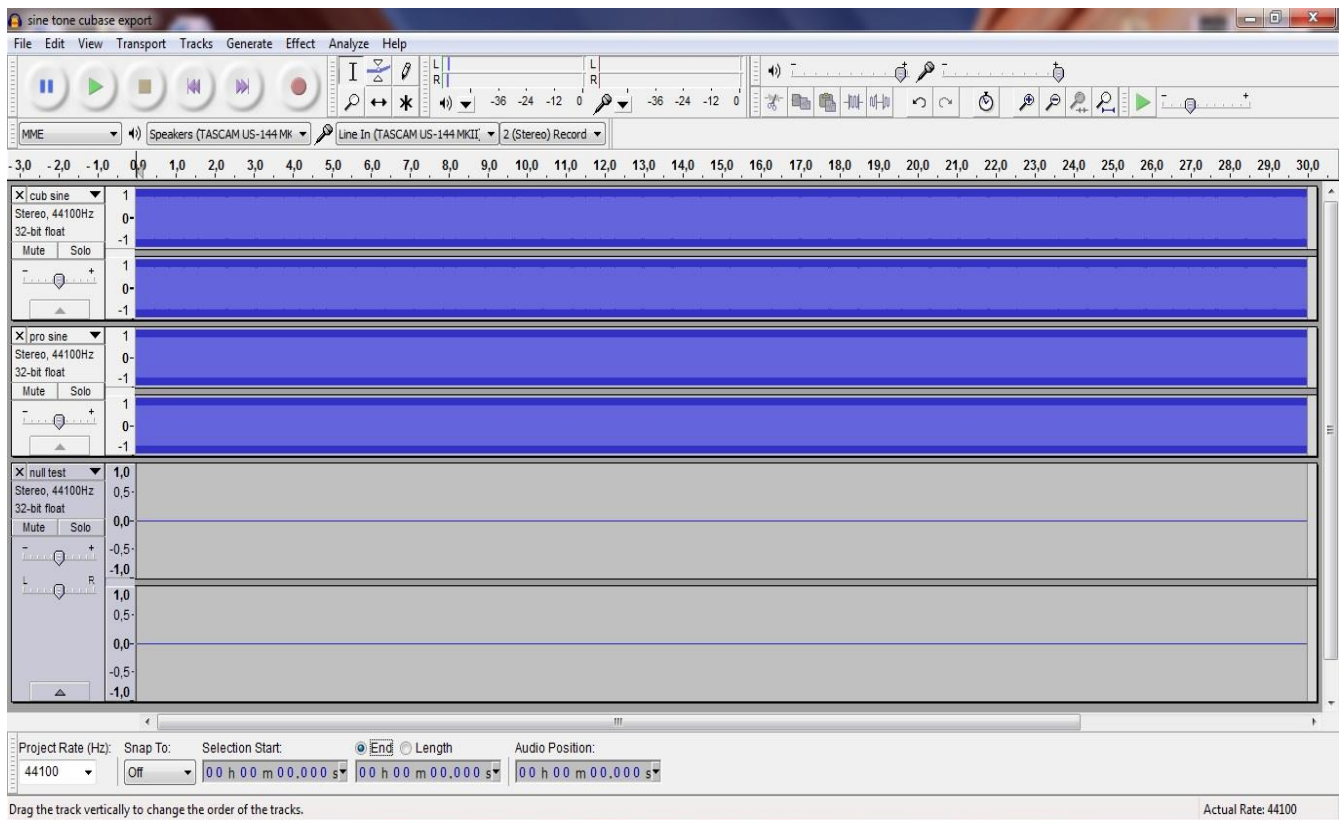
Τέλος και για το ζευγάρι Logic-Pro tools παρατηρούμε ότι οι ηχητικές διαφορές είναι αρκετές με την χρήση plug-ins. Η κυματομορφή της διαφοράς των δυο export βλέπουμε να είναι παρόμοια με την κυματομορφή στην εικόνα 6.5 έτσι επιβεβαιώνεται το συμπέρασμα που είπαμε και παραπάνω ότι τα Pro tools σε σχέση με Logic και το Cubase έχουν μεγαλύτερη διαφορά στον τρόπο που διαχειρίζονται τα plug-ins και την DSP επεξεργασία.

6.1.7 Null test για ημίτονο στα 440Hz (export σε Logic, Pro tools, Cubase)

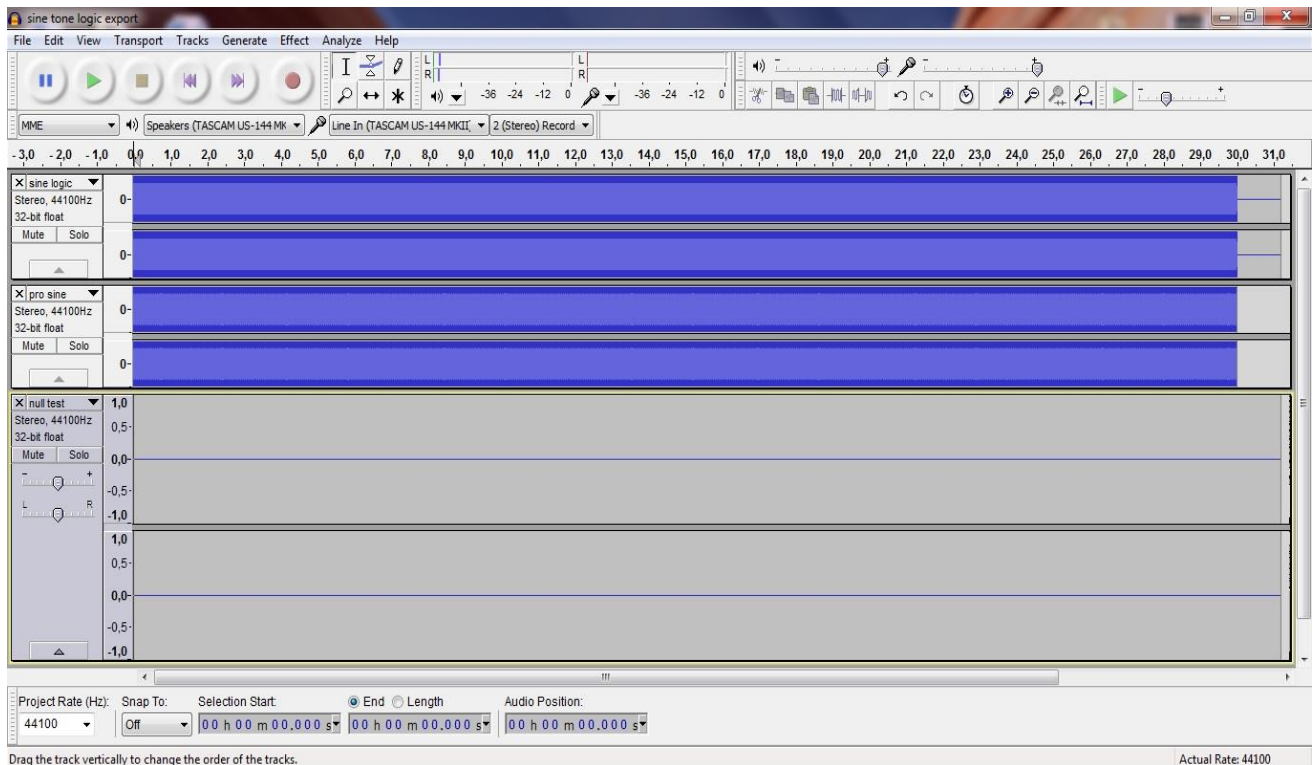
Για την συγκεκριμένη διαδικασία χρησιμοποιήσαμε ένα ημίτονο που δημιουργήσαμε στο audacity και το κάναμε export στα τρία DAW που μελετάμε χωρίς την χρήση plug-ins και πραγματοποιήσαμε την ίδια διαδικασία (null test) με τα παραπάνω δείγματα. Να σημειώσουμε ότι και στα τρία DAW εισάγαμε το ημίτονο σε 3 διαφορετικά κανάλια ώστε να δουλέψει η αθροιστική μηχανή των λογισμικών και μειώσαμε την στάθμη στα -5.2dBfs για να μην κλιπάρει το master out. Το pan law ήταν όπως και στα προηγούμενα δείγματα στα -3dB και για τα τρία DAW.



Εικόνα 6.7 Null test για ημίτονο με το ζευγάρι Cubase-Logic (χωρίς plug-ins).



Εικόνα 6.8 Null test για ημίτονο με το ζευγάρι Cubase-pro tools (χωρίς plug-ins).

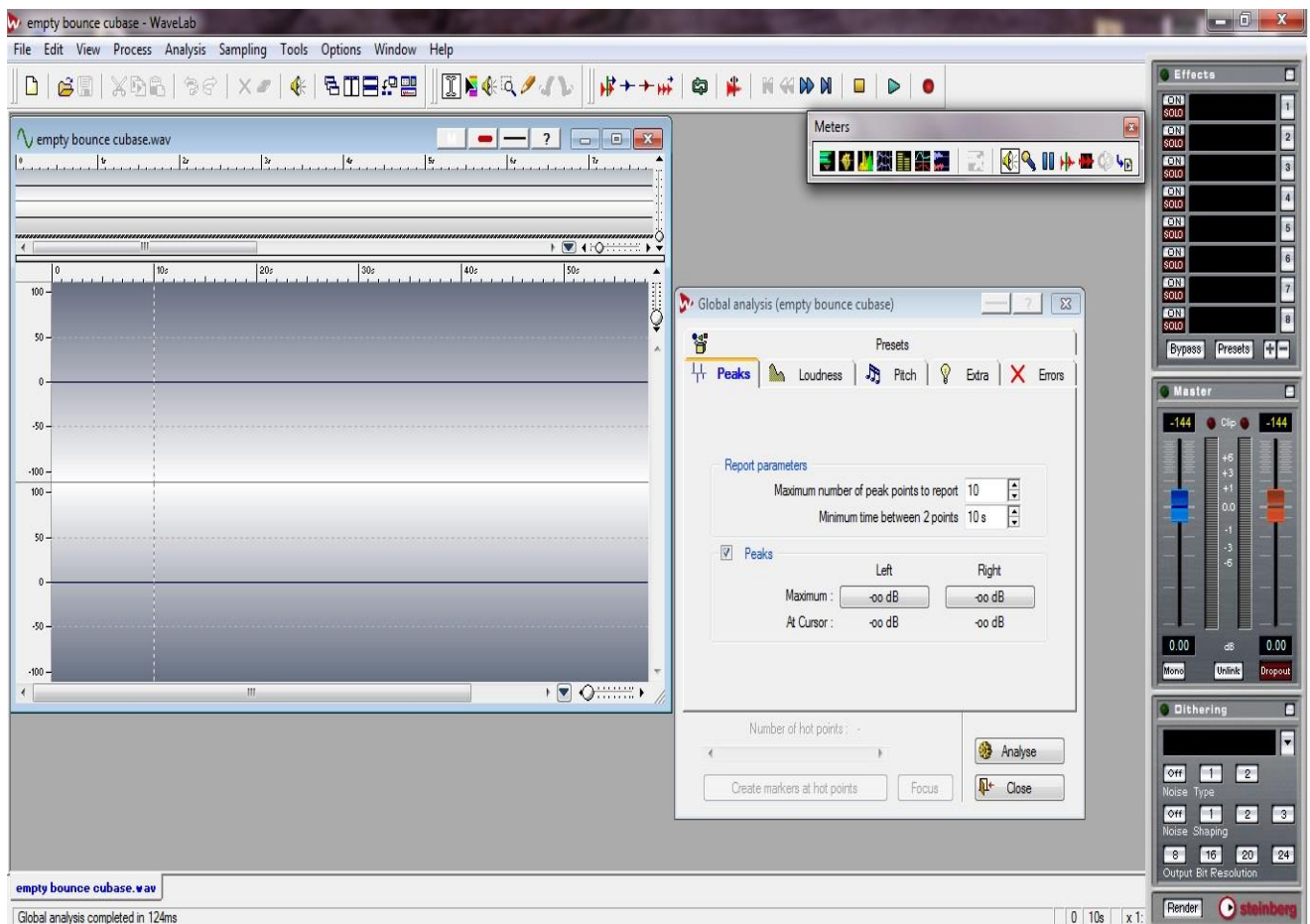


Εικόνα 6.9 Null test για ημίτονο με το ζευγάρι Logic-Pro tools (χωρίς plug-ins).

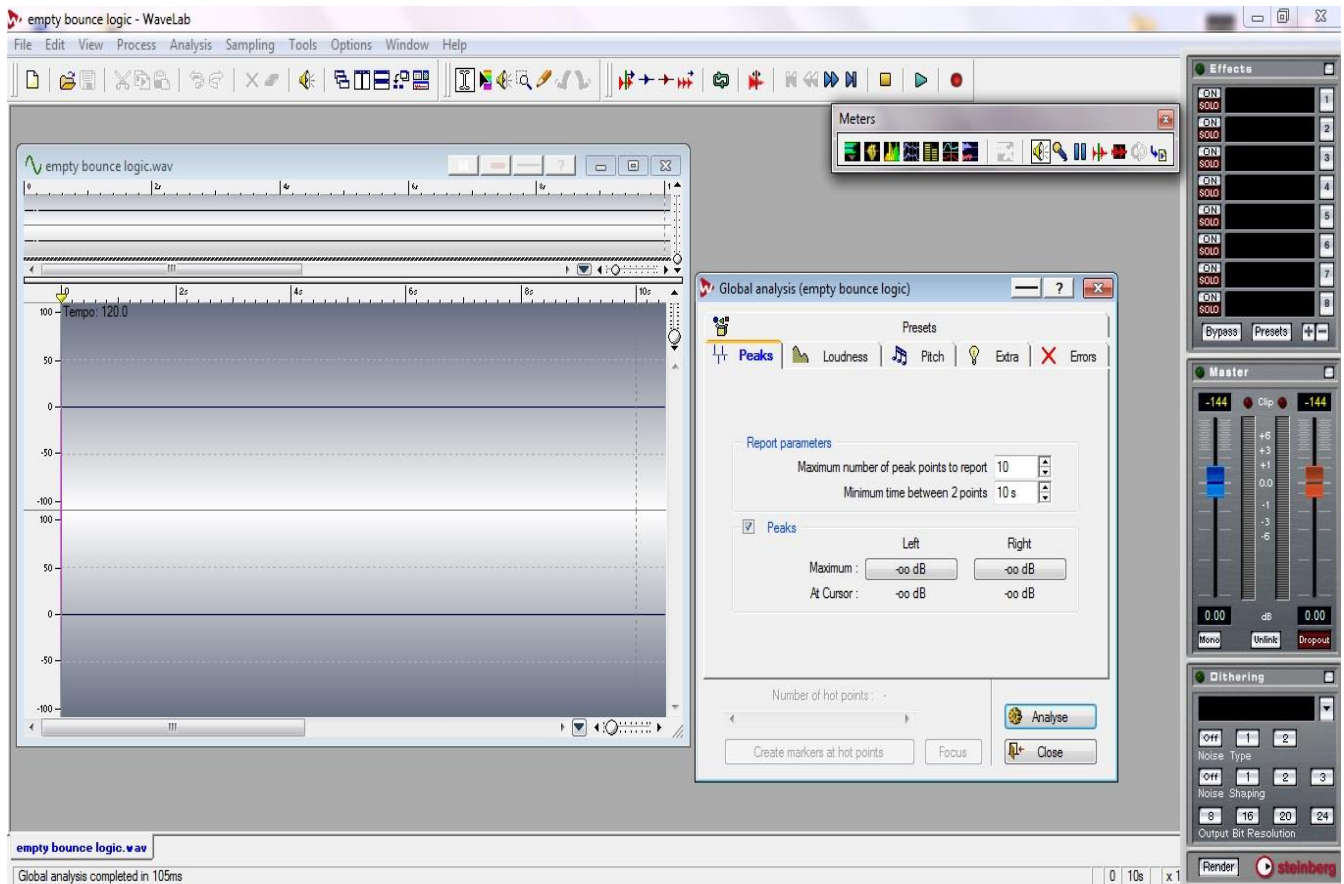
Από τις εικόνες 6.7, 6.8, 6.9 καταλήγουμε πάλι στο συμπέρασμα ότι οι μηχανές ήχου των DAW που μελετάμε δεν έχουν καμία διαφορά στο ηχητικό αποτέλεσμα με την άθροιση απλών σημάτων χωρίς την χρήση plug-ins. Και στις τρεις εικόνες βλέπουμε πως με αναστροφή φάσης το ενός σήματος και πρόσθεσης των σημάτων τα δείγματα από τα DAW αλληλοαναιρούνται, εύκολα καταλαβαίνουμε δηλαδή ότι δεν παρουσιάζουν διαφορές.

6.1.8 Global analysis κενών export των τριών DAW στο wavelab

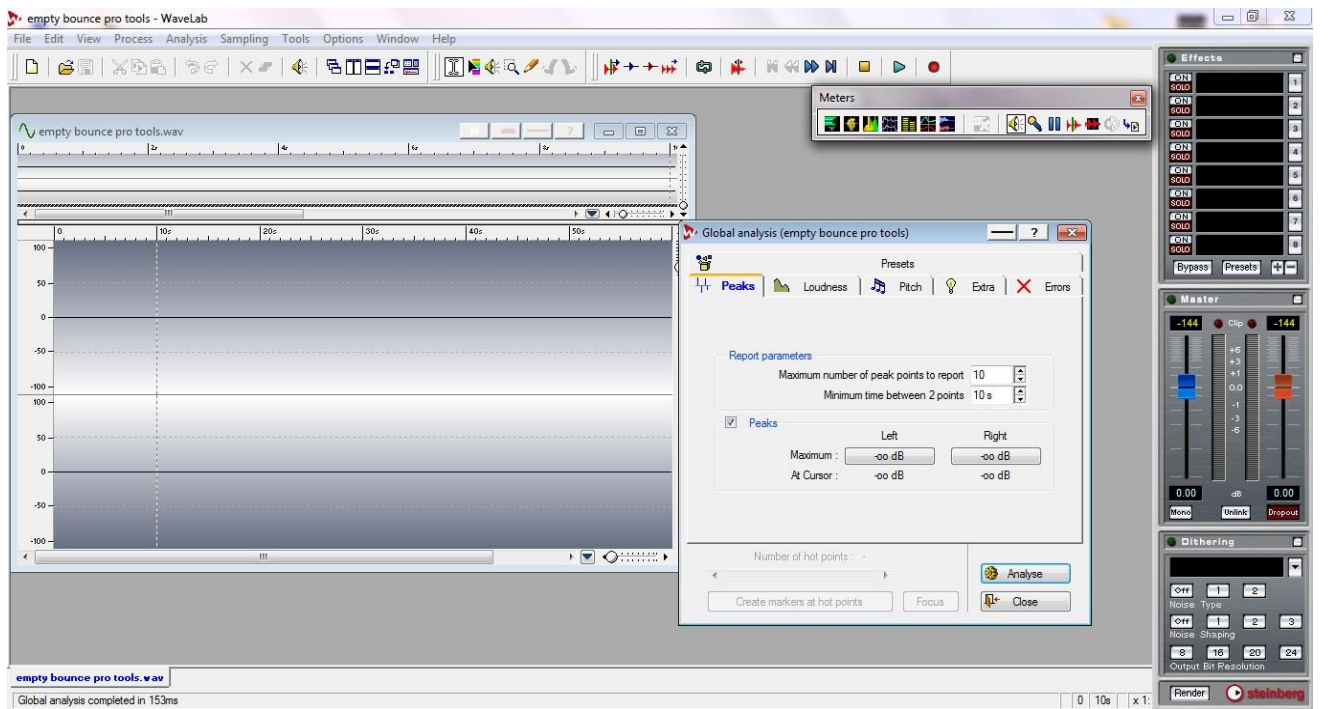
Το σκεπτικό της συγκεκριμένης ανάλυσης είναι να δούμε στα DAW που εξετάζουμε όταν κάνουμε κενό export αν περνάνε χρήσιμες πληροφορίες που με ανάλυση μπορούν να μας δώσουν κάποια στοιχεία για περαιτέρω διαφορές ανάμεσα σε αυτά τα λογισμικά που εξετάζουμε. Κατά το export τα DAW περνάνε μέσα στο σήμα κάποια μορφή θορύβου ή dither. Με την βοήθεια της λειτουργίας global analysis του wavelab θα δούμε αν μπορούμε να εξάγουμε κάποια συμπεράσματα.



Εικόνα 6.10 Global analysis για empty export από Cubase.



Εικόνα 6.11 Global analysis για empty export από Logic.



Εικόνα 6.12 Global analysis για empty export από Pro tools.

Από την παραπάνω ανάλυση και των τριών DAW και από τις εικόνες 6.10, 6.11, 6.12 δεν μπορούν να προκύψουν κάποια χρήσιμα συμπεράσματα. Βλέπουμε τόσο σε ανάλυση peaks αλλά όσο και σε ανάλυση loudness, pitch, errors (δεν φαίνονται στην εικόνα) ότι τα αποτελέσματα είναι ίδια. Σε peaks και loudness και για τα τρία DAW η στάθμη βρίσκεται στο $-\infty$. Το μόνο χρήσιμο συμπέρασμα είναι ότι και τα τρία DAW έχουν την ίδια συμπεριφορά σε peak και loudness levels όταν πραγματοποιούμε κενό export.

6.2 Υποκειμενικό μέρος πειραματικής διαδικασίας (listening tests με την μέθοδο abx double blind)

Όπως αναφέραμε και στα κεφάλαια 4-5 με την μεθοδολογία abx double blind test, ο ακροατής έχει πρόσβαση σε τρεις πηγές που ονομάζονται *A*, *B*, και *X*. *A* και *B* είναι οι αναφορές. Για παράδειγμα, *A* το αρχείο ήχου από το Logic και *B* το αρχείο ήχου από τα Pro Tools. *X* είναι η άγνωστη πηγή. Μπορεί να είναι ή *A* ή *B*. Ο ακροατής πρέπει να επιλέξει ποιά από τις δύο αναφορές (*A* ή *B*) αναπαράγεται κάθε φορά στην *X* επανάληψη.

Στην συγκεκριμένη πειραματική διαδικασία για να πραγματοποιήσουμε το συγκεκριμένο τεστ χρησιμοποιήσαμε το λογισμικό foobar2000 και την λειτουργία abx comparator. Αρχικά θα θυμίσουμε του περιορισμούς και τι πρέπει να προσέξουμε για να πραγματοποιηθεί σωστά η συγκεκριμένη πειραματική ακρόαση.

1. Το σύνολο των υποκειμένων (ακροατών) που πρέπει να συμμετάσχει σε τέτοιου είδους πειράματα πρέπει να είναι από 20 και πάνω. Στην συγκεκριμένη διαδικασία χρησιμοποιήσαμε 20 υποκείμενα ($N=20$).
2. Ο αριθμός των επαναλήψεων για κάθε ζευγάρι δειγμάτων πρέπει να είναι πάνω από 10 και το πολύ μέχρι 20 για να είναι τα αποτελέσματα μας αξιόπιστα. Επειδή ο όγκος των δειγμάτων ήταν μεγάλος και ήθελε αρκετό χρόνο για να ολοκληρωθεί η εργασία κάτι που θα οδηγούσε στο να κουραστούν τα υποκείμενα κρίθηκε σκόπιμο ο αριθμός των επαναλήψεων να είναι 12.
3. Το πρόγραμμα foobar2000 από την λειτουργία του (by default) χρησιμοποιεί binomial κατανομή (confidence level 95% - p-value 0,05). Άρα σύμφωνα με τους πίνακες 4.1-4.2 ο αριθμός των σωστών απαντήσεων που μπορούμε να μιλάμε για στατιστικά σημαντικά αποτελέσματα (statistically significant result) είναι από 10/12 και άνω.

Στην συνέχεια θα παραθέσουμε πίνακες για το κάθε ζευγάρι δειγμάτων και τις σωστές απαντήσεις που έδωσαν οι ακροατές. Έτσι θα δούμε αν κατάφεραν να πιάσουν το όριο του 10/12 σωστών απαντήσεων.

6.2.1 Διατύπωση της μηδενικής υπόθεσης (null hypothesis)

Πριν ξεκινήσουμε να παραθέτουμε τα αποτελέσματα πρέπει να διατυπώσουμε την μηδενική υπόθεση για την συγκεκριμένη ακρόαση. Η μηδενική υπόθεση αναφέρεται σε μια γενική δήλωση ή θέση ότι δεν υπάρχει καμία σχέση μεταξύ των δύο φαινομένων που μετρούνται. Απόρριψη ή ανασκευή αυτής της υπόθεσης – και οδηγούμαστε στο συμπέρασμα ότι υπάρχουν λόγοι να πιστεύουμε ότι υπάρχει μια σχέση μεταξύ δύο φαινομένων. Στην συγκεκριμένη πειραματική διαδικασία η μηδενική υπόθεση αναφέρεται στο ότι “*οι μηχανές ήχου (audio engines) των συγκεκριμένων DAW που μελετάμε δεν παρουσιάζουν διαφορές στο ηχητικό αποτέλεσμα έτσι τα δείγματα που συγκρίνουμε δεν έχουν διαφορές*”. Ο μόνος τρόπος για να ακυρωθεί η μηδενική υπόθεση είναι οι ακροατές να δώσουν σωστές απαντήσεις από 10 και άνω για να μιλάμε για στατιστικά σημαντικά αποτελέσματα (statistically significant result).

6.2.2 Πρώτο ζευγάρι: Cubase-Logic (χωρίς plug-ins)

p-value (%) (πιθανότητα να δόθηκαν οι απαντήσεις στην τύχη)	Αριθμός σωστών απαντήσεων	Αριθμός ακροατών που απάντησαν
100%	1/12	-
99.7%	2/12	-
98.1%	3/12	2
92.7%	4/12	5
80.6%	5/12	2
61.3%	6/12	4
38.7%	7/12	2
19.4%	8/12	5
7.3%	9/12	-
1.9%	10/12	-
0.3%	11/12	-
0.0%	12/12	-
Σύνολο (N)		N=20

Πίνακας 6.1 Ο αριθμός των σωστών απαντήσεων που έδωσαν οι ακροατές για το ζευγάρι Cubase-Logic (χωρίς plug-ins).

Στον παραπάνω πίνακα βλέπουμε ότι

- 2 ακροατές απάντησαν σωστά 3 στις 12 απαντήσεις
- 5 ακροατές απάντησαν σωστά 4 στις 12 απαντήσεις
- 2 ακροατές απάντησαν σωστά 5 στις 12 απαντήσεις
- 4 ακροατές απάντησαν σωστά 6 στις 12 απαντήσεις
- 2 ακροατές απάντησαν σωστά 7 στις 12 απαντήσεις
- 5 ακροατές απάντησαν σωστά 8 στις 12 απαντήσεις

Με βάση τα παραπάνω στοιχεία βλέπουμε ότι κανένας ακροατής στο συγκεκριμένο ζευγάρι δειγμάτων (Cubase-Logic χωρίς plug-ins) δεν μπόρεσε να απορρίψει την μηδενική υπόθεση με το ποσοστό των απαντήσεων του. Αρκετά υποκείμενα έδωσαν 8/12 σωστές απαντήσεις αλλά αυτό δεν είναι αρκετό σύμφωνα με όσα έχουμε αναφέρει στην θεωρία.

6.2.3 Δεύτερο ζευγάρι: Cubase-Pro tools (χωρίς plug-ins)

p-value (%) (πιθανότητα να δόθηκαν οι απαντήσεις στην τύχη)	Αριθμός σωστών απαντήσεων	Αριθμός ακροατών που απάντησαν
100%	1/12	-
99.7%	2/12	-
98.1%	3/12	-
92.7%	4/12	4
80.6%	5/12	2
61.3%	6/12	6
38.7%	7/12	6
19.4%	8/12	2
7.3%	9/12	-
1.9%	10/12	-
0.3%	11/12	-
0.0%	12/12	-
Σύνολο (N)		N=20

Πίνακας 6.2 Ο αριθμός των σωστών απαντήσεων που έδωσαν οι ακροατές για το ζευγάρι Cubase-pro tools (χωρίς plug-ins).

Στον παραπάνω πίνακα βλέπουμε ότι

- 4 ακροατές απάντησαν σωστά 4 στις 12 απαντήσεις
- 2 ακροατές απάντησαν σωστά 5 στις 12 απαντήσεις
- 6 ακροατές απάντησαν σωστά 6 στις 12 απαντήσεις
- 6 ακροατές απάντησαν σωστά 7 στις 12 απαντήσεις
- 2 ακροατές απάντησαν σωστά 8 στις 12 απαντήσεις

Τα ίδια αποτελέσματα έχουμε και γι αυτό το ζευγάρι δειγμάτων. Κανένα από τα 20 υποκείμενα δεν κατάφερα να απορρίψει την μηδενική υπόθεση. Το μεγαλύτερο ποσοστό απαντήσεων κυμάνθηκε γύρω από το 6/12 και 7/12 μικρότερο σε σχέση με το πρώτο ζευγάρι.

6.2.4 Τρίτο ζευγάρι: Logic-Pro tools (χωρίς plug-ins)

p-value (%) (πιθανότητα να δόθηκαν οι απαντήσεις στην τύχη)	Αριθμός σωστών απαντήσεων	Αριθμός ακροατών που απάντησαν
100%	1/12	-
99.7%	2/12	-
98.1%	3/12	1
92.7%	4/12	1
80.6%	5/12	5
61.3%	6/12	7
38.7%	7/12	2
19.4%	8/12	1
7.3%	9/12	2
1.9%	10/12	1
0.3%	11/12	-
0.0%	12/12	-
Σύνολο (N)		N=20

Πίνακας 6.3 Ο αριθμός των σωστών απαντήσεων που έδωσαν οι ακροατές για το ζευγάρι Logic-pro tools (χωρίς plug-ins).

Στον παραπάνω πίνακα βλέπουμε ότι

- 1 ακροατής απάντησε σωστά 3 στις 12 απαντήσεις
- 1 ακροατής απάντησε σωστά 4 στις 12 απαντήσεις
- 5 ακροατές απάντησαν σωστά 5 στις 12 απαντήσεις
- 7 ακροατές απάντησαν σωστά 6 στις 12 απαντήσεις

- 2 ακροατές απάντησαν σωστά 7 στις 12 απαντήσεις
- 1 ακροατής απάντησε σωστά 8 στις 12 απαντήσεις
- 2 ακροατές απάντησαν σωστά 9 στις 12 απαντήσεις
- 1 ακροατής απάντησε σωστά 10 στις 12 απαντήσεις

Στο συγκεκριμένο ζευγάρι συναντάμε το πρώτο υποκείμενο που κατάφερε να απορρίψει την μηδενική υπόθεση. Με 10/12 και με 1.9% πιθανότητας να απάντησε στην τύχη ο συγκεκριμένος ακροατής δείχνει ότι στο συγκεκριμένο ζευγάρι (logic-pro tools χωρίς plug-ins) κατάλαβε διαφορές στο ηχητικό αποτέλεσμα κάτι που συνεπάγεται ότι αυτά τα δύο DAW παρουσιάζουν διαφορές στην άθροιση. Αυτό το συμπέρασμα όμως θα μπορούσε να είναι βάσιμο αν μεγαλύτερο ποσοστό ακροατών είχε δώσει σωστά 10 απαντήσεις στις 12. Βλέπουμε όμως ότι μόνο 1 στους 20 το κατάφερε και ότι το μεγαλύτερο ποσοστό απάντησε 6/12. Με όλα τα παραπάνω είναι πολύ δύσκολο να προκύπτουν διαφορές στο συγκεκριμένο ζευγάρι.

6.2.5 Τέταρτο ζευγάρι: Cubase-Logic (με plug-ins)

p-value (%) (πιθανότητα να δόθηκαν οι απαντήσεις στην τύχη)	Αριθμός σωστών απαντήσεων	Αριθμός ακροατών που απάντησαν
100%	1/12	-
99.7%	2/12	-
98.1%	3/12	1
92.7%	4/12	3
80.6%	5/12	3
61.3%	6/12	6
38.7%	7/12	3
19.4%	8/12	4
7.3%	9/12	-
1.9%	10/12	-
0.3%	11/12	-
0.0%	12/12	-
Σύνολο (N)		N=20

Πίνακας 6.4 Ο αριθμός των σωστών απαντήσεων που έδωσαν οι ακροατές για το ζευγάρι Cubase-Logic (με plug-ins).

Στον παραπάνω πίνακα βλέπουμε ότι

- 1 ακροατής απάντησε σωστά 3 στις 12 απαντήσεις
- 3 ακροατές απάντησαν σωστά 4 στις 12 απαντήσεις
- 3 ακροατές απάντησαν σωστά 5 στις 12 απαντήσεις
- 6 ακροατές απάντησαν σωστά 6 στις 12 απαντήσεις
- 3 ακροατές απάντησαν σωστά 7 στις 12 απαντήσεις
- 4 ακροατές απάντησαν σωστά 8 στις 12 απαντήσεις

Περνώντας τώρα στα export με plug-ins τα πράγματα δεν αλλάζουν. Κανένας ακροατής δεν κατάφερε με το ποσοστό του να απορρίψει την μηδενική υπόθεση. Αυτή είναι και η μεγάλη διαφορά με το αντικειμενικό μέρος. Στο αντικειμενικό μέρος είδαμε ότι στα δείγματα με plug-ins αρχίζουν να προκύπτουν διαφορές. Βλέπουμε όμως ότι στο υποκειμενικό μέρος το ανθρώπινο σύστημα ακοής δεν μπορεί να τις καταλάβει και να τις εντοπίσει με αποτέλεσμα το μεγαλύτερο ποσοστό των υποκειμένων να απαντά 6 στις 12 απαντήσεις σωστά με πιθανότητα να απάντησε στην τύχη 61.3%.

6.2.6 Πέμπτο ζευγάρι: Cubase-Pro tools (με plug-ins)

p-value (%) (πιθανότητα να δόθηκαν οι απαντήσεις στην τύχη)	Αριθμός σωστών απαντήσεων	Αριθμός ακροατών που απάντησαν
100%	1/12	-
99.7%	2/12	-
98.1%	3/12	1
92.7%	4/12	3
80.6%	5/12	3
61.3%	6/12	5
38.7%	7/12	5
19.4%	8/12	3
7.3%	9/12	-
1.9%	10/12	-
0.3%	11/12	-
0.0%	12/12	-
Σύνολο (N)		N=20

Πίνακας 6.5 Ο αριθμός των σωστών απαντήσεων που έδωσαν οι ακροατές για το ζευγάρι Cubase-Pro tools (με plug-ins).

Στον παραπάνω πίνακα βλέπουμε ότι

- 1 ακροατής απάντησε σωστά 3 στις 12 απαντήσεις
- 3 ακροατές απάντησαν σωστά 4 στις 12 απαντήσεις
- 3 ακροατές απάντησαν σωστά 5 στις 12 απαντήσεις
- 5 ακροατές απάντησαν σωστά 6 στις 12 απαντήσεις
- 5 ακροατές απάντησαν σωστά 7 στις 12 απαντήσεις
- 3 ακροατές απάντησαν σωστά 8 στις 12 απαντήσεις

Τα ίδια αποτελέσματα έχουμε και για το πέμπτο ζευγάρι. Κανένα υποκείμενο δεν κατάφερε να απορρίψει την μηδενική υπόθεση δηλαδή ότι οι μηχανές ήχου (audio engines) των συγκεκριμένων DAW που μελετάμε δεν παρουσιάζουν διαφορές στο ηχητικό αποτέλεσμα έτσι τα δείγματα που συγκρίνουμε δεν έχουν διαφορές.

6.2.7 Έκτο ζευγάρι: Logic-Pro tools (με plug-ins)

p-value (%) (πιθανότητα να δόθηκαν οι απαντήσεις στην τύχη)	Αριθμός σωστών απαντήσεων	Αριθμός ακροατών που απάντησαν
100%	1/12	-
99.7%	2/12	-
98.1%	3/12	1
92.7%	4/12	3
80.6%	5/12	1
61.3%	6/12	6
38.7%	7/12	5
19.4%	8/12	3
7.3%	9/12	1
1.9%	10/12	-
0.3%	11/12	-
0.0%	12/12	-
Σύνολο (N)		N=20

Πίνακας 6.6 Ο αριθμός των σωστών απαντήσεων που έδωσαν οι ακροατές για το ζευγάρι Logic-Pro tools (με plug-ins).

Στον παραπάνω πίνακα βλέπουμε ότι

- 1 ακροατής απάντησε σωστά 3 στις 12 απαντήσεις
- 3 ακροατές απάντησαν σωστά 4 στις 12 απαντήσεις

- 1 ακροατής απάντησε σωστά 5 στις 12 απαντήσεις
- 6 ακροατές απάντησαν σωστά 6 στις 12 απαντήσεις
- 5 ακροατές απάντησαν σωστά 7 στις 12 απαντήσεις
- 3 ακροατές απάντησαν σωστά 8 στις 12 απαντήσεις
- 1 ακροατής απάντησε σωστά 9 στις 12 απαντήσεις

Στο τελευταίο ζευγάρι βλέπουμε τέλος κανένας να μην μπορεί να απορρίψει την μηδενική υπόθεση αλλά βλέπουμε ένα μικρό ποσοστό των ακροατών να δίνει 8 και 9 στις 12 σωστές απαντήσεις. Δεν έπιασαν το ποσοστό των 10 σωστών αλλά αυτό μπορεί να μας πληροφορεί ότι ίσως σε αυτό το ζευγάρι με περισσότερες επαναλήψεις οι ακροατές να κατάφεραν να απορρίψουν την μηδενική υπόθεση και να δείξουν ότι καταλαβαίνουν διαφορές.

Κεφάλαιο 7^ο: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΙ ΤΗΣ ΕΡΕΥΝΑΣ

Εισαγωγή 7^{ου} Κεφαλαίου

Στο κεφάλαιο αυτό, απαντώνται τα ερευνητικά ερωτήματα που τέθηκαν ερμηνεύοντας τα αποτελέσματα που παρουσιάστηκαν στο προηγούμενο κεφάλαιο και εξάγονται τα συμπεράσματα αν τελικά έχουν διαφορές οι μηχανές ήχου των DAW.

Πρέπει να σημειωθεί ότι τα συμπεράσματα της συγκεκριμένης έρευνας είναι δύσκολο να γενικευτούν για ολόκληρο πληθυσμό, μιας και υπόκεινται σε μια σειρά περιορισμών όπως αυτοί θα αναλυθούν παρακάτω.

7.1 Συμπεράσματα

Τα συμπεράσματα που βγαίνουν από την συγκεκριμένη πειραματική διαδικασία είναι αρκετά αλλά πολύ συγκεκριμένα. Σκοπός μας ήταν να προσπαθήσουμε να εμβαθύνουμε μέσα στην φιλοσοφία των DAW, να αναλύσουμε την δομή τους και στην συνέχεια να δούμε αν προκύπτουν διαφορές στις μηχανές ήχου τους και γενικά στην άθροισή τους και στο ηχητικό αποτέλεσμα. Τέλος ελέγξαμε αν το ανθρώπινο σύστημα ακοής είναι σε θέση να καταλάβει τις όποιες διαφορές προκύπτουν. Σε αυτό μας βοήθησε η πειραματική διαδικασία της πτυχιακής εργασίας που χωρίστηκε σε δύο μέρη:

- στο *αντικειμενικό κομμάτι* όπου με την βοήθεια προγραμμάτων edit (audacity-wavelab) και με την χρήση null test αλλά και global analysis μπορέσαμε αντικειμενικά να αναλύσουμε και να δούμε τις όποιες διαφορές προκύπτουν ανάμεσα στα exports από τα DAW που εξετάσαμε (Logic, Cubase, Pro tools)
- και στο *υποκειμενικό κομμάτι* όπου με listening tests και συγκεκριμένα με την μεθοδολογία abx double blind test ελέγξαμε αν ο άνθρωπος (ακροατής) είναι σε θέση να αντιληφθεί αυτές τις διαφορές που προκύπτουν.

7.1.1 Αντικειμενικό κομμάτι

Οι κατηγορίες των δειγμάτων ήταν τρεις: Η πρώτη κατηγορία ήταν τα export από την πολυκάναλη παραγωγή των Radiohead, Nude με και χωρίς plug-ins. Η δεύτερη κατηγορία ήταν export ενός ημιτόνου από τα 3 DAW και η τρίτη κατηγορία ήταν κενά exports.

- Στην πρώτη κατηγορία των δειγμάτων από την παραγωγή των Radiohead τα πράγματα είναι πολύ ξεκάθαρα. Βλέπουμε ότι στην σύγκριση των ζευγαριών με χρήση null test (Κεφάλαιο 6, §6.1), τα ζευγάρια χωρίς την χρήση plug-ins αλληλοακυρώθηκαν δηλαδή το αποτέλεσμα της κυματομορφής ήταν μηδενικό.

Αυτό μας οδηγεί στο συμπέρασμα ότι η απλή άθροιση σημάτων στα DAW γίνεται σε όλα με τον ίδιο τρόπο και δεν παρουσιάζουν καμία ηχητική διαφορά. Στα ζευγάρια με χρήση plug-ins προέκυψαν διαφορές, το αποτέλεσμα από τα null test δεν ήταν μηδενικό, υπήρχε ξεκάθαρο οπτικό και ηχητικό αποτέλεσμα. Το συμπέρασμα που καταλήγουμε είναι ότι τα DAW και πιο συγκεκριμένα τα Cubase, Logic, Pro tools που εξετάσαμε παρουσιάζουν διαφορές στο τρόπο που επεξεργάζονται και χρησιμοποιούν τα plug-ins στην real time επεξεργασία άρα και το ηχητικό αποτέλεσμα της άθροισης είναι διαφορετικό. Είδαμε ότι με ακριβώς το ίδιο plug-in, το ίδιο preset προέκυψαν διαφορές από DAW σε DAW. Αυτό φαίνεται από το 3^ο σήμα στις εικόνες 6.4, 6.5 και 6.6. Βέβαια όπως καταλαβαίνουμε από τις εικόνες το Logic με το Cubase διαφέρουν ελάχιστα στην επεξεργασία των plug-ins σε σχέση με τα Pro tools.

- Στην δεύτερη κατηγορία με την χρήση ημιτόνου χωρίς την προσθήκη plug-ins δεν υπάρχουν ηχητικές διαφορές οι μηχανές των DAW συμπεριφέρονται με τον ίδιο τρόπο και το ηχητικό αποτέλεσμα είναι το ίδιο.
- Τέλος στην τρίτη κατηγορία με την global analysis κενών exports από Cubase, Logic, Pro tools τα δεδομένα που εξάγονται είναι τα ίδια δεν έχουμε δηλαδή ηχητικές διαφορές.

7.1.2 Υποκειμενικό κομμάτι (ακροάσεις)

Με την χρήση της μεθοδολογίας abx double blind test και με την βοήθεια της λειτουργίας abx comparator του foobar2000 player τα συμπεράσματα που εξάγονται είναι απόλυτα και ξεκάθαρα.

- Χρησιμοποιώντας τα 6 ζευγάρια δειγμάτων που έχουμε αναφέρει αναλυτικά ποια είναι στα κεφάλαια 5-6 και σε δείγμα 20 ακροατών μόνο 1 στους 20 ακροατές και σε μόνο ένα ζευγάρι από τα 6 (συγκεκριμένα Logic-Pro tools χωρίς plug-ins) κατάφερε να απορρίψει την μηδενική υπόθεση ότι “οι μηχανές ήχου (audio engines) των συγκεκριμένων DAW που μελετάμε δεν παρουσιάζουν διαφορές στο ηχητικό αποτέλεσμα έτσι τα δείγματα που συγκρίνουμε δεν έχουν διαφορές”. Ο συγκεκριμένος ακροατής στα υπόλοιπα ζευγάρια και με plug-ins και χωρίς plug-ins δεν κατάφερε να πιάσει το όριο των σωστών απαντήσεων για να καταρρίψει την μηδενική υπόθεση. Από όλα αυτά συμπεραίνουμε ότι υποκειμενικά ο άνθρωπος δεν μπορεί να αντιληφθεί τις όποιες ηχητικές διαφορές προκύπτουν από τις μηχανές ήχου των DAW. Αναλυτικότερα είδαμε ότι αντικειμενικά τα δείγματα χωρίς plug-ins δεν έχουν ηχητικές διαφορές. Άρα η προσπάθεια δημιουργίας διαφορών από το ανθρώπινο αυτί είναι αποτέλεσμα ψυχοακουστικών παραγόντων. Στην συνέχεια στα δείγματα με plug-ins εκεί που αντικειμενικά

δημιουργούνται διαφορές ναι μεν το ποσοστό των σωστών απαντήσεων ήταν μεγαλύτερο αλλά με βάση την μεθοδολογία abx double blind test και της στατιστικής της ανάλυσης κανένας δεν μπόρεσε να απορρίψει την μηδενική υπόθεση άρα σαν αποτέλεσμα κανένας δεν μπόρεσε να δείξει ότι υπάρχουν ηχητικές διαφορές ανάμεσα στα DAW.

7.2 Περιορισμοί της έρευνας

Στην συνέχεια παρατίθενται ορισμένοι περιορισμοί που αφορούν στον τρόπο διεξαγωγής της πειραματικής διαδικασίας.

- Όπως αναφέρθηκε και σε προηγούμενο κεφάλαιο, το δείγμα που επιλέχτηκε για τη διεξαγωγή των listening tests, δεν ήταν τυχαίο, αλλά μη πιθανοτικό. Αυτό σημαίνει, πως το δείγμα δεν επιλέγεται τυχαία, αλλά με βάση ορισμένα κριτήρια που θέτει ο ερευνητής (τεχνικές-μουσικές γνώσεις, ηλικία) και συνεπώς τα συμπεράσματα της έρευνας δεν μπορούν να γενικευτούν σε ολόκληρο τον πληθυσμό.
- Το μέγεθος του δείγματος (20 άτομα) και των επαναλήψεων (12 επαναλήψεις), είναι μεν επαρκές για έρευνα με την συγκεκριμένη μεθοδολογία, αλλά είναι και οριακό (σύμφωνα με την μεθοδολογία το δείγμα πρέπει να είναι $N \geq 20$ και οι επαναλήψεις πάνω από 10). Πιθανώς κάποια αποτελέσματα να μην είναι τόσο σωστά και να υπάρχουν σφάλματα.
- Ο χώρος στον οποίο διεξήχθη η πειραματική διαδικασία διαθέτει ηχομόνωση και σχετικά χαμηλή στάθμη θορύβου, όχι όμως σε επίπεδο που να μπορεί να εξασφαλιστεί η απολυτή σιγή. Πιθανώς τα υποκείμενα να έδωσαν εσφαλμένες απαντήσεις οι οποίες να προήλθαν από κάποιο εξωγενή θόρυβο.
- Η μέση διάρκεια συμπληρώσεως των ερωτηματολογίων όπως αναφέρθηκε και σε προηγούμενο κεφάλαιο ήταν 90 λεπτά. Πιθανώς ορισμένα υποκείμενα να κουραστήκαν από μια τόσο μακροσκελή ακρόαση και να μην έδωσαν προσεκτικές απαντήσεις.

7.3 Γενικό συμπέρασμα υποκειμενικό πτυχιακής εργασίας

Για να είμαστε αντικειμενικοί, όλοι όσοι λένε πως διαφορετικά DAW έχουν διαφορετικό αποτέλεσμα δεν εννοούν τελικά ακριβώς ως σύγκριση των μηχανών ήχου των DAW. Και σε τέτοια περίπτωση, μπορεί να έχουν δίκιο. Συγκεκριμένα σήμερα είναι πολλοί εκείνοι που γράφουν μουσική μόνο με εικονικά - software όργανα και εφέ. Για αυτούς το DAW δεν

είναι μόνο το sequencing και mixing των audio αρχείων. Όταν λένε DAW εννοούν και τα επί μέρους instruments και effects που το απαρτίζουν. Έτσι, είναι προφανές ότι κάθε DAW θα έχει διαφορετικά αποτελέσματα αφού το κάθε DAW έχει διαφορετικά synths, samplers, EQs, compressors, reverbs, delays μέχρι τους πιο υποκειμενικούς και ψυχολογικούς όπως το γενικότερο workflow του προγράμματος ή ακόμα και κάτι τόσο επιφανειακό όπως απλά η "όψη" του, τα χρώματά του, η οπτική δομή του που μπορεί να εμπνέουν έναν χρήστη να δουλέψει με αυτό ή αντίθετα να τον αποθαρρύνουν.

Και πάλι όμως όλα αυτά είναι υποκειμενικά και δε θα μπορούσαν να σταθούν ως κριτήρια για την μηχανή ήχου ενός DAW.

Έν κατακλείδι, αυτό που πρέπει να μας μείνει είναι πως τα DAW είναι "διάφανα" (transparent) και δεν έχουν καμία σχέση με hardware υλοποιήσεις εγγραφής όπου η ποιότητα κατασκευής του μέσου επηρέαζε, είτε θετικά είτε αρνητικά, το τελικό ηχητικό αποτέλεσμα.

Σαν γενικό συμπέρασμα μπορούμε να πούμε ότι τα software DAW το μόνο που κάνουν όταν τους εισάγουμε audio αρχεία είναι να πραγματοποιούν μαθηματικές πράξεις (0 και 1) και να βγάζουν ένα τελικό αποτέλεσμα που εξαρτάται καθαρά και μόνο από το τι του "εισάγαμε". Υπό αυτήν την έννοια, όχι απλά δεν έχουν διαφορές στον ήχο τους, αλλά δεν έχουν χρωματίζουν καν τον ήχο. Οι όποιες ηχητικές διαφορές προκύπτουν από την real-time επεξεργασία των plugins και των format που χρησιμοποιούν τα DAW γι αυτά, όπως είδαμε και στο αντικειμενικό κομμάτι του πειραματικού μέρους και στην θεωρία.

Από την άλλη μεριά όμως, αν διαπιστώνουμε ότι με κάποιο DAW μπορούμε να πετύχουμε "καλύτερο" ηχητικό αποτέλεσμα, απλά συνεχίζουμε να χρησιμοποιούμε εκείνο το DAW που μας βοηθά να φτάσουμε στο επιθυμητό αποτέλεσμα πιο εύκολα και πιο γρήγορα. Απλά πρέπει να θυμόμαστε να το εκφράζουμε σωστά διότι δεν έχει το X DAW καλύτερο ήχο από το Y, εμείς μπορούμε να βγάλουμε καλύτερο ήχο στο X DAW από ότι στο Y.

7.4 Περαιτέρω προτάσεις της έρευνας

Όπως είδαμε και στο Κεφάλαιο 3 και πιο συγκεκριμένα στην παράγραφο 3.2.1 για fixed και floating point συστήματα είδαμε ότι προκύπτουν διαφορές στο πως επεξεργάζονται τα δεδομένα. Μία πρόταση για έρευνα είναι η εκτενής ανάλυση αυτών των δύο κατηγοριών της ψηφιακής επεξεργασίας του σήματος και τις διαφορές που επιφέρουν στα δεδομένα που επεξεργάζονται και ποιο είδος DSP επεξεργασίας είναι κατάλληλο για κάποιο είδος μουσικής και γιατί.

Βιβλιογραφία

Ελληνική

Παπαδάκης, Ν. (2009). Πανεπιστημιακές σημειώσεις για το μάθημα « Εισαγωγή στην Ψυχοακουστική». ΑΤΕΙ Κρήτης, Παράρτημα Ρεθύμνου, Τμήμα Μηχανικών Μουσικής Τεχνολογίας και Ακουστικής.

Κεφαλογιάννης, Ν. (2007). Πανεπιστημιακές σημειώσεις για το μάθημα « Πολυκάναλα συστήματα Ψηφιακής Επεξεργασίας Ήχου». ΑΤΕΙ Κρήτης, Παράρτημα Ρεθύμνου, Τμήμα Μηχανικών Μουσικής Τεχνολογίας και Ακουστικής.

Τραγανίτης, Α. (2004). Πανεπιστημιακές σημειώσεις για το μάθημα « Ψηφιακές επικοινωνίες». Πανεπιστήμιο Κρήτης, Τμήμα Επιστήμης Υπολογιστών.

Ξενόγλωσση

Aviva Petrie; Paul Watson (28 February 2013). [Statistics for Veterinary and Animal Science](#). Wiley. pp. 130–131. [ISBN 978-1-118-56740-1](#)

David Clark, 1982 High-Resolution Subjective Testing Using a Double-Blind Comparator, ABX Company, p.333

Gannon Kashiwa. The Pro Tools 48bit Mixer. Technical White Paper. 2-6

Moore, David; McCabe, George (2003). Introduction to the Practice of Statistics (4 ed.). New York: W.H. Freeman and Co. p. 438. [ISBN 9780716796572](#).

Meier, Kenneth J.; Brudney, Jeffrey L.; Bohte, John (2011). Applied Statistics for Public and Nonprofit Administration (3rd ed.). Boston, MA: Cengage Learning. pp. 189–209. [ISBN 1-111-34280-6](#).

Devore, Jay L. (2011). Probability and Statistics for Engineering and the Sciences (8th ed.). Boston, MA: Cengage Learning. pp. 300–344.

Healy, Joseph F. (2009). The Essentials of Statistics: A Tool for Social Research (2nd ed.). Belmont, CA: Cengage Learning. pp. 177–205. [ISBN 0-495-60143-8](#).

Leonard, Brett; Buttner-Schnirer, Padraig (2012). Subjective Differences in Digital Audio Workstation Math. McGill University

Recommendation sector of UTI (2014). Methods for the subjective assessment of small impairments in audio systems. 4

Chung-Hsien Wu, Yi-Chin Huang, Chung-Han Lee, and Jun-Cheng Guo (2014). Synthesis of Spontaneous Speech With Syllable Contraction Using State-Based Context-Dependent Voice Transformation. IEEE/ACM. 593-594.

Takeda, S et al (1992). Age variation in the upper limit of hearing. European journal of applied physiology. 65, 403-408

Manuj Yadav, Densil Cabrera, Ralph Collins and William L. Martens (2011). Detection of headtracking in room acoustic simulations for one's own voice. University of Sydney, 3

Schatzoff, Martin (1981), "Design of Experiments in Computer Performance Evaluation", IBM Journal of Research and Development, Vol. 25 No. 6,

William H. Beyer. (1968). Handbook of Tables for Probability and Statistics [Second 2nd Edition]

Pohlmann, K. C. (2011). Principles of digital audio. 6th edition. 28-33. 103-106. 342.

John Watkinson (2002). Introduction to Digital Audio, Second Edition. 109-122.

Lars Eriksson (1989). Algorithms for Automatic Segmentation of Speech. Lund University, Dept. of Linguistics. Working Papers 35. 53

ΠΑΡΑΡΤΗΜΑ Α

Οδηγίες χρήσης για το foobar2000 (ABX double blind test)

Καλώς ήρθατε στην πειραματική διαδικασία ακρόασης της πτυχιακής εργασίας με τίτλο «Σύγκριση Audio Engines Digital Audio Workstations (D.A.W) με χρήση listening tests» (μέθοδος ABX). Σκοπός της συγκεκριμένης εργασίας είναι να δούμε αν όντως παρατηρούνται διαφορές στο ήχο κάθε audio engine ενός DAW (pro tools, logic, cubase) και αυτό θα το πετύχουμε κάνοντας export μία συγκεκριμένη πολυκάναλη παραγωγή σε 3 διαφορετικά DAW στα pro tools, στο logic και στο cubase και χρησιμοποιώντας την μέθοδο ABX double blind test θα μπορέσουμε να δούμε αν όντως μπορούμε να καταλάβουμε διαφορές στον ήχο με αποτέλεσμα να επιλέξουμε την σωστή απάντηση.

Σ' ένα ABX listening test ο ακροατής έχει πρόσβαση σε τρεις πηγές που ονομάζονται A, B, και X. A και B είναι οι αναφορές. Για παράδειγμα, A το αρχείο από το logic και B το αρχείο από τα pro tools. X είναι η άγνωστη πηγή. Μπορεί να είναι ή A ή B. Ο ακροατής πρέπει να μαντέψει ποιά από τις δύο αναφορές (A ή B) αναπαράγεται κάθε φορά στην X επανάληψη.

Στην συγκεκριμένη πειραματική διαδικασία έχουν δημιουργηθεί δύο κατηγορίες δειγμάτων από τρία ζευγάρια η κάθε μια. Η πρώτη κατηγορία περιλαμβάνει τα export της πολυκάναλης παραγωγής χωρίς την προσθήκη plug ins σε κάθε κανάλι ενώ η δεύτερη περιλαμβάνει και ορισμένα plug ins σε κάθε κανάλι. Τα τρία ζευγάρια σε κάθε κατηγορία είναι **Pro tools με logic, Pro tools με cubase και logic με cubase** και αυτά θα χρησιμοποιήσουμε. Γενικά ένας αποδεκτός αριθμός επαναλήψεων για το κάθε ζευγάρι είναι πάνω από 10 στις συγκεκριμένες ακροάσεις θα χρησιμοποιήσουμε 12 επαναλήψεις. Δηλαδή για κάθε ζευγάρι θα πρέπει να απαντήσουμε στον άγνωστο X 12 φορές για να είναι τα αποτελέσματα μας αξιόπιστα.

Για να πραγματοποιηθούν οι συγκεκριμένες ακροάσεις και για να βγάλουμε σωστά και έγκυρα στατιστικά αποτελέσματα θα χρησιμοποιήσουμε το λογισμικό foobar2000.

Η λειτουργία του προγράμματος είναι απλή επιλέγουμε τις αναφορές μας πχ. A pro tools και B logic και συνεχίζουμε στην ακρόαση. Οι αναφορές μας καθ' όλη την διάρκεια των επαναλήψεων δεν αλλάζουν. Αυτό που αλλάζει είναι από κάτω το X και το Y. Αυτά αλλάζουν θέσεις κατά την διάρκεια των επαναλήψεων και εμείς πρέπει να αντιστοιχίσουμε αν το A είναι το X ή το B είναι το Y ή το αντίθετο αν το B πήγε στην θέση X και το A στην θέση Y. Αυτές είναι και οι δυο απαντήσεις που μία από τις δύο πρέπει να επιλέγουμε σε κάθε επανάληψη δηλαδή (**X is A, Y is B** ή **Y is A, X is B**). Να σημειώσουμε ότι τις αναφορές A και B μπορούμε να τις ακούμε εκείνη την ώρα μαζί με τα X και Y για να έχουμε κάποιο μέτρο σύγκρισης για να επιλέξουμε την σωστή απάντηση. Περισσότερες λεπτομέρειες και οδηγίες θα δοθούν κατά την διάρκεια των ακροάσεων.

ΠΑΡΑΡΤΗΜΑ Β

Αρχεία με αποτελέσματα από τις ακροάσεις της abx διαδικασίας.

