



Ψηφιακός Ήχος

Ενότητες

- Η φυσική του ήχου
- Ψηφιοποίηση & κωδικοποίηση ψηφιακού ήχου
- Αρχή ηχητικής σκίασης
- Απωλεστική συμπίεση ψηφιακού ήχου & το πρότυπο mp3
- Ιστορικά στοιχεία για το mp3 & άλλα πρότυπα
- Ήχος midi & υποσύστημα ήχου

Η Φυσική Του Ήχου

■ Ήχος

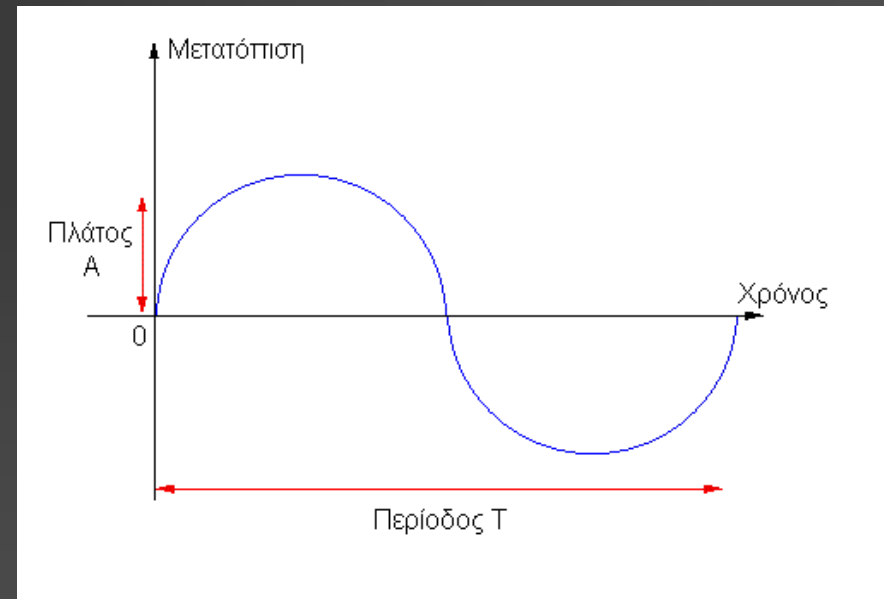
- Δόνηση των μορίων ενός μέσου
- Διάδοση ακουστικού κύματος στο μέσο όπου συμβαίνει η δόνηση.
- Μεταβολές πίεσης (πυκνώματα και αραιώματα) που μεταδίδονται με μια ορισμένη ταχύτητα

■ Συχνότητα (pitch)

- Δονήσεις (κύκλοι)/sec [Hz]
- Ανθρώπινη ακοή 20 – 20000 Hz

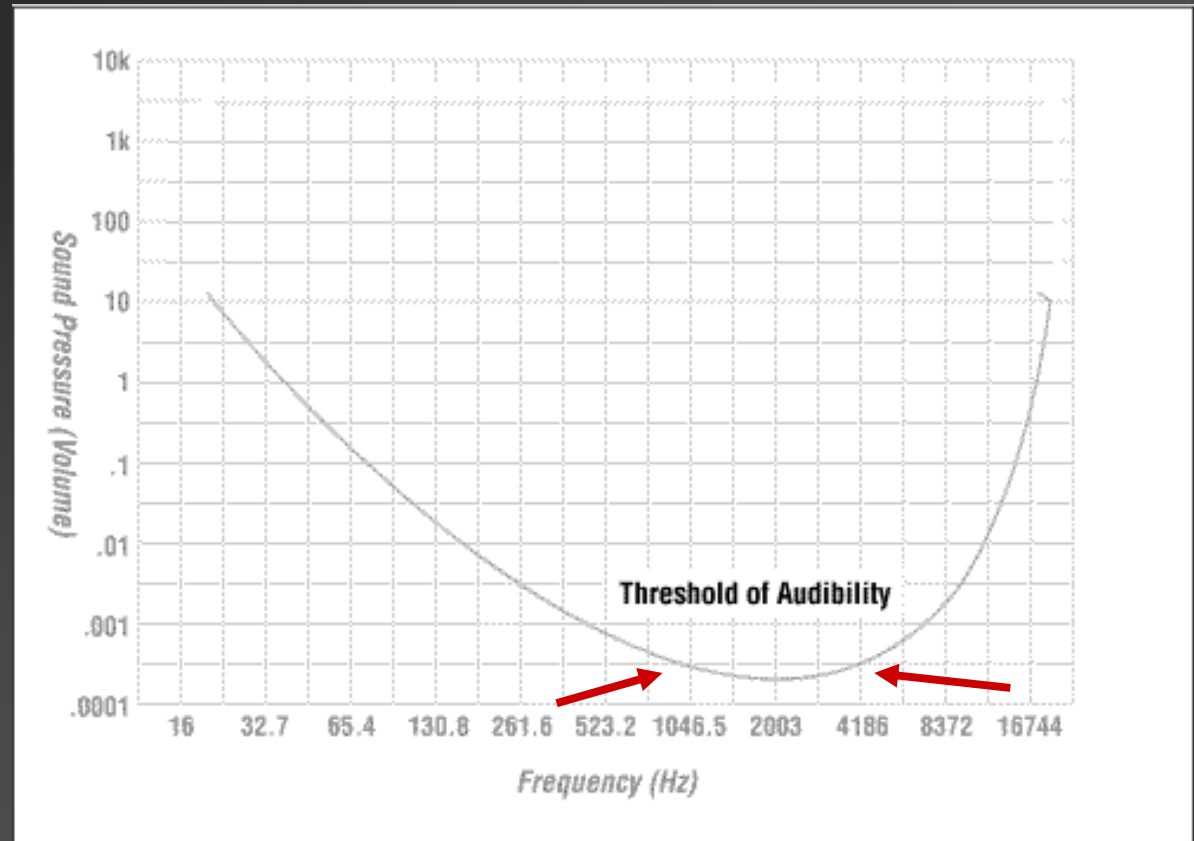
■ Ένταση

- Ισχύς της δόνησης
- Καθορίζεται από το πλάτος της δόνησης



Ευαισθησία της ανθρώπινης ακοής

- Ακουστ(ικ)ότητα
 - Audibility
- Μεγαλύτερη ευαισθησία στην περιοχή 1-4 kHz

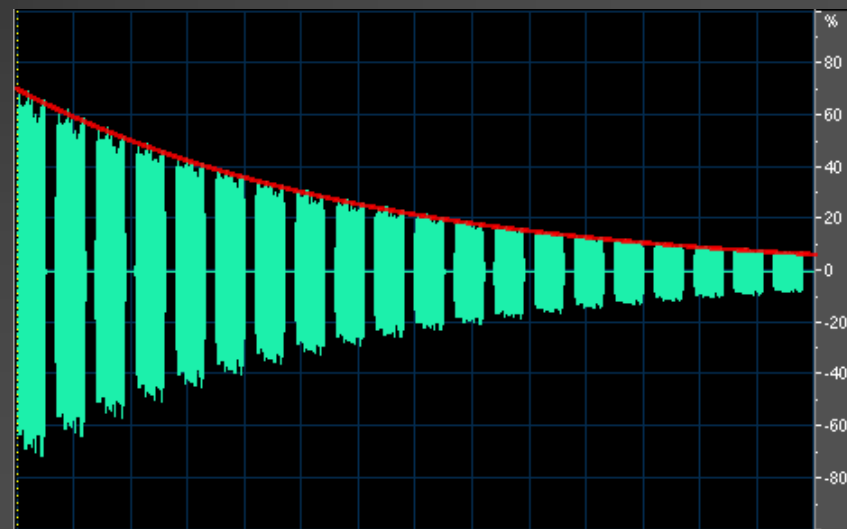
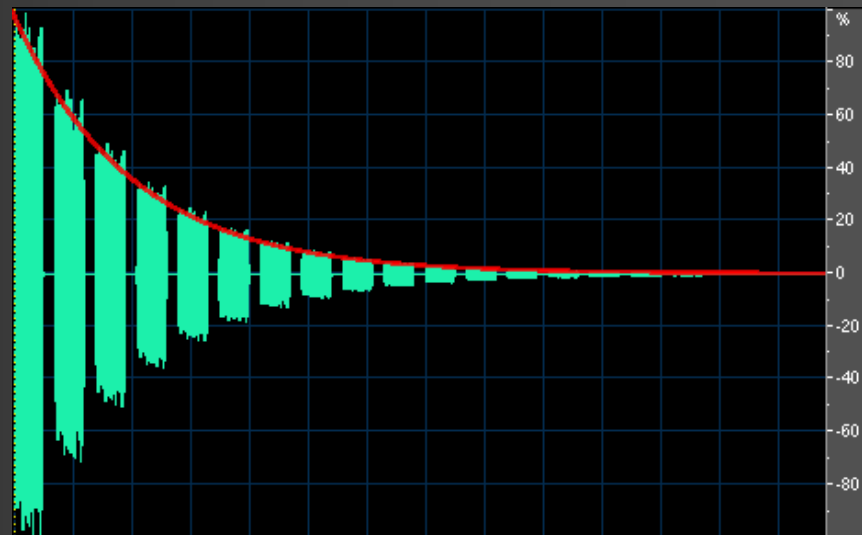
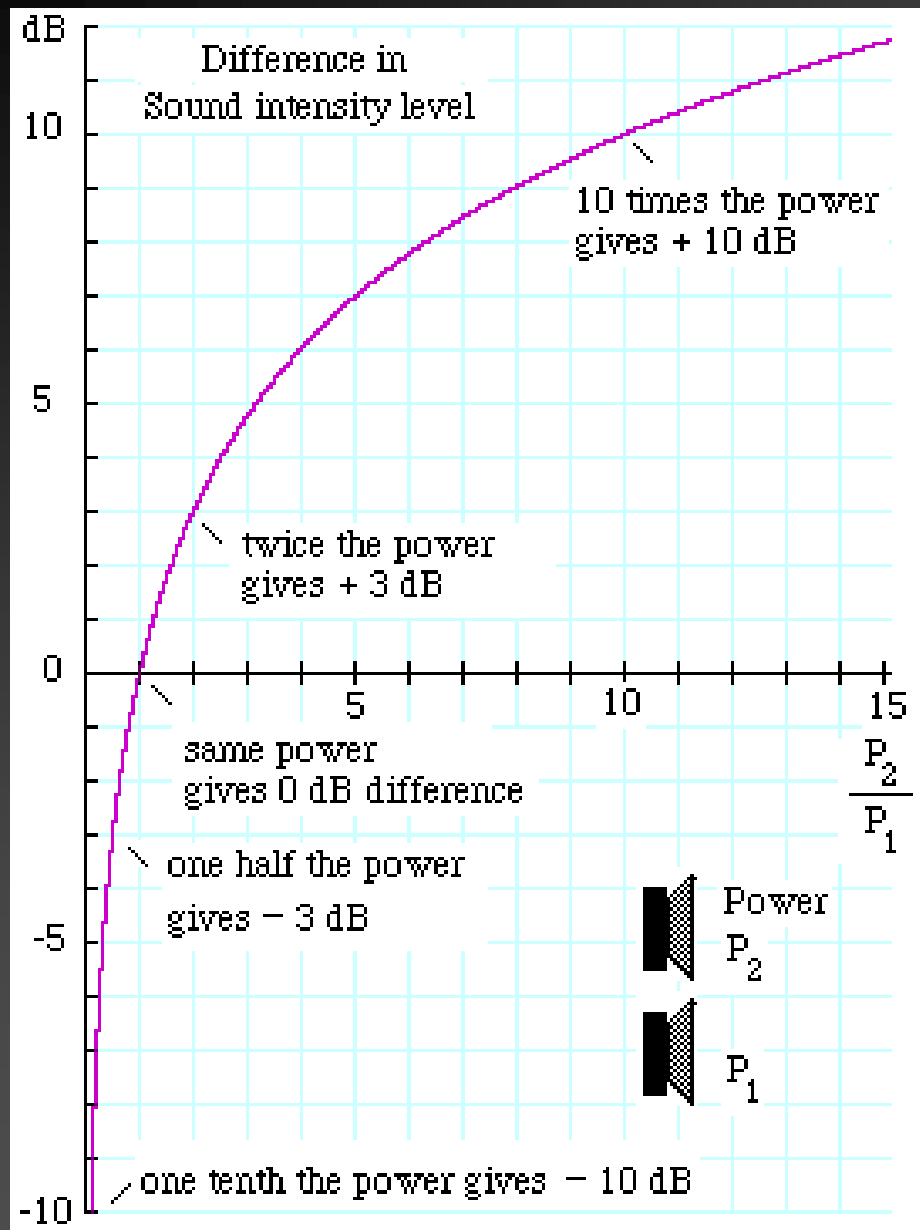


«Λογαριθμική» Ακοή

- Η ανθρώπινη ακοή είναι «**λογαριθμική**» & όχι γραμμική
- Η υποκειμενική αίσθηση μεταξύ των εντάσεων δύο ήχων δεν εξαρτάται από την απόλυτη τιμή της διαφοράς τους αλλά από το **λόγο** τους
- Πχ. αύξηση της έντασης από $2 \rightarrow 4 \mu\text{Watt/m}^2$ είναι το ίδιο με την αύξηση από $5 \rightarrow 10 \mu\text{Watt/m}^2$

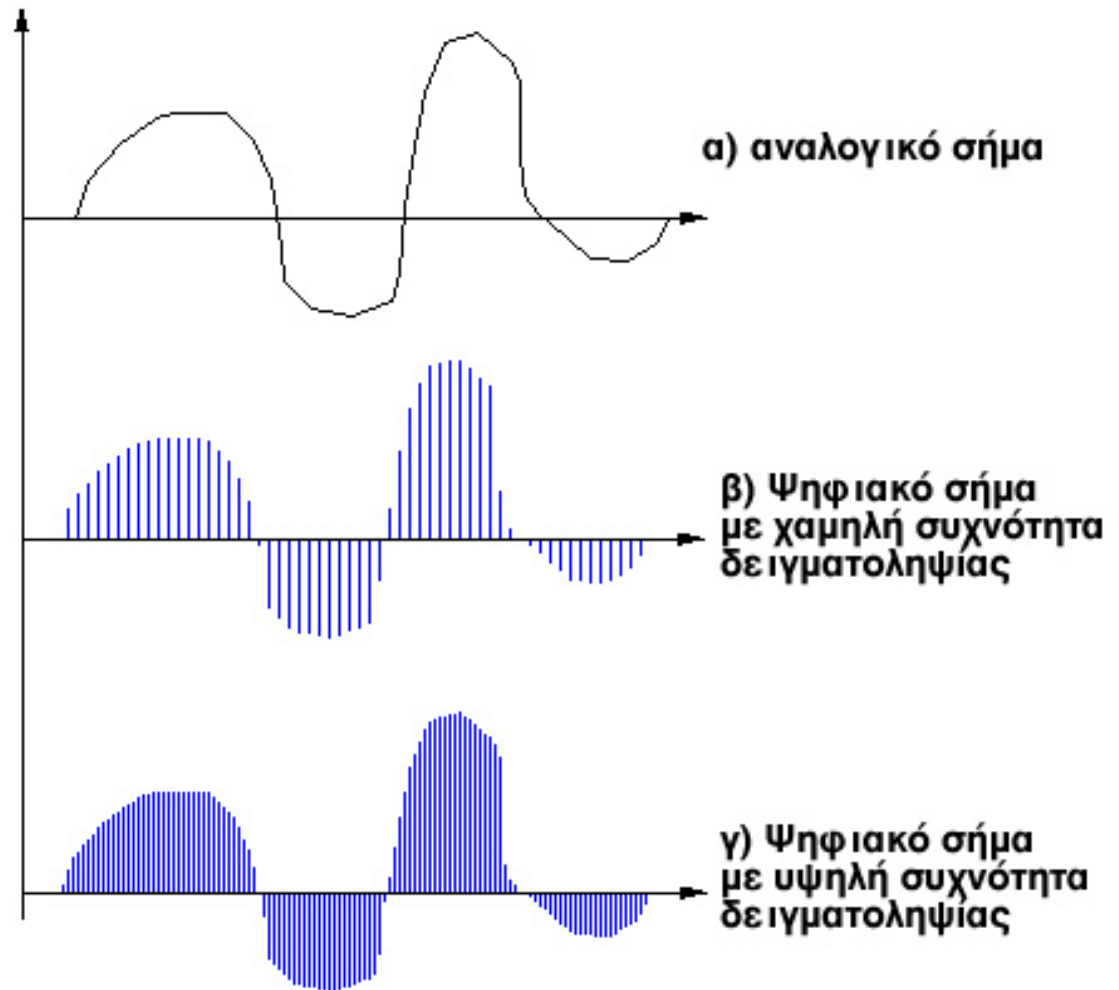
Μονάδα decibel (db)

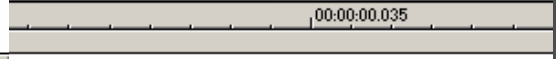
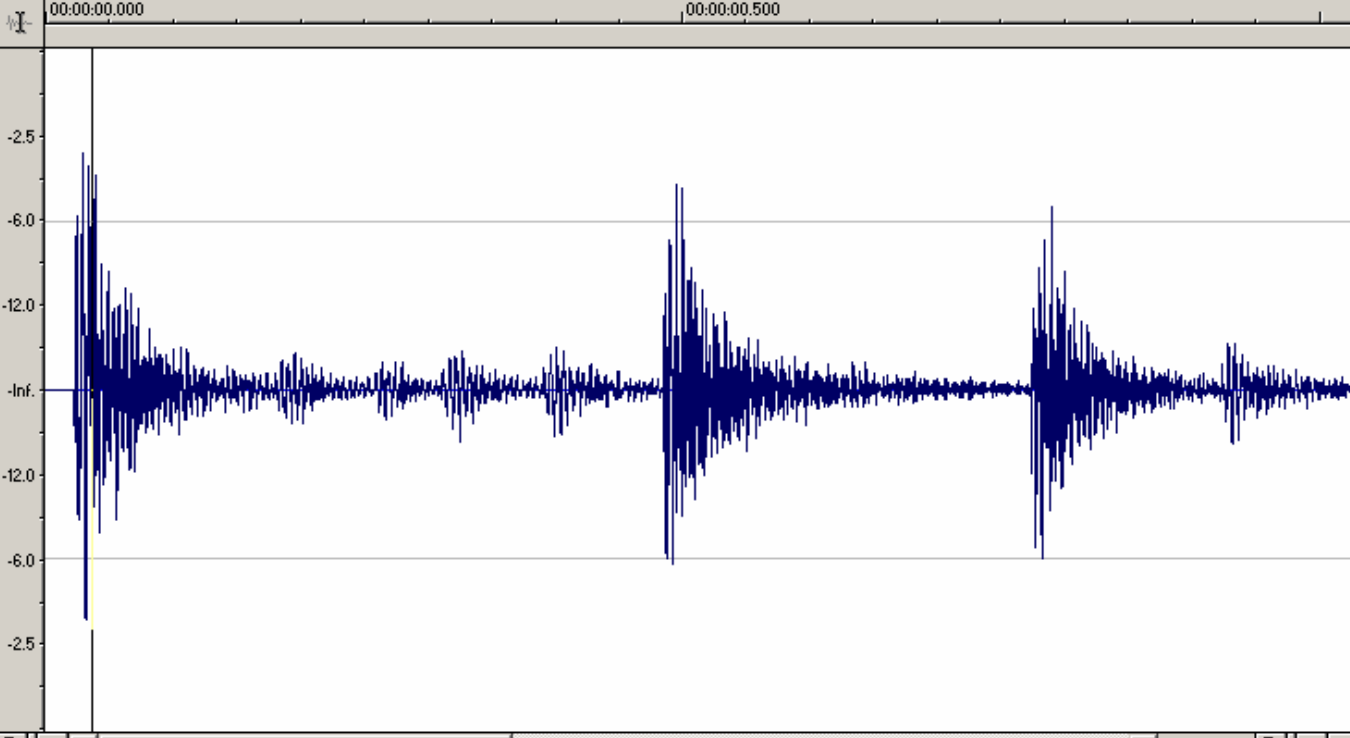
- Σχετική διαφορά ισχύος δύο ήχων = $10 \log (P/ P_0)$ σε db
 - P η τιμή ισχύος του ήχου
 - P_0 τιμή αναφοράς (ήχος αναφοράς)
- Το decibel είναι «σχετική» μονάδα → συγκρίνει το μετρούμενο μέγεθος (P) με μία **στάθμη αναφοράς** (P_0)
- Ίδια ισχύς $P = P_0 \rightarrow$ **Διαφορά = 0 db**
- Διπλασιασμός της ισχύος $P = 2 P_0$ σημαίνει αύξηση κατά **3 db**



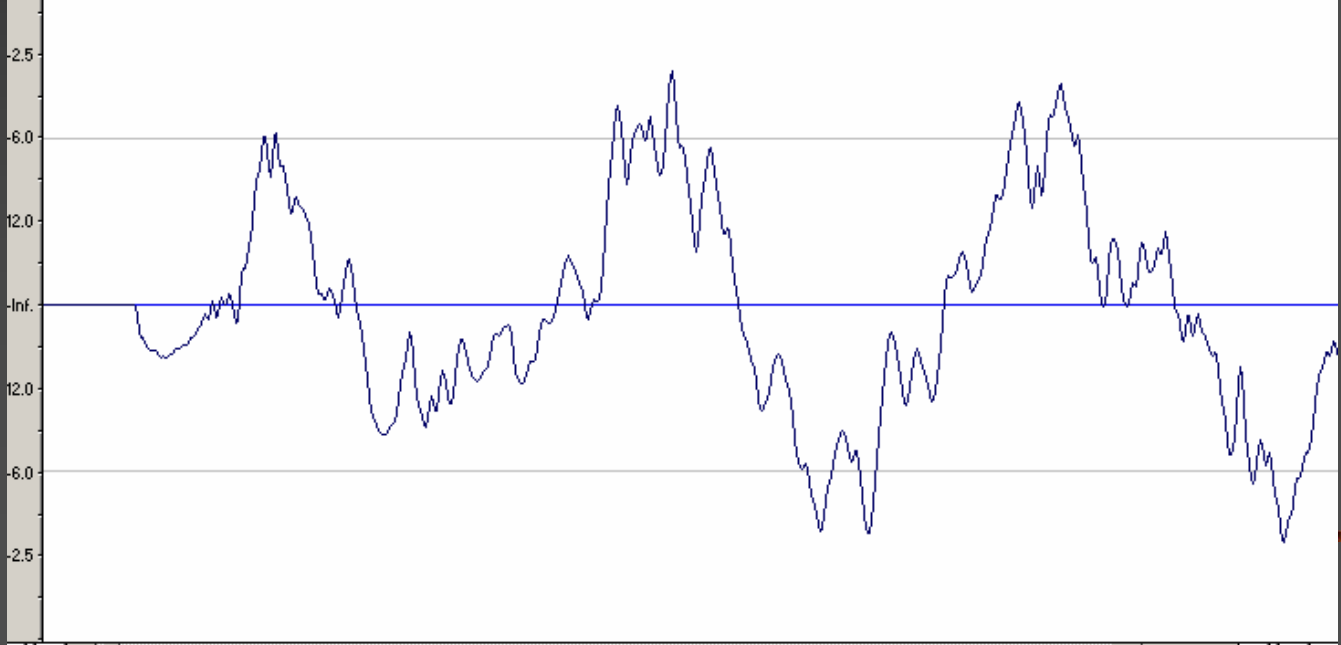
Ψηφιοποίηση Του Ήχου

- Δειγματοληψία
- Κβαντοποίηση
- Κωδικοποίηση





Q | Q | ◀ ▶ 1:64 Q | Q | ◀ ▶



Q | Q | ◀ ▶ 1:1 Q | Q | ◀ ▶

Μέγεθος αρχείου

- Μέγεθος Αρχείου (bits)=
 - Συχνότητα Δειγματοληψίας (Hz) x Μέγεθος δείγματος (bits) x Διάρκεια (sec)
 - Μέγεθος Αρχείου = $[44.100 \text{ (Hz)} \times 16 \text{ (bits)} \times 1 \text{ (sec)}] = 88.200$ byte για μονοφωνικό
 - Ή το διπλάσιο 176.4KB για στερεοφωνικό



Μέθοδοι Κωδικοποίησης

PCM

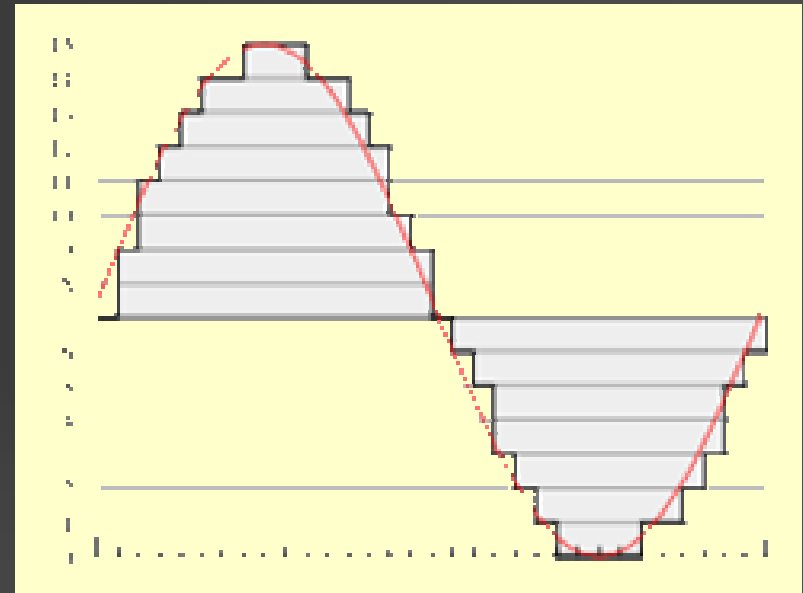
μ-Law & A-Law PCM

DPCM

ADPCM

PCM (Pulse Code Modulation - παλμοκωδική διαμόρφωση)

- Αποθηκεύει ένα προς ένα τα δείγματα σε ψηφιακή μορφή
- **Γραμμική** (ομοιόμορφη) κωδικοποίηση
 - το πλήθος των σταθμών κβάντωσης κατανέμεται εξίσου σε όλο το εύρος του πεδίου τιμών του σήματος
 - Δεν περιλαμβάνει συμπίεση



4-bit κωδικοποίηση
ημιτονοειδούς σήματος

DPCM (Differential Pulse Code Modulation – διαφορική παλμοκωδική κωδικοποίηση)

- Κωδικοποιεί τη **διαφορά** μεταξύ της πραγματικής τιμής του δείγματος και της προβλεπόμενης τιμής του δείγματος (predicted value)
 - Η πρόβλεψη βασίζεται στις τιμές των προηγούμενων δειγμάτων
- Η βασική ιδέα της DPCM (κωδικοποίηση της διαφοράς) είναι εφαρμόσιμη γιατί τα περισσότερα ηχητικά σήματα παρουσιάζουν σημαντικό **βαθμό συσχέτισης** (correlation) μεταξύ διαδοχικών δειγμάτων
 - Η DPCM εκμεταλλεύεται τον πλεονασμό στις τιμές των δειγμάτων και κωδικοποιεί με μικρότερο bit rate (4 bit / sample).

μυ-Law PCM και A-Law PCM

- Λογαριθμική αντιστοίχιση
- Αντιστοιχεί **περισσότερες** στάθμες κβάντωσης στις χαμηλές συχνότητες και **λιγότερες** στις υψηλές.
 - Στις χαμηλές συχνότητες περιέχεται το μεγαλύτερο ποσοστό της πληροφορίας που γίνεται **αντιληπτή** από το ανθρώπινο αυτί.
- Με λογαριθμική κωδικοποίηση μυ-Law 8-bit κωδικοποιείται το ίδιο δυναμικό εύρος ήχου όπως με 14-bit γραμμική PCM



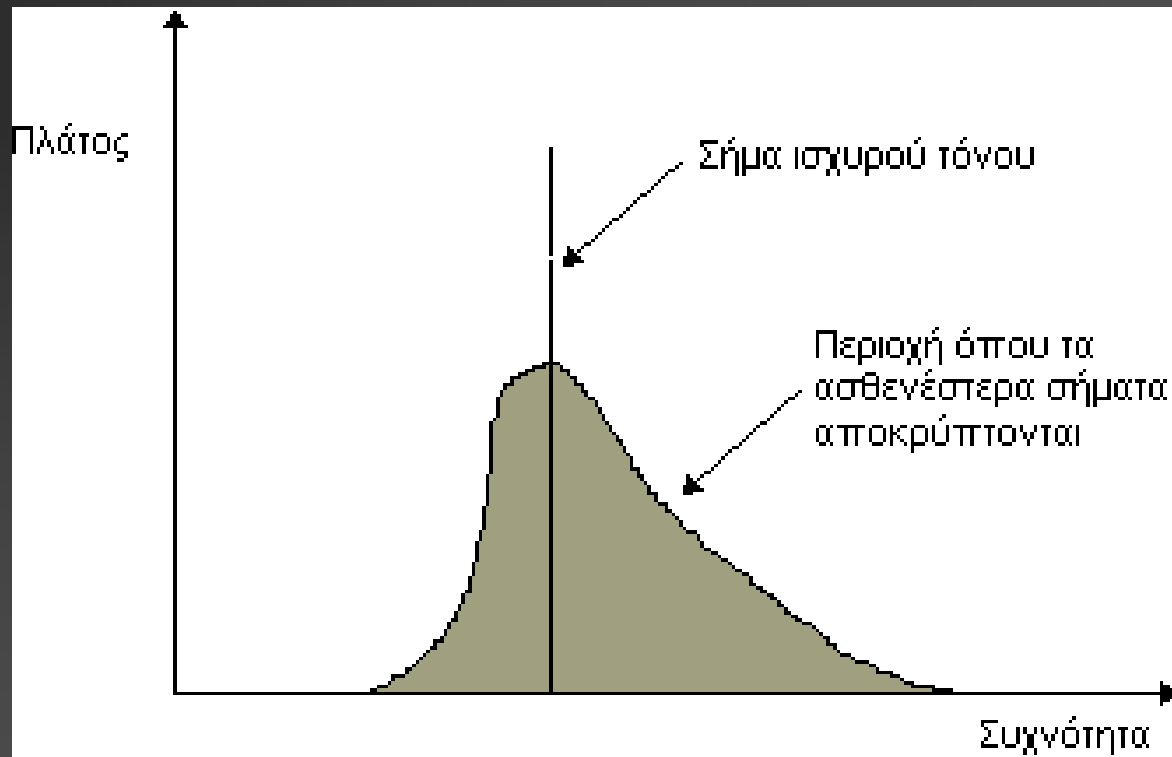
Η αρχή της ηχητικής σκίασης

Συμπύεση πληροφορίας ψηφιακού ήχου

- Κατά τη συμπίεση του ήχου χρησιμοποιούνται **πολύπλοκοι** αλγόριθμοι συμπίεσης που συνδυάζουν τόσο **απωλεστικές** όσο και **μη απωλεστικές** τεχνικές συμπίεσης
- Στις σύγχρονες μεθόδους συμπίεσης εφαρμόζονται διάφορα **ψυχο-ακουστικά** μοντέλα βάσει των οποίων **απορρίπτονται** οι συχνότητες που έτσι και αλλιώς δεν μπορούσαν να γίνουν αντιληπτές από το ανθρώπινο αυτί.

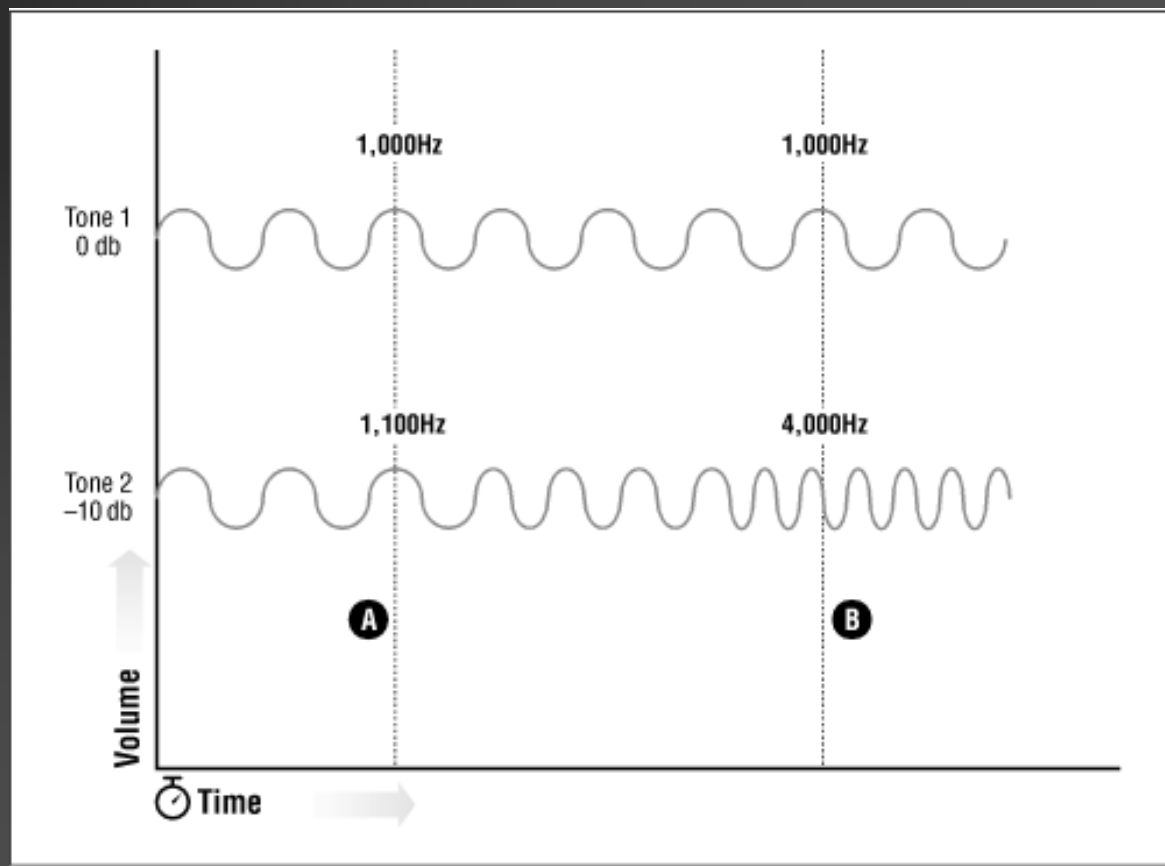
Ηχητική Σκίαση (auditory masking)

- Συχνότητες χαμηλής έντασης **κοντά** σε μία συχνότητα υψηλής έντασης (κυρίαρχη) δεν ακούγονται
- Το κατώφλι ακουστότητας μιας συχνότητας **ποικίλει** ανάλογα με τη συχνότητα και τον ακροατή



Ηχητική Σκίαση συχνοτήτων

- **Ισχυρή** συχνότητα
1000 Hz
- **Ασθενής** συχνότητα
(-10 db) 1100 Hz
- Δεν γίνεται αντιληπτή
παρά μόνον αν η
συχνότητα αυξηθεί
σημαντικά πχ. 4000
Hz



Ηχητική σκίαση στο χρόνο (temporal masking)

- Ταυτόχρονη αναπαραγωγή ισχυρής και ασθενούς συχνότητας → η ασθενής δεν γίνεται αντιληπτή
- Η ασθενής μπορεί να γίνει αντιληπτή αν «καθυστερήσει» σε σχέση με την ισχυρή
- Για καθαρούς τόνους (συχνότητες) η χρονική διάρκεια της διαφοράς αναπαραγωγής είναι ~ 5 msec

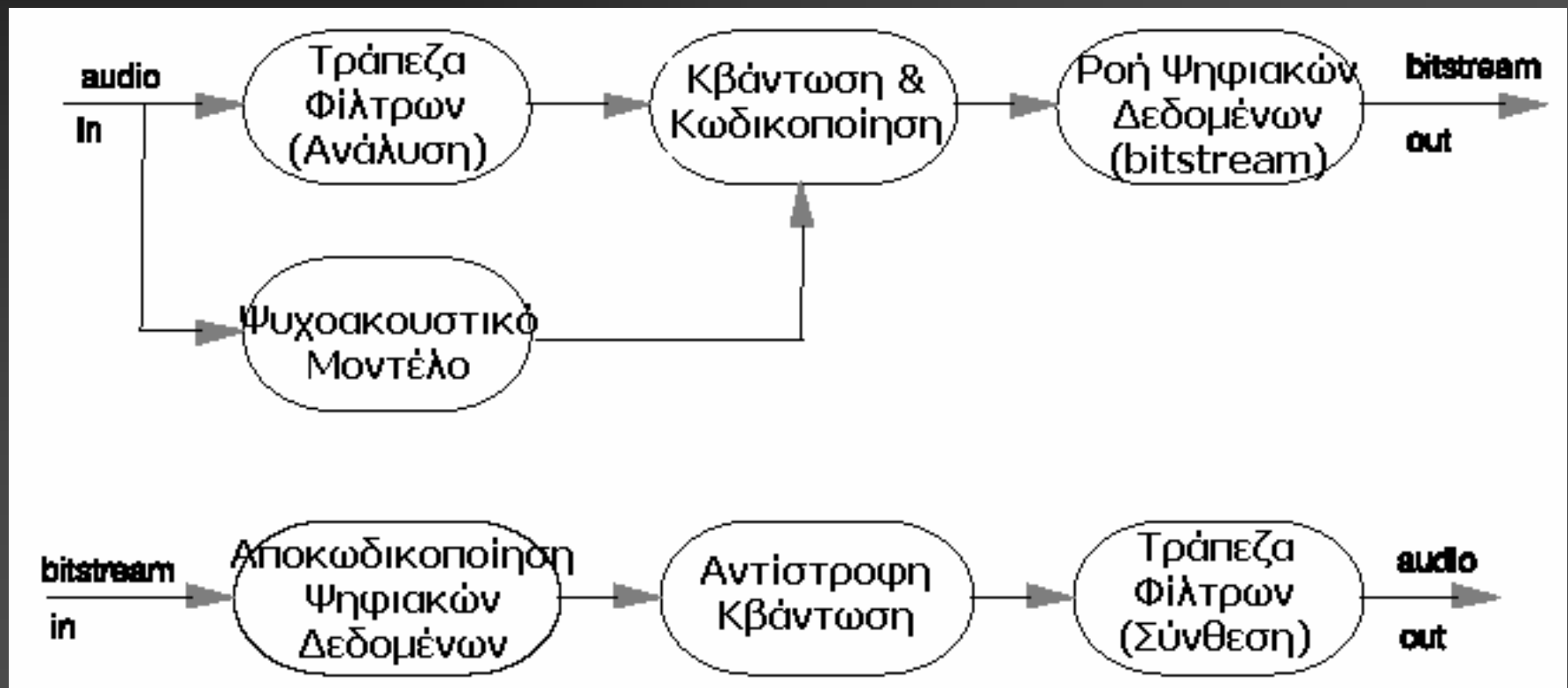


Απωλεστική Συμπίεση Ήχου & Συμπίεση κατά mp3

MPEG

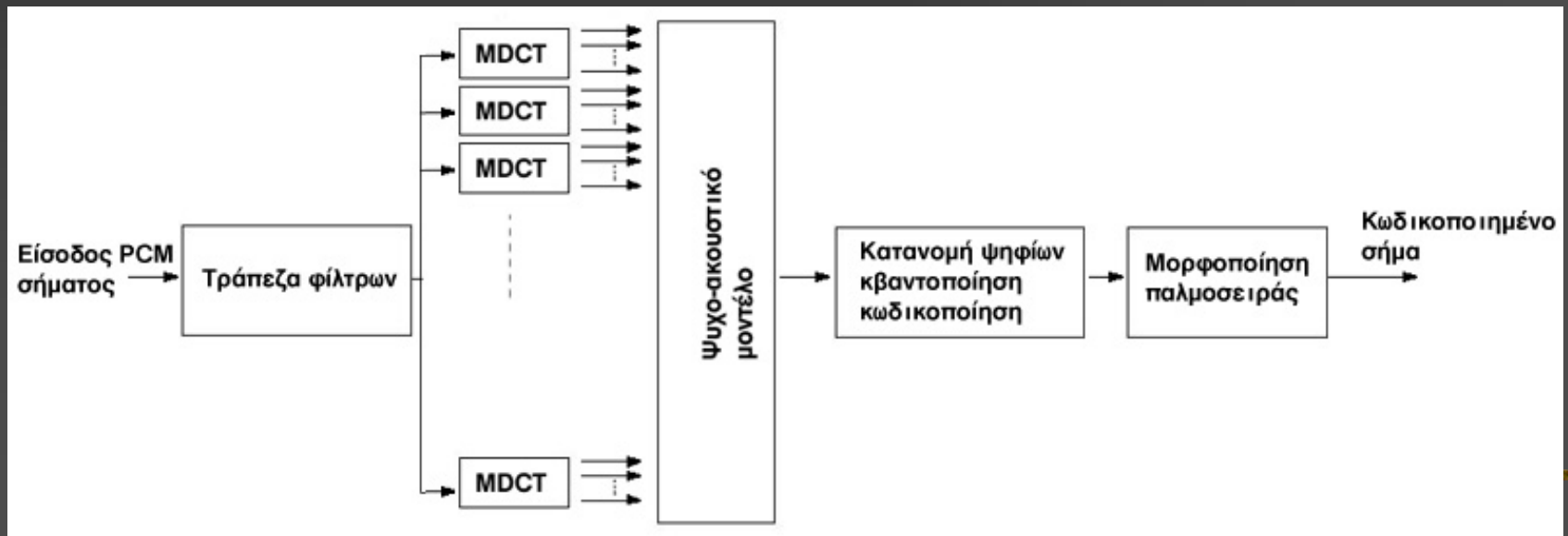
- MPEG (Motion Picture Expert Group) / 1987
 - **Οικογένεια** συμπιεστών (Codecs)
 - Ανάπτυξη: Συνεργασία πανεπιστημίων, ερευνητικών ινστιτούτων και εταιρειών
 - Υπό την αιγίδα του Διεθνούς Οργανισμού Προτύπων (International Standards Organization (ISO))
 - Αναπτύσσει διεθνή πρότυπα για τη συμπίεση εικόνας, ήχου και βίντεο.
-

Βασική δομή ενός ηχητικού κωδικοποιητή



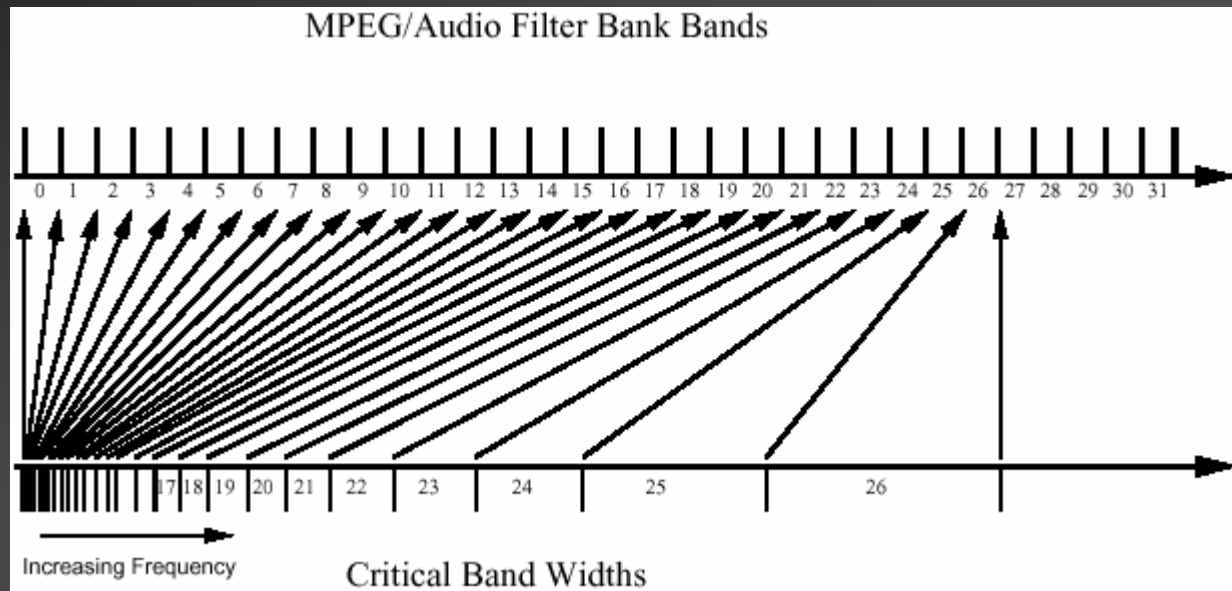
Μρ3 – Τα βήματα

- 1) Τράπεζα Φίλτρων
- 2) Εφαρμογή MDCT (Modified DCT)
- 3) Εφαρμογή Ψυχοακουστικού Μοντέλου
- 4) Επιμερισμός διαθέσιμων bits (ανάλογα με το bit rate)
- 5) Εφαρμογή κωδικοποίησης εντροπίας



(1) Τράπεζα φίλτρων

- Το ακουστικό σήμα διαιρείται σε 32 ζώνες συχνοτήτων ίσου μεγέθους
- Το ανθρώπινο αυτί έχει περιορισμένη διακριτική ικανότητα
- Μέσα σε μια περιορισμένη κριτική περιοχή εύρους ζώνης **συγχέει** τις συχνότητες που ακούει
- Οι κριτικές περιοχές αντιστοιχούνται σε 32 ζώνες

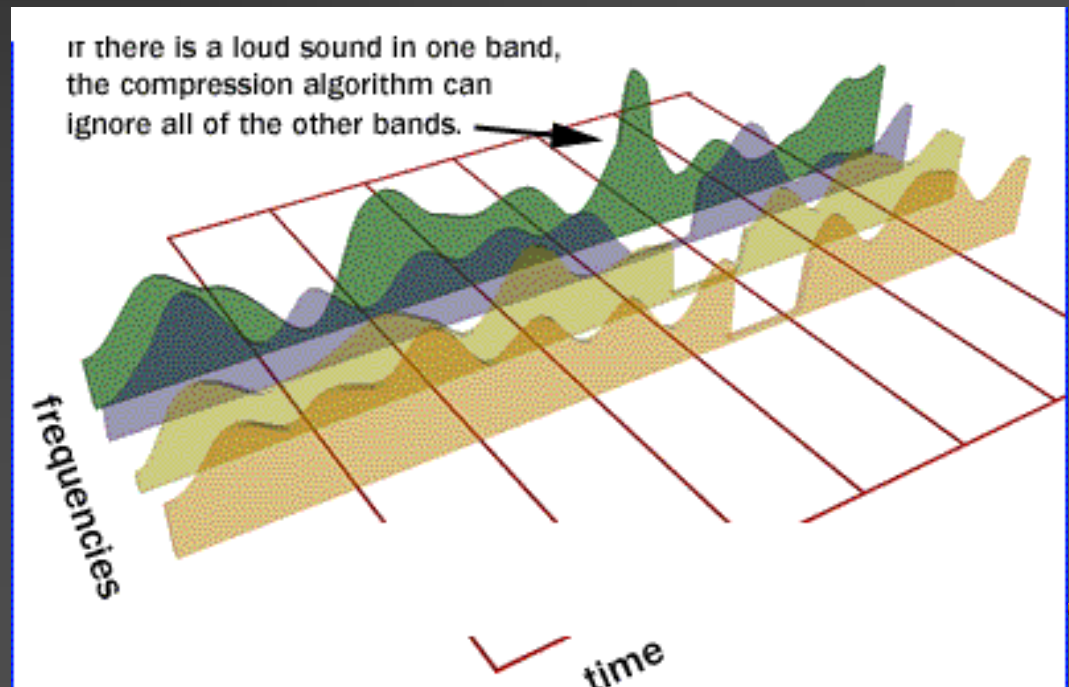
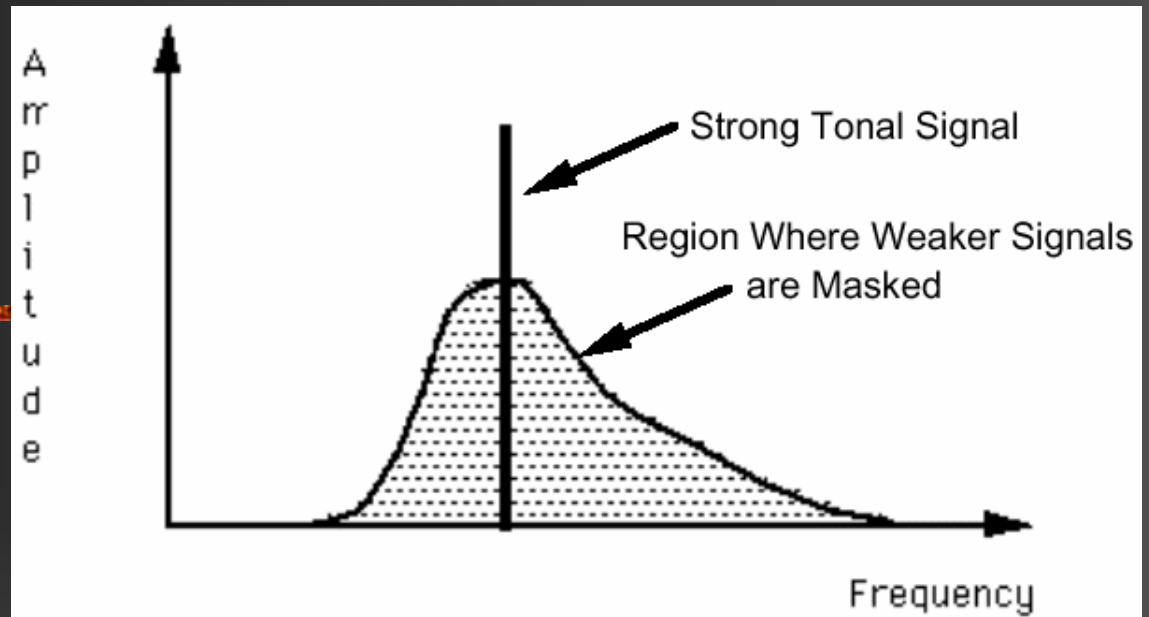


(2) Εφαρμογή MDCT

- MDCT = Modified Discrete Cosine Transform
- Υποδιαιρείται επιπλέον το φάσμα συχνοτήτων με στόχο την καλύτερη διακριτικότητα
- Χρησιμοποιούνται 18 συντελεστές MDCT
- Τελικά δημιουργούνται συνολικά $32 \times 18 = 576$ ζώνες στο φάσμα συχνοτήτων

(3) Εφαρμογή Ψυχοακουστικού Μοντέλου

- Εφαρμόζεται το ψυχοακουστικό μοντέλο ηχητικής σκίασης σε καθεμιά από τις ζώνες
- Δηλ. υπολογίζεται το κατώφλι ακουστότητας σε κάθε ζώνη για κάθε δεδομένη στιγμή (frame)
- Υπολογίζεται έτσι ποιες συχνότητες μπορούν να γίνουν αντιληπτές



(4) Επιμερισμός διαθέσιμων bit

- Καθορίζεται το πλήθος των bits που θα διατεθεί σε κάθε ζώνη
 - (με βάση το ψυχοακουστικό μοντέλο)

(5) Κβάντωση & Κωδικοποίηση

- Κβάντωση & Κωδικοποίηση Δεδομένων
 - Προσθήκη συνοδευτικών δεδομένων
 - Κωδικοποίηση κατά Huffman

MPEG-1 Layer 3 – Ρυθμός Μετάδοσης

- Το μέγεθος και η ποιότητα των αρχείων ήχου καθορίζονται από τον ρυθμό μετάδοσης των δεδομένων (bit rate).
- Το πρότυπο MP3 υποστηρίζει αρκετούς προκαθορισμένους ρυθμούς μετάδοσης δεδομένων.
- Τα αρχεία με μεγαλύτερο bit rate θα ακούγονται καλύτερα από αυτά με μικρότερο bit rate.

Συμπίεση

■ MPEG-1 Layer 3 – Ρυθμός Μετάδοσης

bandwidth	mode	Bit rate	Λόγος Συμπίεσης	Ποιότητα
2.5 kHz	Mono	8 kbps	96:1	Ήχος τηλεφώνου
4.5 kHz	Mono	16 kbps	48:1	Βραχεία
11 kHz	Stereo	56..64 kbps	26 .. 24:1	FM
15 kHz	Stereo	96 kbps	16:1	Σχεδόν σαν CD
>15 kHz	Stereo	112..128 kbps	14..12:1	CD

Χαρακτηριστικά του MPEG-1 Layer 3

- Συμπίεση αρχείων ήχου μπορεί να φτάσει και το **85%**
- Με ρυθμό μετάδοσης (bit rate) άνω των **128 Kbps** η διαφορά από το πρωτότυπο CD δεν γίνεται αντιληπτή.
- Το αποτέλεσμα είναι να έχουμε μουσικά αρχεία 4-6 λεπτών που κανονικά θα καταλάμβαναν χώρο 40-70 MB, να πιάνουν τώρα χώρο μόλις 3-7MB.



Ιστορικά Στοιχεία

MPEG

- Motion Picture Expert Group
- Προσπάθεια συμπίεσης ψηφιακού ήχου: 1987, Ινστιτούτο Fraunhofer
- MPEG-1, MPEG-1 Audio
- Επίπεδα συμπίεσης: MPEG-1 Layer 1, 2, 3

Επίπεδα συμπίεσης στο MPEG-1

- MPEG-1 Layer 1
 - Κασέτα DCC της Philips
 - Συμπίεση 4:1
 - Ποιότητα μέτρια
 - Ρυθμός μετάδοσης 384 Kbps/κανάλι
- MPEG-1 Layer 2
 - Ψηφιακό ραδιόφωνο DAB, VideoCD & DVD
 - Συμπίεση 6:1 & 8:1
 - Ποιότητα εφάμιλλη του CD
 - Ρυθμός μετάδοσης βελτιστοποιημένος στην περιοχή 192..256 Kbps/κανάλι
- MPEG-1 Layer 3 (mp3)
 - Αρχεία mp3 στο διαδίκτυο
 - Συμπίεση 10:1 & 12:1
 - Ποιότητα εφάμιλλη του CD
 - Ρυθμός μετάδοσης 112..128 Kbps για στερεοφωνικό ήχο (64 kbps/κανάλι)

Άλλα πρότυπα

Type	Extensions	Codec
AIFF (Mac)	.aif, .aiff	*PCM
AU (Sun/Next)	.au	*u-law
CD audio (CDDA)	N/A	PCM
MP3	.mp3	MPEG Audio Layer-III
Windows Media Audio	.wma	Proprietary (Microsoft)
QuickTime	.qt	Proprietary (Apple Computer)
RealAudio	.ra, .ram	Proprietary (Real Networks)
WAV	.wav	*PCM

Άλλα πρότυπα (1/3)

- **WAVE (*.wav)**

- Αποτελεί το πρότυπο αποθήκευσης ψηφιακού ήχου και είναι υποσύνολο του πρότυπου RIFF.
- Εφαρμόζουν κωδικοποίηση PCM format
 - Δηλ. είναι ασυμπίεστα και μεγάλου μεγέθους

- **AIFF (*.aif)**

- Δημιουργήθηκε από την Apple αλλά χρησιμοποιείται σε όλες τις πλατφόρμες. Υποστηρίζει δειγματοληψία 32 bit.

Άλλα πρότυπα (2/3)

■ WMA

- Windows Media Audio
- Αναπτύχθηκε από την Microsoft με στόχο την υποστήριξη απωλεστικής συμπίεσης και ως «απάντηση» στο mp3 & mp3 pro
- Προσφέρει ίδια ποιότητα στο μισό μέγεθος των αρχείων και σε χαμηλότερα bit rates (πχ. στα 64 Kbps)

■ MP3 Pro

- Συνδυάζει ένα mp3 αρχείο χαμηλού bitrate και τεχνικές υψηλής συμπίεσης στις υψηλότερες συχνότητες
 - Προσφέρει εξαιρετική ποιότητα στις χαμηλές (bass) και μεσαίες συχνότητες (treble).
-

Άλλα πρότυπα (3/3)

- OGG – Vorbis

- Ποιότητα συγκρίσιμη με το mp3
- Ελεύθερος codec
- Ενσωμάτωση σε πληθώρα συσκευών (από φορητές συσκευές μέχρι συστήματα Linux)

- RealAudio

- Προορίζεται για την άμεση αναπαραγωγή ήχων μέσω του Διαδικτύου.

Ήχος Midi

- Musical Instrument Digital Interface
- Αποθηκεύει **τέσσερις** παραμέτρους για κάθε νότα που παίζεται:
 - Θεμελιώδης συχνότητα (ποια **νότα** είναι),
 - Αρμονικές (ποιο **όργανο**, ηχόχρωμα),
 - Δυναμική (πόσο **δυνατά** παίζεται η νότα)
 - Διάρκεια (**χρόνος**)

Υλικό

- Κάρτα ήχου.
 - Η εταιρεία Ad Lib κατασκεύασε την πρώτη κάρτα ήχου, για προσωπικούς υπολογιστές περίπου στις αρχές της δεκαετίας του 90.
 - Μερικά χρόνια αργότερα εμφανίσθηκε και η πρώτη SoundBlaster της Creative με δυνατότητα ηχογράφησης.

Υλικό

■ Κάρτα ήχου

- **DSP** (Digital Signal Processor)
 - ADC / DAC μετατροπείς
 - **Synthesizer**
 - Wavetable ROM
 - Μνήμη RAM
 - Έξοδος ηχείων
 - Είσοδος μικροφώνου
 - MIDI interface
 - Στερεοφωνικός ήχος, (16bit/ 44,1 kHz ή ακόμη και 24/96)
 - Πολυκάναλος ήχος
 - «Τρισδιάστατος» ήχος, πχ. 5.1
-

Λογισμικό

- Προγράμματα επεξεργασίας ψηφιοποιημένου ήχου
 - Sound Forge της Sonic,
 - SoundEdit της Macromedia,
 - WaveStudio της Creative,
 - Cool Edit της Syntrillium

 - Προγράμματα επεξεργασίας ήχων MIDI
 - Cakewalk
 - Cubase Audio της Steinberg,
 - MidiSoft Studio,
 - Master Tracks της Passport Design
-

Ήχος και Εφαρμογές Πολυμέσων

- Ήχοι περιεχομένου
 - οι αφηγήσεις κειμένου,
 - οι εκφωνήσεις οδηγιών,
 - τα ηχητικά ντοκουμέντα
 - η μουσική όταν αποτελεί μέρος του αντικειμένου παρουσίασης.
- Ήχοι περιβάλλοντος
 - Τα διάφορα ηχητικά εφέ (ήχοι της φύσης, φωνές ζώων, ήχοι πόλης)
 - η μουσική επένδυση

Ενότητες

- Η φυσική του ήχου
- Ψηφιοποίηση & κωδικοποίηση ψηφιακού ήχου
- Αρχή ηχητικής σκίασης
- Απωλεστική συμπίεση ψηφιακού ήχου & το πρότυπο mp3
- Ιστορικά στοιχεία για το mp3 & άλλα πρότυπα
- Ήχος midi & υποσύστημα ήχου