



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

# Τεχνολογίες Πολυμέσων

Ενότητα # 6: Ήχος

Γιώργος Καρυδάκης

Τμήμα Πολιτισμικής Τεχνολογίας και Επικοινωνίας



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ  
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

# Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



# Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ  
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ  
*επένδυση στην κοινωνία της γνώσης*  
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ  
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ  
πρόγραμμα για την ανάπτυξη

# Σύνοψη προηγούμενης διάλεξης

- Δομή κειμένου
- Κωδικοποίηση
- Γραμματοσειρές
- Μορφοποίηση
- Αναγνωσιμότητα/Προσβασιμότητα
- Γλώσσες επισήμανσης/Μεταγλώσσες

# Περιεχόμενα ενότητας

- Φυσικά χαρακτηριστικά του ήχου
- Αντίληψη του ήχου
- Ψηφιοποίηση ήχου
- MIDI
- Συμπίεση αρχείων ήχου
- Επεξεργασία ήχου
- Streaming Audio

# Ενότητα # 6

Ήχος

# Ήχος στα πολυμέσα

- Περιεχομένου – παρέχει ουσιαστική πληροφορία στον χρήστη και είναι ισοδύναμοι με τον διάλογο που υπάρχει στο θέατρο ή τον κινηματογράφο
  - Αφηγήσεις
  - Εκφωνήσεις οδηγιών
  - Ηχητικά ντοκουμέντα
  - Μουσική που αποτελεί μέρος του αντικειμένου παρουσίασης
- Περιβάλλοντος – δεν παρέχει κάποια ουσιώδη πληροφορία αλλά χρησιμοποιείται για να ενισχύσει τα μηνύματα και να βελτιώσει την παρουσίαση της πληροφορίας

# Ήχος

- Ο ήχος κατάλληλα συνδυασμένος με τα άλλα είδη πληροφορίας ( κυρίως με video ή animation) μπορεί να κάνει μια εφαρμογή πολυμέσων πιο αποτελεσματική
- Σε όλες τις εφαρμογές πολυμέσων ο σχολιασμός των όσων παρουσιάζονται στην οθόνη, διάλογοι, αλλά και η κατάλληλη μουσική υπόκρουση, βοηθά στην μετάδοση του μηνύματος και καθιστά την παρακολούθηση της εφαρμογής πιο ευχάριστη
- Υπάρχει video χωρίς ήχο;



# Ήχος

- Το μοναδικό χαρακτηριστικό του ήχου να γίνεται αντιληπτός χωρίς να έχουμε την προσοχή μας εστιασμένη, καθιστά τα ηχητικά σήματα αναντικατάστατα στην απόσπαση της προσοχής του χρήστη
- Υπάρχουν ορισμένες κατηγορίες εφαρμογών όπου ο ήχος αποτελεί την καρδιά του συστήματος (π.χ. ραδιόφωνο).
- Εφαρμογές που προορίζονται για ανθρώπους με προβλήματα όρασης κάνουν εκτενή και αποτελεσματική χρήση του ήχου

# Χαρακτηριστικά πολυμεσικών εφαρμογών με ήχο

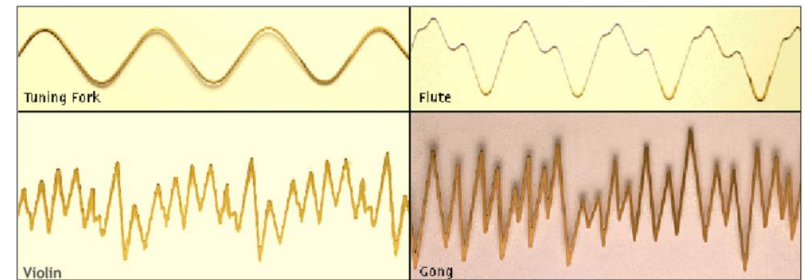
- Πλεονεκτήματα χρήσης ήχου
  - Κυριολεξία (λόγος)
  - Πληροφορίες
  - Ευχάριστος
  - Αποσπά την προσοχή κάποιου
  - Προσθέτει αληθοφάνεια– ρεαλιστικότητα
- Μειονεκτήματα χρήσης ήχου
  - Υπερβολή μπορεί να κουράσει

# Εφαρμογές ήχου στα Πολυμέσα

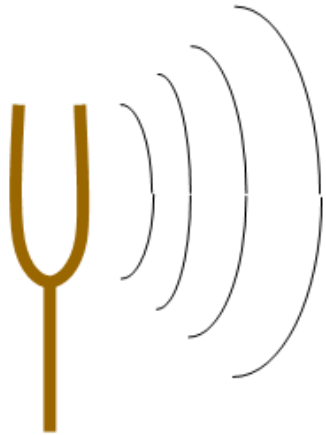
- Ψυχαγωγικές Εφαρμογές
- Εκπαιδευτικές Εφαρμογές
- Information kiosks
- Εφαρμογές με αντικείμενο τη μουσική
- Εφαρμογές που προορίζονται για ανθρώπους με προβλήματα όρασης
- Αναγνώριση και σύνθεση ομιλίας σε επαγγελματικές εφαρμογές

# Η φύση του ήχου

- Τι είναι ήχος
  - Ο ήχος παράγεται όταν δονείται ένα αντικείμενο
  - Αυτές οι δονήσεις παράγουν ηχητικά κύματα
  - Όταν τα ηχητικά κύματα φθάνουν στο αυτί μετατρέπονται σε σήματα που ερμηνεύονται από τον εγκέφαλό μας
  - Οι απλούστερες μορφές υγιών δονήσεων, είναι καθαροί τόνοι



# Φύση ηχητικού σήματος



διαπασών



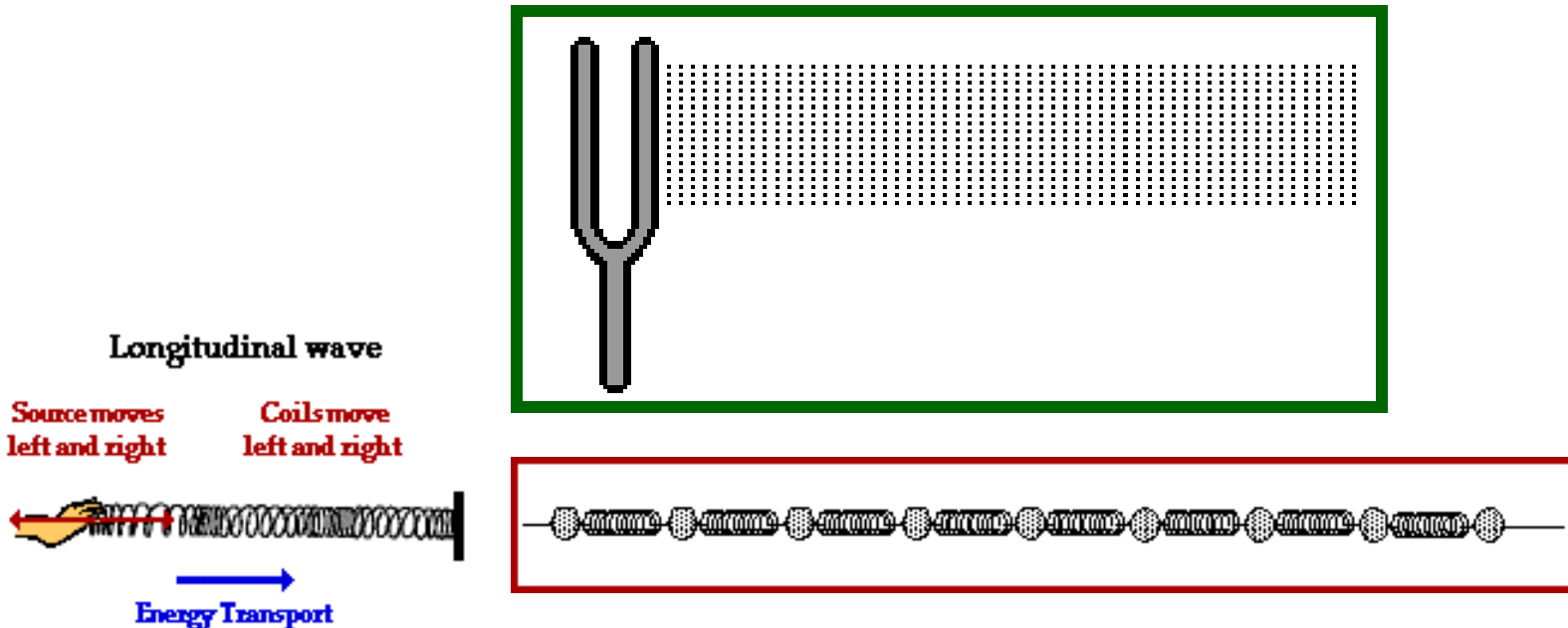
αυτί



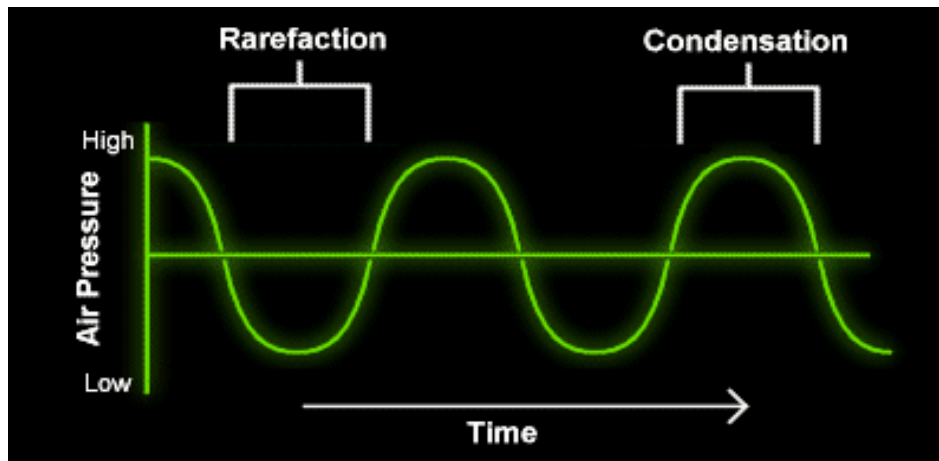
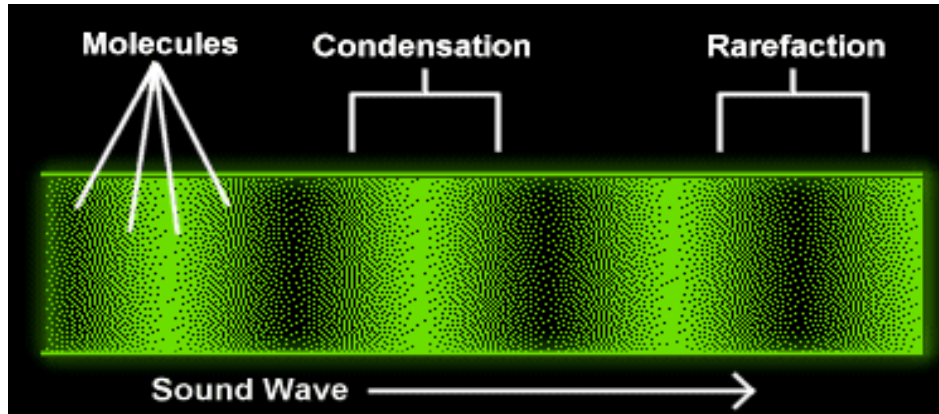
ηχητικό κύμα

# Πώς διαδίδεται ο ήχος;

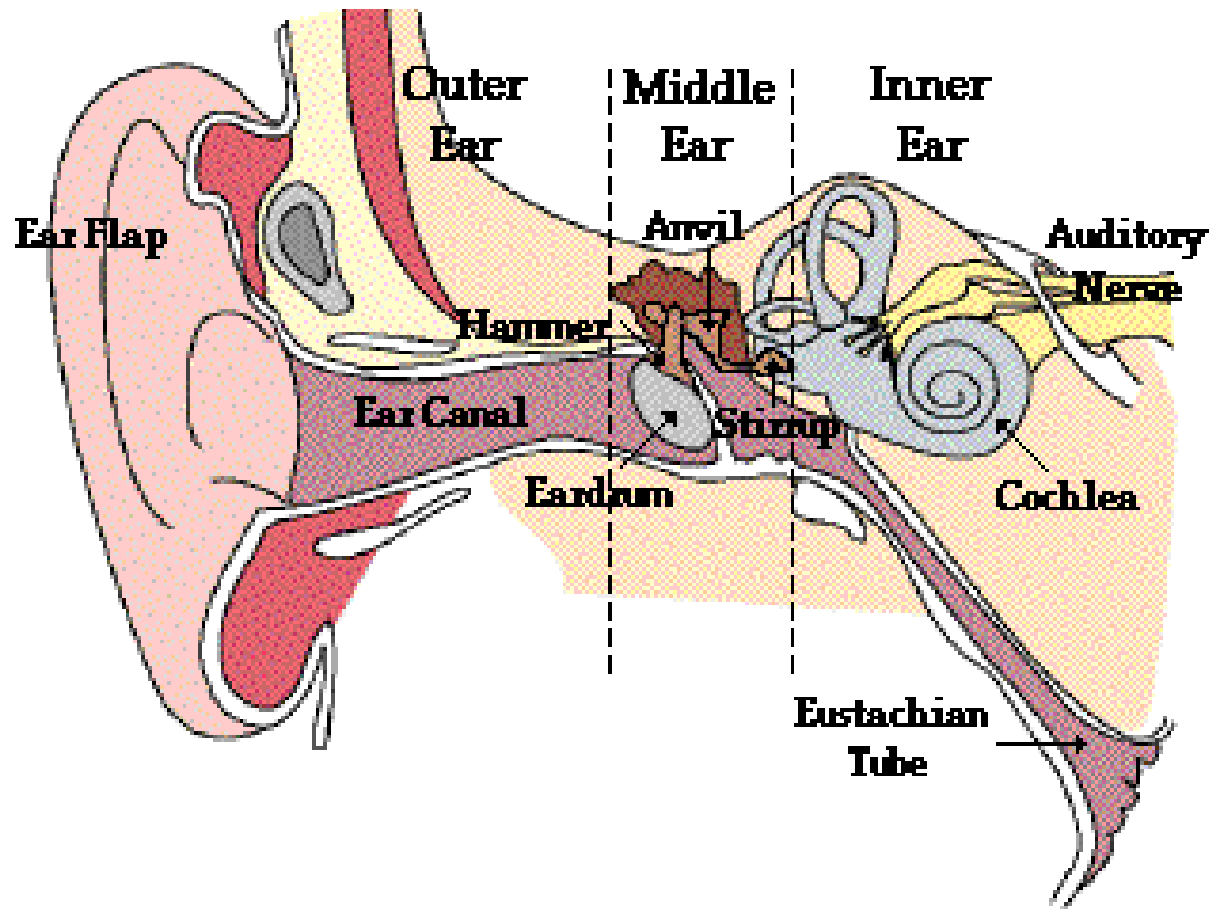
- Τα ηχητικά κύματα απαιτούν ένα υλικό μέσο για τη διάδοση, πχ. αέρας, νερό, κτλ.



# Πώς διαδίδεται ο ήχος;



# Πως ακούμε;



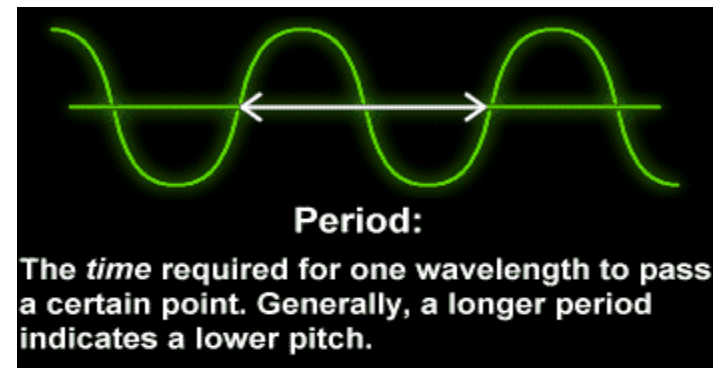
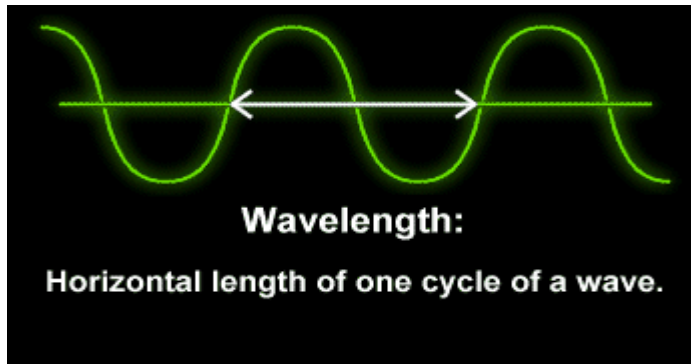


# Χαρακτηριστικά του ήχου

- Αντικειμενικά χαρακτηριστικά
  - Συχνότητα
  - Ένταση
- Υποκειμενικά χαρακτηριστικά
  - Ακουστικότητα
  - Ύψος
  - Χροιά

# Συχνότητα ήχου

- Η συχνότητα έχει να κάνει με το πόσο γρήγορα ή αργά πάλλεται το σώμα που προκαλεί τον ήχο
- Η συχνότητα ορίζει τον αριθμό των ολοκληρωμένων δονήσεων (κύκλων) στη μονάδα του χρόνου και μετράται σε κύκλους ανά δευτερόλεπτο ή Hertz (Hz)



# Συχνότητα ήχου (συνέχεια)

- Ο άνθρωπος μπορεί να ακούσει ήχους από 15Hz μέχρι 20 kHz (1kHz=1000Hz).
- Ήχοι με συχνότητα κάτω των 15Hz είναι γνωστός ως υπόηχοι (infrasound) ενώ ήχοι άνω των 20 kHz είναι γνωστός ως υπέρηχος (ultrasound).

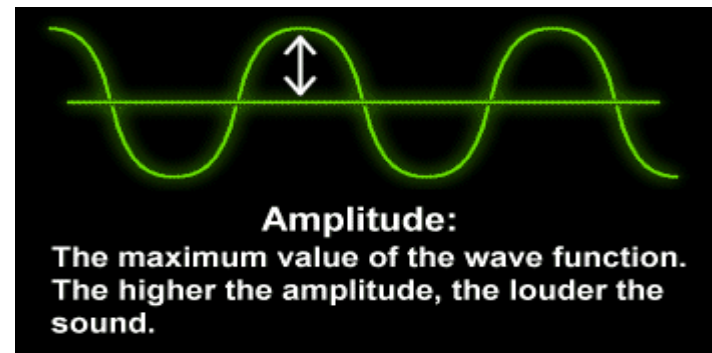
# Ένταση ήχου

- Η ένταση ενός ήχου καθορίζεται από το πλάτος της δόνησης. Όσο μεγαλύτερο το πλάτος, τόσο ισχυρότερα ακούγεται ο ήχος
- Η φυσική μονάδα μέτρησης της έντασης του ήχου είναι το decibels (db).

$$\text{Ένταση ήχου} = 10 \log (P/P_0) \text{db}$$

Οι τιμές  $P$  = τιμή μέτρησης  
και  $P_0$  = τιμή αναφοράς αναφέρονται σε μονάδες  
 $\text{Watt/m}^2$  .

Η τιμή αναφοράς είναι  $P_0=10^{-12} \text{ Watt/m}^2$



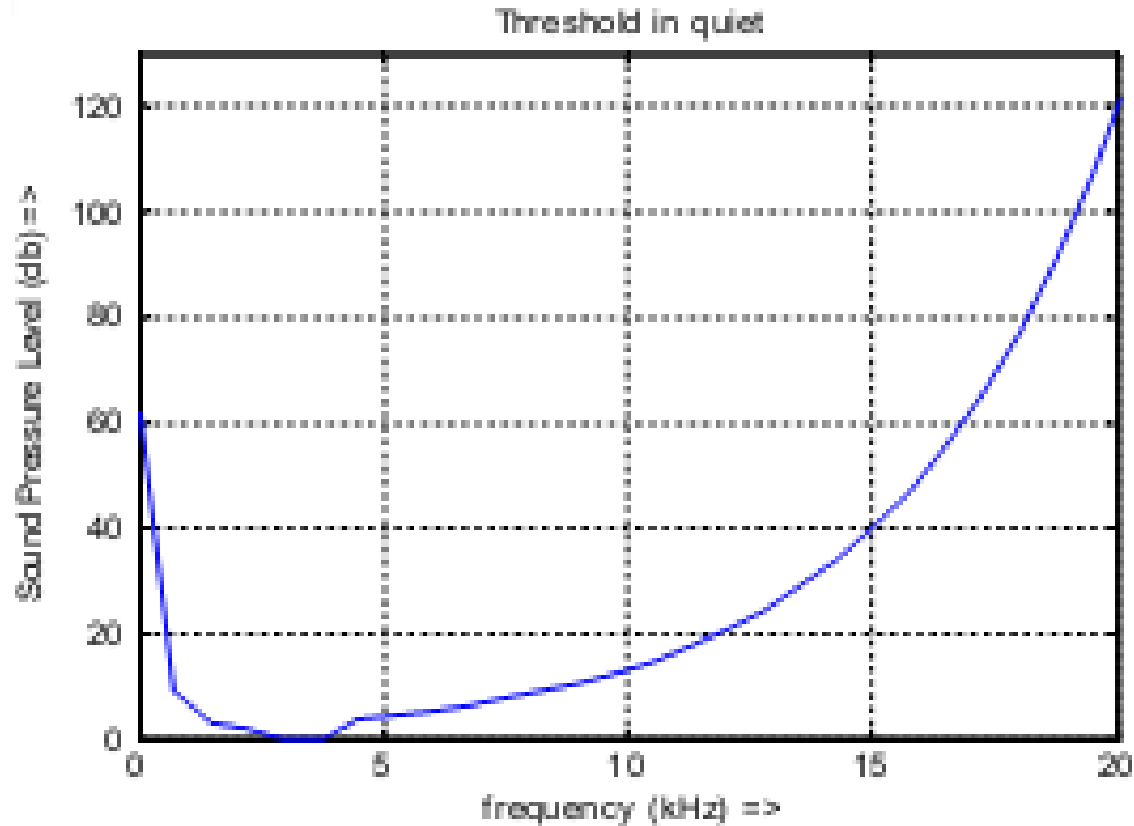
# Ψυχοακουστική - psychoacoustics

- Εξέταση των δυνατοτήτων του ανθρώπινου ακουστικού συστήματος:
  - Ακουστική ικανότητα σε σχέση με τη συχνότητα
  - Διακριτική ικανότητα αυτιού σε σχέση με τη συχνότητα
  - Φασματική κάλυψη
  - Χρονική κάλυψη

# Ακουστικότητα

- Η ακουστικότητα έχει σχέση με την ένταση του ήχου (ήχος με ίδια ένταση θα έχει μικρότερη ακουστικότητα για ένα άτομο με προβληματική ακοή)
- Ανάλογα με την ακουστικότητα διακρίνουμε τους ήχους σε ασθενείς και ισχυρούς
- Η ακουστικότητα μετριέται σε Phon
- Ήχος μόλις ακούγεται έχει ακουστικότητα **1Phon** ενώ ήχος ακουστικότητας **130Phon** προκαλεί πόνο στο αυτί.

# Ακουστική ικανότητα ανά συχνότητα



# Ύψος

- Το ύψος συνδέεται με την συχνότητα του ήχου
- Ανάλογα με το ύψος διακρίνουμε τους ήχους σε οξείς (μεγάλης συχνότητας) και βαρείς (μικρής συχνότητας)
- Η συχνότητα του παραγόμενου ήχου εξαρτάται (στην περίπτωση των έγχορδων) από το πάχος και το μήκος της χορδής. Όσο μεγαλύτερη η διάμετρος τόσο πιο μπάσο (μικρής συχνότητας) ήχο παράγει. Όσο μειώνεται το πάχος και το μήκος τόσο αυξάνεται και η συχνότητα.

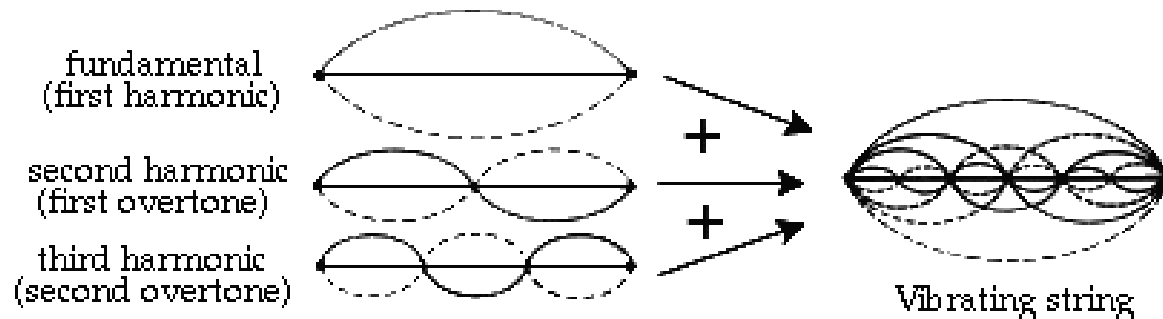
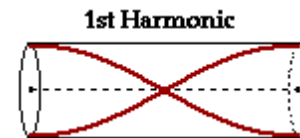
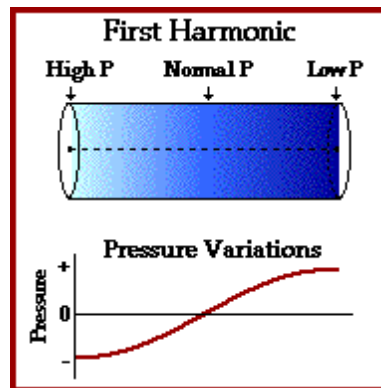


# Χροιά

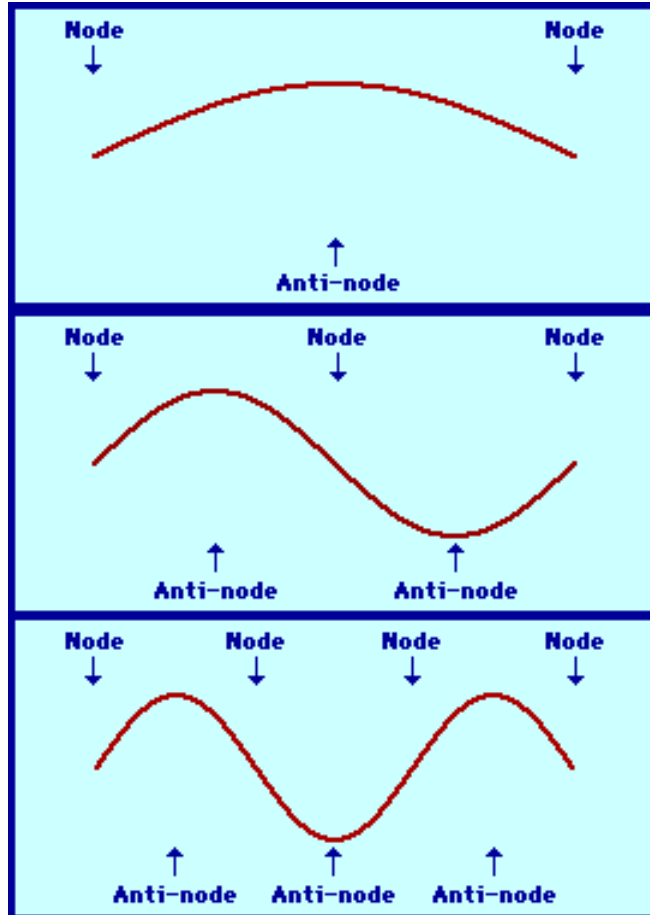
- Χροιά(timbre)
  - Η χροιά αναφέρεται στους σύνθετους ήχους και εξαρτάται από τους απλούς ήχους που αποτελούν το σύνθετο
  - Κάθε περιοδικός ήχος μπορεί να αναλυθεί σε ένα άθροισμα συνημιτονικών συναρτήσεων με διάφορα πλάτη και φάσεις και με συχνότητες ακέραια πολλαπλάσιες μιας θεμελιώδους συχνότητας (fundamental frequency)
  - Οι συχνότητες αυτές αποκαλούνται αρμονικές (harmonic frequency) και το πλήθος και το σχετικό τους πλάτος είναι σε μεγάλο βαθμό υπεύθυνες για το υποκειμενικό αίσθημα της χροιάς ενός ήχου

# Χροιά (1)

- Η χροιά αναφέρεται στους σύνθετους ήχους και εξαρτάται από τους απλούς ήχους που αποτελούν το σύνθετο.



# Χροιά (2)



# Μουσική, ομιλία, θόρυβος

- Ανάλογα με το φάσμα των συχνοτήτων ενός ήχου το ανθρώπινο αυτί κάνει το διαχωρισμό ανάμεσα σε:
  - **Μουσική**: περιέχει ένα σύμπλεγμα από νότες. Κάθε μουσική νότα περιέχει πέρα από τη θεμελιώδη συχνότητα υψηλότερες συχνότητες που είναι ακέραια πολλαπλάσια της θεμελιώδους συχνότητας
  - **Ομιλία**: περιέχει ένα σύμπλεγμα ήχων οι οποίοι (οι περισσότεροι αλλά όχι όλοι) είναι σε αρμονική σχέση μεταξύ τους
  - **Θόρυβος**: περιλαμβάνει ένα μίγμα ήχων με συχνότητες άσχετες μεταξύ τους

# Διακριτική ικανότητα

- Το αυτί μπορεί να διακρίνει συχνότητες που
  - Απέχουν  $>100$  Hz μεταξύ τους για συχνότητες  $<500$  Hz ανεξάρτητα την ισχύ που έχει καθεμία από αυτές (εφόσον φυσικά η ένταση τους ξεπερνά την ελάχιστη ακουστή τιμή)
    - π.χ. μπορεί να διακρίνει τις συχνότητες 450 Hz και 330 Hz
- Δεν μπορεί να διακρίνει
  - συχνότητες 450 Hz και 400 Hz (διαφέρουν λιγότερο από 100 Hz) και ακούει τη ισχυρότερη από αυτές
- Για συχνότητες  $> 500$  Hz η διακριτική ικανότητα του αυτιού επιδεινώνεται
- Η διακριτική ικανότητα του αυτιού σε σχέση με τις συχνότητες οδήγησε στη χρήση φίλτρων ανάλυσης σε ζώνες (subband filters) στο πλαίσιο της κωδικοποίησης με βάση την αντίληψη

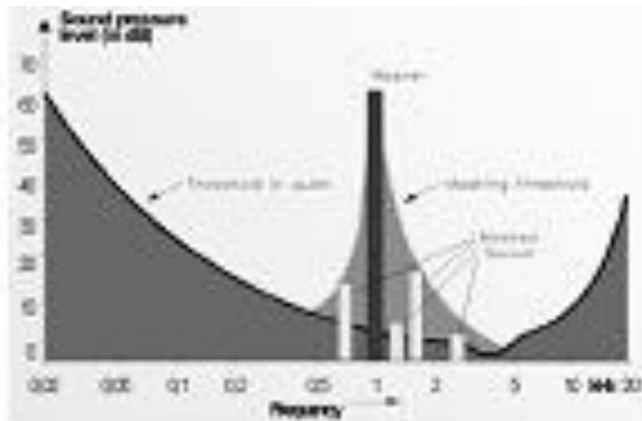
# Αισθήσεις και παραισθήσεις (1)

- Ο ήχος υπάρχει μόνο σαν αίσθηση στον ανθρώπινο εγκέφαλο και η αντίληψη του ήχου δεν είναι απλά καταγραφή των φυσικών φαινομένων
- Κάθε άνθρωπος αντιδρά διαφορετικά σε διαφορετικούς ήχους
  - λίγοι ενήλικες ακούν συχνότητες στα 20 kHz, ενώ τα περισσότερα παιδιά μπορούν
- Κάποιος ακούει να καλείται το όνομά του από την άλλη άκρη ενός συνωστισμένου δωματίου, ενώ δεν μπορεί να αντιληφθεί τα υπόλοιπα λόγια

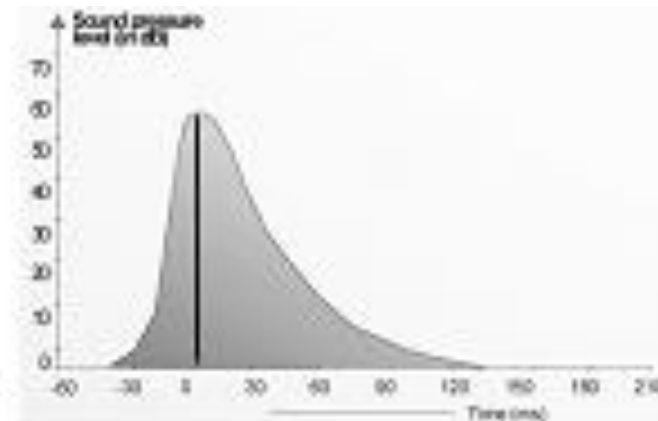
## Αισθήσεις και παραισθήσεις (2)

- Το ανθρώπινο αυτί έχει την ικανότητα να 'αναπληρώνει' κάποιες αρμονικές συχνότητες ακόμα κι αν αυτές δεν υπάρχουν στον μεταδιδόμενο ήχο
- Ή να αγνοεί κάποιες συχνότητες ακόμα κι όταν αυτές υπάρχουν
- Οι ιδιότητες αυτές συχνά καλύπτουν τις αδυναμίες του εξοπλισμού που χρησιμοποιείται για την παραγωγή και μετάδοση του ήχου

# Αρμονικές και φασματική κάλυψη



Φασματική κάλυψη (spectral masking)



Χρονική κάλυψη (temporal masking)

–  $a_n$  και  $b_n$ : αρμονικές (harmonics) του σήματος

$$g(t) = \frac{1}{2}c + \sum_{n=1}^{\infty} a_n \sin(2\pi nft) + \sum_{n=1}^{\infty} b_n \cos(2\pi nft)$$



# Στερεοφωνικός ήχος

- Από τις μεγαλύτερες ψευδαισθήσεις στην αντίληψη του ήχου είναι η στερεοφωνία
  - **στερεοφωνικός ήχος** – δύο κανάλια ηχητικής πληροφορίας
  - **τριδιάστατος ήχος** – δύο στερεοφωνικές εξόδους οι οποίες μπορούν να οδηγήσουν τέσσερα ηχεία που αν διαταχθούν περιμετρικά δίνουν την ψευδαίσθηση ότι ο ήχος έρχεται πράγματι από όλες τις διαστάσεις
- Το μυαλό αντιλαμβάνεται την πηγή ενός ήχου βάσει των διαφορών ανάμεσα στα σήματα που λαμβάνει το αριστερό και δεξί αυτί
- Αν πανομοιότυπα σήματα φτάσουν και στα 2 αυτιά τότε ο εγκέφαλος το μεταφράζει σαν η πηγή να βρίσκεται ακριβώς μπροστά.

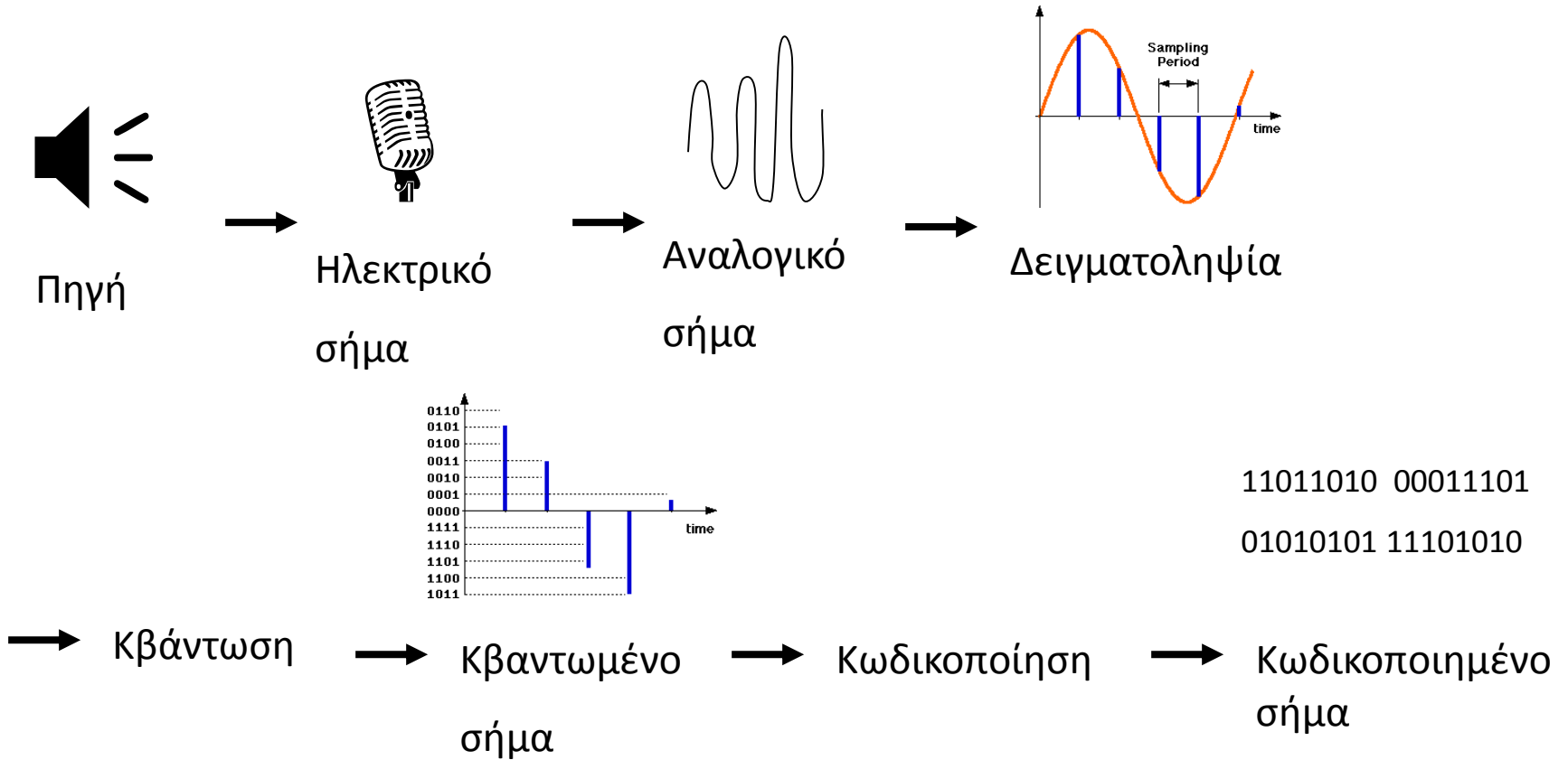


# Πολυκαναλικός

- Ο ήχος έχει 3Δ πληροφορία η οποία μπορεί να προσομοιωθεί με περισσότερα κανάλια (πηγές)
- Σχετική θέση πηγής
- Πολυκαναλικός χρησιμοποιείται στον κινηματογράφο



# Ο ήχος στους υπολογιστές



# Ψηφιοποίηση

- Αριθμός bits ανά δείγμα (sampling size)
  - Σχετίζεται με την ποιότητα του ήχου
  - Τυπικές τιμές
    - 8 bits /sample (256 τιμές)
    - 16 bits /sample (65536 τιμές)
- Το σφάλμα κβαντισμού αποτελείται από τη διαφορά μεταξύ των σημάτων εισόδου και εξόδου του κβαντιστή
- Η μέγιστη στιγμιαία τιμή αυτού του σφάλματος είναι το μισό ενός κβάντου

# Μέγεθος αρχείου

- Χωρητικότητα (bits)=συχνότητα δειγματοληψίας (Hz) x ανάλυση πληροφορία ή μέγεθος δείγματος (bits) x διάρκεια (sec)

Ποιότητα	Mono/Stereo	Συχνότητα δειγματοληψίας	Κβαντοποίηση	Μέγεθος
CD	Stereo	44100Hz	16	172.26KBps
Radio	Mono	22050Hz	8	21.53KBps
Τηλέφωνο	Mono	11025Hz	8	10.7KBps

Για στερεοφωνικό ήχο πολλαπλασιάζω X 2  
Διάρκεια = 1sec, 1byte=8bit & 1Kb=1024bytes

Δηλ. 2 λεπτά ψηφιοποιημένου ήχου ποιότητας τηλεφώνου είναι:  
 $10,7 \times 60 \times 2 = 1284 \text{Kb} = 1.25 \text{MBps}$

# Θόρυβος (φύσημα) ή παραμόρφωση

- Ο θόρυβος προέρχεται από την πεπερασμένη διακριτικότητα της ψηφιακής αναπαράστασης και είναι γνωστός και ως θόρυβος κβαντοποίησης
- Ο θόρυβος μετριέται σε αρνητικά decibel
- Signal-to-Noise Ratio (SNR) δείχνει την αναλογία μεταξύ ηχητικού σήματος και του θορύβου

# Μέθοδοι κωδικοποίησης (1)

- PCM – Pulse Code Modulation Παλμοκωδική κωδικοποίηση
  - Η PCM αποθηκεύει ένα προς ένα τα δείγματα σε ψηφιακή μορφή και χρησιμοποιεί γραμμική (ομοιόμορφη) κωδικοποίηση – το πλήθος των σταθμών (επιπέδων κβάντωσης) κατανέμονται εξίσου σε όλο το εύρος του πεδίου τιμών του σήματος
  - Δεν πραγματοποιείται συμπίεση και το αρχείο ήχου που προκύπτει έχει μεγάλο μέγεθος
  - Ένα αρχείο κωδικοποιημένο κατά PCM μπορεί να αναπαραχθεί χωρίς καμία απώλεια

# Μέθοδοι κωδικοποίησης (2)

- Mu-Law PCM και A-Law PCM
  - Η Mu-Law PCM χρησιμοποιεί λογαριθμική κωδικοποίηση – αντιστοιχεί περισσότερες στάθμες στις χαμηλές συχνότητες (όπου περιέχεται μεγαλύτερος ποσοστό της πληροφορίας που γίνεται αντιληπτή από το ανθρώπινο αυτί ενώ στις υψηλότερες παρουσιάζεται το πρόβλημα θορύβου) και λιγότερες στις υψηλές
  - Καλύτερη αναπαράσταση του σήματος και συμπίεση των δεδομένων
  - Η Mu-Law χρησιμοποιείται κυρίως στα δίκτυα ISDN στη βόρεια Αμερική και Ιαπωνία
  - Η A-Law χρησιμοποιείται κυρίως στα δίκτυα ISDN των υπόλοιπων χωρών



# Μέθοδοι κωδικοποίησης (3)

- DPCM (Differential Pulse Code Modulation – διαφορική παλμοκωδική κωδικοποίηση)
  - Στην PCM το κάθε δείγμα κωδικοποιείται ανεξάρτητα από τα υπόλοιπα
  - Στην DPCM το κάθε δείγμα κωδικοποιείται σε σχέση με τα γειτονικά δείγματα
  - Στην DPCM αποθηκεύεται πλήρως το πρώτο δείγμα ενώ για τα υπόλοιπα αποθηκεύονται μόνο οι διαφορές τους από τα προηγούμενα

# Μέθοδοι κωδικοποίησης (4)

- ADPCM (Adaptive Differential Pulse Code Modulation – προσαρμοσμένη διαφορική παλμοκωδική κωδικοποίηση)
  - Η ADPCM αποτελεί επέκταση της DPCM. Επειδή τα γειτονικά δείγματα είναι πιθανό να είναι όμοια είναι δυνατό να προβλεφθεί η τιμή ενός δείγματος με βάση την τιμή του γειτονικού του.
  - Η ADPCM προβλέπει την τιμή του δείγματος και στην συνέχεια υπολογίζει την διαφορά μεταξύ πραγματικής και προβλεπόμενης τιμής και κωδικοποιείται.
  - Η διαφορά είναι συνήθως μικρή και απαιτεί λίγα bit για την κωδικοποίησή της

# Μέθοδοι κωδικοποίησης (5)

- LPC (Linear Predictive Coding – γραμμική προβλεπόμενη κωδικοποίηση)
  - Έχει σχεδιαστεί για την ανθρώπινη ομιλία
  - Συγκρίνει τα δείγματα με ένα αναλυτικό μοντέλο φωνής που έχει αποθηκευμένο
  - Τα χαρακτηριστικά που ταιριάζουν στο μοντέλο μεταδίδονται
  - Ο αποκωδικοποιητής χρησιμοποιεί τα χαρακτηριστικά αυτά για να ανασυνθέσει τα φωνητικά σήματα
  - Η LPC χρησιμοποιείται στην μετάδοση ομιλίας στην κινητή τηλεφωνία

# Συμπίεση (1)

- Εξαρτάται από το είδος του ήχου
  - Έχουν αναπτυχθεί διαφορετικοί αλγόριθμοι για ομιλία και μουσική
- Πρότυπα Τηλεφωνίας
  - Οι εταιρίες τηλεφώνου χρησιμοποιούν ψηφιακή τεχνολογία ήχου από το 1960
  - Το μικρό εύρος των τηλεφωνικών γραμμών οδήγησε στη δημιουργία τεχνικών συμπίεσης
- Πρότυπα MPEG
  - Ξεκίνησαν σαν πρότυπα για την συμπίεση κινούμενης εικόνας
  - Επειδή το video συμπεριλαμβάνει και ήχο συμπεριλαμβάνουν και τεχνικές για συμπίεση ήχου

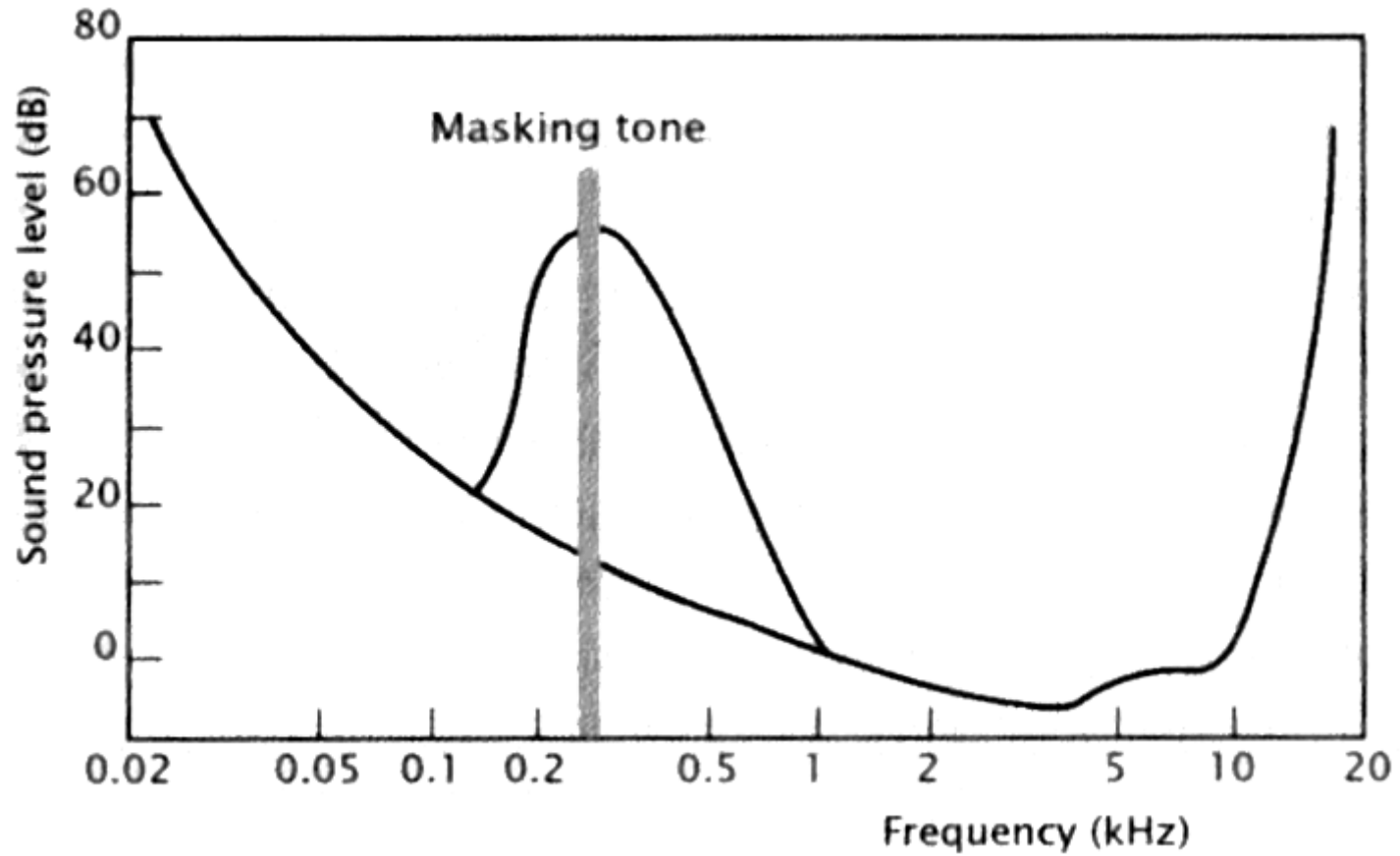
## Συμπύεση (2)

- Μη-απωλεστικοί
- Απωλεστικοί – ψυχο-ακουστικά μοντέλα βάση των οποίων απορρίπτονται ήχοι που ούτως ή άλλως δεν γίνονται αντιληπτοί από τον άνθρωπο MPEG και AC3
- Οι σημαντικότερες μέθοδοι συμπίεσης αρχείων ήχου στηρίζονται στο φαινόμενο της ακουστικής σκίασης (auditory masking)

# Ακουστική Σκίαση (Auditory Masking) (1)

- Σε ένα συγκεκριμένο φάσμα συχνοτήτων το ανθρώπινο αυτί έχει την ικανότητα να συλλαμβάνει ορισμένες 'κυρίαρχες συχνότητες'
- Άλλοι ήχοι που επίσης υπάρχουν αλλά η συχνότητά τους πλησιάζει αυτές των κυρίαρχων συχνοτήτων δεν γίνονται αντιληπτοί
- Έτσι το σήμα χωρίζεται σε ζώνες συχνοτήτων όπου ανιχνεύονται και κωδικοποιούνται μόνο οι κυρίαρχες συχνότητες.

## Ακουστική Σκίαση (2)



# Συμπίεση MPEG

- Η προσπάθεια του ψηφιακού ήχου ξεκίνησε από το 1987 στα πλαίσια της ευρωπαϊκής Eureka, πρόγραμμα EU147 για την τυποποίηση του ψηφιακού ραδιοφώνου (DAB, Digital Audio Broadcasting)
- Ψυχοακουστικό μοντέλο
  - Εκτίμηση σημασίας κάθε ζώνης/συντελεστή
  - Κατανομή των διαθέσιμων bit ανάλογα με τη σημασία



# Κωδικοποίηση εντροπίας

- MP1/2
  - PCM
- MP3
  - Huffman
- Στερεοφωνικά σήματα
  - Ανεξάρτητη ή εξαρτημένη κωδικοποίηση (joint stereo)
- Τρία επίπεδα με προς τα πίσω συμβατότητα
  - Πιο δημοφιλές το επίπεδο 3 (MP3 = MPEG Audio Layer 3)
- Τελική κωδικοποίηση
  - Επίπεδα 1 και 2: σταθερός ρυθμός bit
  - Επίπεδο 3: προαιρετικά και μεταβλητός ρυθμός bit
- Ρυθμός δεδομένων: τουλάχιστον 32 Kbps
  - Μέχρι 448 Kbps (1), 384 Kbps (2), 320 Kbps (3)

# Αντιληπτική κωδικοποίηση

- Κωδικοποίηση γενικών ήχων εκτός φωνής
  - Η κωδικοποίηση πηγής δεν είναι εφικτή
  - Δεν γνωρίζουμε τα χαρακτηριστικά της πηγής
  - Γνωρίζουμε όμως τα χαρακτηριστικά της ακοής
- Αντιληπτική κωδικοποίηση
  - Χρήση ψυχοακουστικού μοντέλου της ακοής
  - Εντοπισμός των μη αντιληπτών τμημάτων του ήχου
  - Κωδικοποίησή τους με μικρή ακρίβεια ή καθόλου
- Δυναμικό εύρος της ακοής
  - Πλέον ισχυρό / πλέον ασθενές αντιληπτό σήμα
  - Εξαρτάται από την περιοχή συχνοτήτων
  - Μεγιστοποιείται στην περιοχή 2-5 KHz
  - Σήματα που ακούγονται εκεί, δεν ακούγονται αλλού

# Αντιληπτική κωδικοποίηση

- Απόκρυψη συχνοτήτων
  - τα ισχυρά σήματα μειώνουν τοπικά το δυναμικό εύρος
  - Στις γειτονικές συχνότητες ανεβαίνει το κάτω όριο
  - Το εύρος της επίδρασης εξαρτάται από τη συχνότητα
- Απόκρυψη χρόνου
  - Τα ισχυρά σήματα αποκρύπτουν τα ασθενή για λίγη ώρα
  - Το εύρος της επίδρασης εξαρτάται από τη συχνότητα
- Αξιοποίηση των αποκρύψεων
  - Αναλύουμε το σήμα σε περιοχές συχνοτήτων
  - Σε κάθε περιοχή εντοπίζουμε τα ισχυρά σήματα
  - Υπολογίζουμε την επίδρασή τους στα ασθενή σήματα
- Σε κάθε περιοχή έχουμε διαφορετική επίδραση
  - Κωδικοποιούμε τα ασθενή σήματα με μικρότερη ακρίβεια

# MPEG-1 (1)

- **MPEG-1 Audio** - έχει τα εξής χαρακτηριστικά:
- Συχνότητα δειγματοληψίας: 32kHz, 44.1kHz, ή 48kHz
- Υποστήριξη ακουστικών καναλιών: το πρότυπο προβλέπει ένα ή δύο ηχητικά κανάλια
  - Απλό μονοφωνικό
  - Διπλό μονοφωνικό για δύο ανεξάρτητα ακουστικά κανάλια
  - Από στερεοφωνικό - συνδυασμένο στερεοφωνικό το οποίο εκμεταλλεύεται είτε το άθροισμα είτε την διαφορά φάσεων των καναλιών ή και τα δύο
- Ρυθμός μετάδοσης δεδομένων (bit rate) από 32 – 224 kbps ανά κανάλι

# MPEG-1 (2)

- Επίπεδα συμπίεσης
  - MPEG-1 Layer 1:
    - Χρησιμοποιήθηκε στο σύστημα συμπίεσης της ψηφιακής κασέτας DCC της Phillips
    - Προσφέρει συμπίεση 4:1
    - Μέτρια ηχητική ποιότητα
    - Ο ρυθμός μετάδοσης που απαιτείται είναι 192 ή 256 Kbps
  - MPEG-1 Layer 2
    - MP2 που χρησιμοποιήθηκε στο ψηφιακό ραδιόφωνο DAB, στο σύστημα DVB (στην ψηφιακή τηλεόραση) και στον ήχο Video CD και DVD
    - Προσφέρει συμπίεση 6:1 ως 8:1
    - Ο ρυθμός μετάδοσης που απαιτείται είναι 96 ως 128 Kbps για κάθε κανάλι

# MPEG-1 (3)

- Επίπεδα συμπίεσης
  - MPEG-1 Layer 3
    - Λόγος συμπίεσης 12:1
    - Ποιότητα που πλησιάζει τα μουσικά CD
    - Απαιτεί ρυθμό μετάδοσης 64Kbps
    - MP3 – προορίζονται για μεταφορά και φόρτωση μέσω δικτύου καθώς και για αναπαραγωγή και ανάκληση από σκληρό δίσκο

# MPEG-2

- MPEG-2 είναι μια επέκταση του MPEG-1 όσον αφορά τους αλγορίθμους συμπίεσης και άλλες δυνατότητες όπως πολυκάναλη ηχητική κωδικοποίηση
- Αποτέλεσε την βάση του DVD Video
- Το πρότυπο παρέχει δειγματοληψία ελαττωμένη (16kHz, 22.05kHz, 24kHz) και παρέχει βελτιωμένη ποιότητα για ρυθμούς μετάδοσης σε 64KBits/sec

# MPEG-4

- MPEG-4 στηρίζεται στην fractal γεωμετρία, οπτικοποίηση και τεχνητή νοημοσύνης
- Στόχος του MPEG-4 είναι η υψηλή ποιότητα του ψηφιοποιημένου σήματος με πολύ χαμηλούς ρυθμούς μετάδοσης από 6 ως 64 Kbits/sec



# Κωδικοποιητές ήχου (Codecs)

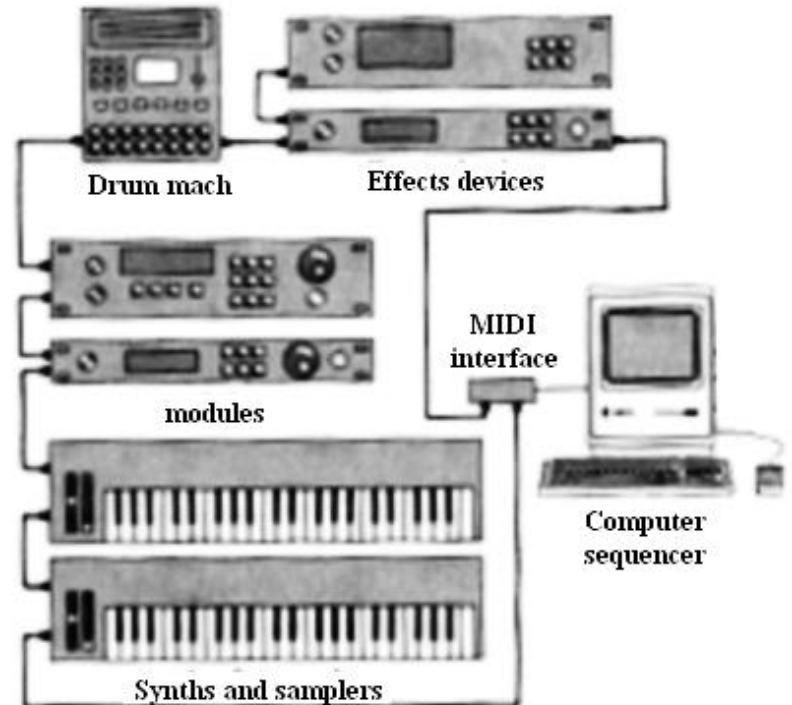
- Codec προέρχεται από τις λέξεις Coder/Decoder (κωδικοποιητής/αποκωδικοποιητής) και αναφέρεται στον μηχανισμό κωδικοποίησης και αποκωδικοποίησης ενός αρχείου
- Ο Audio Codec υλοποιεί κάποιες μεθόδους συμπίεσης
- Για την αναπαραγωγή του αρχείου θα πρέπει το οποιοδήποτε πρόγραμμα να χρησιμοποιεί το συγκεκριμένο Codec

# Τύποι αρχείων

- Υπάρχουν δύο τύποι αρχείων που μπορούν να αποθηκευτούν σε δίσκο ή CD-ROM
  - **Κυματομορφής** (digital audio) – παρέχουν μία ακριβής ψηφιακή εικόνα της κυματομορφής του ήχου
  - **MIDI** (Musical Instrument Digital Interface) – δεν καταγράφουν ηχητικό σήμα αλλά κάθε ενέργεια που γίνεται για να παραχθεί αυτό
  - Έτσι αντί να αποθηκεύσουμε τους ήχους μετατρέποντας το ηχητικό σήμα σε ψηφιακό, αποθηκεύονται 4 αριθμοί:
    - ποια νότα είναι
    - ποιο όργανο
    - πόσο δυνατά παίζει το όργανο
    - για πόσο χρόνο παίζει

# Απαιτήσεις για MIDI

- Κάρτα ήχου
  - Κυκλώματα παραγωγής ήχου
- Εφαρμογή MIDI sequencer
  - Ένα πρόγραμμα που 'λέει' στην κάρτα ήχου ποιες νότες να παίξει.
- Κατευθείαν σύλληψη ήχου από όργανα με δυνατότητα εξόδου midi ήχου



# Τεχνικές σύνθεσης ήχου

- Υπάρχουν δύο βασικές τεχνικές σύνθεσης ήχου
  - Η σύνθεση **FM (Frequency Modulation)** – η μετατροπή των αρχείων ήχου γίνεται από ειδικά τσιπ που λέγονται τσιπ σύνθεσης FM – η βασική αρχή της FM είναι η αναπαραγωγή μιας κυματομορφής χρησιμοποιώντας μια ποικιλία απλών κυματομορφών μέσω συγκεκριμένων αλγορίθμων που καθορίζουν τις σχέσεις διαμόρφωσης μεταξύ του συνόλου αυτών των απλών κυματομορφών
  - Και η τεχνική **Wavetable** – για την αναπαραγωγή MIDI αρχείων ήχου – γίνεται η αναπαραγωγή ήχου με τη χρήση δειγμάτων φυσικών ήχων

# Παραγωγή ήχου

- Εξαρχής ηχογράφηση με χρήση
  - Μικρόφωνο Η/Υ: κακή ποιότητα
  - Ψηφιακή ηχογράφηση σε στούντιο χρησιμοποιώντας DAT (Digital Audio Tape) recorder
  - MIDI
- Από Audio CD με τη χρήση ειδικού λογισμικού
  - Προσοχή: η αντιγραφή και χρήση του περιεχομένου των audio-CD αποτελεί κλοπή πνευματικής ιδιοκτησίας!
- Από το διαδίκτυο
  - Παρότι επιτρέπεται το 'κατέβασμα' ήχου από νόμιμες πηγές η χρήση των αρχείων ήχου χωρίς άδεια μπορεί να είναι παράνομη!

# Επεξεργασία ψηφιακού ήχου

- Καθορισμός επιπέδου έντασης για ψηφιακή ηχογράφηση
  - Ανάμεσα σε -3db και -10db είναι η συνιστώμενη επιλογή
- Συνδυασμός πολλών ηχογραφήσεων σε μία (multiple track mix)
- Αποκοπή (trimming):
- Αφαίρεση κενών ανάμεσα σε ηχογραφήσεις ή στο τέλος τους
- Ρύθμιση έντασης:
  - Καθορισμός ενιαίας μέγιστης ή μέσης έντασης για όλες τις ηχογραφήσεις
- Μετατροπή τύπου (format conversion)
- Επαναδειγματοληψία ή υποδειγματοληψία
- Παρεμβολή (interpolation) για αύξηση των δειγμάτων με τεχνικό τρόπο
- Αφαίρεση δειγμάτων ανά τακτά χρονικά διαστήματα

# Επεξεργασία ψηφιακού ήχου

- Fade in / Fade out:
  - Βαθμιαία εισαγωγή / ολοκλήρωση ηχογράφησης
  - Ομαλή εναλλαγή ανάμεσα σε ηχογραφήσεις
- Εξισορρόπηση (equalization):
  - Ρύθμιση της ισχύος για διάφορες συχνότητες (
- Χρονική σύμπτυξη:
  - Αλλαγή τέμπου(tempo)
- Ψηφιακά effect:
  - Chorus (για δημιουργία χορωδίας)
  - Reverb (για προσομοίωση ακουστικών χώρων – π.χ. Club, κενού δωματίου κλπ)

# Κάρτες ήχου

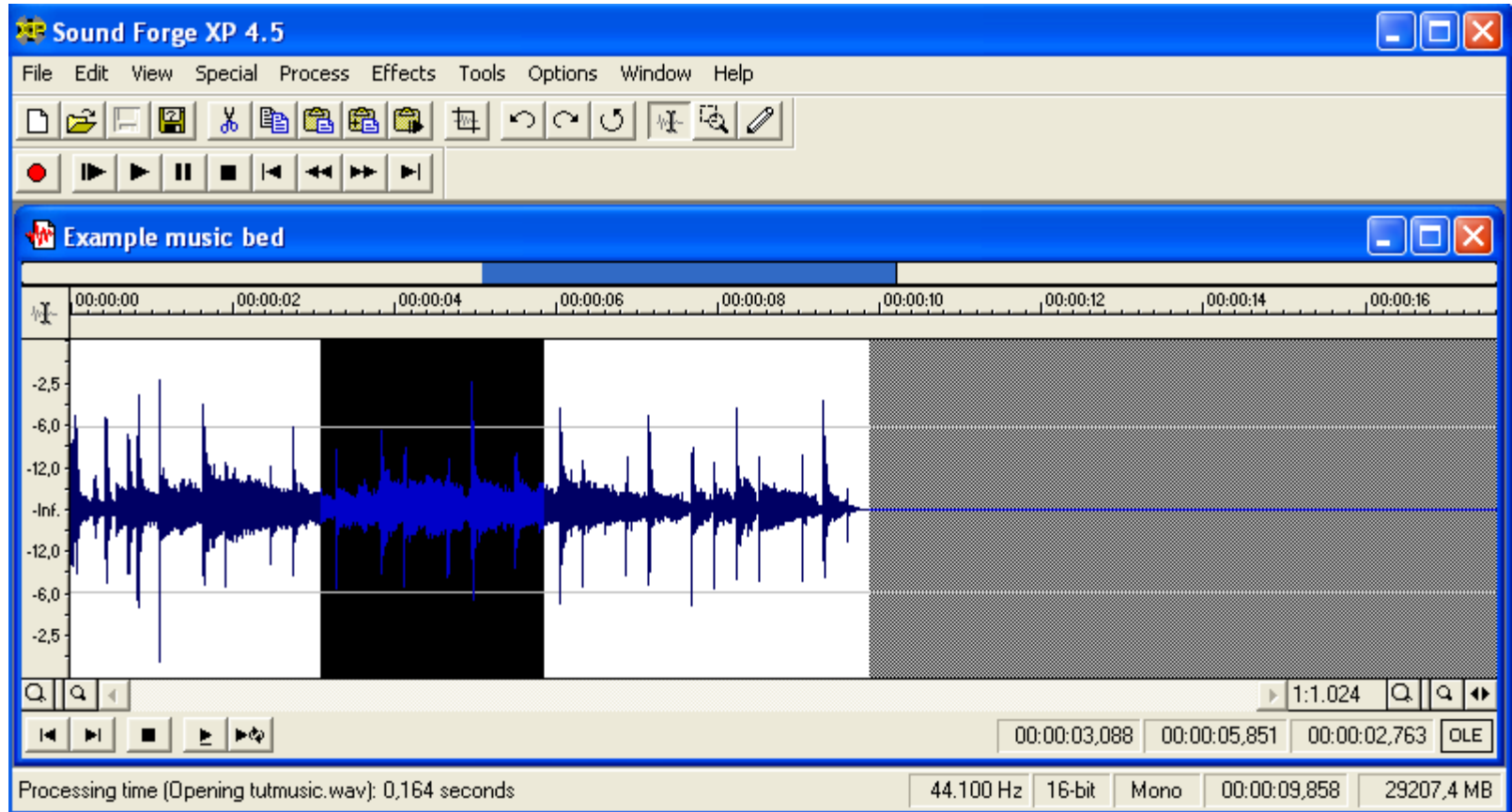
- Η λειτουργία της κάρτας ήχου είναι σχετικά απλή και στηρίζεται στην ύπαρξη:
  - ενός εξειδικευμένου επεξεργαστή του Digital Signal Processor (DSP)
  - κατάλληλου μετατροπέα (ADC)
  - κατάλληλου μετατροπέα (DCA)
  - Synthesizer
  - μνήμη RAM
  - μνήμη wavetable ROM
  - στο πίσω μέρος της κάρτας υπάρχουν δύο έξοδοι ήχου και μία θύρα MIDI



# Λογισμικό (1)

- Προγράμματα επεξεργασίας ήχου κυματομορφής. Οι βασικές δυνατότητες αυτών των εργαλείων είναι:
  - εισαγωγή ήχου
  - αντιγραφή, αποκοπή, επικόλληση ηχητικού τμήματος
  - ρύθμιση έντασης και συχνότητας
  - αλλαγή επιπέδων κβάντωσης
  - αλλαγή χρονικής διάρκειας
  - μείξη ήχων
  - απαλοιφή θορύβου
  - εισαγωγή διαφόρων εφέ
  - συμπίεσης

# Λογισμικό (1)



# Λογισμικό (2)

- Προγράμματα επεξεργασίας ήχου MIDI.
  - Αναπαριστούν τον ήχο ως παρτιτούρα και εξυπηρετούν ένα μουσικό όπως ένα σχεδιαστικό πρόγραμμα εξυπηρετεί ένα γραφίστα
  - Δίνουν την δυνατότητα να τοποθετηθούν νότες, παύσεις και άλλα σύμβολα της μουσικής σημειογραφίας πάνω στο πεντάγραμμο και να καθορίσουν τον τρόπο με που θέλει να παιχτεί κάθε μία νότα
- Δυνατότητα να ακούσει άμεσα ο δημιουργός την σύνθεσή του
- Δυνατότητας ηχογράφησης και επεξεργασίας ηχητικού σήματος από περιφερειακές πηγές
- Για την δημιουργία και επεξεργασία αυτών των αρχείων MIDI χρειάζονται ένα ειδικό λογισμικό που λέγεται MIDI χρονιστής (MIDI sequencer) - Cakewalk, Cubase Audio, MidiSoft Studio, Master Tracks κ.α.

# Λογισμικό (2)

The screenshot displays the Music MasterWorks software interface. The main window title is "Music MasterWorks" and the file path is "C:\Program Files\MusicMasterWorks\AnthemAustralia.mid". The interface is divided into two main sections, each with its own control bar and musical staff.

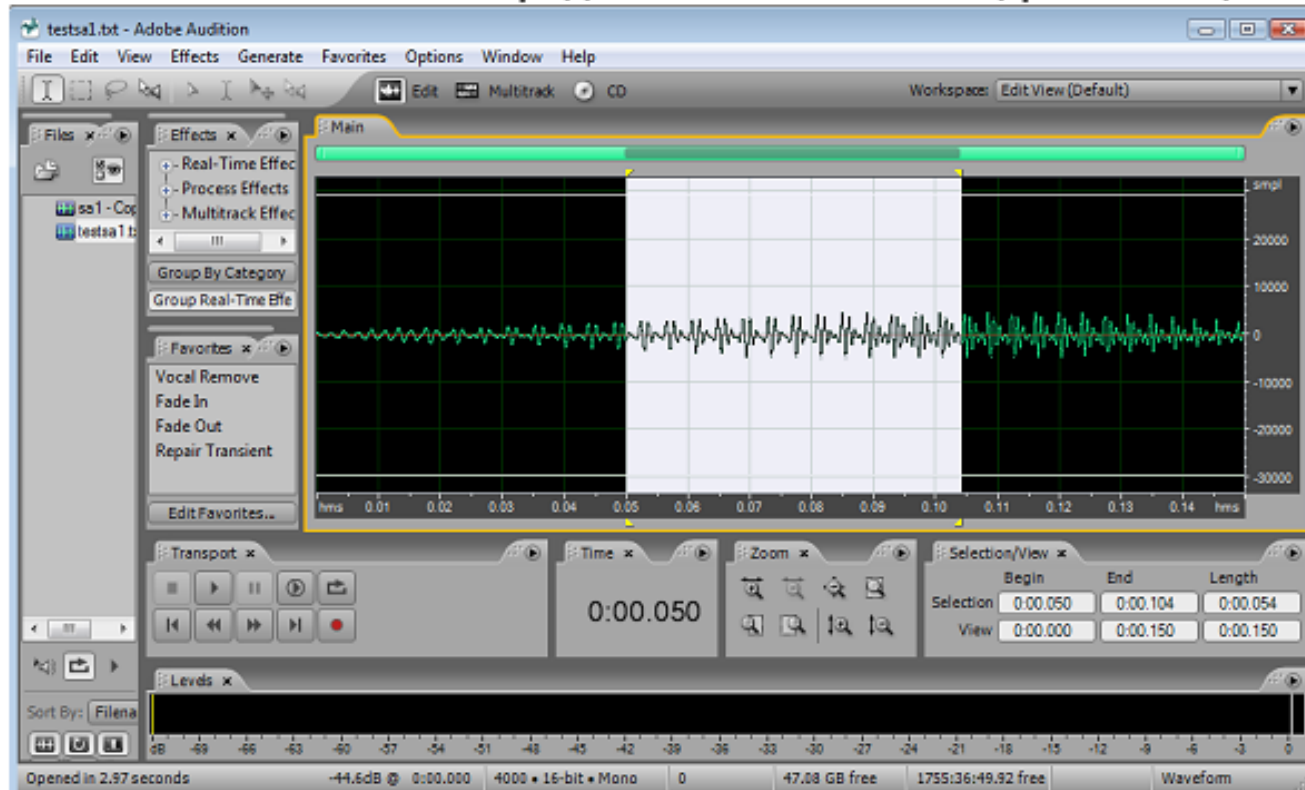
The top section is for "Track 1" (Trombone). The control bar includes a "Staff" dropdown, a "Filter: Track 1" field, and a "Trombone" instrument selection dropdown. The musical staff shows a score in 4/4 time, with measures 1 through 6. The bottom staff of this section shows the bass line.

The bottom section is for "Track 2" (String Ensemble 1). The control bar includes a "Staff" dropdown, a "Filter: Track 2" field, and a "String Ensemble 1" instrument selection dropdown. The musical staff shows a score in 4/4 time, with measures 1 through 6. The bottom staff of this section shows the bass line.

At the bottom of the window is a playback control bar with various icons for play, stop, and other functions. There are also two "Click to Buy" buttons and a "G4" label.

# Adobe Audition

– Adobe audition: <http://www.adobe.com/products/audition/>



# Τύποι αρχείων ήχου

<b>.rif</b>	<b>RIFF</b>	Microsoft υποστηρίζει αρχεία ψηφιακού ήχου WAV, MIDI
<b>.wav</b>	<b>Wave</b>	Αποτελεί πρότυπο αποθήκευσης ψηφιακού ήχου και είναι υποσύνολο του προτύπου RIFF
<b>.mid</b>	<b>MIDI</b>	Διεθνές πρότυπο για την αποθήκευση μουσικών αρχείων MIDI
<b>.aif</b>	<b>AIFF</b>	Από την Apple αλλά υποστηρίζεται και από άλλες πλατφόρμες, υποστηρίζει δειγματοληψία 32 bit
<b>.rmi</b>	<b>RMI</b>	Από την Microsoft για την υποστήριξη αρχείων MIDI
<b>.mp3</b>	<b>MPEG-layer3</b>	Πρότυπο συμπίεσης αρχείων ήχου που χρησιμοποιείται κυρίως για την διακίνηση αρχείων μουσικής στο Διαδίκτυο
<b>.wma</b>	<b>WMA</b>	Windows Media Audio από την Microsoft με στόχο την υποστήριξη απωλεστικής συμπίεσης
<b>.ra</b>	<b>Real Audio</b>	Προορίζεται για την άμεση αναπαραγωγή ήχων μέσω του Διαδικτύου

# Streaming Audio

- Το streaming audio προσομοιάζει την ραδιοφωνική μετάδοση:
  - Ο ήχος μεταδίδεται μέσω του τηλεπικοινωνιακού δικτύου και εκτελείται καθώς φτάνει στον προορισμό του, χωρίς να απαιτείται η αποθήκευσή του
  - Η Real Networks ήταν η πρώτη που υλοποίησε επιτυχώς streaming audio κωδικοποίηση (format) την Real Audio

# Περιεχόμενα ενότητας

- Φυσικά χαρακτηριστικά του ήχου
- Αντίληψη του ήχου
- Ψηφιοποίηση ήχου
- MIDI
- Συμπίεση αρχείων ήχου
- Επεξεργασία ήχου
- Streaming Audio



# Τέλος Ενότητας



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης