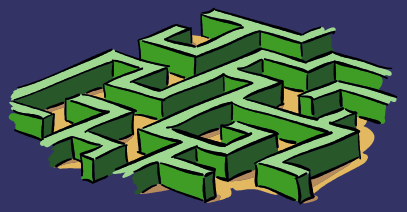
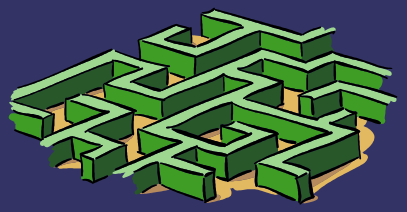
Από τις τριγωνομετρικές συναρτήσεις στο Mp3



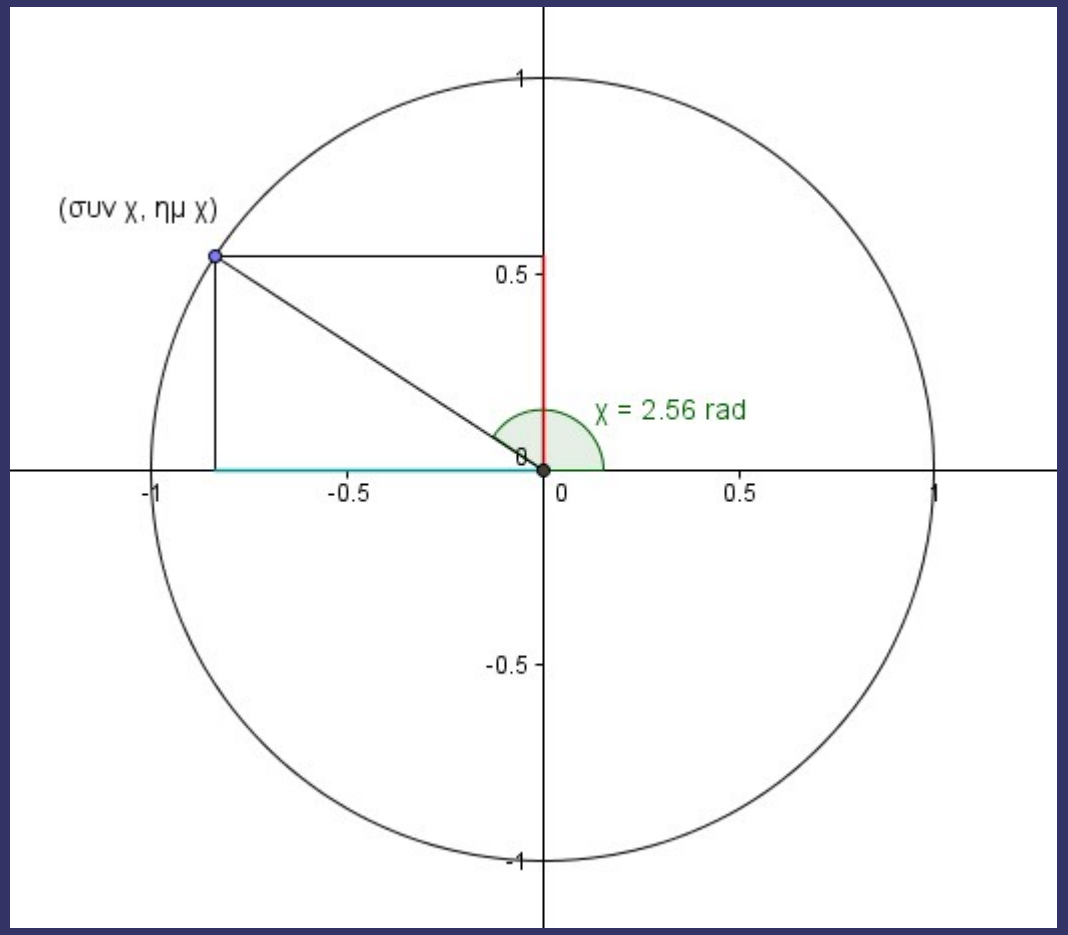
Επαμεινώνδας Διαμαντόπουλος, 1ο ΓΕΛ Ξάνθης, 2013 - 2014

Εισαγωγή

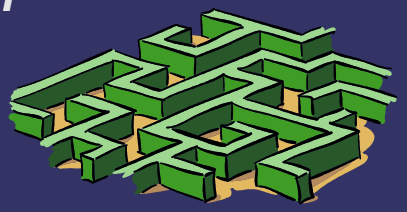
Οι περισσότεροι μαθητές δεν γνωρίζουν πως μία από τις περισσότερο αγαπημένες τους συνήθειες που είναι η ανταλλαγή τραγουδιών στο διαδίκτυο (ή και στο κινητό τους τηλέφωνο) βασίζεται σε τεχνολογία που αναπτύχθηκε πάνω σε ένα από τα μαθησιακά αντικείμενα της Β' Λυκείου που τους δυσκολεύει πολύ : Στις τριγωνομετρικές συναρτήσεις!



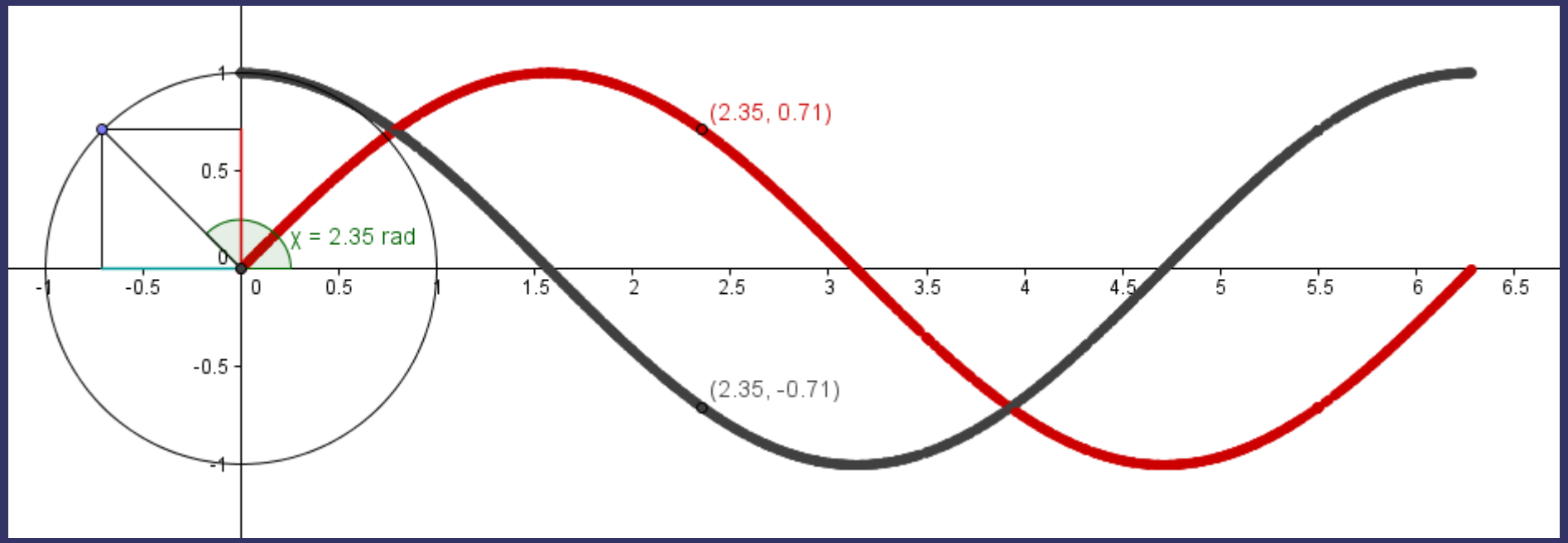
Το ημίτονο μίας γωνίας



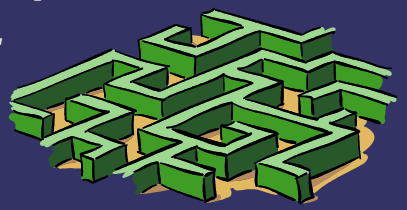
Ορισμός : Το ημίτονο και το συνημίτονο μίας γωνίας προσδιορίζεται από τις προβολές της μοναδιαίας ακτίνας στους άξονες y'y και x'x αντίστοιχα.



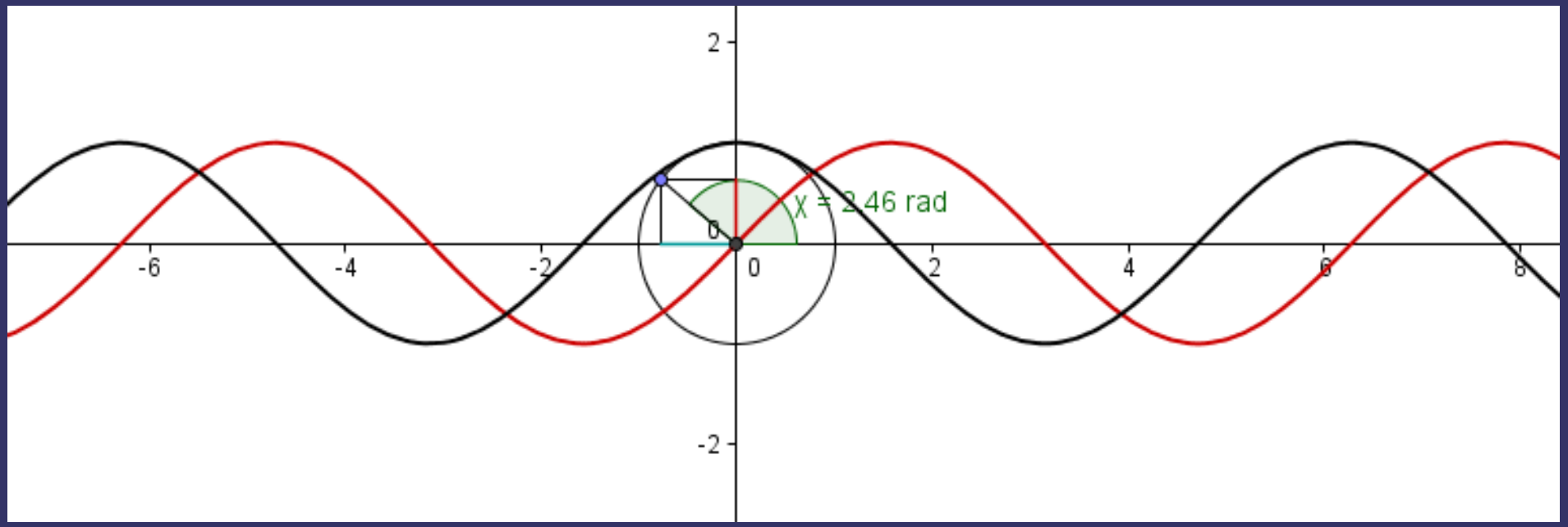
Η γραφική παράσταση της συνάρτησης του ημιτόνου



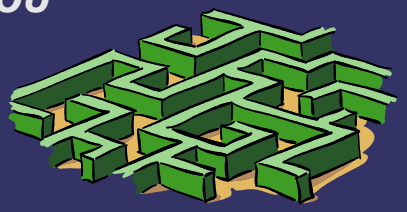
Η γωνία x αναλαμβάνει το ρόλο της τετμημένης ενώ η τιμή του ημιτόνου/συνημιτόνου το ρόλο της τεταγμένης και παίρνουμε τις γραφικές παραστάσεις των συναρτήσεων f(x) = ημ x και g(x) = συν x...



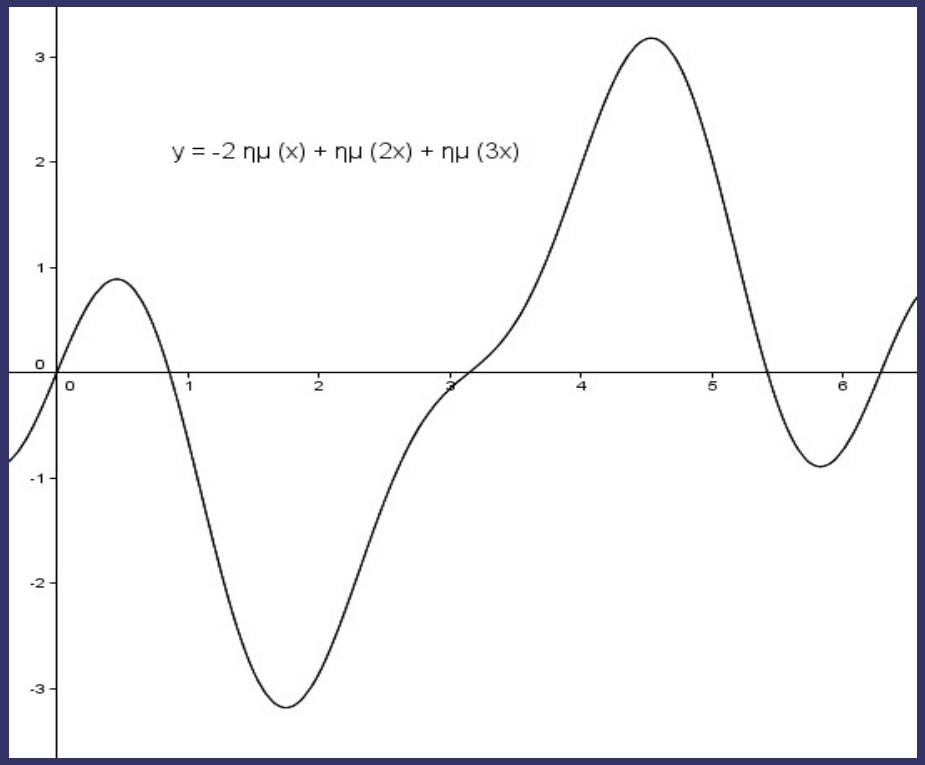
Η γραφική παράσταση της συνάρτησης του ημιτόνου

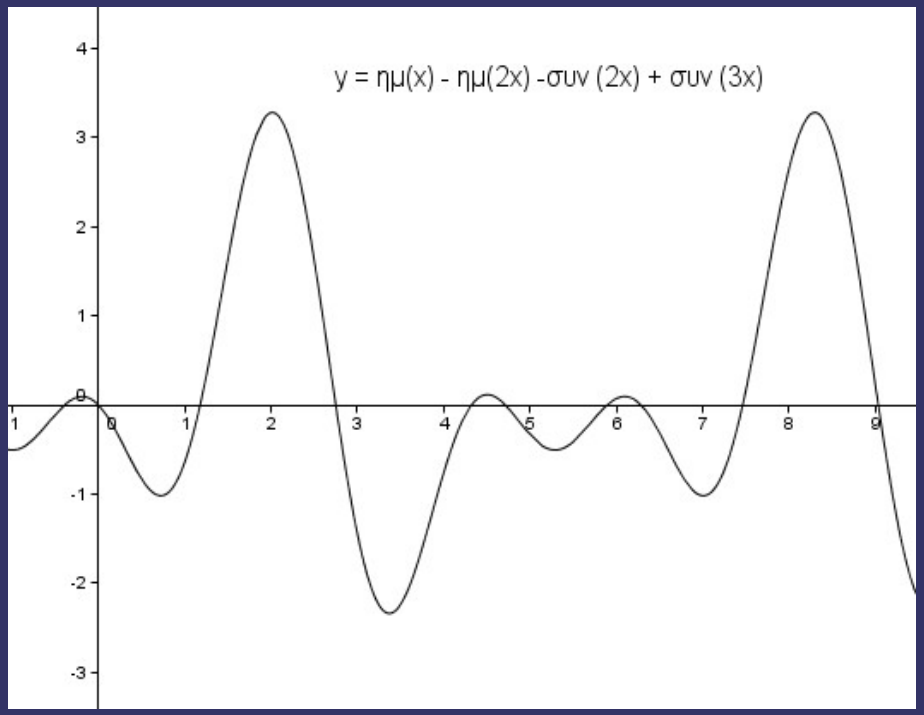


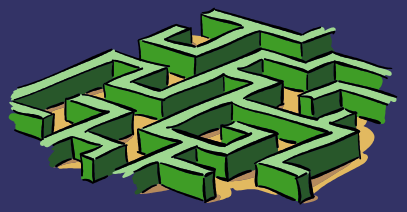
... οι οποίες ορίζονται για όλα τα πραγματικά x ενώ λόγω του ορισμού τους, επαναλαμβάνονται περιοδικά με περίοδο 2π.



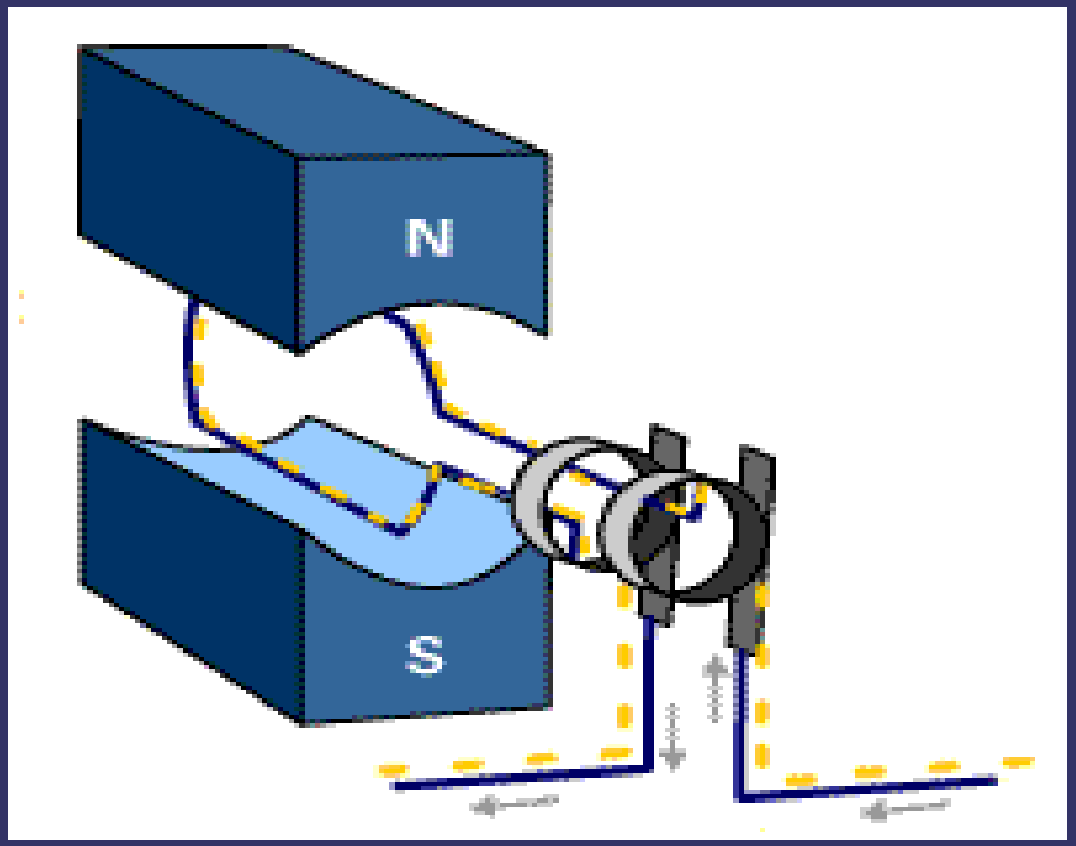
Με λίγη περισσότερη προσπάθεια μπορεί να γίνει η γραφική παράσταση ενός γραμμικού συνδυασμού συναρτήσεων ημιτόνου ή/και συνημιτόνου με διάφορους συντελεστές και διάφορες φάσεις...





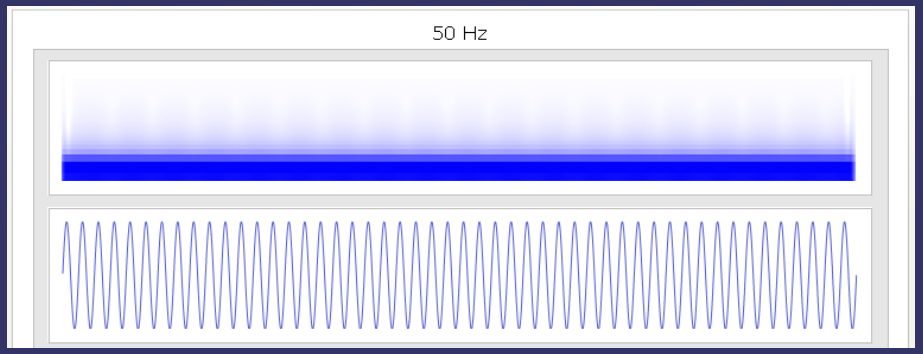


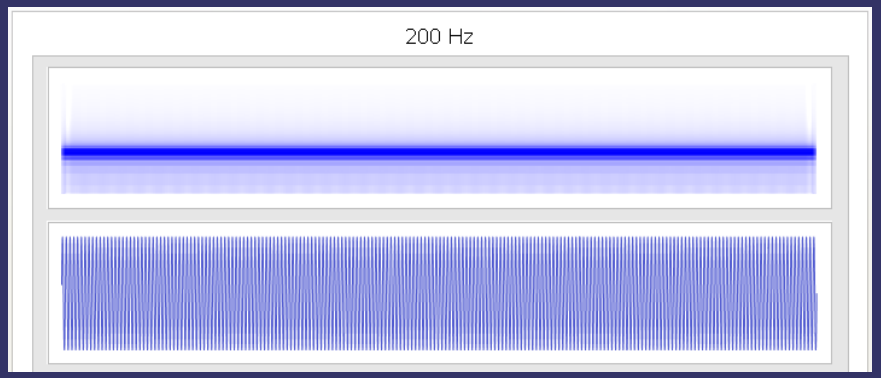
Εφαρμογές των τριγωνομετρικών συναρτήσεων

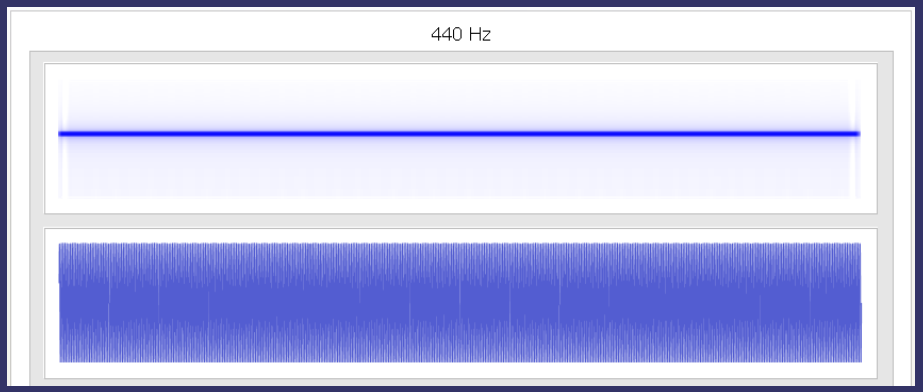


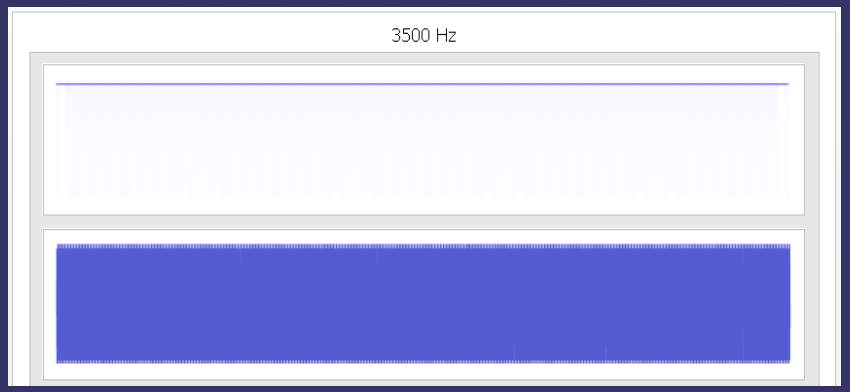
Τέλη του 19ου αιώνα, ήταν σαφές πως η συνάρτηση του ημιτόνου είναι η κατάλληλη για την περιγραφή περιοδικών φαινομένων όπως για παράδειγμα το εναλλασσόμενο ρεύμα. Πράγματι, η θέση/ταχύτητα των ηλεκτρονίων στην παραπάνω διάταξη περιγράφεται από τη συνάρτηση του ημιτόνου. (x είναι ο χρόνος, μόνο θετικός)

Εφαρμογές των τριγωνομετρικών συναρτήσεων

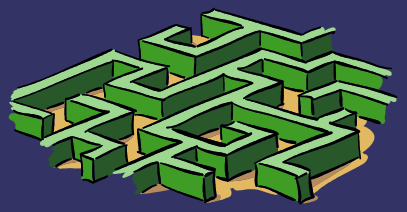




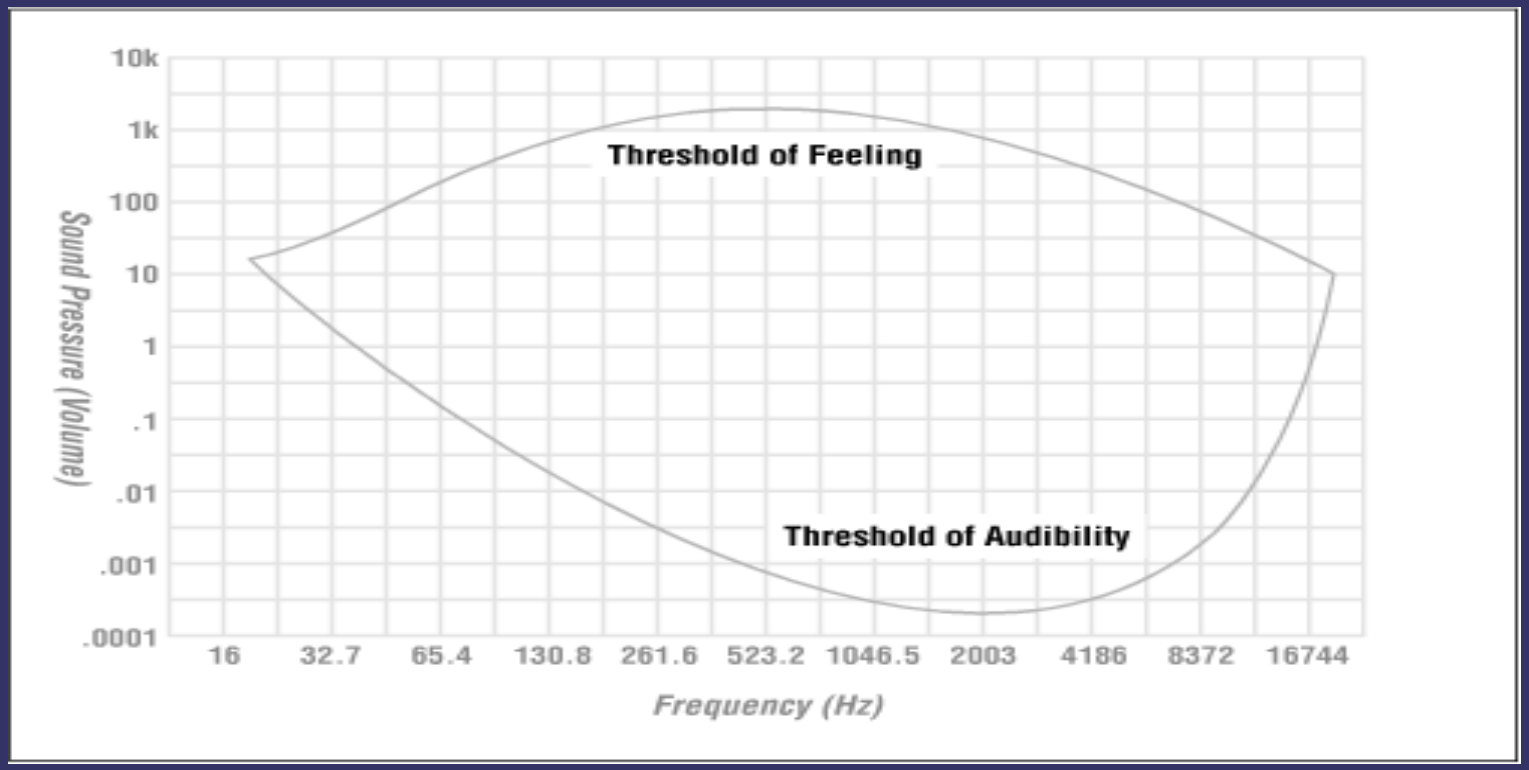


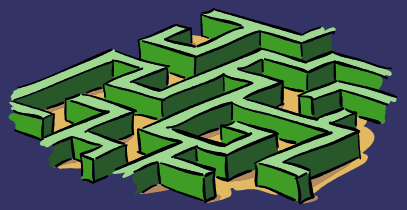


Μία ακόμα φυσική εφαρμογή ήταν η αναπαράσταση των ηχητικών κυμάτων! Πράγματι, ο ήχος είναι ένα κύμα το οποίο κάνει το τύμπανο του αυτιού να πάλλεται, γεγονός που δημιουργεί νευρικές ώσεις στον εγκέφαλο, ώσεις που αναγνωρίζονται ως ήχος.



Αν και οι συχνότητες που ακούστηκαν είχαν ίδια ένταση, δεν δημιουργούν την ίδια εντύπωση στο ανθρώπινο αυτί. Το γεγονός αυτό δεν είναι τυχαίο. Η ψυχοακουστική είναι ο επιστημονικός τομέας που ασχολείται με την έρευνα στο πεδίο αυτό.

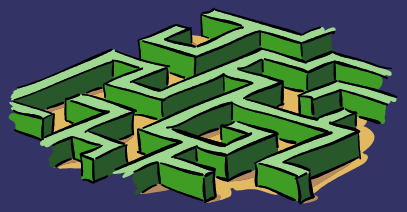


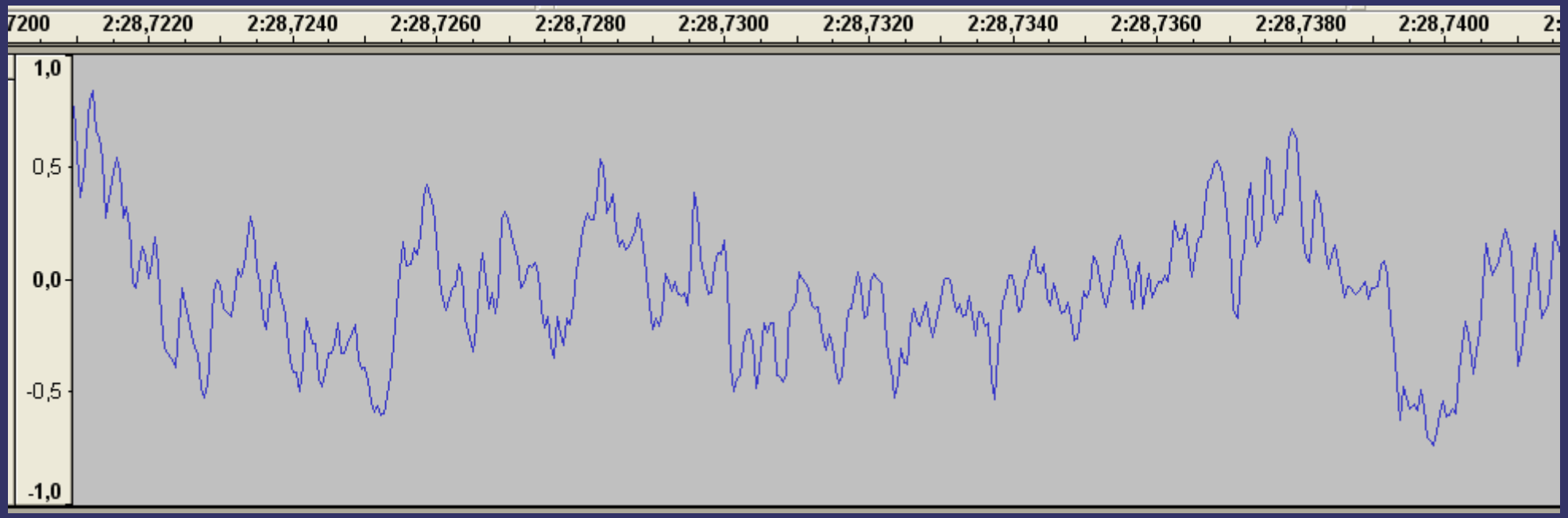


Περαιτέρω, το 1894, ο Alfred Marshall Mayer ανάφερε πως ένας ήχος μπορεί να γίνει μη αισθητός στο ανθρώπινο αυτί αν ταυτόχρονα εκπέμπεται ένας τόνος μικρότερης συχνότητας. Το 1959, ο Richard Ehmer περιέγραψε ένα πλήρες σύνολο από ακουστικές καμπύλες σχετικά με το παραπάνω φαινόμενο.

Βρέθηκε πως ο άνθρωπος δεν μπορεί να διαχωρίσει ήχους με μικρή διαφορά στη συχνότητά τους και πως όλο το ακουστικό εύρος μπορεί να χωριστεί σε 24 τέτοιες ζώνες.

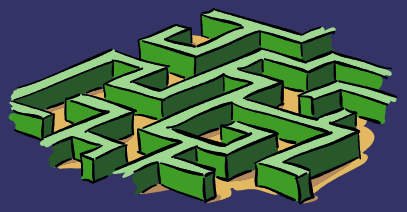
Σε κάθε μία τέτοια ζώνη μπορεί να επιλεχθεί μία συχνότητα που κυριαρχεί και καλύπτει όλες τις υπόλοιπες της αντίστοιχης ζώνης. Αν αυτή παρουσιαστεί ως αντιπρόσωπος όλων των υπολοίπων της αντίστοιχης ζώνης, τότε ο ακροατής δεν θα καταλάβει τη διαφορά.



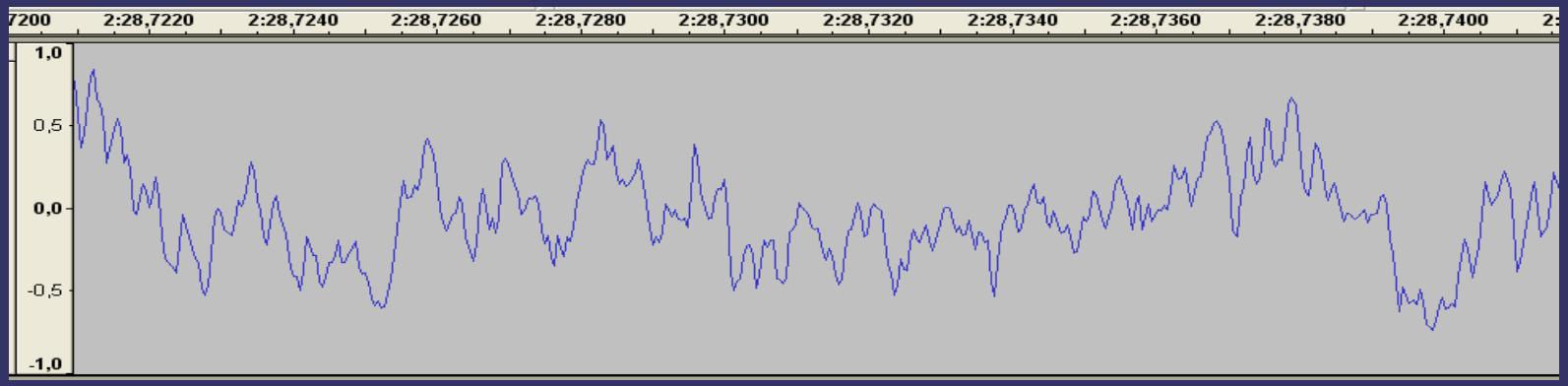


Εντωμεταξύ είχαν ήδη καταφέρει να καταγράψουν το γράφημα ενός ηχητικού κύματος,

... ενώ το ζήτημα που εκκρεμούσε ήταν η αποθήκευση του ήχου/κύματος σε κάποιο μέσο.

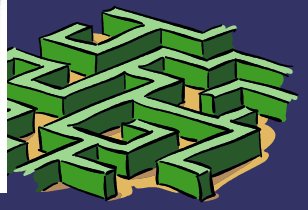


Εποχή του πικ – απ (1920 - 1990)...

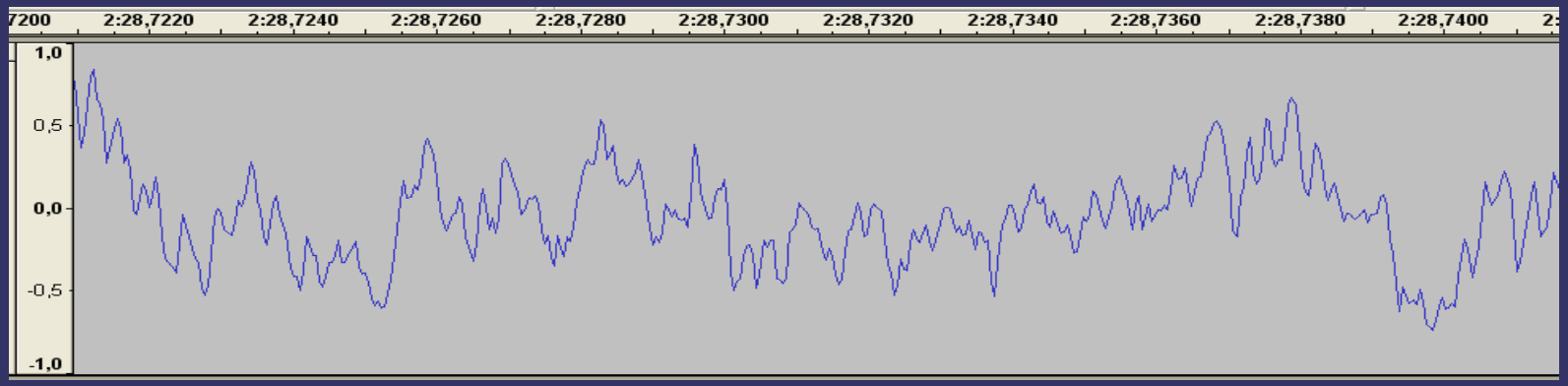


Η καταγραφή ενός ηχητικού σήματος όπως το παραπάνω έγινε καταγράφοντας το παραπάνω κύμα ως αυλάκι μεταβλητού βάθους σε έναν δίσκο βινυλίου.

Για την επανεκτέλεση του αρκούσε ένα πικ-απ!



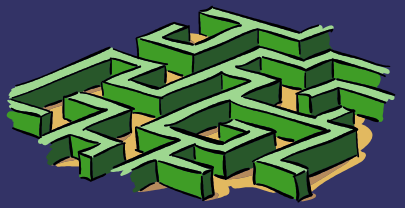
... και της μαγνητικής κασέτας (1962 - σήμερα)



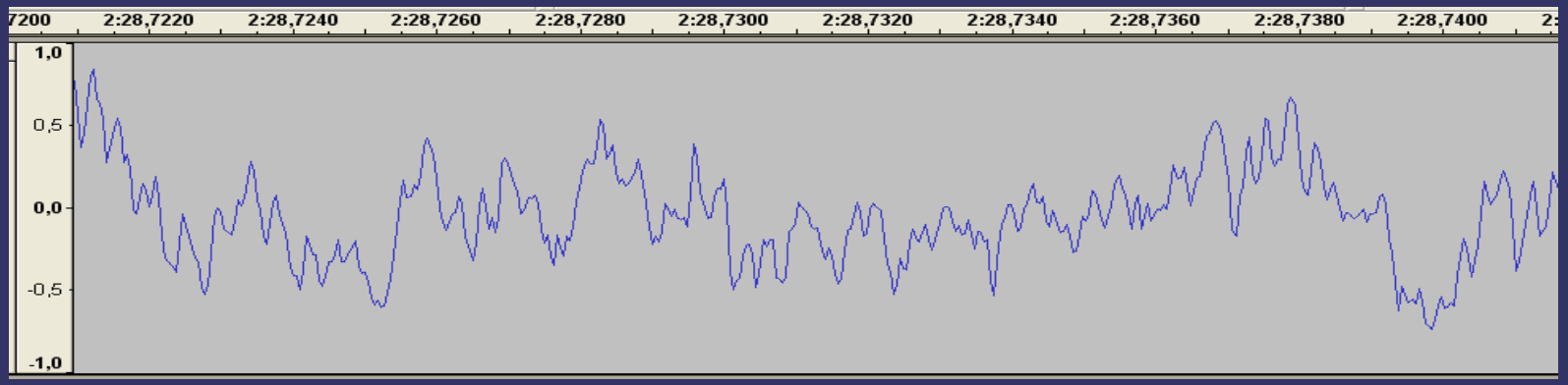
Λίγο αργότερα από το δίσκο βινυλίου, εμφανίστηκε η μαγνητική κασέτα.

Η κασέτα, όπως τη γνωρίζουμε πατενταρίστηκε το 1964. Το πλάτος (δηλαδή το ύψος) του κύματος καταγραφόταν πάνω στην επιφάνεια της με έναν μικρό ηλεκτρομαγνήτη καθώς περνούσε από μπροστά του η ταινία με σταθερή ταχύτητα.

Η ανάγνωση του κύματος γινόταν με ένα κασετόφωνο!



Η ψηφιακή εποχή ... (1980 - σήμερα)

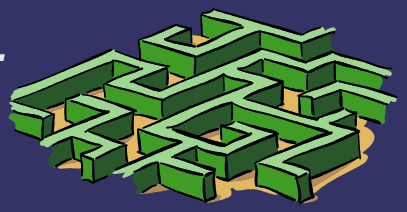


Το 1976 η Sony, το 1979 η Philips και στο τέλος του έτους και οι δύο μαζί ανέπτυξαν το CDROM, δηλαδή την ψηφιακή (διακριτή) αναπαράσταση ενός συνεχούς ηχητικού σήματος, αλγόριθμος που έγινε γνωστός με τα αρχικά PCM (Pulse Code Modulation).

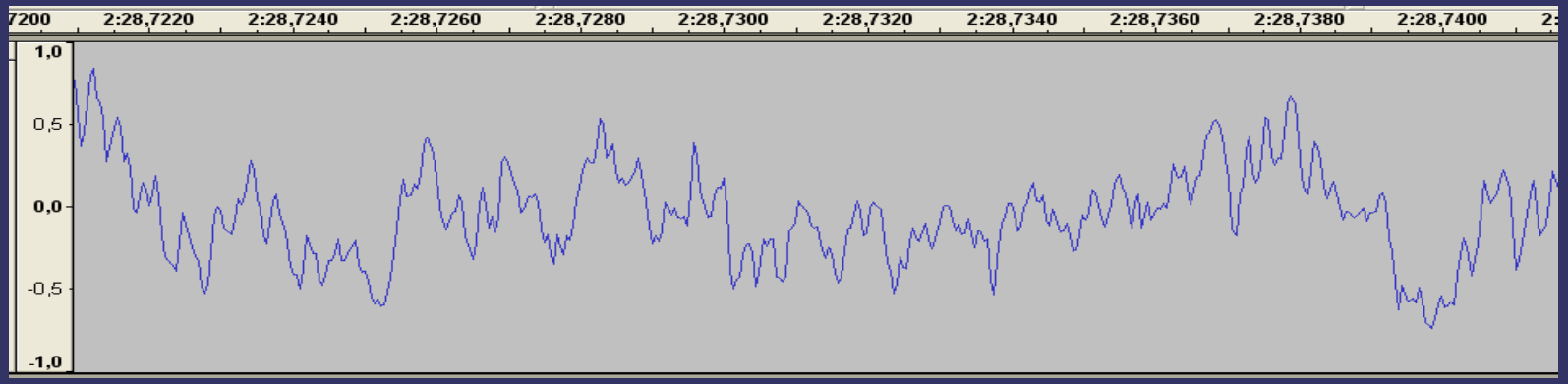
Με την τεχνολογία αυτή γίνεται δειγματοληψία στο κύμα με ρυθμό 44,1 KHz (δηλαδή 44.100 φορές το δευτερόλεπτο) και από κάθε ένα τέτοιο δείγμα καταγράφεται και αποθηκεύεται το ύψος του κύματος με ακρίβεια 16bit = 4 byte.

Ένας απλός υπολογισμός δείχνει πως ένα λεπτό μουσικής με αυτόν τον τρόπο απαιτεί 60* 44.100 * 4 = 10.584.000 bytes ή απλά 10,6 Mb αποθηκευτικού χώρου.

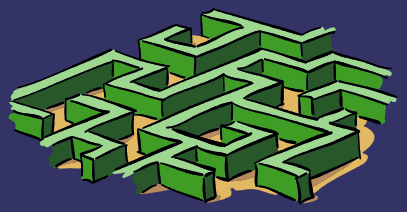
Έτσι, σε ένα CDROM (700 Mb) αποθηκεύονται περίπου 67 λεπτά ήχου.



Το κίνητρο

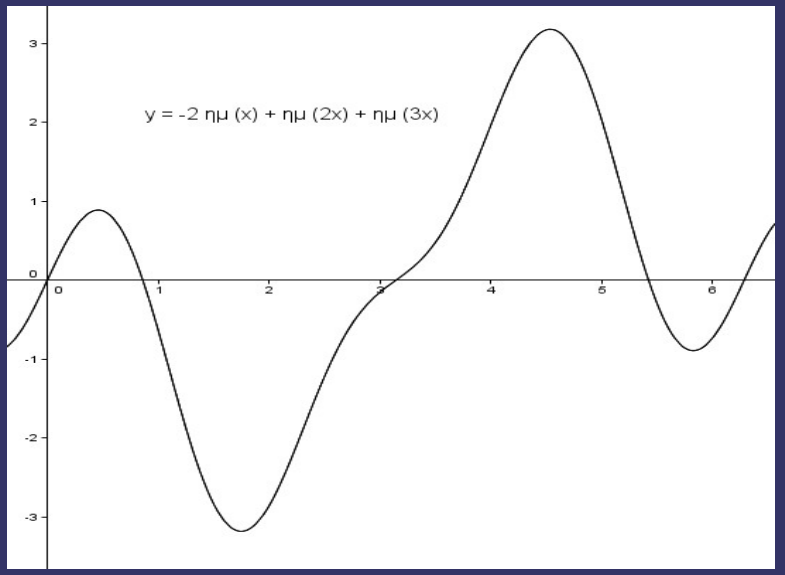


Ο διαμοιρασμός αρχείων μουσικής από το δίκτυο επέβαλε την ανάπτυξη κάποιου αλγορίθμου συμπίεσης μουσικής. Οι παραδοσιακές μέθοδοι (π.χ. PK-ZIP και Stuffit) είχαν μικρό αποτέλεσμα (συμπίεση γύρω στο 10%), ενώ αλγόριθμοι που δημιουργήθηκαν ειδικά για αρχεία που είχαν δημιουργη-θεί για τη μέθοδο PCM έφταναν έως και το 60%. Ωστόσο, το κίνητρο ήταν ακόμα μεγάλο για μία καλύτερη μέθοδο.

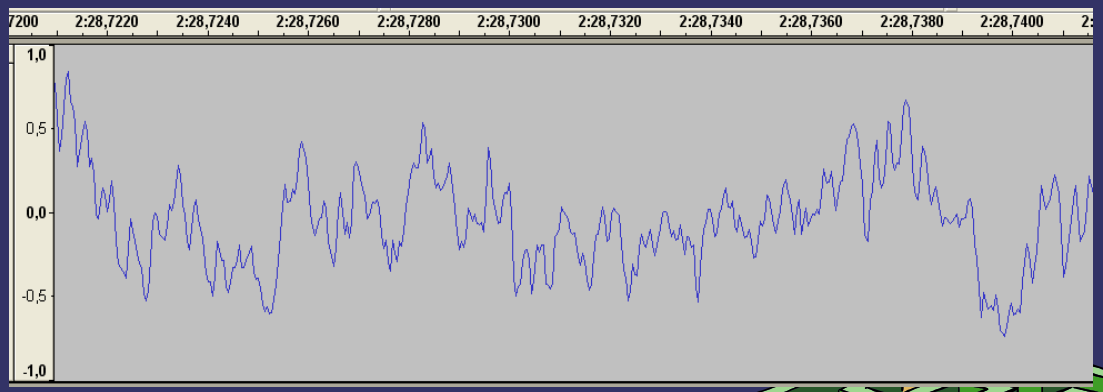


Κάπου και κάπως έτσι ιδέα δεν άργησε να έρθει...

Μήπως χρησιμοποιώντας γραμμικούς συνδυασμούς του ημιτόνου...



... μπορούμε να αναπαραστήσουμε ένα οποιοδήποτε ηχητικό κύμα;



Οι δυνατότητες του ημιτόνου και του συνημιτόνου

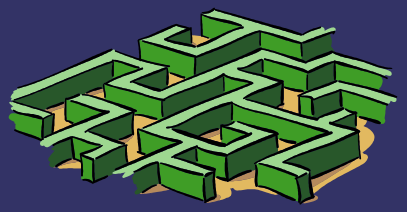
Ο Γάλλος μαθηματικός Joseph Fourier στις αρχές του 19ου αιώνα απέδειξε πως κάθε περιοδική συνάρτηση f μπορεί να αναλυθεί σε ένα άπειρο άθροισμα από όρους ημιτόνου και συνημιτόνου :

$$f(x)=\frac{1}{2}\beta_0+\sum_{\nu=1}^{\infty}\alpha_\nu \eta\mu(\nu x)+\beta_\nu \sigma\upsilon\nu(\nu x)=\frac{1}{2}\beta_0+\alpha_1\eta\mu(x)+\beta_1\sigma\upsilon\nu(x)+\alpha_2\eta\mu(2x)+\beta_2\sigma\upsilon\nu(2x)+\ldots$$

όπου

$$\alpha_\nu=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,\eta\mu\,\nu x\,dx \qquad \beta_\nu=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,\sigma\upsilon\nu\,\nu x\,dx$$

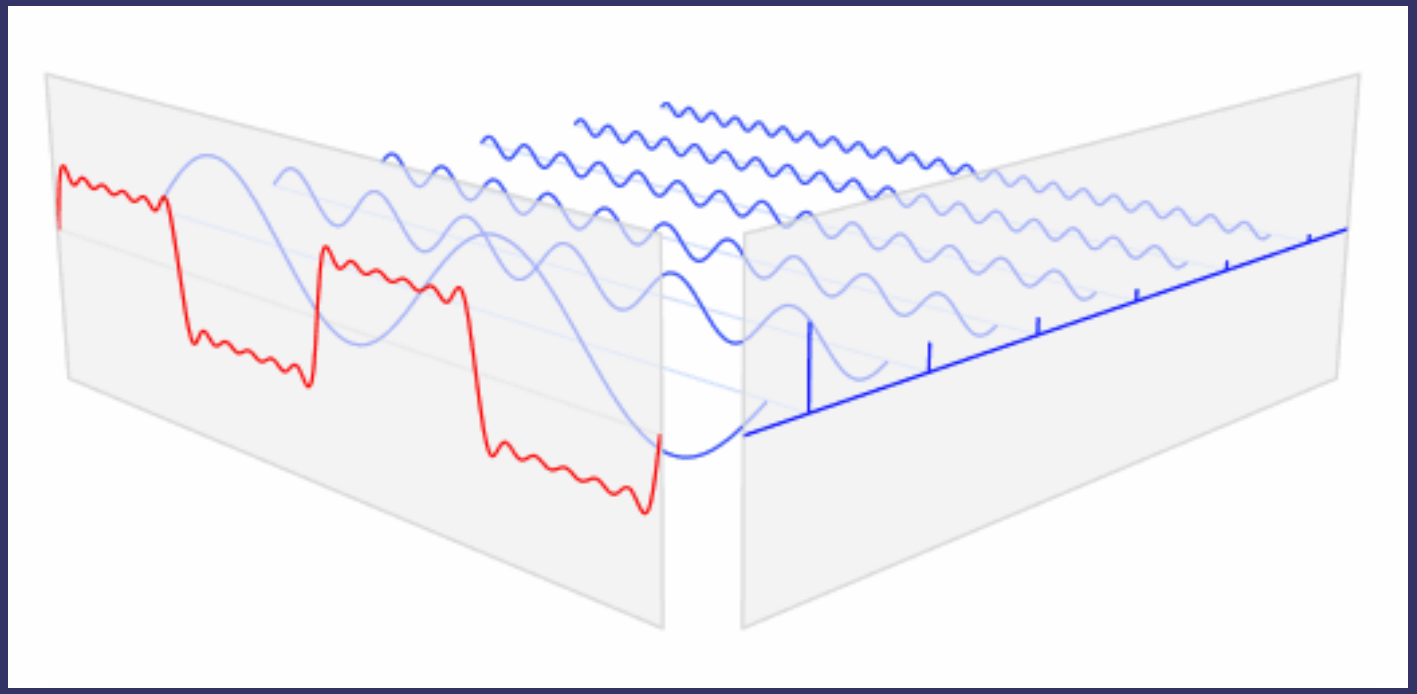
Ισοδύναμα, μπορεί να πει κάποιος πως κάθε περιοδική συνάρτηση αντιπροσω-πεύεται πλήρως από τους συντελεστές $\{(\alpha_\nu, \beta_\nu), \nu=0,1,\ldots\}$ του αναπτύγματος Fourier.



Οι δυνατότητες του ημιτόνου και του συνημιτόνου

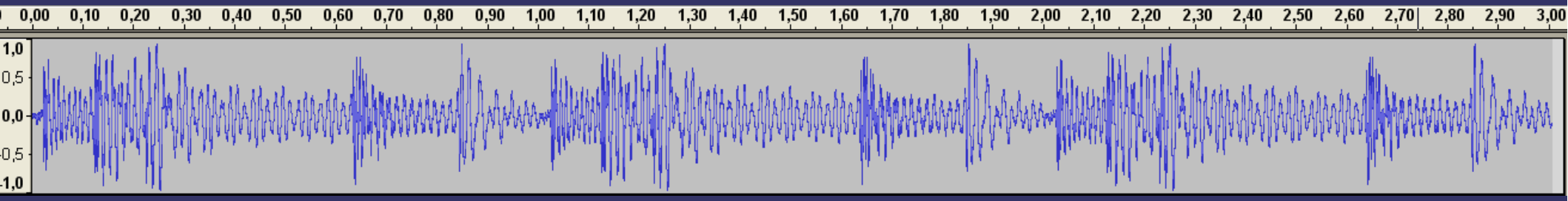
Παραστατικά :

(πηγή: http://nautil.us/blog/the-math-trick-behind-mp3s-jpegs-and-homer-simpsons-face



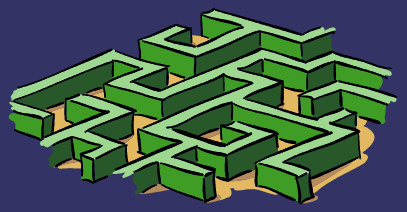
Οι δυνατότητες του ημιτόνου και του συνημιτόνου

Μεταφρασμένο σε όρους ηχητικών σημάτων, ο Fourier έδειξε πως κάθε περιοδικό ηχητικό σήμα οποιασδήποτε πολυπλοκότητας...



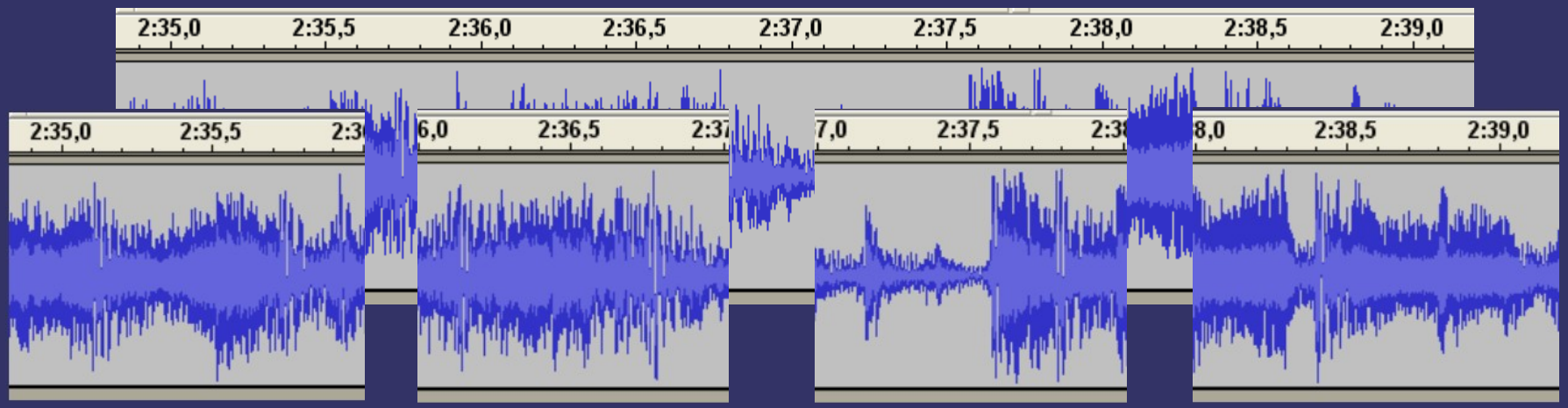
... αναλύεται σε ένα γραμμικό συνδυασμό απλών ημιτονοειδών σημάτων και αναπαριστάται πλήρων από τους αντίστοιχους συντελεστές.

Στην ιστοσελίδα http://falstad.com/fourier/e-sawtooth.html ο ενδιαφερόμενος αναγνώστης μπορεί να βρει μία εντυπωσιακή εφαρμογή αναπαράστασης!

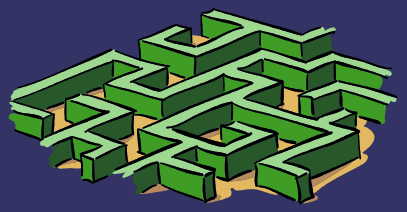


Οι δυνατότητες του ημιτόνου και του συνημιτόνου

Η γενίκευση από τα περιοδικά κύματα σε κάθε ηχητικό κύμα έρχεται φυσικά : Κάθε ηχητικό σήμα μπορεί να θεωρηθεί ως μία συνεχόμενη ακολουθία από περιοδικά ηχητικά σήματα.



Με τον τρόπο αυτό, κάθε μέρος του αρχικού ηχητικού σήματος αναπαρίσταται από ένα σύνολο συντελεστών $\{(\alpha_v, \beta_v), v=0,1,...\}$, και με τη σειρά τους όλοι οι συντελεστές αυτοί αναπαριστούν το αρχικό ηχητικό κύμα.



Αλγόριθμός Mp3

Το αρχικό κύμα διαχωρίζεται σε frames, διάρκειας 26 ms.

Για κάθε ένα frame ...

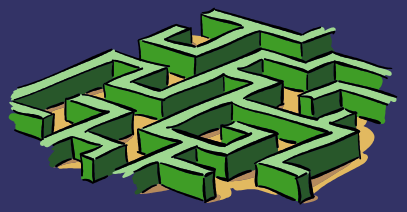
... υπολογίζεται η ανάλυση σε σειρά Fourier,

... διαιρείται το εύρος συχνοτήτων σε 32 μπάντες, κάθε μια από τις οποίες το ανθρώπινο αυτί τις ακούει ξεχωριστά και οι οποίες όλες μαζί αναπαριστούν το αρχικό κύμα.

Για κάθε μία από τις συνιστώσες αυτές κρατούνται μόνο οι 18 περισσότερο ισχυρές συχνότητες, βάσει γνωστών μοντέλων ψυχοακουστικής που είναι ενσωματωμένα στο λογισμικό κωδικοποίησης!

Παράδειγμα αυτής της επιλογής είναι η εγκεφαλική αναγνώριση δύο παράλληλων ηχητικών κυμάτων ως ένα μοναδικό που έχει συχνότητα το μέγιστο κοινό διαιρέτη των δύο αρχικών κυμάτων!

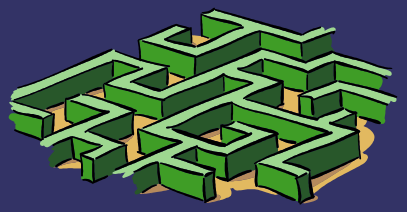
Επιπλέον, η ποιότητα σε κάθε ένα frame περιορίζεται από την απαίτηση για κάθε δευτερόλεπτο ήχου να αφιερώνονται 128Kbit = 16 Kbytes μνήμης.



Αλγόριθμός Mp3

Το κύμα που προκύπτει συμπιέζεται περαιτέρω με τον αλγόριθμό του Huffmann, μία τεχνική η οποία αναπαριστά επαναλαμβανόμενες συχνότητες με λιγότερα δεδομένα. (μειώνοντας για παράδειγμα το χώρο που καταλαμβάνουν τα διαστήματα σιωπής).

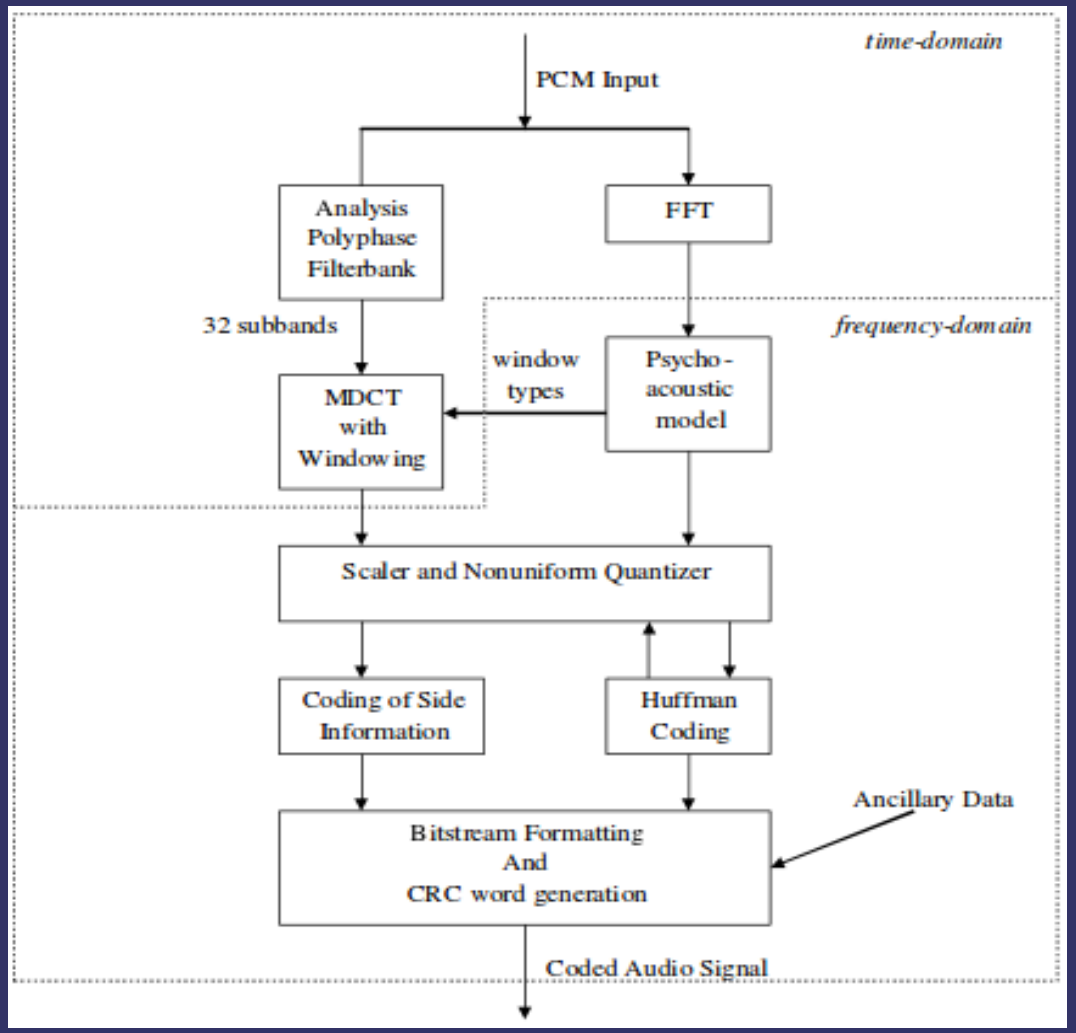
Το συνολικό αποτέλεσμα είναι για κάθε ένα λεπτό μουσικής να απαιτείται περίπου 0,5 με 1 Mbyte, κάτι που σημαίνει σχεδόν υποδεκαπλασιασμός του μεγέθους του αρχείου.

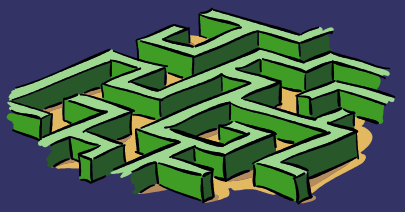


Αλγόριθμός Mp3

Η διαδικασία περιγράφεται στο παρακάτω διάγραμμα...

Πηγή: The Theory Behind Mp3, Rassol Raissi, December 2002





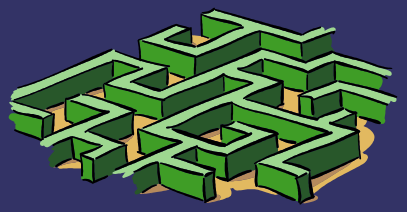
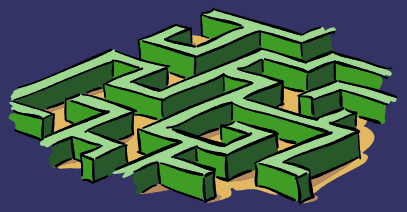
Αλγόριθμός Mp3

Ιστοσελίδες αναφοράς :

http://oreilly.com/catalog/mp3/chapter/ch02.html

http://www.guardian.co.uk/technology/2002/apr/04/internetnews.maths

http://el.wikipedia.org/wiki/MP3



Σας ευχαριστώ για την προσοχή σας!