Επεξεργασία Ήχου Φωνής
4η Διάλεξη

ΨΗΦΙΑΚΟ ΣΤΟΥΝΤΙΟ

ΓΕΝΙΚΕΣ ΔΟΜΕΣ ΚΑΙ ΚΑΤΗΓΟΡΙΕΣ ΣΥΣΤΗΜΑΤΩΝ

- Θα ασχοληθούμε με τις τεχνικές και τεχνολογίες που επιτρέπουν την κωδικοποίηση επεξεργασία και διανομή στον τελικό χρήστη – ακροατή ηχητικών δεδομένων, οι οποίες βασίζονται σε μεθόδους Ψηφιακής Επεξεργασίας Σήματος.

Αριθμητική αναπαράσταση σταθερού σημείου και βήμα κβαντισμού

- Ο ήχος έχει μετατραπεί σε σειρά αριθμών πως όμως κωδικοποιείται; Δύο τρόποι
- Με αριθμητική σταθερού σημείου (fixed-point number representation)
- Με αριθμητική κινητής υποδιαστολής (floating-point number representation).

Αναπαράσταση με Αριθμητική Σταθερού Σημείου

- Σήμα από $-S_{max}$ μέχρι S_{max} κανονικοποιήται στο -1 – 1

$$\hat{s}(k) = \frac{\hat{s}_d(k)}{\left|s_{\max}\right|}$$

και έχουμε σειρά της μορφής

$$\hat{s}(k) = -b_0 + b_1 2^{-1} + b_2 2^{-2} + \ldots + b_{N-1} 2^{-(N-1)}$$

$$-1 \le \hat{s}(k) \le 1 - 2^{-(N-1)}$$

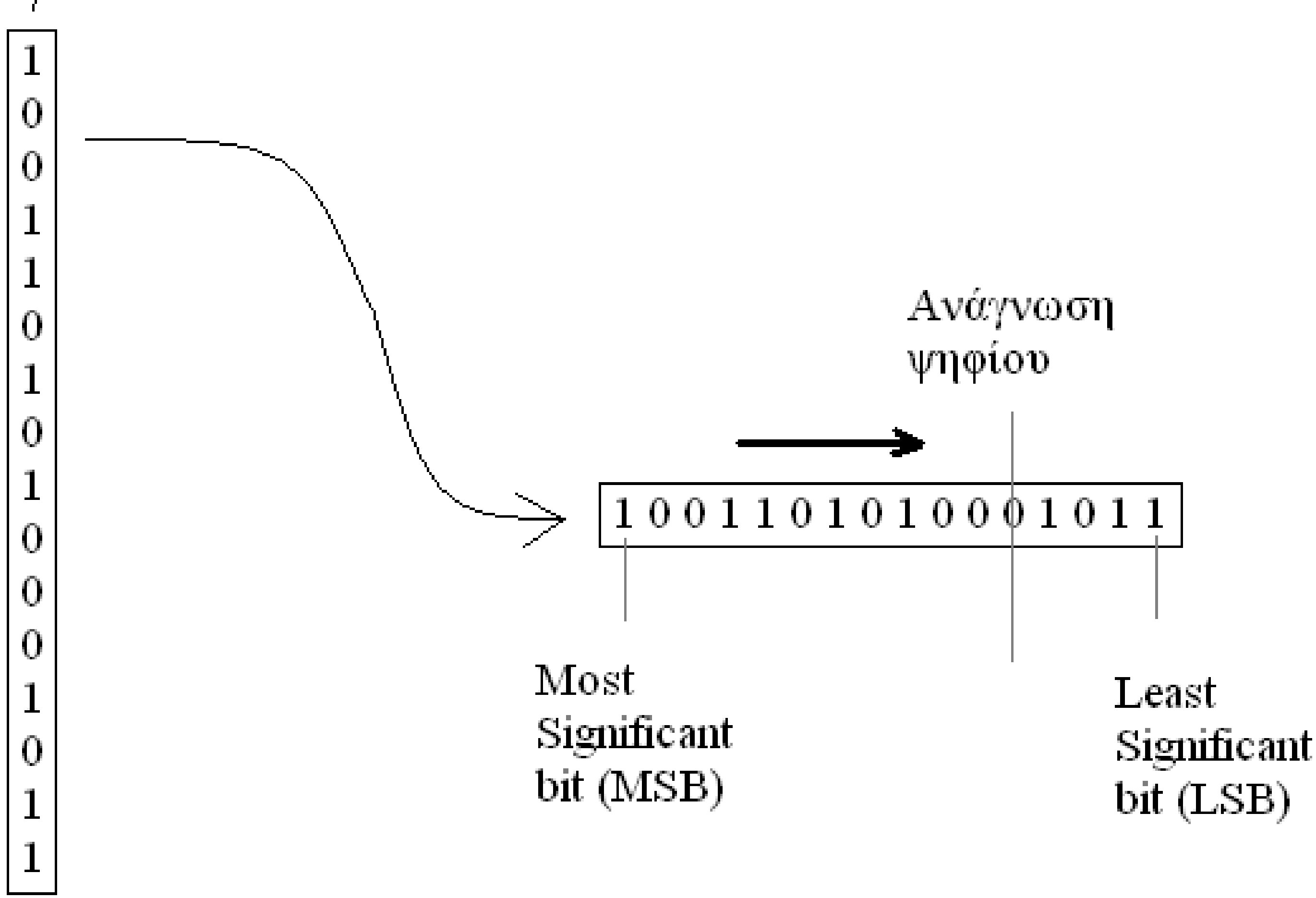
θα είναι δηλαδή ανάμεσα σε [1000…000] και [0111…111]

Το εύρος τιμών δηλαδή θα είναι $2\text{-}2^{-(N-1)} \sim 2$

Η διακριτική ικανότητα – βήμα κβαντισμού

$$\Delta = \frac{2}{2^N} = 2^{-(N-1)}$$

$$\hat{s}(k) = -1 + 0 \cdot 2^{-1} + 0 \cdot 2^{-2} + \ldots + 1 \cdot 2^{-15}$$



Παράδειγμα

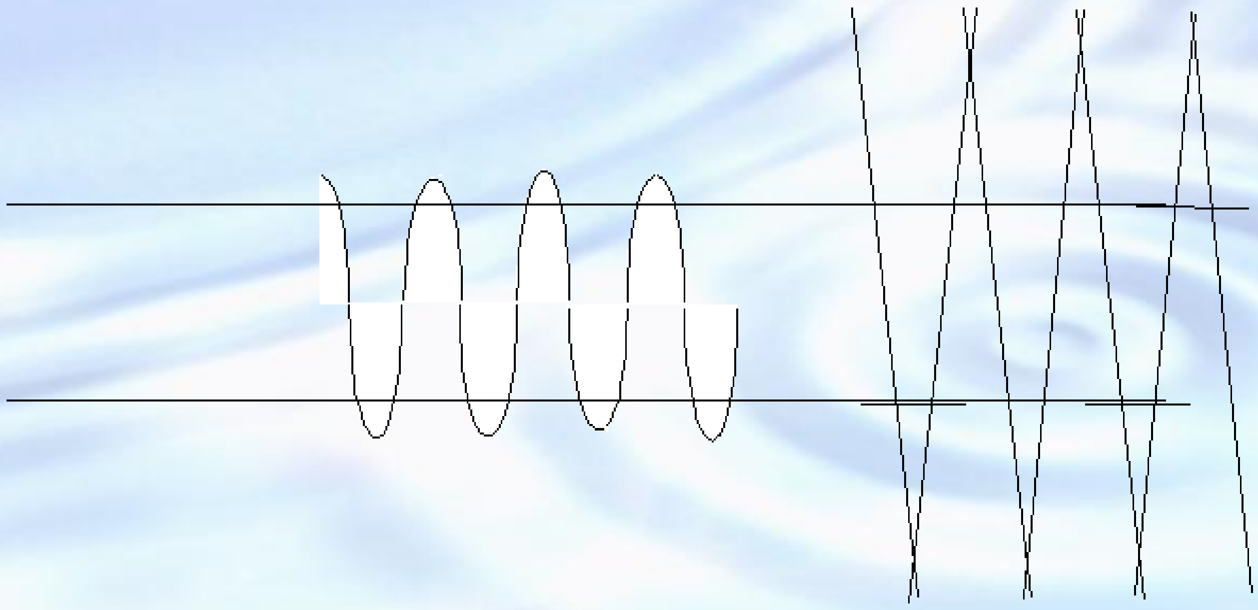
- Κάρτα ήχου δουλεύει στα ±5 V αν κάνουμε ψηφιοποίηση στα 8 bit. Πόσες στάθμες υπάρχουν στα 0-5 V.

a. Ποια η ελάχιστη μεταβολή τάσης για την αλλαγή του LSB.
b. Αν έχω σήμα πλάτους 1.5 V πόσες στάθμες εκμεταλλεύεται;
c. Αν έχω στιγμιαία τάση -1.2 V πως θα είναι αυτή αποθηκευμένη ψηφιακά ; Πώς θα ήταν αυτό αν είχαμε 16 bit;
d. Αν είχα δειγματοληψία 44 kHz Πόσο χώρο θα καταλάμβαν 102,4 sec στην πρώτη περίπτωση και πόσα στην δέυτερη.

a. -5 – 5 :256 , 0-5 :128 καταστάσεις άρα Δ=5/128=0.39 (βήμα κβάντισης)
b.1.5/0.39*2=77 (βήμα κβάντισης * 1.5)
c.-1.2 V →-1.2/5=0.24 για 8 bit 1 1 1 0 0 0 0 1.
Για 16 bit 1 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1
b.X=44000*102,4*8 (bits) Αν τα θέλω σε kB x/8/1024=4400 kB ~4.3 Mb

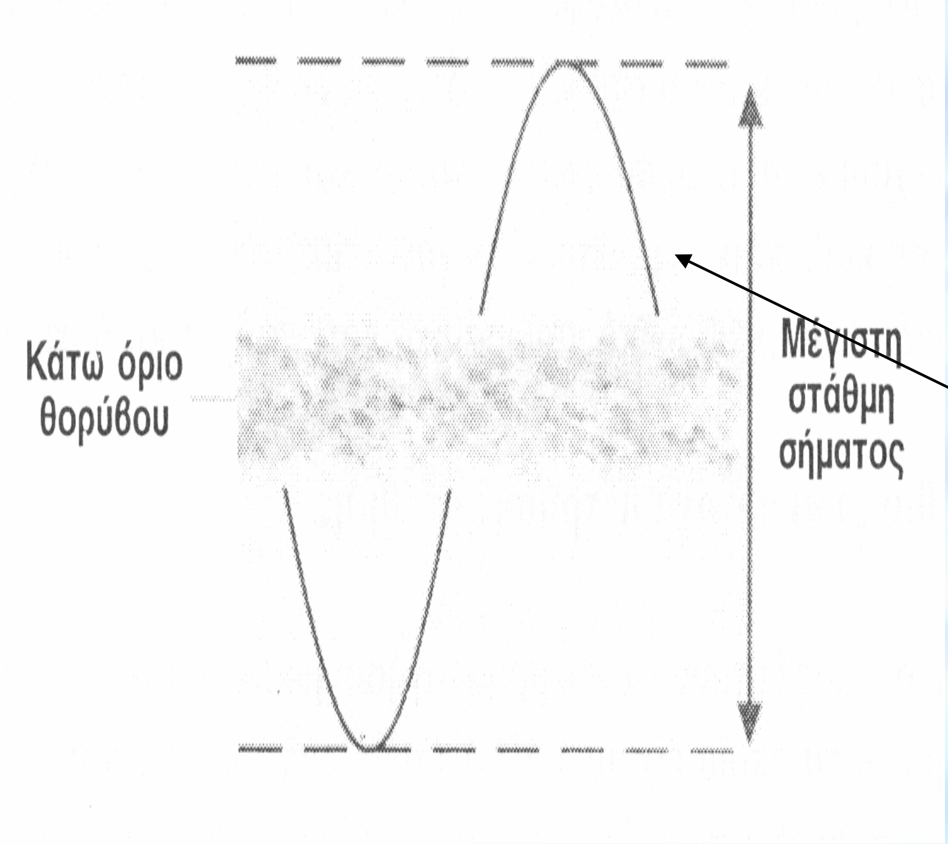
Ψαλιδισμός (clipping)

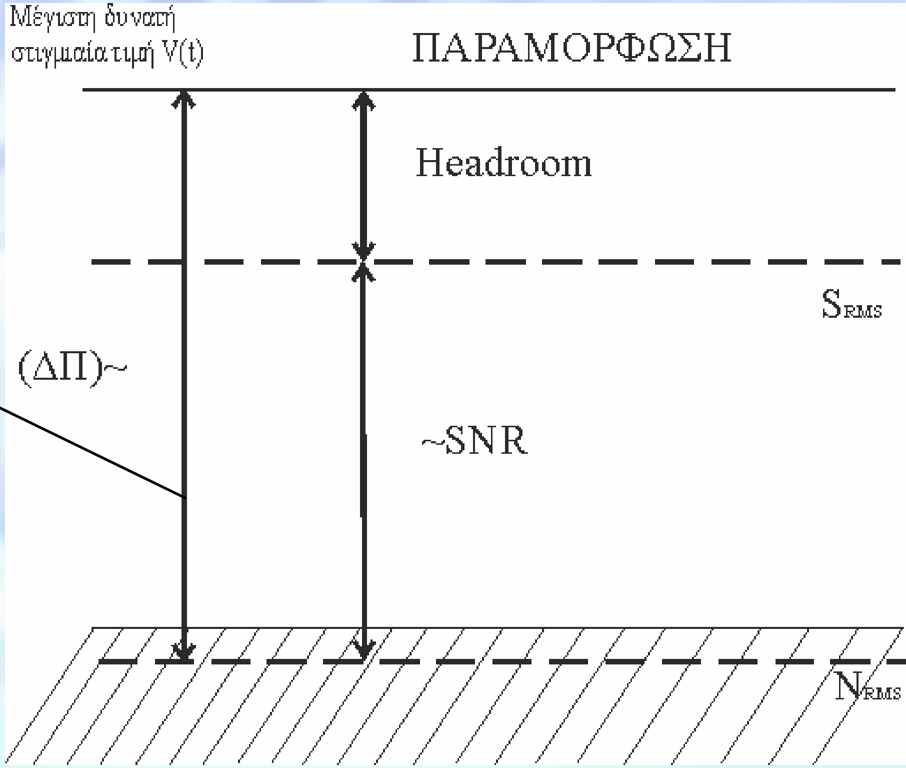
- Κατά την εγγραφή του σήματος η μέγιστη στάθμη του σήματος εισόδου υπερβαίνει την μέγιστη στάθμη κβαντισμού οπότε και «ψαλιδίζονται» οι άκρες του



Δυναμική Περιοχή Ηχητικού Συστήματος

•Η διαφορά, σε dB, μεταξύ των πιο δυνατών και πιο ήσυχων περασμάτων του προγράμματος που μπορούν να αποδοθούνε με ευκρίνεια από το σύστημα όπως φαίνεται στο σχήμα





SNR

- Σ' ένα αναλογικό ηχητικό σύστημα, η μέση τετραγωνική τιμή $s(t)$ ορίζεται από τη σχέση

$$s_{RMS} = \left(\frac{1}{T} \int_0^T s^2(t)dt \right)^{1/2}$$

- Ορίζεται η σχέση σήματος προς θόρυβο (Signal to Noise Ratio, SNR)

$$SNR_{(dB)} = 20\log_{10} \frac{S_{RMS}}{N_{RMS}}$$

S_{RMS} και N_{RMS} η μέση τετραγωνική τιμή του σήματος εισόδου και του θορύβου αντίστοιχα

- Αυτή η σχέση δίνει τιμές σε dB μικρότερες από αυτές της δυναμικής περιοχής, γιατί η μέση τετραγωνική τιμή ενός σήματος είναι πάντα μικρότερη από την μέγιστη απόλυτη τιμή

$$S_{RMS} < \max\{|s(t)|\}$$

Πλάτος ή μέση τετραγωνική τιμή

- Ένα μέγεθος που έχει σημασία είναι ο λόγος της μέγιστης στιγμιαίας τιμής προς τη μέση τετραγωνική τιμή του σήματος

$$\frac{\max\{|s(t)|\}}{S_{RMS}}$$

- Ο λόγος αυτός εκφράζει ένα μέγεθος το οποίο σχετίζεται με τη μεταβλητότητα του σήματος,
- για ηχητικά σήματα ομιλίας είναι γύρω στο 8:1
- για κλασσική μουσική γύρω στο 30:1
- Διαφορά στη στατιστική μεταξύ των ηχητικών σημάτων → διαφορετική σχέση μεταξύ δυναμικής περιοχής και RMS.

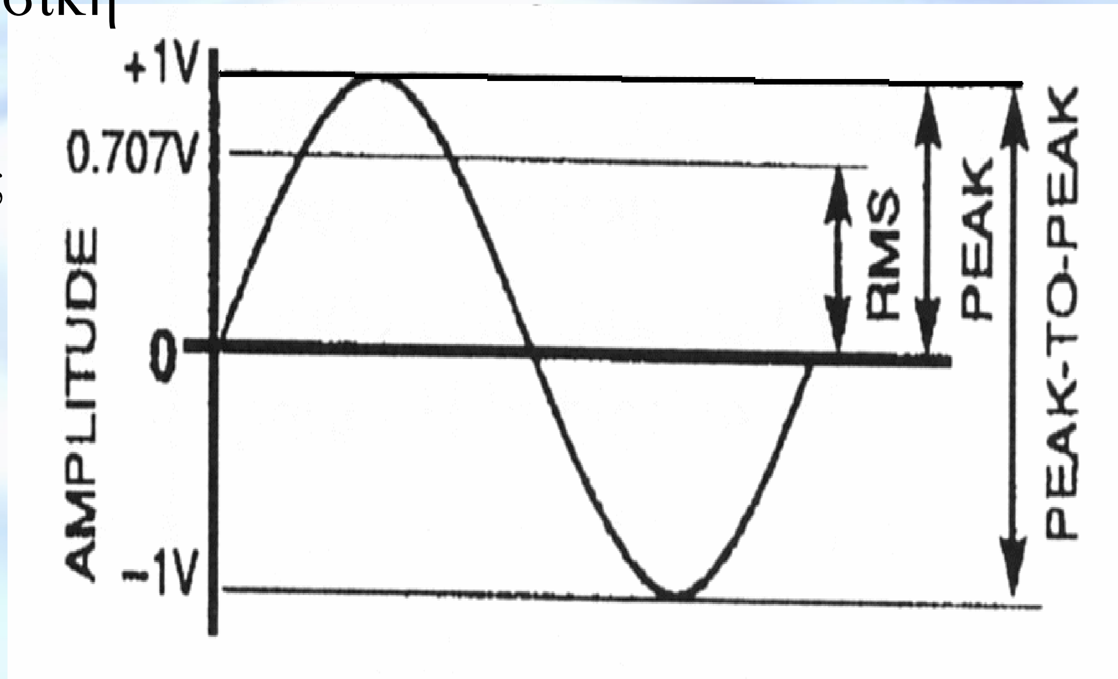
Headroom

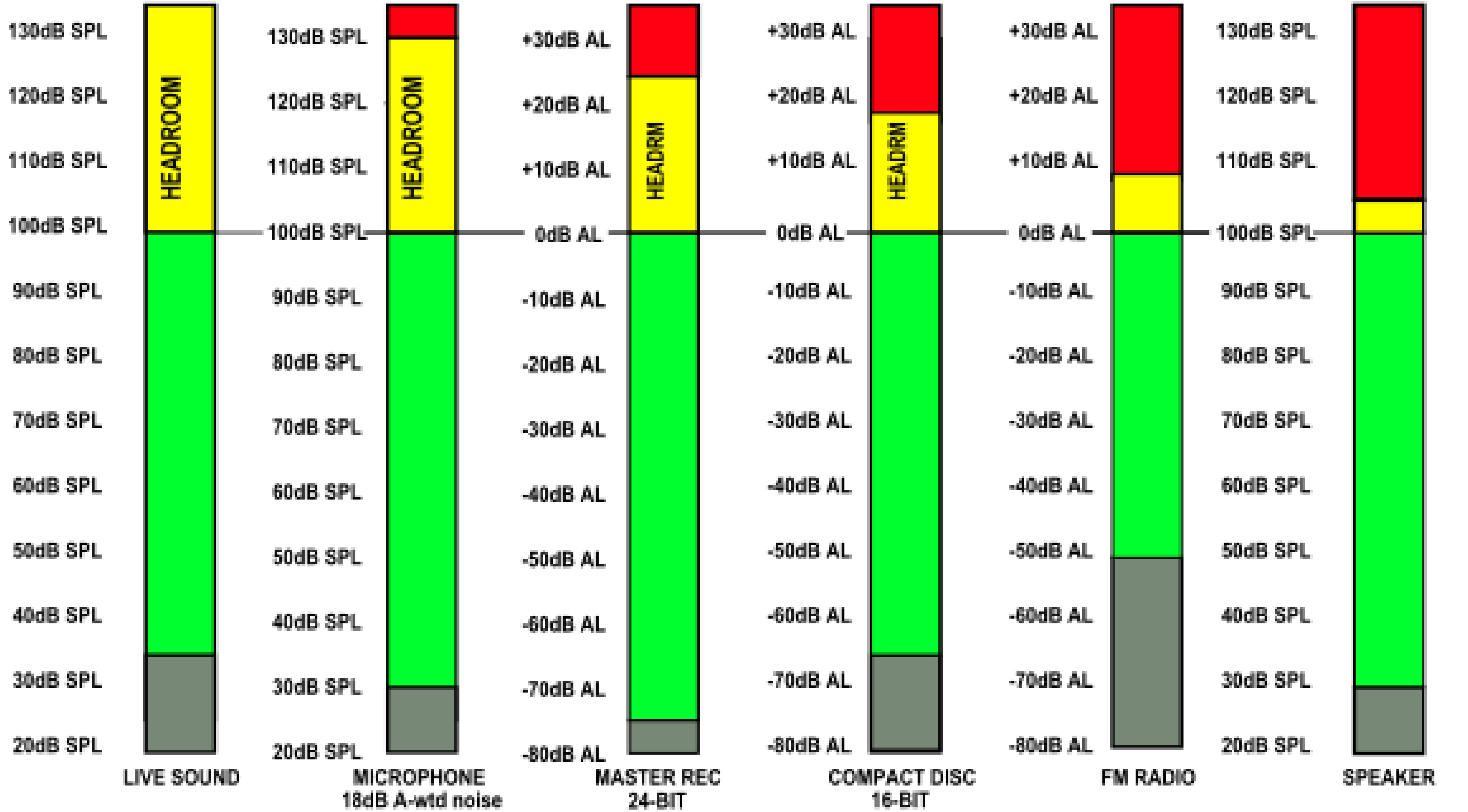
- Θεωρούμε ένα μέγεθος το οποίο το ονομάζουμε headroom του οποίου η τιμή σε dB είναι τέτοια ώστε $(\Delta\Pi)_{dB}$=SNR+Headroom αποδεικνύεται

$$Headroom = 20\log_{10}\left(\frac{\max\{|s(t)|\}}{S_{RMS}}\right)$$

- Ο διαχωρισμός $\max\{|s(t)|\}$ *και* S_{RMS} φαίνεται και σε επαγγελματικά ηχητικά συστημάτα τα οποία ορίζουν το peak level και *νόμιμη στάθμη* λειτουργίας (nominal level) που ισούται με το S_{RMS}
- Σε επαγγελματικό σύστημα HR ~ 20dB, πράγμα που επιτρέπει λόγο 10:1 απαραίτητο για τη μουσική

Σχέση μεταξύ μεγίστης στιγμιαίας τιμής (peak), πλάτους από κορυφή σε κορυφή (peak to peak, p-p) και μέσης τετραγωνικής τιμής (RMS) για ένα ημιτονικό σήμα. ($I_{rms} = I_0/\sqrt{2}$)





PROGRAMME LEVEL CAPABILITY AT TYPICAL STAGES OF THE AUDIO PROCESS

LEVELS are Peak RMS Equivalent

NOISE is measured ITU-R 468 weighted for true subjective validity

MASTER recording assumes 24-bit with 24dB of headroom assigned (alternative EBU standard)

SPEAKERS are typical hi-end 100W per channel 86dB for 1W sensitivity, pair at 2m in listening room

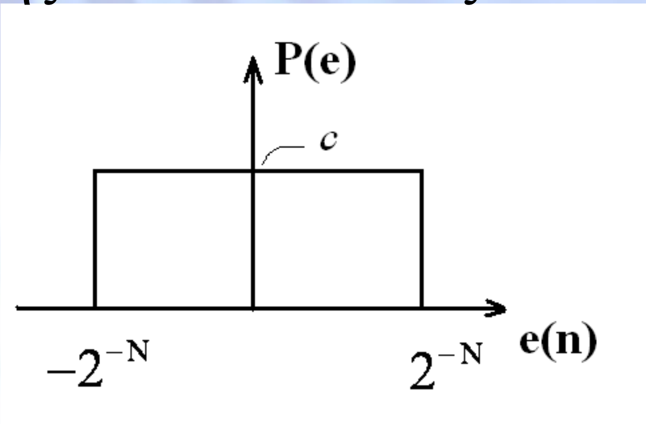
NOTE HOW, WITH AN ASSUMED ALIGNMENT LEVEL OF 100dB SPL, (NEEDED FOR REALISTIC LEVELS) MOST SYSTEMS CANNOT COPE

Παράδειγμα

- Ένα αναλογικό σύστημα ήχου λειτουργεί σε στάθμες V_{pp}=10. Ποια η νόμιμη στάθμη λειτουργίας σε Volt δεδομένου ότι χρειαζόμαστε headroom 20dB;
- Καταρχάς, παρατηρούμε ότι τα 10 V ορίζονται από κορυφή σε κορυφή (p-p) που σημαίνει ότι το σήμα μας θα κινείται μεταξύ ±5V (το πλάτος θα είναι δηλαδή 5V). Από παραπάνω βλέπουμε ότι $20\log_{10}(5/V_{RMS})=20$ dB επομένως τελικά $V_{RMS}=0.5$ V. Δηλαδή, η νόμιμη στάθμη λειτουργίας είναι ίση με το 10% της μέγιστης απόλυτης στάθμης.

Θόρυβος Κβάντοποίησης

- Το σφάλμα κβάντισης δημιουργεί θόρυβο $e(n) = \hat{s}(n) - s(n)$
- Μέγιστο *e(n)* όταν η πραγματική τιμή του σήματος είναι στη μέση του διαστήματος κβάντισης άρα $|e(n)| \le \Delta/2$ άρα άρα $|e(n)| \le 2^{-N}$
- Αν θεωρήσουμε ότι η πιθανότητα εμφάνισης Σ.Κ είναι όπως στο σχήμα, αποδεικνύεται εύκολα ότι $c=2^{N-1}$



- Η E_{RMS} θα είναι $E_{RMS} = \dfrac{\Delta}{2\sqrt{3}} = \dfrac{2^{-N}}{\sqrt{3}}$
- Για το S_{RMS} Έστω *s(n) ημιτονικό*

  άρα $S_{RMS} = 1/\sqrt{2}$ άρα

$$SNR(dB) = 20\log_{10}\frac{1/\sqrt{2}}{2^{-N}/\sqrt{3}} = 20\log_{10} 2^N + 20\log_{10}\frac{\sqrt{3}}{\sqrt{2}} = 6,02N + 1,76\ dB$$

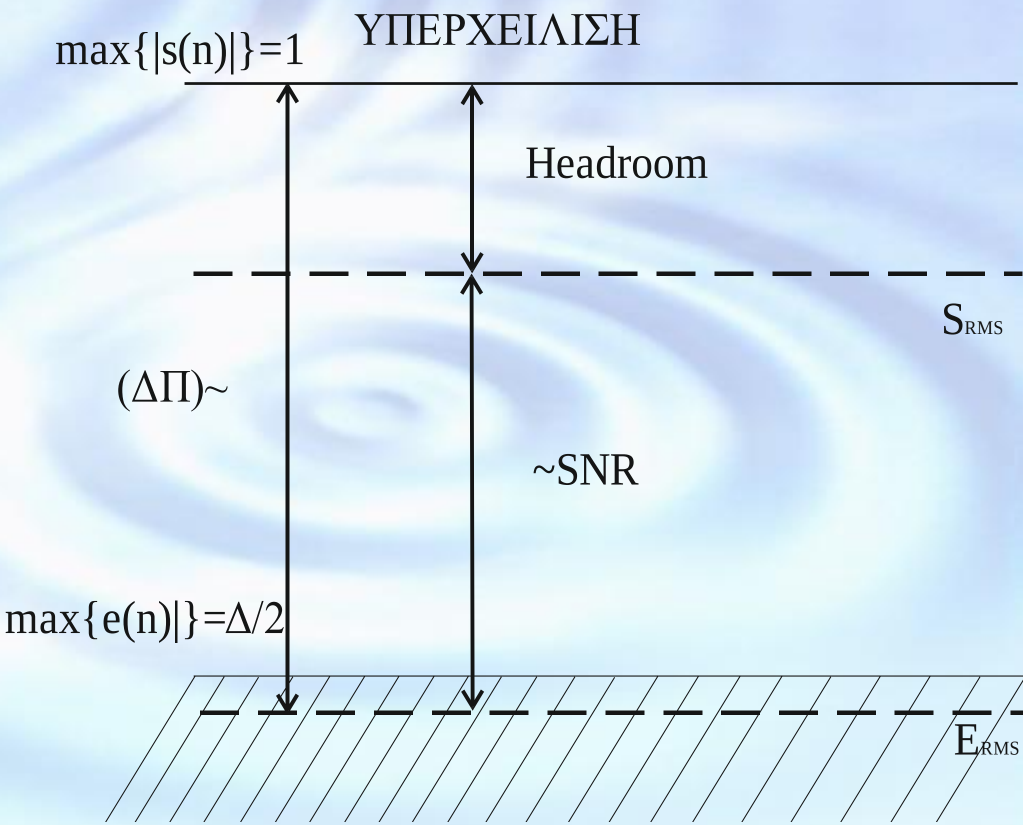
με N την τάξη κβαντιστή.

- Αύξηση της τάξης κβαντιστή κατά 1 αύξηση του SNR κατά 6 db υποδιπλασιασμό της στάθμης θορύβου

Θόρυβος Κβάντοποίησης

- Στον πίνακα αριστερά οι ενδεικτικές τιμές SNR αντιστοιχούν σε διαφορετικές τάξεις κβαντιστή
- Στο σχήμα δεξιά η σχέση μεταξύ Δυναμικής περιοχής (ΔΠ),. Headroom και SNR σε ένα ψηφιακό σύστημα.

Τάξη N(bits)	SNR (db)
8	49.8
12	73.8
16	97.8
18	109.8



Ασκήσεις – Παραδείγματα

Παράδειγμα: Η τάξη κβαντιστή για μετάδοση ημιτονικού σήματος είναι N=16. Πόσο το μέγιστο SNR υπό την παραδοχή ότι το σήμα είναι κλασσικής μουσικής (θεωρήστε ότι ο λόγος της μέγιστης απόλυτης τιμής προς τη μέση τετραγωνική τιμή του σήματος για κλασσική μουσική είναι ίσος με 30:1).

Απάντηση: Το SNR για ημιτονικό σήμα 16 bit είναι **SNR_1=6,02*16+1.76=98.08 dB** το Headroom για κλασσική μουσική είναι **HR=20$\log_{10}$30=29,54dB** όμως το SNR_1 έχει υπολογιστεί για **HR=3**, για να έχουμε λόγο της μέγιστης απόλυτης τιμής προς τη μέση τετραγωνική τιμή 30, και αφού το μέγιστο δεν μπορεί να αλλάξει πρέπει να υπολογίσουμε το SNR με **HR==29.54dB**. Από το ημιτονικό σήμα έχουμε μόνο 3 άρα χρειαζόμαστε άλλα **29,54-3=26,54 dB** άρα το **SNR** θα πρέπει να γίνει **SNR=98.08-26,54= 71,54 dB.**

- **Παράδειγμα**: Θέλουμε να ηχογραφήσουμε και να μεταδώσουμε στερεοφωνικό ψηφιακό σήμα με συχνότητα δειγματοληψίας *Fs*=48kHz. Το μέγιστο bit rate που μας επιτρέπει το κανάλι μετάδοσης είναι 1940kbps. Πόσο το μέγιστο SNR που μπορώ να έχω για ημιτονικό σήμα δεδομένου ότι έχω κβαντιστή μεταβλητής τάξης *N*; Πόσο το μέγιστο SNR υπό την παραδοχή ότι το σήμα είναι κλασσικής μουσικής (θεωρήστε ότι ο λόγος της μέγιστης απόλυτης τιμής προς τη μέση τετραγωνική τιμή του σήματος για κλασσική μουσική είναι ίσος με 30:1).
- **Απάντηση:** Αν *N* η μέγιστη τάξη του κβαντιστή για τα παραπάνω δεδομένα θα ισχύει: **48000x2x*N*=1940000**. Άρα, ***N*=1940000/96000=20,21**. Επειδή η τάξη του κβαντιστή πρέπει να είναι ακέραιος αριθμός συμπεραίνουμε ότι *N*=20. Για ημιτονικό σήμα, βλέπουμε ότι **SNR=6,02*N*+1.76=122,16dB**.
- Για σήμα κλασσικής μουσικής βλέπουμε ότι χρειαζόμαστε **HR=20$\log_{10}$30=29,54dB**. Δεδομένου ότι για ημιτονικό σήμα το headroom είναι μόνο **3 dB**, αυτό σημαίνει ότι χρειαζόμαστε **29,54-3=26,54 dB** περισσότερο headroom, άρα η τιμή του SNR θα μειωθεί ανάλογα SNR=122,16-26,54=96,52 dB.

Παράδειγμα: Για μουσική ικανοποιητικής πιστότητας απαιτούμε *SNR*=90 dB. Αν η συχνότητα δειγματοληψίας είναι *Fs*=44,1kHz, προσδιορίστε τη μνήμη που απαιτείται για διαφύλαξη ενός, μέσης διάρκειας 3 min στερεοφωνικού μουσικού κομματιού. Θεωρείστε ένα headroom 10dB.
Για ένα headroom ίσο με 10dB, σημαίνει ότι θα έχω 7dB λιγότερα από ότι μου υπαγορεύει η σχέση που δίνει το SNR(ημιτονικό σήμα) . Θα είναι επομένως SNR=6,02*N*+1.76-7=90=6,02*N*-5.24 και επομένως $N=\frac{90+5.24}{6.02}=15{,}82$. Άρα το N θα πρέπει να είναι ίσο με 16. Για τα τρία λεπτά (180 sec) θα χρειαζόμαστε επομένως Μνήμη=44100·16·2·180=254016000=254016kbits~31 MB (CD).

Ορισμός dB-Full Scale

- Το μέγεθος αυτό ορίζεται με τέτοιον τρόπο ώστε η τιμή 0 dB-FS να αντιστοιχεί στο μέγιστο δυνατό πλάτος του σήματος, όταν δηλαδή το σήμα αξιοποιεί όλο το δυναμικό εύρος του συστήματος.
- Για παράδειγμα, μια ημιτονική κυματομορφή χαρακτηρίζεται πλήρους πλάτους (full scale) όταν το πλάτος V_0 είναι ίσο με 1, δεδομένου ότι χρησιμοποιείται κανονικοποιημένη κλίμακα

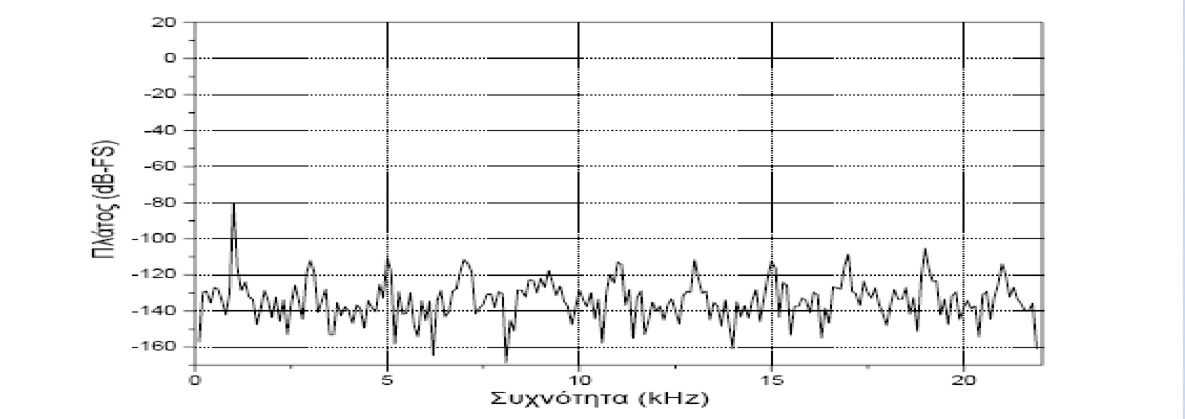
Παράδειγμα: Η είσοδος μιας ψηφιακής κάρτας ήχου λειτουργεί για στάθμες αναλογικού σήματος. Εκφράστε σε dB-FS τις τιμές του πλάτους ±5Vκαι της μέσης τετραγωνικής τιμής του ψηφιακού σήματος που θα προκύψει, θεωρώντας ότι στην είσοδο έχουμε ημιτονικό σήμα πλάτους 0.5 V. Από πόσους σταθμούς κβαντισμού διέρχεται το ψηφιοποιημένο ημιτονικό σήμα αν *N*=16 bits;
Για το πλάτος θα ισχύει προφανώς $20\log_{10}(0,5/5)= 20\log_{10}(1/10)$=-20dB-FS και η RMS τιμή του σήματος είναι -20-3 = -23dB-FS. Το ημιτονικό σήμα αξιοποιεί το 1/10 του δυναμικού εύρους του κβαντιστή, άρα διέρχεται από $2^{16}/10$ =6553 σταθμούς κβαντισμού.

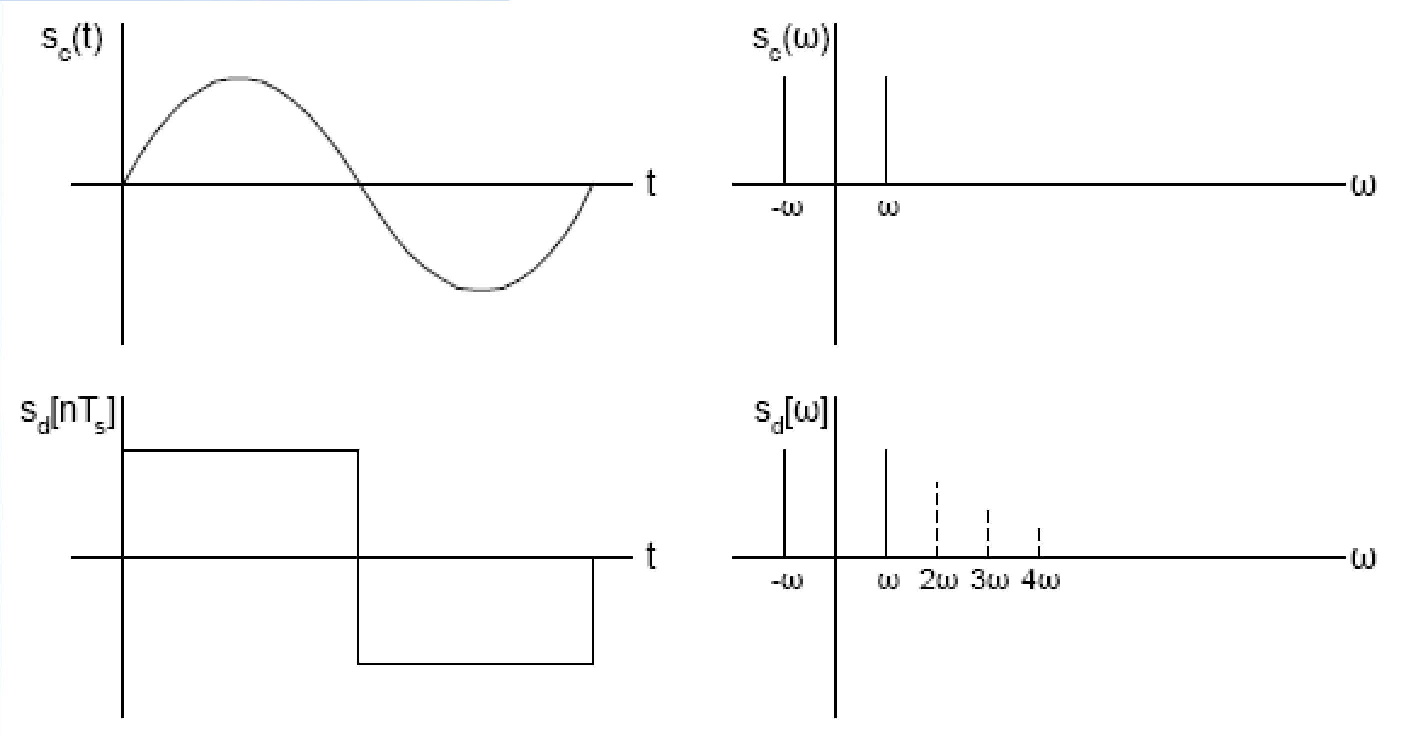
DITHER

- Σε χαμηλές στάθμες σήματος το σφάλμα κβαντισμού δεν έχει τη μορφή λευκού θορύβου και δεν ισχύουν όσα έχουμε κουβεντιάσει μέχρι τώρα.
- Αν το πλάτος του σήματος είναι μικρότερο της ελάχιστης στάθμης κβαντισμού μπορεί να εμφανιστούν και αλλά φαινόμενα παραμορφώσεων όπως aliasing κλπ.
- Ο ελλιπής κβαντισμός περιορίζει επίσης την ακουστική ευκρίνεια στην αναπαραγωγή μουσικών ήχων.
- Συμπερασματικά, ο θόρυβος που παράγουν τα ψηφιακά ηχητικά συστήματα εμφανίζεται κατά το στάδιο του κβαντισμού (ή της αναπαραγωγής) και έχει ιδιάζοντα και ανεπιθύμητο χαρακτήρα.
- Λύσεις :
- Χρήση μη γραμμικών μετατροπέων.
- Αύξηση της τάξης του κβαντιστή
- Dithering Προσθήκη θορύβου χαμηλής στάθμης με τιμές πλάτους συγκρίσιμες με την στάθμη κβαντισμού Δ και χαρακτήρα λευκού θορύβου.

Επίδραση του κβαντισμού σε χαμηλής στάθμης ήχους

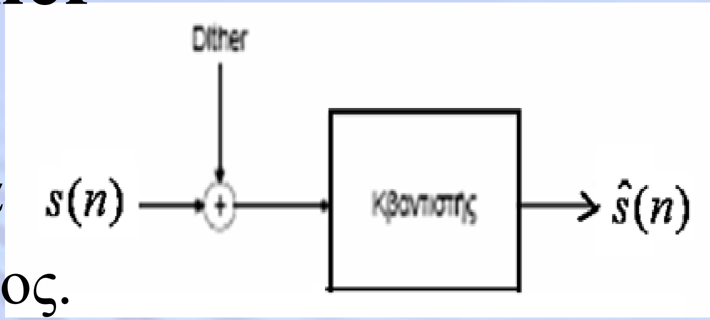
Μορφή φάσματος ημιτονικού σήματος πλάτους –80dB-FS, συχνότητας 1kHz, N=16 bit





Επίδραση του κβαντισμού στο φασματικό περιεχόμενο σημάτων χαμηλής στάθμης

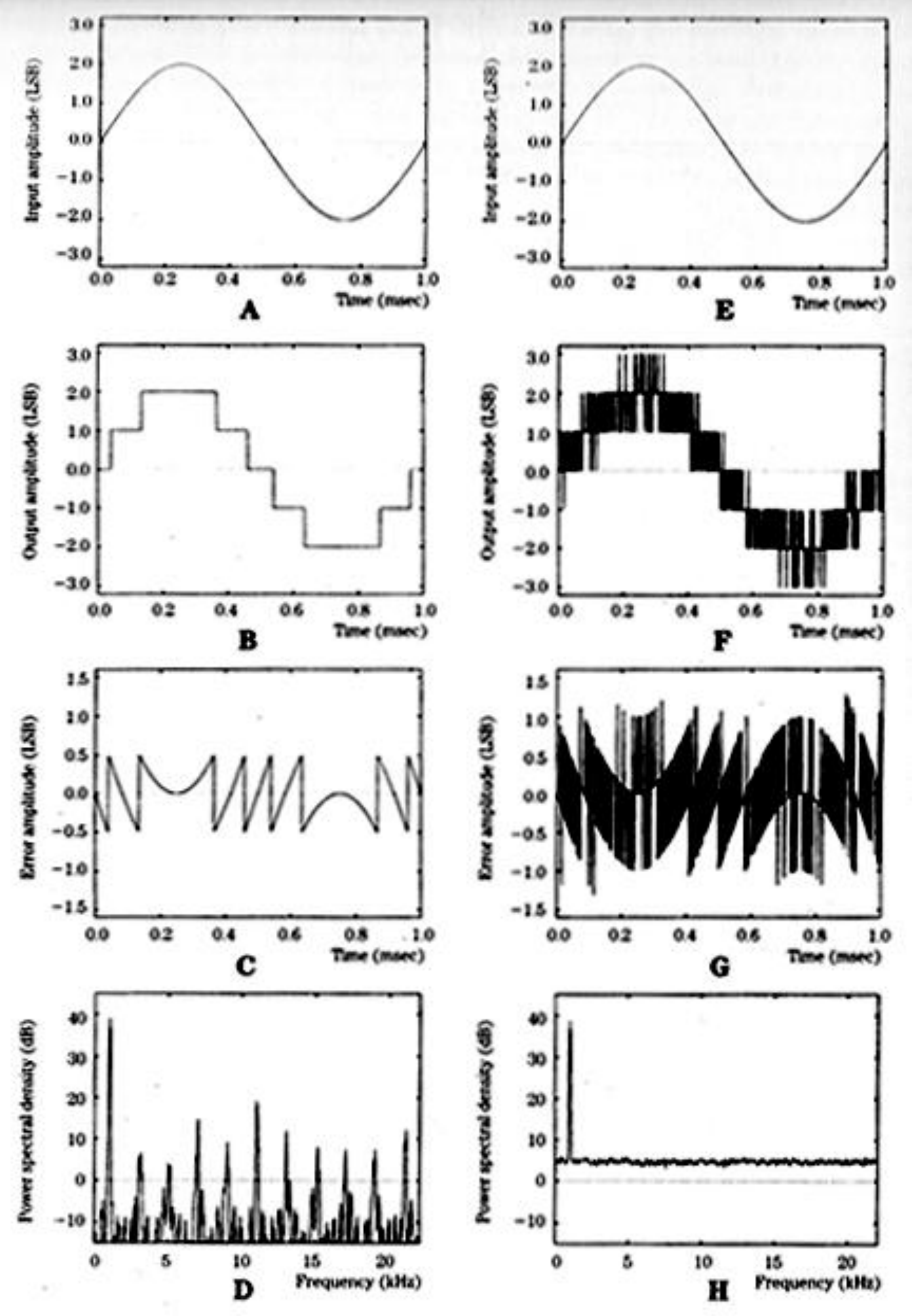
Κβαντισμός και Dither



- Μηχανισμός δράσης dither
- *Με την χρήση Dither βελτιώνεται σημαντικά* η υποκειμενική χροιά του ψηφιακού σήματος.
- Η προσθήκη Dither στην είσοδο ενός κβαντιστή προκαλεί τις εξής σημαντικές βελτιώσεις στη συνολική λειτουργία του:
  1. Αλλάζει την μορφή του σφάλματος κβαντισμού και το καθιστά ανεξάρτητο από το σήμα εισόδου,
  2. Αφαιρεί κάθε είδους αρμονικής παραμόρφωσης η οποία εμφανίζεται κατά την μετατροπή σημάτων πολύ μικρού πλάτους, δίνοντας στον θόρυβο κβαντισμού χαρακτηριστικά όμοια με τα χαρακτηριστικά λευκού θορύβου και
  3. Βελτιώνει τη διακριτική ικανότητα του κβαντιστή, αυξάνοντας κατά υποκειμενικό (κι όχι ποσοτικό) τρόπο τη δυναμική του περιοχή.

Επίδραση του Dither στο φάσμα του ψηφιακού σήματος.

- Αριστερά χωρίς Dither δεξιά με Dither
- Α,Ε Ένα ημιτονικό σήμα συχνότητας 1 kHz εισέρχεται σε κβαντιστή με πλάτος 2 LSB.
- Εισέρχεται στον κβαντιστή στο Β μορφή χωρίς Dither. Στο F με Dither πλάτους 1 LSB ο θόρυβος υπερτίθεται πάνω στην ημίτονο και αλλάζει την απόκριση του κβαντιστή.
- Στο διάγραμμα C φαίνεται το σφάλμα κβαντισμού χωρίς Dither. Το σφάλμα κβαντισμού, στο G έχει αποκτήσει και αυτό διαφορετική μορφή.
- Στο D η φασματική πυκνότητα. Όσο λιγοστεύουν οι σταθμοί κβαντισμού το σήμα παίρνει μια μορφή που μοιάζει με τετραγωνικό παλμό, και εμφανίζονται συχνότητες αρμονικές της αρχικής. Υπερισχύει η αρχική συχνότητα του 1 kHz αλλά εμφανίζονται και αρμονικές στα 3, 5, 7, 9 kHz. Στο H για τη φασματική πυκνότητα οι αρμονικές συχνότητες έχουν εξουδετερωθεί. Παρατηρείται ένας ομοιόμορφα κατανεμημένος λευκός θόρυβος, ο οποίος είναι λιγότερο ενοχλητικός από την περίπτωση με τις δυνατές αρμονικές.



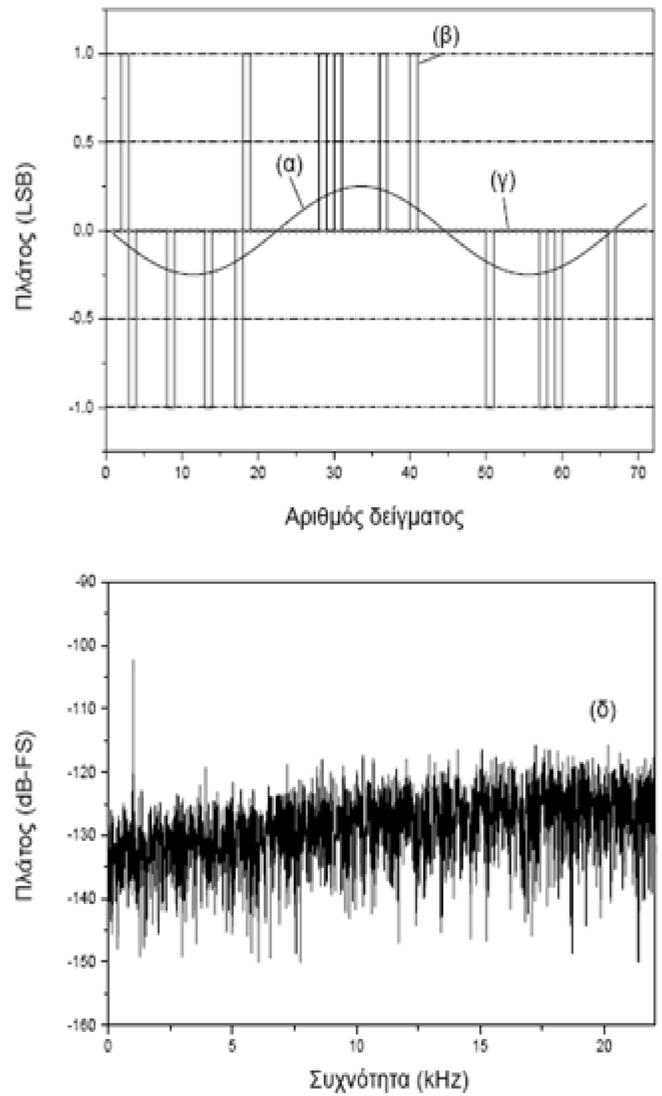
Οπτικό Ανάλογο

Original image using the web-safe color palette with no dithering applied. Note the large flat areas and loss of detail.

Original image using the web-safe color palette with Floyd–Steinberg dithering. Note that even though the same palette is used, the application of dithering gives a better representation of the original.

Παράδειγμα: Η επίδραση του dither στη συνολική απόδοση 16bit κβαντιστή με dither πλάτους 2 LSB p-p. Στην είσοδο του κβαντιστή εφαρμόζεται ένα ημιτονοειδές σήμα πλάτους LSB/4 και συχνότητας 1kHz (2.14(α)). Ένας απλός κβαντιστής (χωρίς dither) (2.14(γ)). δεν μπορεί να κωδικοποιήσει ένα τόσο μικρού πλάτους σήμα, και θα έδινε μηδενική έξοδο. Στην περίπτωση κβαντιστή με dither, οι νέες τιμές του σήματος εισόδου (αρχικό σήμα + dither – (2.14(β))), υποχρεώνει την έξοδο του κβαντιστή να είναι μη μηδενική σε κάποιες περιόδους δειγματοληψίας.

Λόγω των στατιστικών χαρακτηριστικών του dither, η έξοδος του κβαντιστή περιέχει σε μεγάλο βαθμό την πληροφορία του αρχικού σήματος (2.14(δ)).

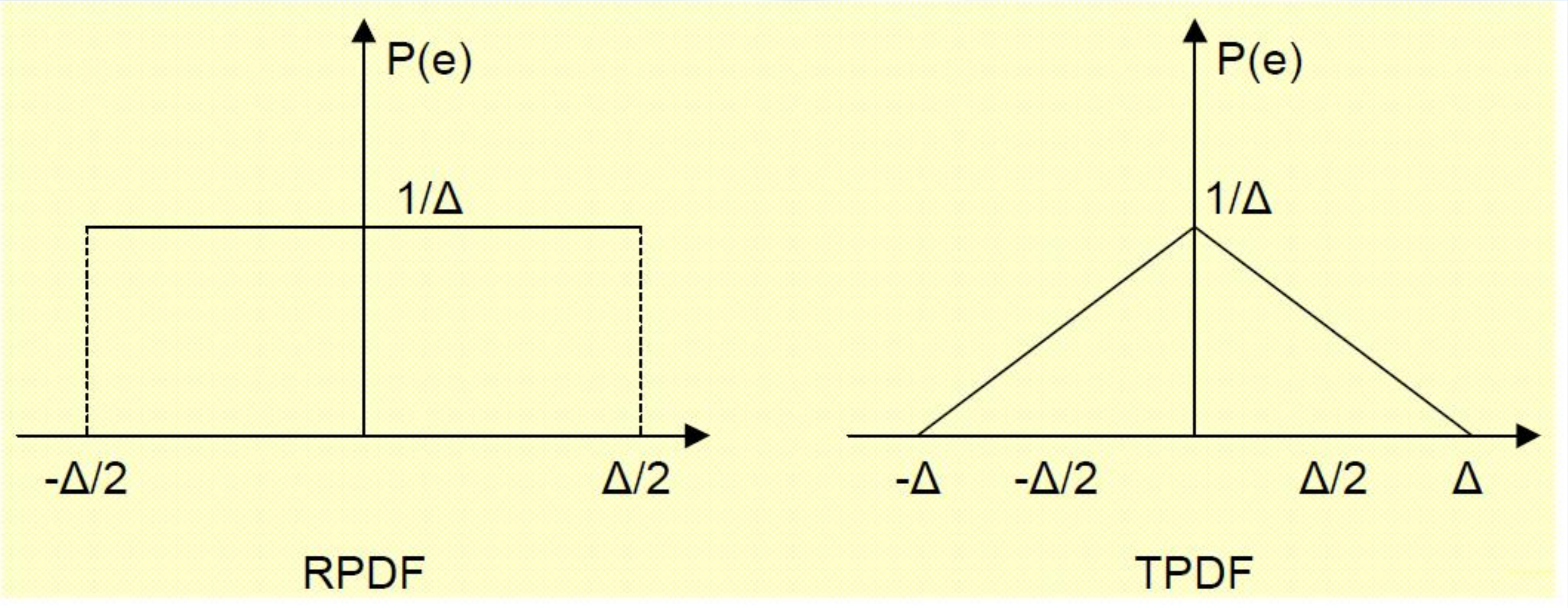


Σχήμα 2.14: Παράδειγμα επίδρασης του θορύβου dither σε ημιτονικό σήμα 1kHz πλάτους –102.3dB-FS

- Από τα παραπάνω φαίνεται η αναγκαιότητα του Dither.
- Τα περισσότερα ψηφιακά συστήματα λειτουργούν σε συχνότητα δειγματοληψίας 44,1kHz με κβαντισμό τουλάχιστο16bit με προσθήκη dither κατά την ηχογράφηση.

Είδη - Χαρακτηριστικά dither

- Τα διάφορα είδη dither δημιουργούνται με κριτήριο τη συνάρτηση πυκνότητας πιθανότητας (Probability Density Function, PDF).
- Ομοιόμορφη (rectangular) κατανομή (RPDF) πλάτους 1LSB p-p
- Τριγωνική (triangular) κατανομή (TPDF) πλάτους 2LSB p-p
- Έχει αποδειχθεί ότι dither με TPDF παρουσιάζει την καλύτερη στατιστική συμπεριφορά και ελαχιστοποιεί τη συσχέτιση του θορύβου κβαντισμού με το σήμα, προσθέτει θόρυβο τάξης 6 dB.



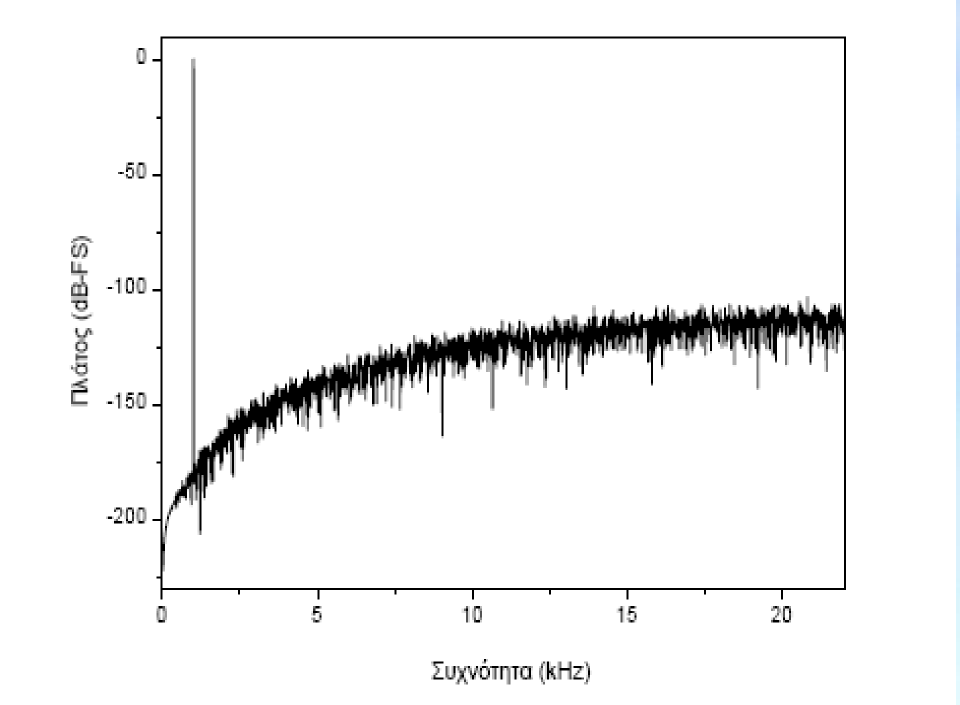
Μορφοποίηση Θορύβου

• Η χρήση dither έχει σαν αποτέλεσμα την αποδιαμόρφωση του σφάλματος κβαντισμού και την μείωση της αρμονικής παραμόρφωσης που δημιουργείται κατά τον κβαντισμό σημάτων μικρού πλάτους με ταυτόχρονη αύξηση το επιπέδου του θορύβου κβαντισμού.

• Για σήματα χαμηλής έντασης, το επίπεδο του θορύβου ξεπερνά το ελάχιστο κατώφλι ακοής και ο θόρυβος κβαντισμού καθίσταται ακουστός.

• Αν ο θόρυβος «μετατοπίζονταν» σε μεγαλύτερες των 4 kHz συχνότητες τότε δεν θα ήταν ακουστός από τον άνθρωπο. Το γεγονός αυτό πραγματοποιείται με την χρήση διαφόρων φίλτρων με τεχνικές μορφοποίησης θορύβου (Noise-Shaping).

Φάσμα εξόδου κβαντιστή για είσοδο ημιτονικού σήματος πλάτους 0dB-FS, συχνότητας 1kHz και χρήση Noise-Shaping τεχνικής.



Υπερδειγματοληψία

- Η αύξηση της συχνότητας δειγματοληψίας (ή της ροής δεδομένων) κατά R φορές από την ελάχιστη συχνότητα που απαιτείται από το θεώρημα N-S .
- Τα κύρια πλεονεκτήματα αυτής (της θεωρητικά περιττής) αύξησης είναι:
    1. Η ελάττωση των παραμορφώσεων από το αναλογικό antialiasing φίλτρο εισόδου ή εξόδου του μετατροπέα.
    2. Η ελάττωση του θορύβου σφάλματος κβαντισμού, πράγμα που διευκολύνεται με τη χρήση τεχνικών μορφοποίησης φάσματος του σφάλματος (Noise-Shaping).
    3. Η δυνατότητα ελάττωσης της τάξης του κβαντιστή του μετατροπέα, χωρίς ταυτόχρονη ελάττωση της σχέσης SNR.
    4. Η απλούστευση της κατασκευής και ρύθμισης του μετατροπέα και η ελάττωση του κόστους κατασκευής του.

Υπερδειγματοληψία

- Με υπερδειγματοληψία R-φορές, λαμβάνονται δείγματα του σήματος με μια νέα συχνότητα $f'_s=Rf_s$ (Hz), π.χ. εάν R=4 και f_s=44,1kHz, τότε f'_s= 176,4kHz.
- Η νέα μέγιστη συχνότητα η οποία είναι δυνατόν να εμφανίζεται f'_{max} θα είναι επίσης αυξημένη κατά R φορές, δηλαδή $f'_{max}=Rf_{max}$.
- Επιτρέπεται η χρήση ομαλότερων αναλογικών anti-aliasing φίλτρων και δε χρειάζεται <u>τόσο υψηλός ρυθμός αποκοπής.</u>
- Αν ο θόρυβος κβαντισμού είναι λευκός, τότε εμφανίζει και σταθερή φασματική πυκνότητα.
- Στο «υπερδειγματοληπτούμενο» σήμα το φάσμα καλύπτει R-φορές μεγαλύτερη περιοχή και άρα εμφανίζει ελάττωση του πλάτους του κατά R-φορές (αρχή αβεβαιότητας) σε συγκεκριμένη συχνότητα, π.χ. αν R=4, τότε

$$\left(\overline{e}\,'\right)^2 = \frac{\left(\overline{e}\right)^2}{4} \Leftrightarrow 10\log\frac{\left(\overline{e}\right)^2}{\left(\overline{e}\,'\right)^2} = -6\text{dB}$$

με $\left(\overline{e}\,'\right)^2$ και $\left(\overline{e}\right)^2$ είναι η πυκνότητα φασματικής ισχύος του σφάλματος κβαντοποίησης με και χωρίς υπερδειγματοληψία αντίστοιχα.

- Κάθε διπλασιασμός του R αυξάνει την τιμή SNR κατά 3dB περίπου
- Οι περισσότερες περιπτώσεις υπερδειγματοληψίας συνδυάζονται με **Noise-Shaping**, έτσι ώστε το σφάλμα κβαντισμού να ελαχιστοποιείται στην περιοχή ακουστικών συχνοτήτων (20Hz-20kHz)

Αύξηση της «Ενεργού Τάξης Κβαντισμού»

- Αποδεικνύεται ότι ένας κβαντιστής τάξης N χωρίς Υπ/ψία θα έχει το ίδιο SNR με ένα κβνατιστή με Υπ/ψία παράγοντα R με τάξης N' όταν
$$N' - N = \Delta N = 0.5\log_2 R$$

- **Παράδειγμα:** Να βρεθεί η συχνότητα δειγματοληψίας με την οποία ένας κβαντιστής τάξης N=12 bit επιτυγχάνει την ίδια ποιότητα μετατροπής με κβαντιστή τάξης N'=16 bit και συχνότητα δειγματοληψίας 44100 Hz.
- **Απάντηση**: Η τάξη του διαθέσιμου κβαντιστή είναι κατά 4 bit μικρότερη από αυτή που θέλουμε να επιτύχουμε. Από τη παραπάνω σχέση φαίνεται λοιπόν ότι πρέπει να χρησιμοποιηθεί παράγοντας δειγματοληψίας R τέτοιος ώστε $0.5\log_2 R=4$. Επομένως θα ισχύει $\log_2 R=8$ και επομένως $R=2^8=256$. Η συχνότητα δειγματοληψίας εν τέλει θα πρέπει να είναι ίση με $256 \cdot 44100=$ 11,29 MHz περίπου.