

## 2<sup>η</sup> Εργαστηριακή Άσκηση

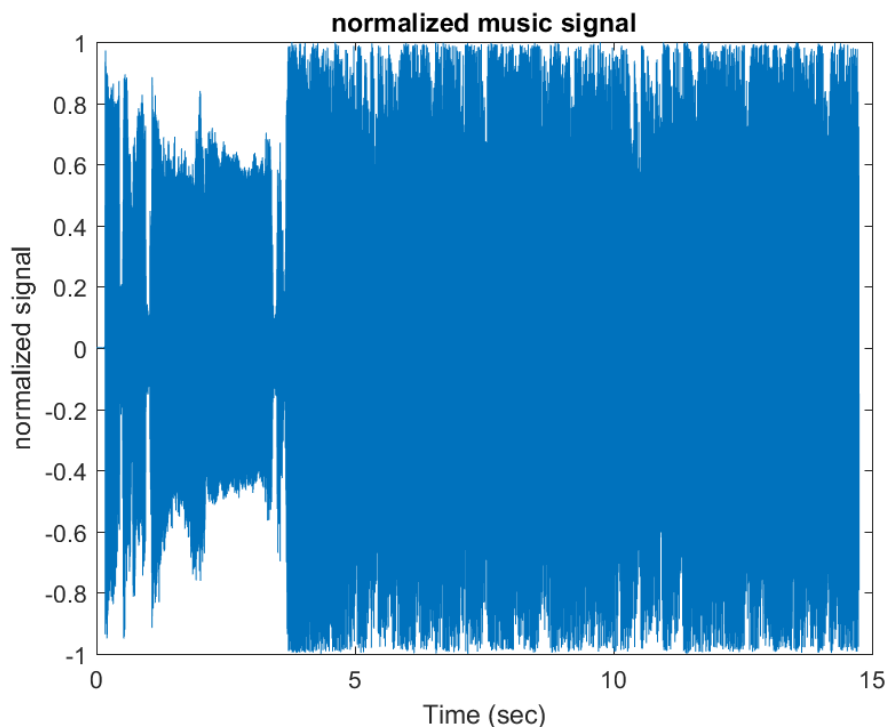
<u>Ονοματεπώνυμο</u>	<u>ΑΜ</u>
Αντωνιάδης Παναγιώτης	03115009
Ποιητάρης Χριστόφορος	03115201

### Θέμα: Κωδικοποίηση σημάτων Μουσικής βάσει ψυχοακουστικού μοντέλου (Perceptual Audio Coding)

#### Μέρος 1: Ψυχοακουστικό Μοντέλο 1

- **Βήμα 1.0: Κανονικοποίηση του σήματος**

Αρχικά, κανονικοποιούμε το σήμα μουσικής **music.wav** έτσι ώστε όλες οι τιμές του να είναι στο διάστημα  $[-1, 1]$ . Το σήμα που προκύπτει φαίνεται παρακάτω:

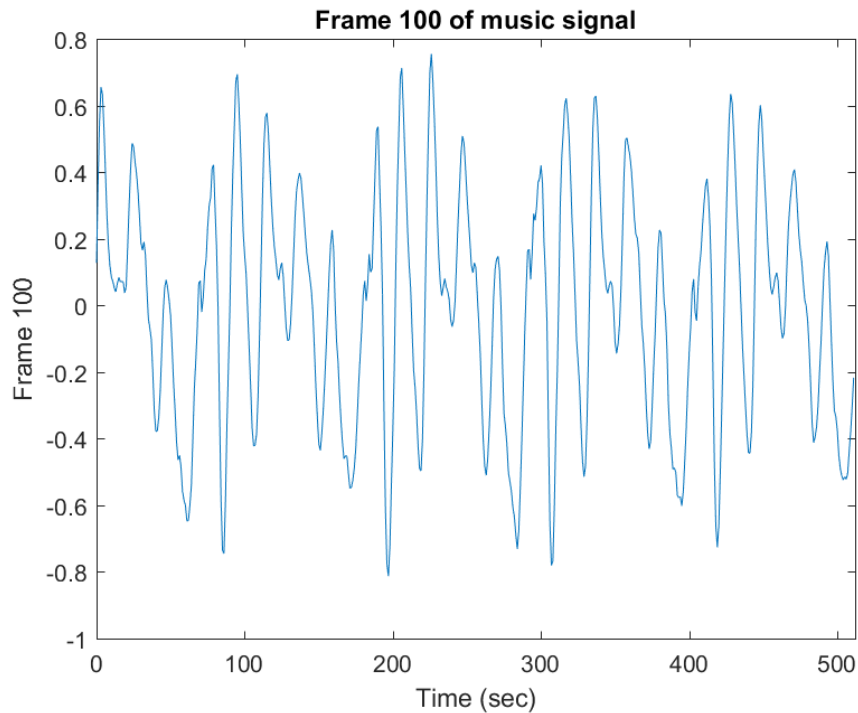


Εικόνα 1. Normalized music signal

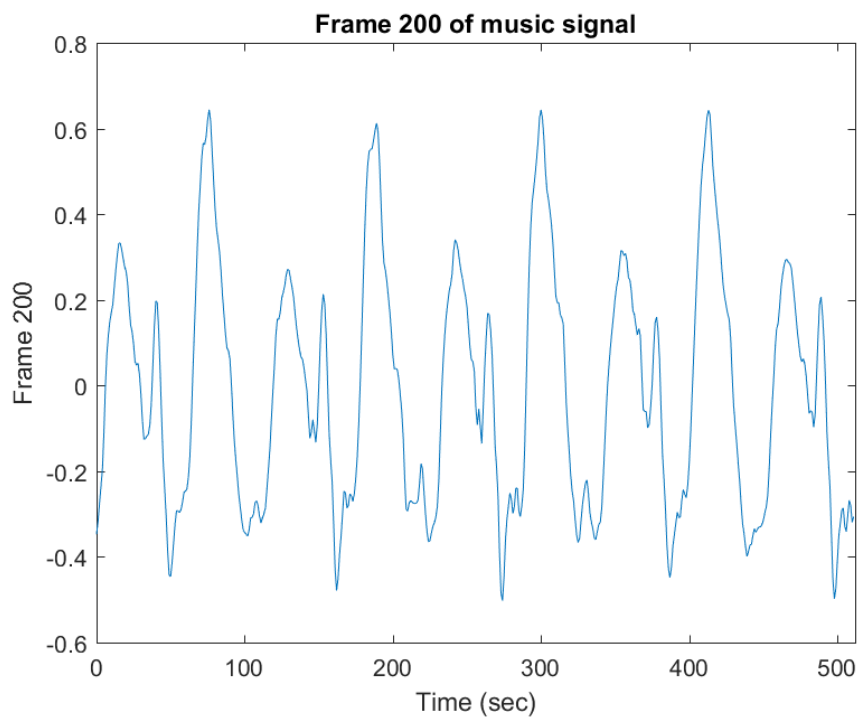
**Σημείωση:** Σε αυτό το σημείο παραθυροποιούμε το σήμα μας σε πλαίσιο μήκους 512 δειγμάτων. Οι υπόλοιποι υπολογισμοί, λοιπόν, γίνονται για κάθε ένα πλαίσιο. Επιλέγουμε το  $100^\circ$  (το οποίο μας δόθηκε και είναι εύκολη η σύγκριση)

## Digital Signal Processing

και το 200<sup>ο</sup> πλαίσιο για να αναπαραστήσουμε τα διάφορα μεγέθη που προκύπτουν ανά πλαίσιο. Οπότε καλό είναι να αναπαραστήσουμε και τα χρονικά διαγράμματα των δύο πλαισίων.



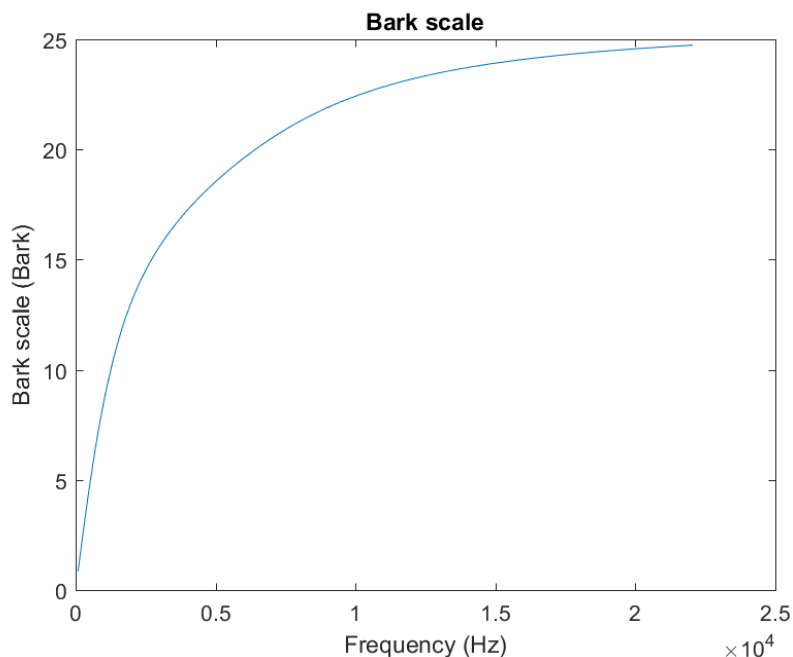
Εικόνα 2. Frame 100 of music signal



Εικόνα 3. Frame 200 of music signal

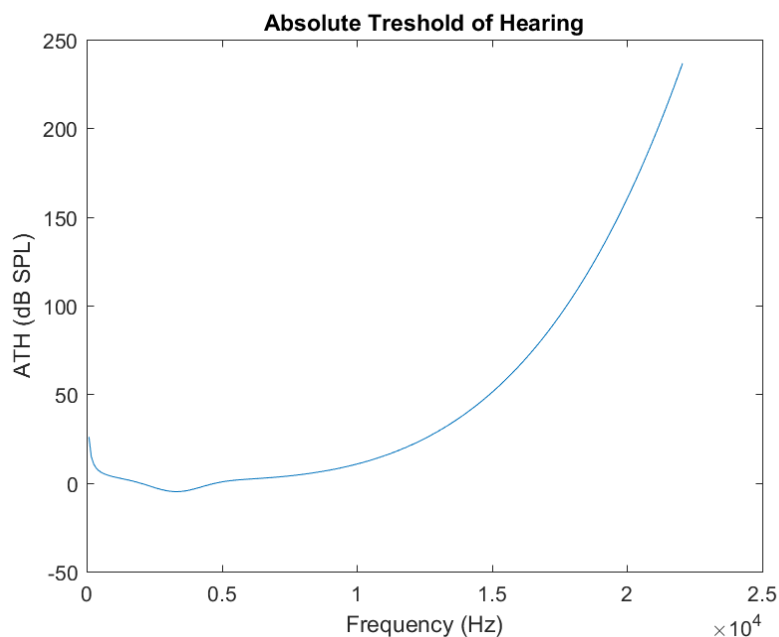
- **Βήμα 1.1: Φασματική Ανάλυση**

Αρχικά, ορίζεται η κλίμακα Bark η οποία φαίνεται παρακάτω:



Εικόνα 4. Bark scale of frequencies

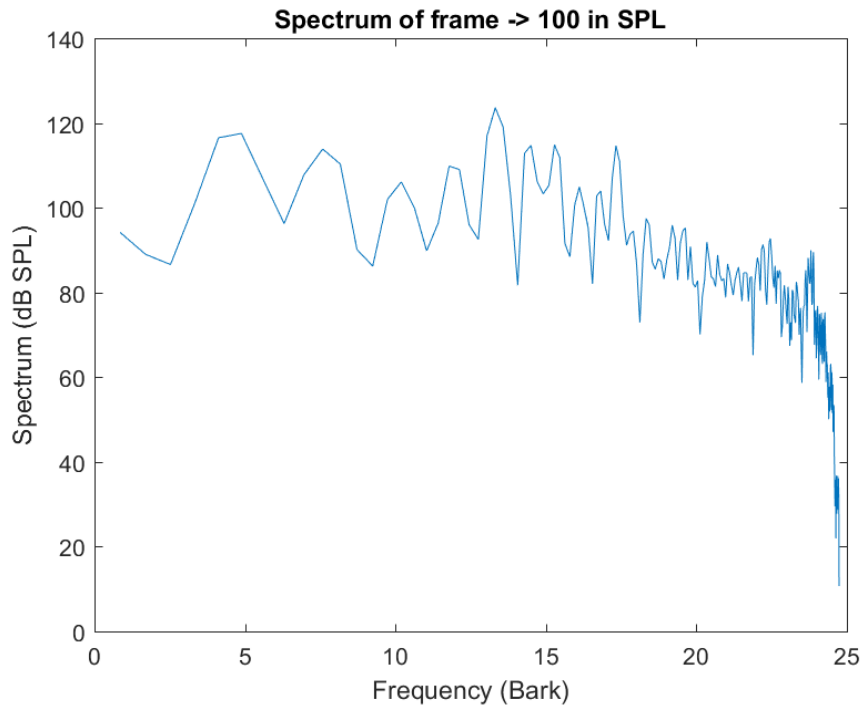
Ακόμη, ορίζεται και το κατώφλι ακοής (Absolute Threshold of Hearing) το οποίο χαρακτηρίζει το ποσό της ενέργειας σε dB - Sound Pressure Level (dB SPL) που πρέπει να έχει ένας τόνος (π.χ. ημίτονο) συχνότητας  $f$  ώστε να γίνει αντιληπτός σε περιβάλλον πλήρους ησυχίας. Παρακάτω βλέπουμε την αναπαράστασή του:



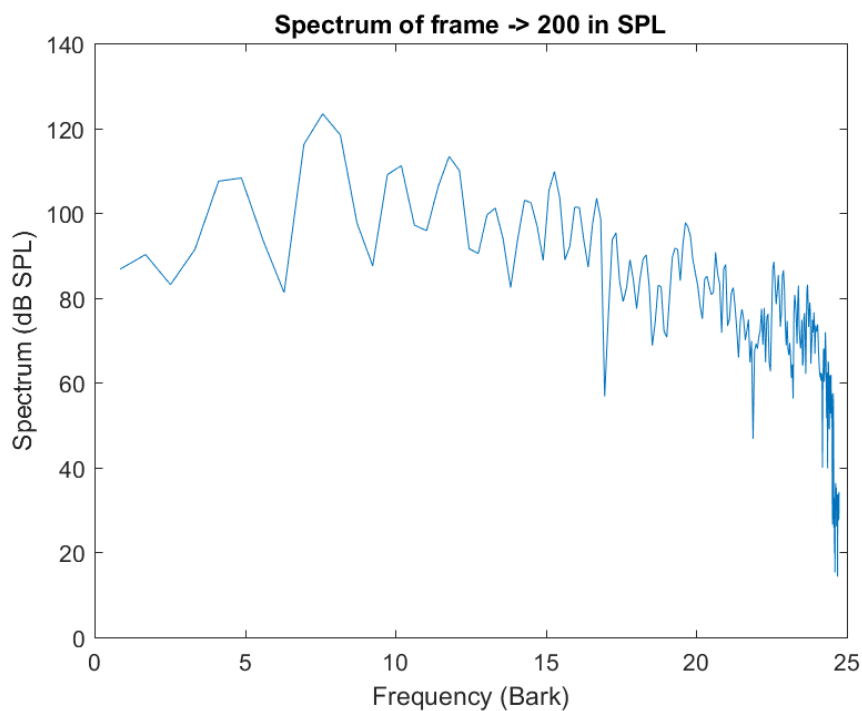
Εικόνα 5. Absolute Threshold of Hearing

## Digital Signal Processing

Τώρα υπολογίζουμε **για κάθε πλαίσιο** το N-σημείων φάσμα ισχύος  $P(k)$  σύμφωνα με τον τύπο που δίνεται στην εκφώνηση. Αλλά, επειδή το αποτέλεσμα του fft είναι συμμετρικό θα κρατήσουμε μόνο τα  $N/2$  πρώτα σημεία (  $N = 512$  δείγματα). Παρακάτω φαίνεται το φάσμα των δύο πλαισίων που έχουμε επιλέξει:



Εικόνα 6. Spectrum of frame 100 in dB SPL

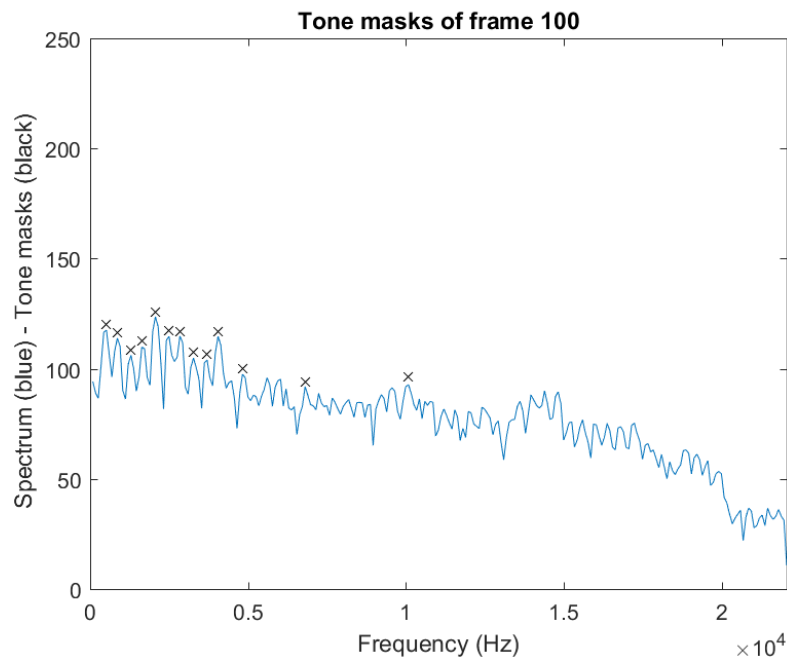


Εικόνα 7. Spectrum of frame 200 in dB SPL

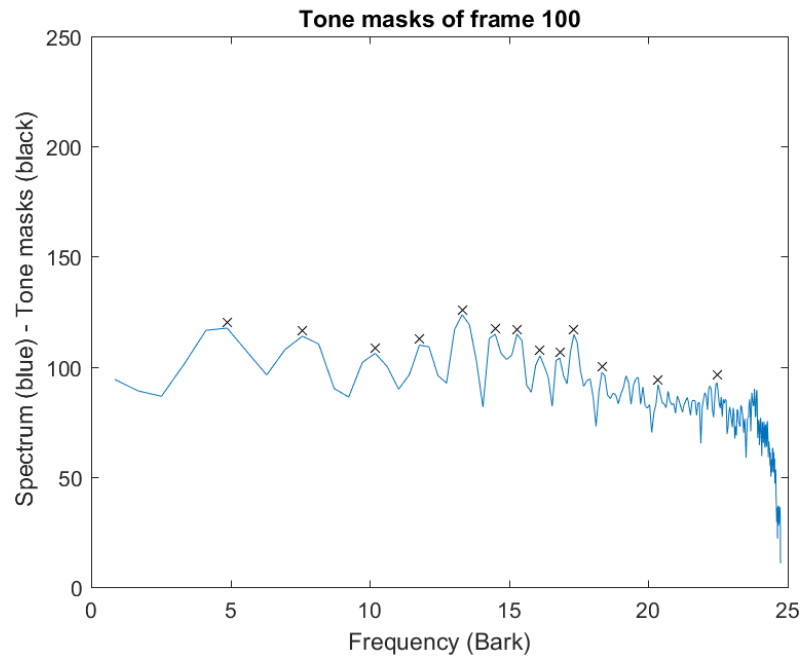
- **Βήμα 1.2: Εντοπισμός μασκών τόνων και θορύβου (Maskers)**

Στο βήμα αυτό θα υπολογίσουμε τις μάσκες τόνων και θορύβου, δηλαδή τοπικά μέγιστα τα οποία είναι μεγαλύτερα από τις γειτονικές τους συχνότητες τουλάχιστον κατά 7dB. Στα παρακάτω διαγράμματα αναπαριστούμε με 'x' τις μάσκες τόνου και με 'o' τις μάσκες θορύβου που εντοπίστηκαν για τα δύο πλαίσια ως προς την συχνότητα σε Hz και σε Bark (έχει προστεθεί σε κάθε γραφική και το φάσμα με χρώμα μπλε).

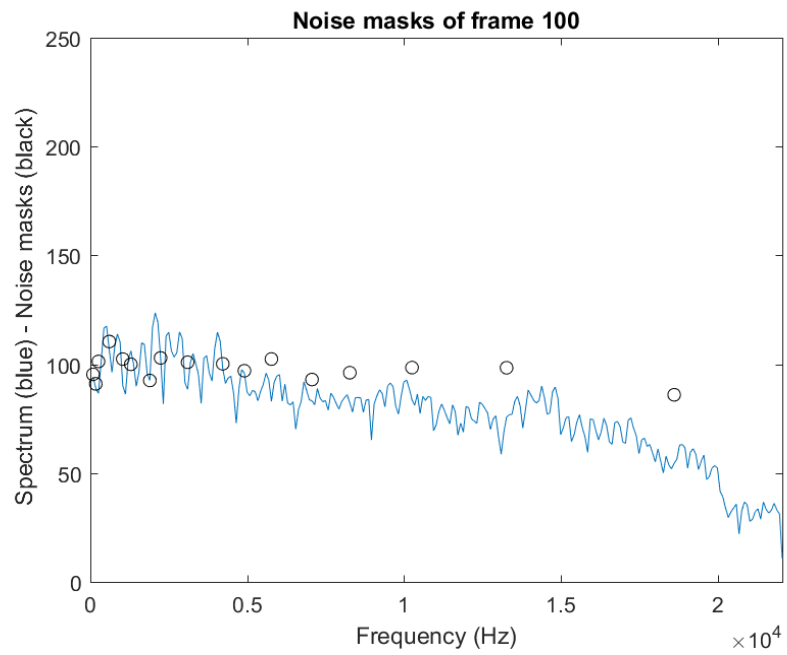
## Πλαίσιο 100



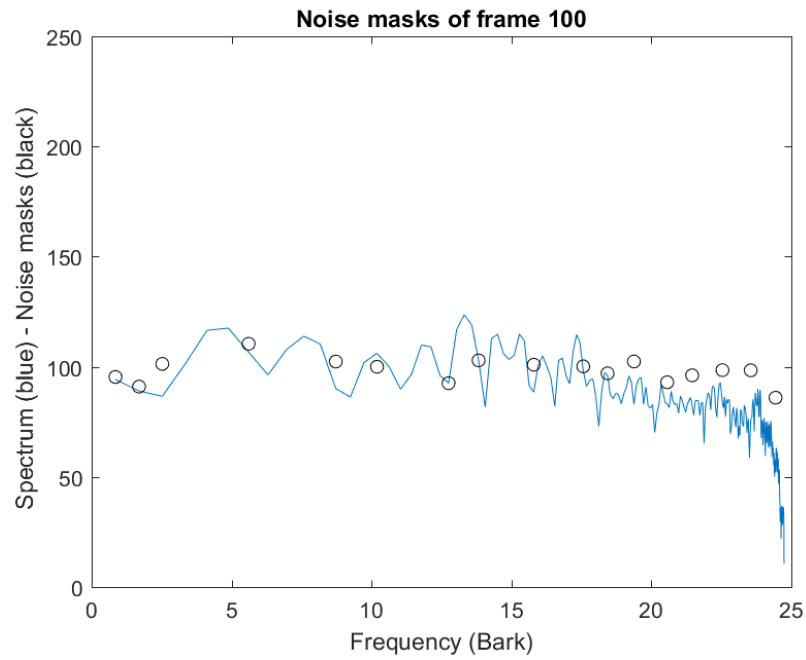
**Εικόνα 8. Tone masks of frame 100 in frequency**



**Εικόνα 9. Tone masks of frame 100 in bark scale**

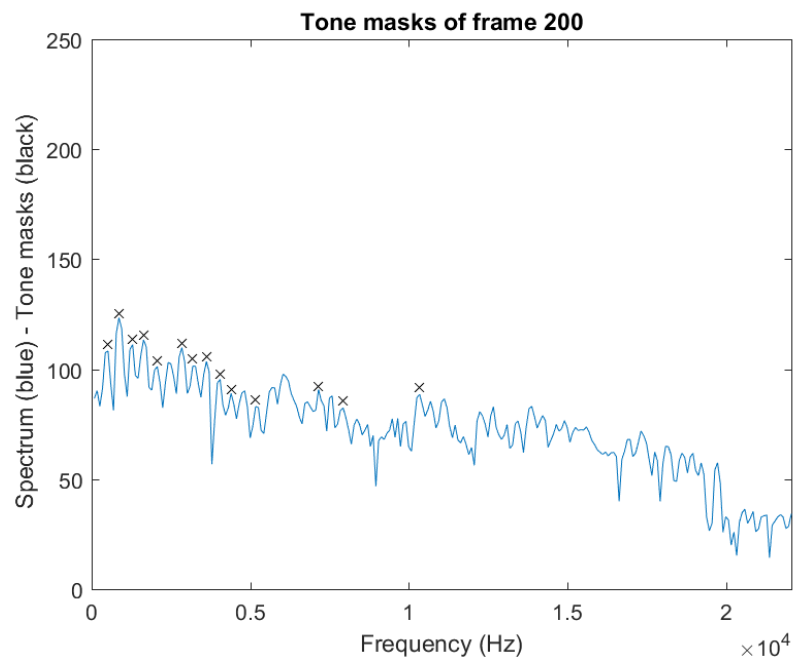


**Εικόνα 10. Noise masks of frame 100 in frequency**

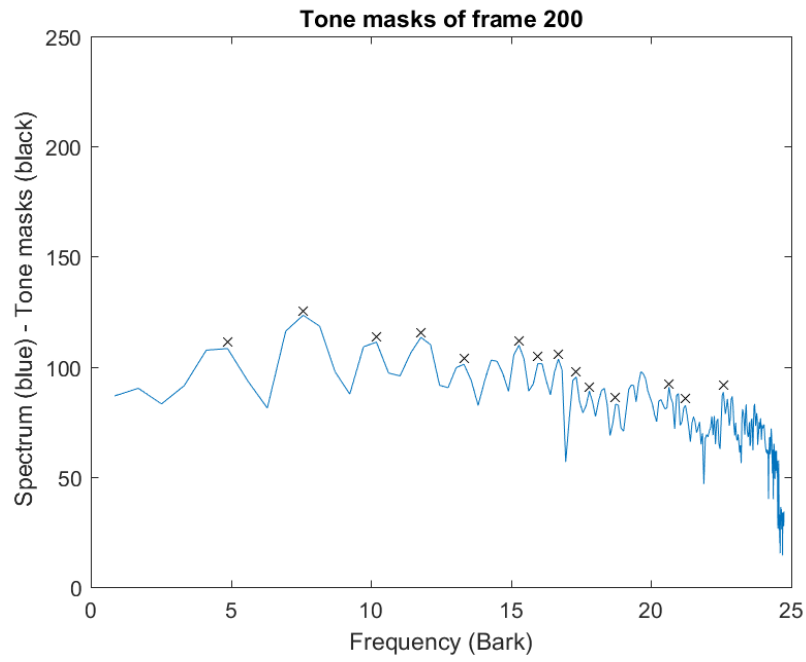


**Εικόνα 11. Noise masks of frame 100 in bark scale**

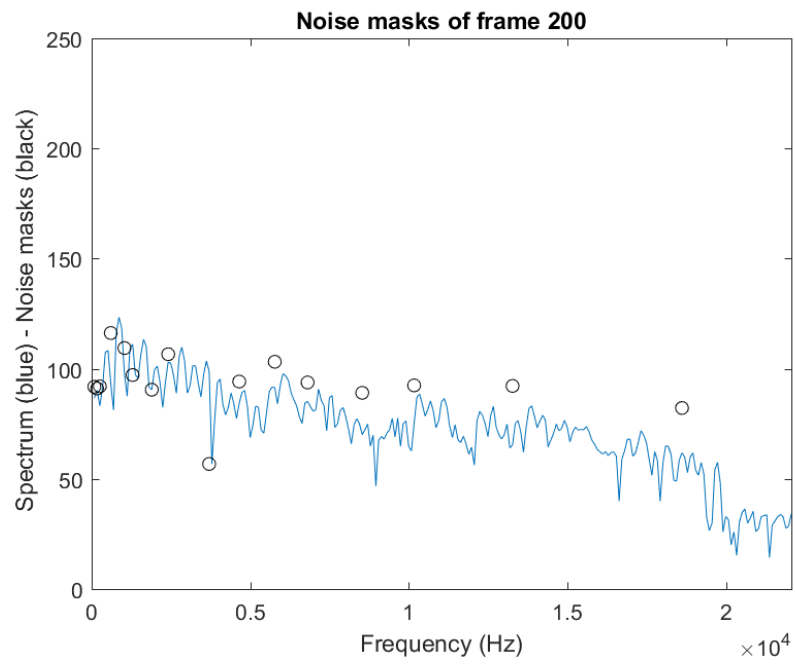
## Πλαίσιο 200



**Εικόνα 12. Tone masks of frame 200 in frequency**

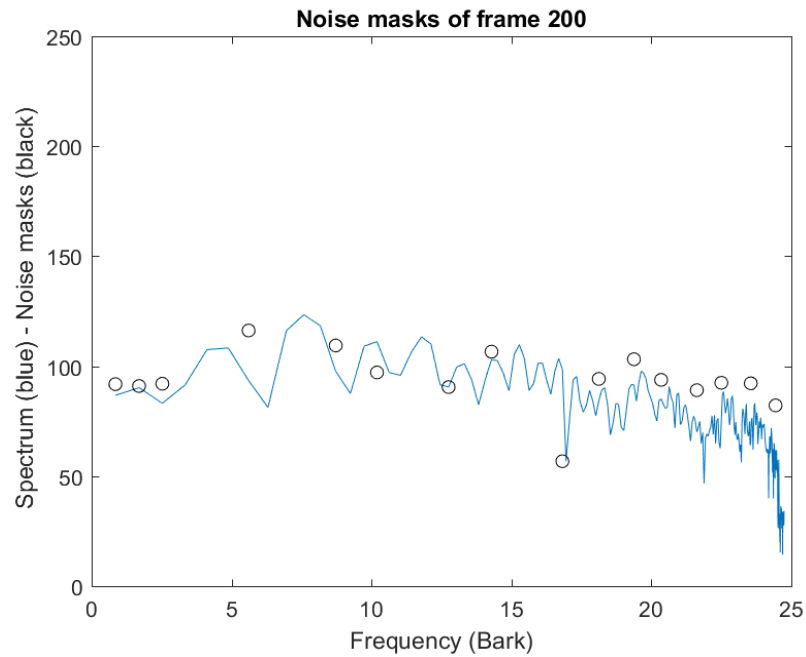


Εικόνα 13. Tone masks of frame 200 in bark scale



Εικόνα 14. Noise masks of frame 200 in frequency



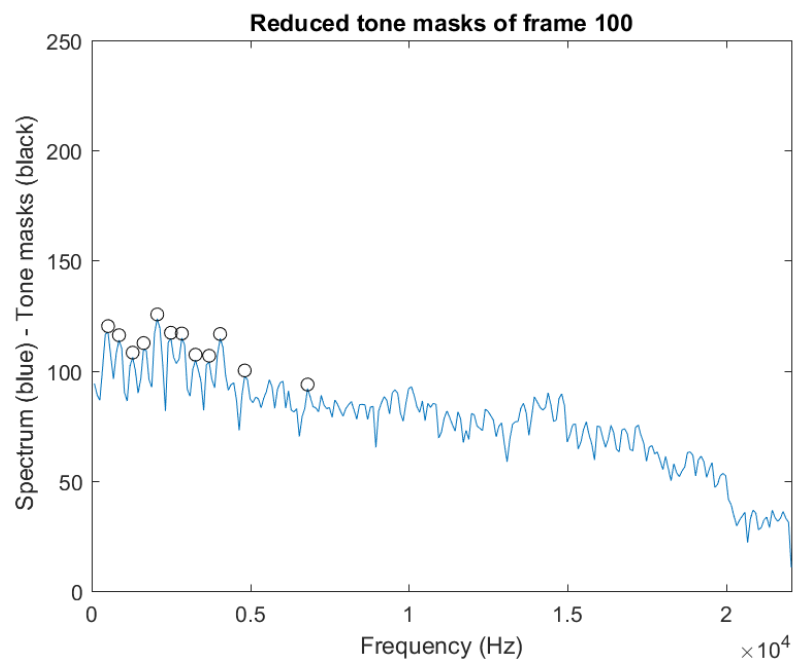


Εικόνα 15. Noise masks of frame 200 in bark scale

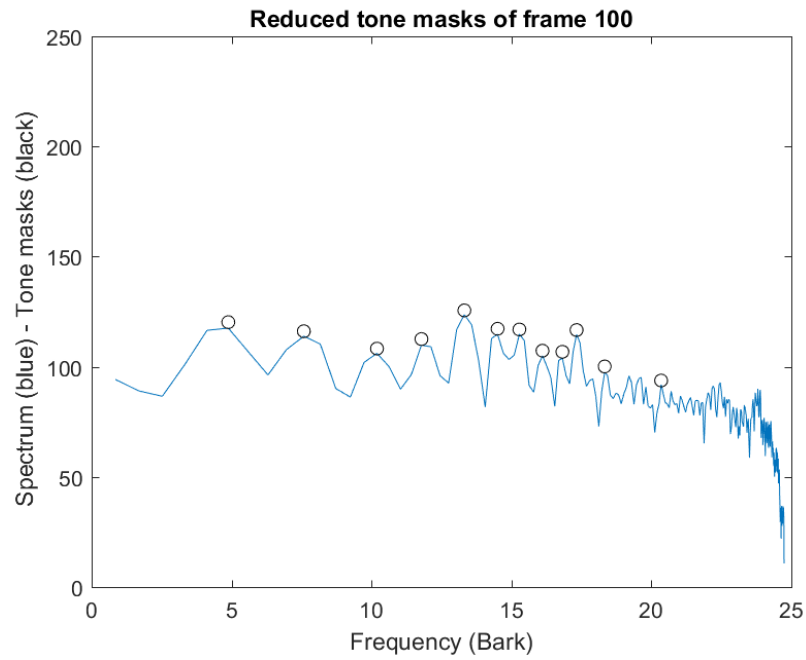
- **Βήμα 1.3: Μείωση και αναδιοργάνωση των μασκών**

Σε αυτό το στάδιο μειώνουμε τον αριθμό των μασκών για κάθε πλαίσιο και έχουμε τα παρακάτω αποτελέσματα:

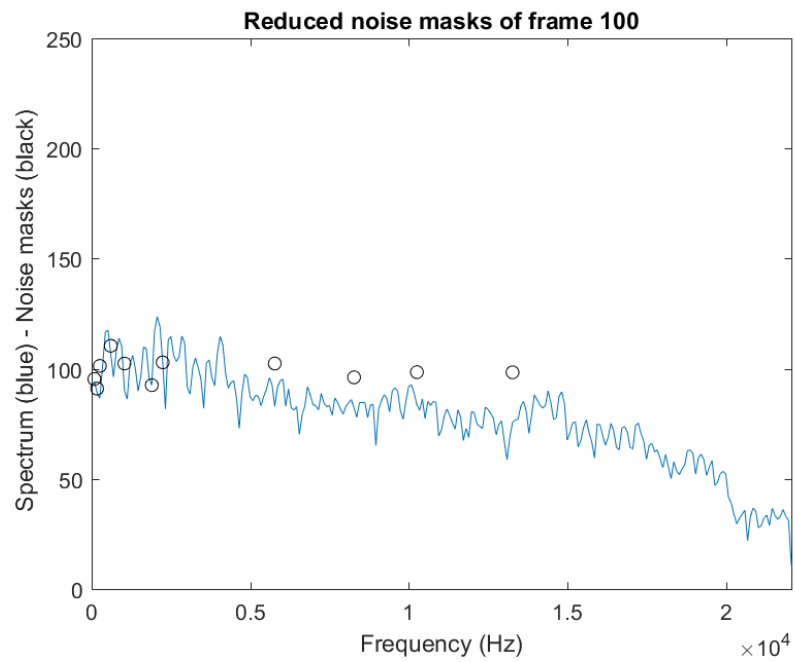
### Πλαίσιο 100



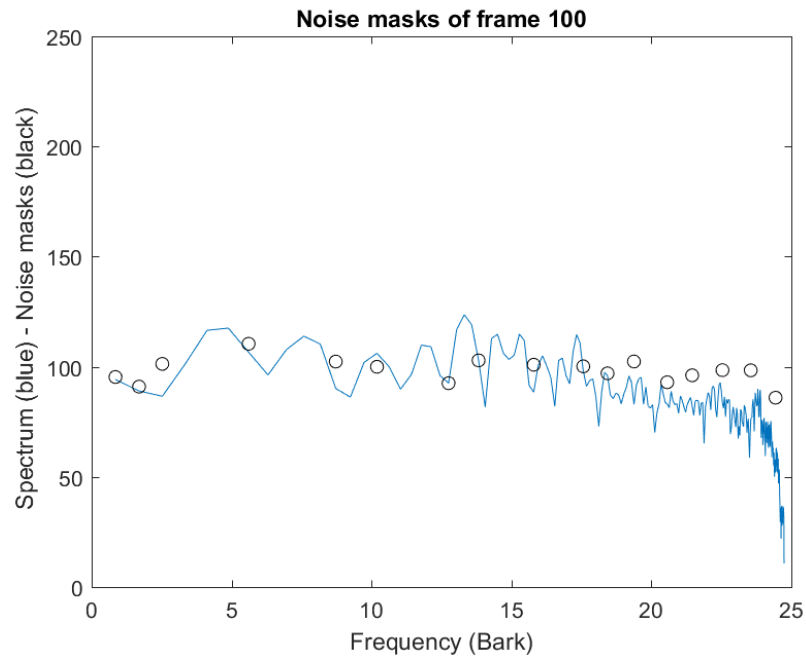
Εικόνα 16.Reduced tone masks of frame 100 in frequency



Εικόνα 17. Reduced tone masks of frame 100 in bark scale

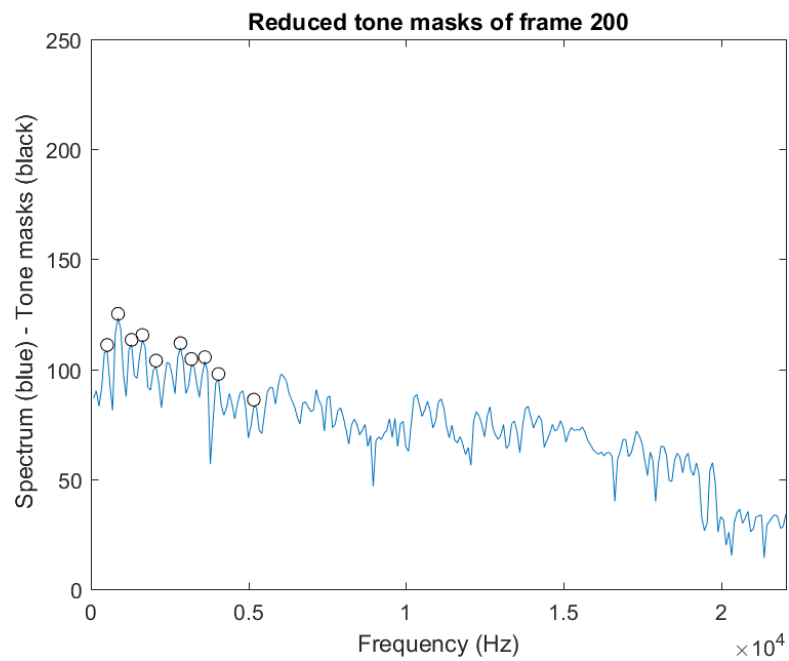


Εικόνα 18. Reduced noise masks of frame 100 in frequency

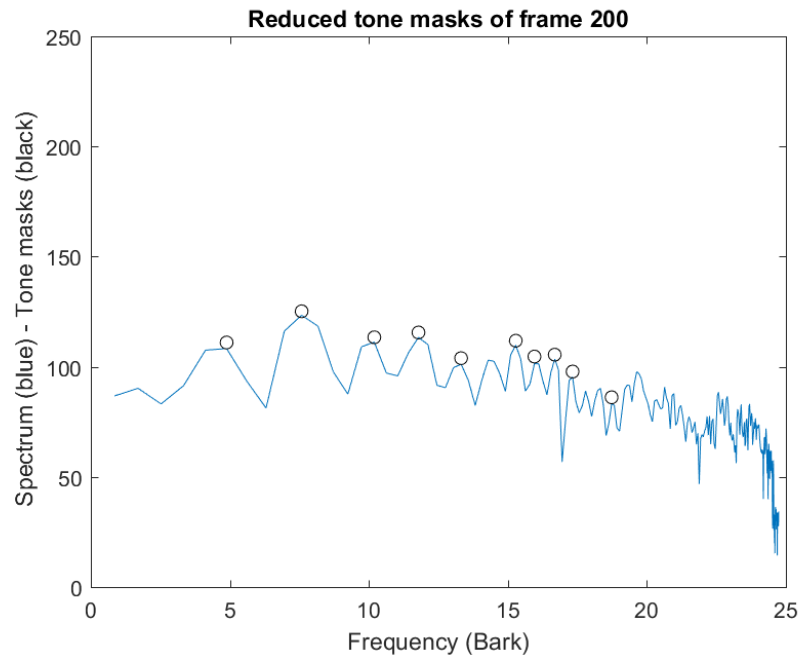


Εικόνα 19. Reduced noise masks of frame 100 in bark scale

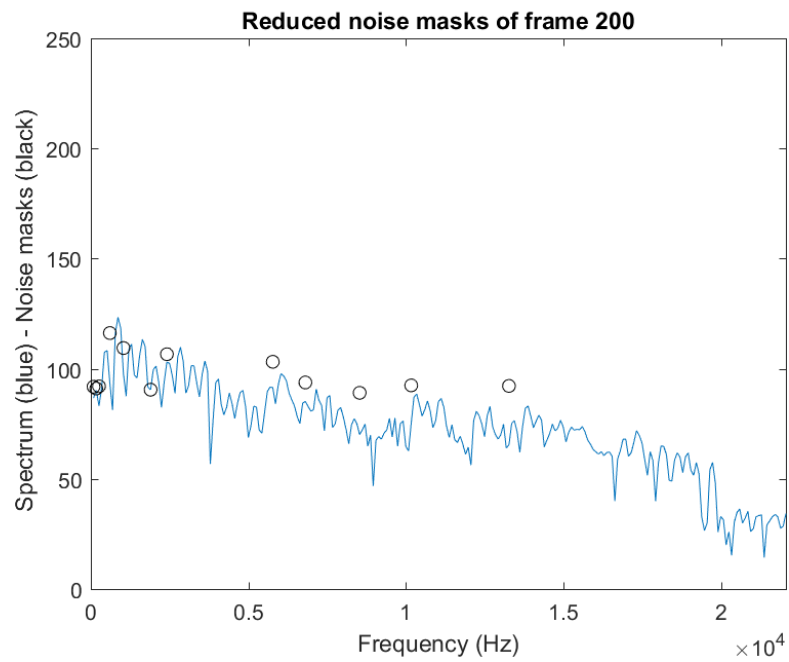
## Πλαίσιο 200



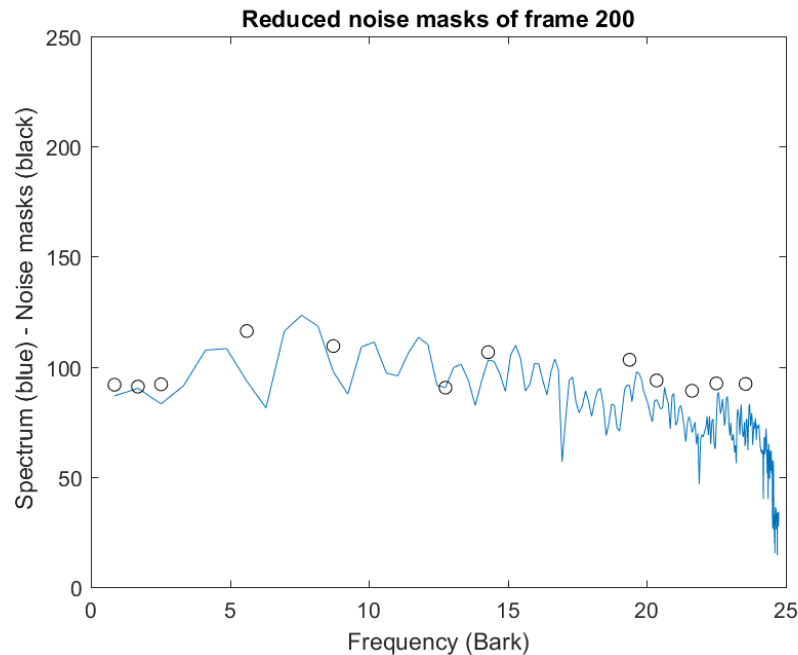
Εικόνα 20. Reduced tone masks of frame 200 in frequency



**Εικόνα 21. Reduced tone masks of frame 200 in bark scale**



**Εικόνα 22. Reduced noise masks of frame 200 in frequency**



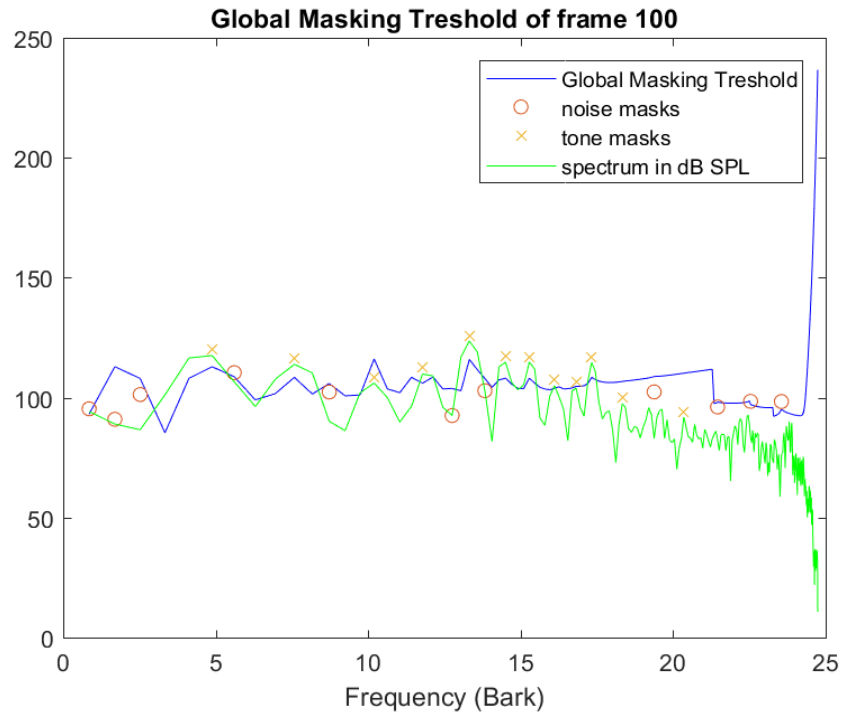
Εικόνα 23. Reduced noise masks of frame 200 in bark scale

- **Βήμα 1.4: Υπολογισμός των δυο διαφορετικών κατωφλίων κάλυψης (Individual Masking Thresholds)**

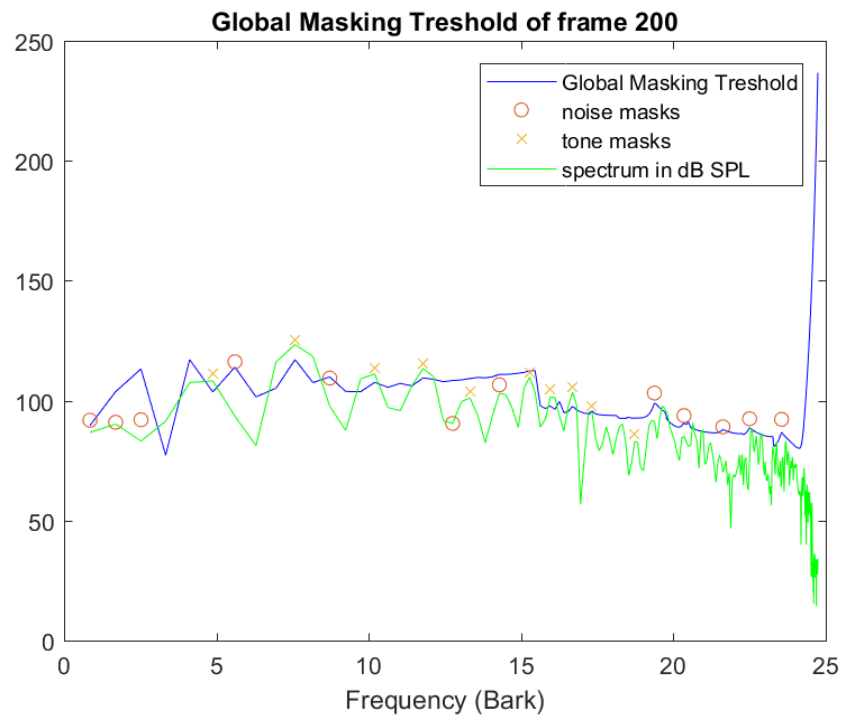
Σε αυτό το στάδιο υπολογίζουμε τα δύο διαφορετικά κατώφλια κάλυψης, τα οποία αντιπροσωπεύουν το ποσοστό κάλυψης στο σημείο  $i$  το οποίο προέρχεται από την μάσκα τόνου ή θορύβου στο σημείο  $j$ .

- **Βήμα 1.5: Υπολογισμός του συνολικού κατωφλίου κάλυψης (Global Masking Threshold)**

Στο τελευταίο στάδιο του ψυχοακουστικού μοντέλου συνδυάζουμε τα ξεχωριστά κατώφλια που υπολογίσαμε προηγουμένως για την δημιουργία του συνολικού κατωφλίου σε κάθε διακριτή συχνότητα ξεχωριστά. Παρακάτω βλέπουμε το συνολικό κατώφλι κάλυψης για το κάθε πλαίσιο ξεχωριστά:



**Εικόνα 24. Global Masking Threshold of frame 100**

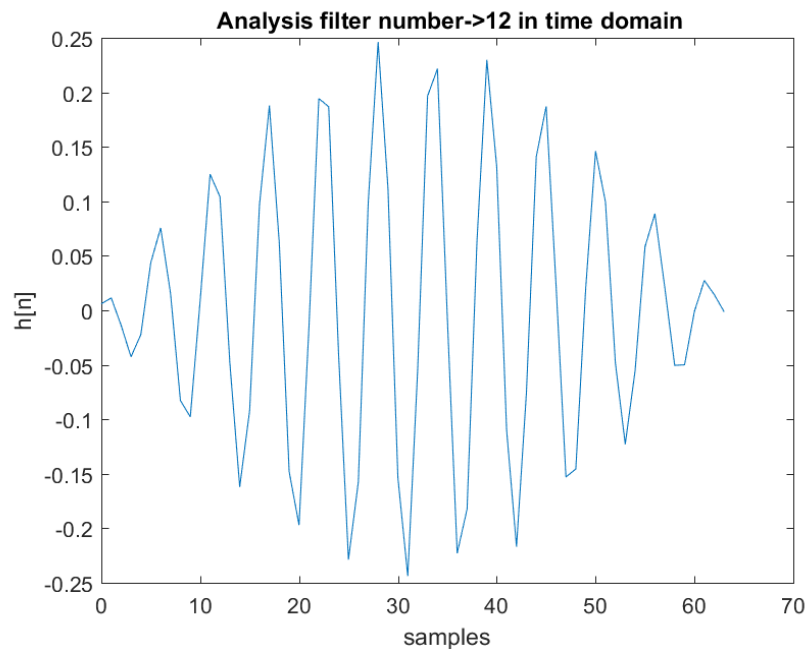


**Εικόνα 25. Global Masking Threshold of frame 200**

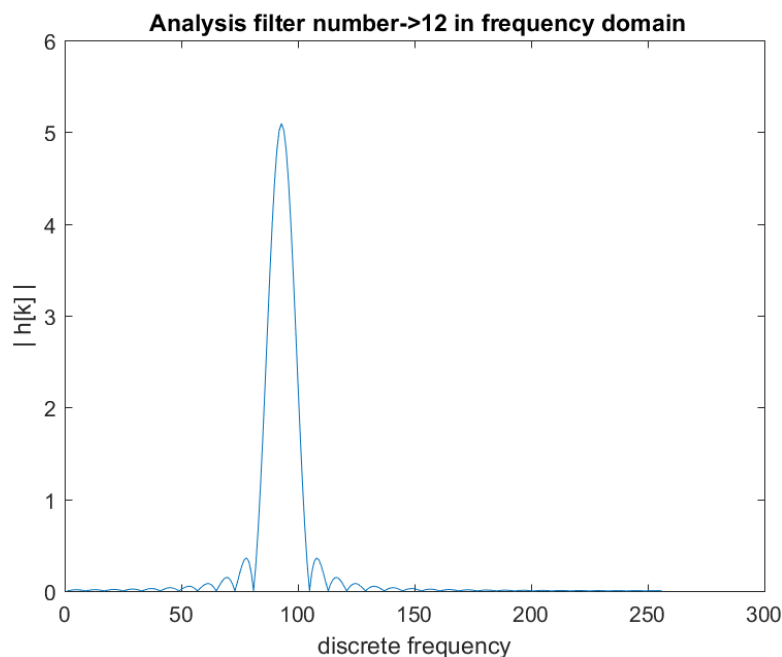
## Μέρος 2: Χρονο-Συχνотική Ανάλυση με Συστοιχία Ζωνοπερατών Φίλτρων

- **Βήμα 2.0: Συστοιχία Ζωνοπερατών Φίλτρων (Filterbank)**

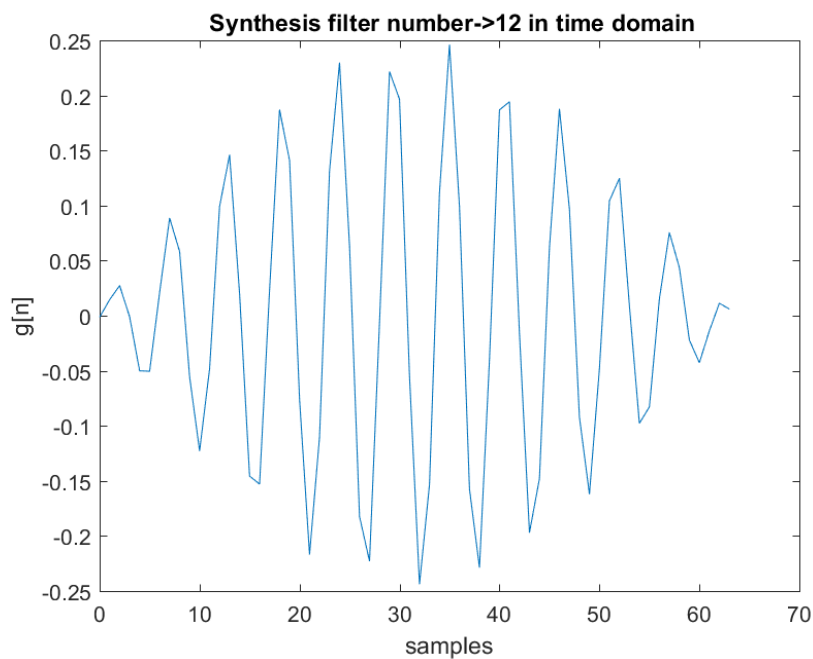
Σε αυτό το στάδιο ορίζουμε μία συστοιχία ζωνοπερατών φίλτρων, η οποία θα χρησιμοποιηθεί για την ανάλυση του σήματος στις κρίσιμες συνιστώσες του, ανά χρονικό πλαίσιο. Παρακάτω βλέπουμε την αναπαράσταση ενός φίλτρου από τα 32 στο χρόνο και την συχνότητα:



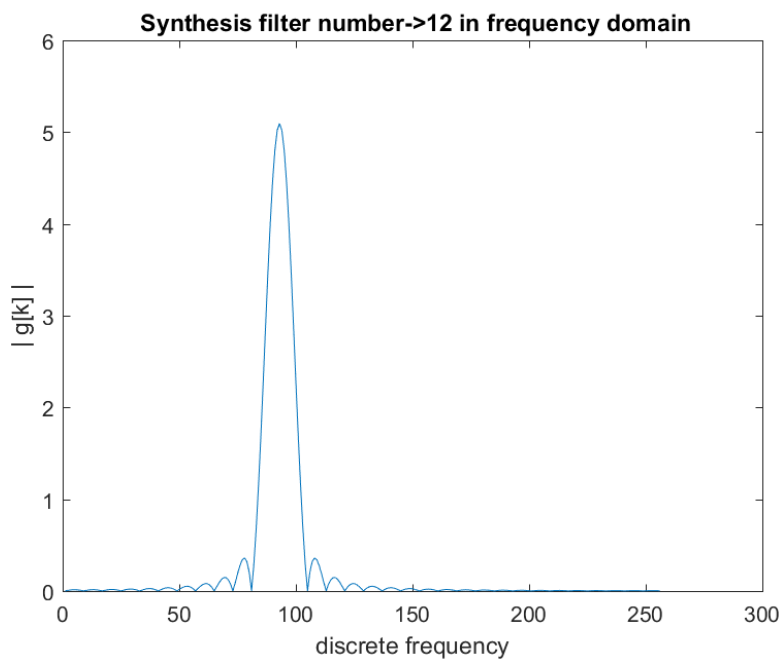
Εικόνα 26. Analysis filter number 12 in time domain



Εικόνα 27. Analysis filter number 12 in frequency domain



Εικόνα 28. Synthesis filter number 12 in time domain



Εικόνα 29. Synthesis filter number 12 in frequency domain

- **Βήμα 2.1: Ανάλυση με Συστοιχία Φίλτρων**

Σε αυτό το στάδιο κάνουμε συνέλιξη του κάθε πλαισίου με τα φίλτρα σύνθεσης και, στη συνέχεια, υποδειγματοληψία κατά παράγοντα 32 για να διαιρεθεί το αρχικό σήμα στις χρονικές του συνιστώσες. Χρησιμοποιήθηκαν οι έτοιμες συναρτήσεις του MATLAB **conv** και **downsample** αντίστοιχα.



### • Βήμα 2.2: Κβαντοποίηση

Σε αυτό το στάδιο υλοποιούμε έναν προσαρμοζόμενο ομοιόμορφο κβαντιστή όπου ο αριθμός των bits κωδικοποίησης ανά δείγμα ακολουθίας εξαρτάται κάθε φορά από την ελάχιστη τιμή του Absolute Threshold of Hearing στα δείγματα που εξετάζουμε. Παράλληλα, κάνουμε την ίδια διαδικασία χρησιμοποιώντας έναν μη προσαρμοζόμενο κβαντιστή με σταθερό αριθμό των 8 bits. Τα συμπεράσματα και οι αντίστοιχες γραφικές παρατίθενται στην ενότητα **Σχολιασμός Αποτελεσμάτων** που βρίσκεται παρακάτω.

### • Βήμα 2.3: Σύνθεση

Στο τελικό στάδιο, για κάθε πλαίσιο αποκωδικοποιούμε τις κβαντισμένες ακολουθίες, τις υπερδειγματοληπτούμε κατά παράγοντα 32, τις κάνουμε συνέλιξη με τα αντίστοιχα φίλτρα σύνθεσης και τις προσθέτουμε. Χρησιμοποιώντας την τεχνική Overlap-Add ενώνουμε τα σήματα ανά πλαίσιο στο τελικό μας σήμα το οποίο είναι μετατοπισμένο κατά 2M σε σχέση με το αρχικό.

## Μέρος 3: Σχολιασμός αποτελεσμάτων

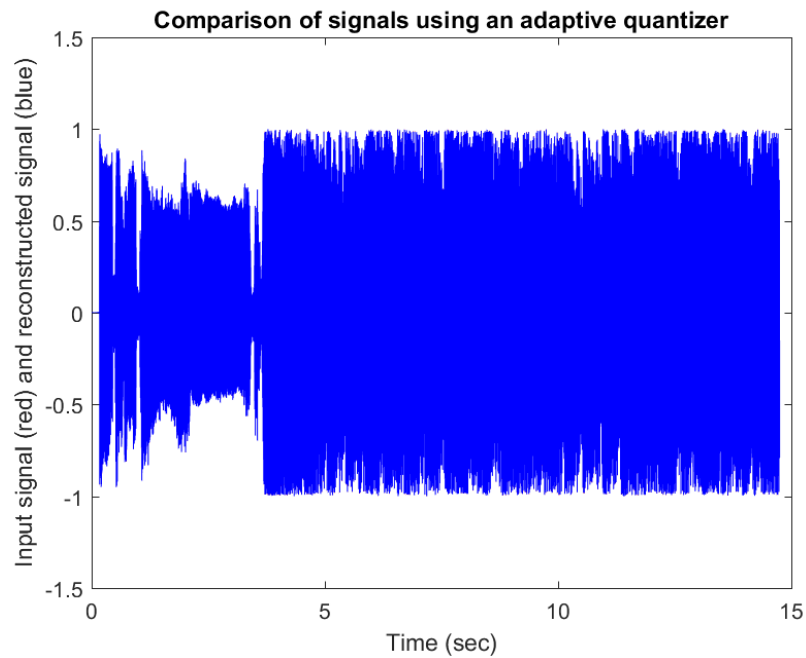
Στον παρακάτω πίνακα βλέπουμε οι τιμές που προέκυψαν χρησιμοποιώντας τις δύο ζητούμενες μεθόδους κβαντισμού. Σημειώνεται ότι ο συνολικός αριθμός bits του αρχικού σήματος ήταν 10406640:

Είδος κβαντιστή	Συνολικός αριθμός bits	Ποσοστό συμπίεσης	Μέσο Τετραγωνικό Λάθος
Μη προσαρμοζόμενος με σταθερό αριθμό bits (8)	5856768	56.2791	0.0548
Προσαρμοζόμενος κβαντιστής	7247016	69.6384	0.000000215

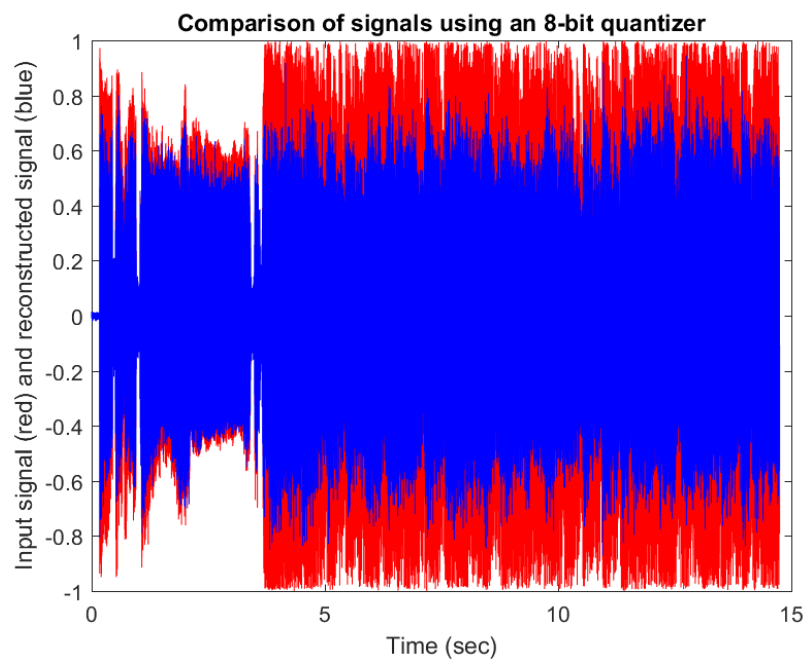
Παρακάτω βλέπουμε και τις γραφικές παραστάσεις κοινές για το σήμα εισόδου και το ανακατασκευασμένο σήμα για τις δύο μεθόδους που χρησιμοποιήσαμε (στην πρώτη γραφική λόγω του πολύ μικρού error η μία

## Digital Signal Processing

γραφική παράσταση πέφτει σχεδόν ακριβώς πάνω στην άλλη με αποτέλεσμα να μην διακρίνουμε το σήμα εισόδου:

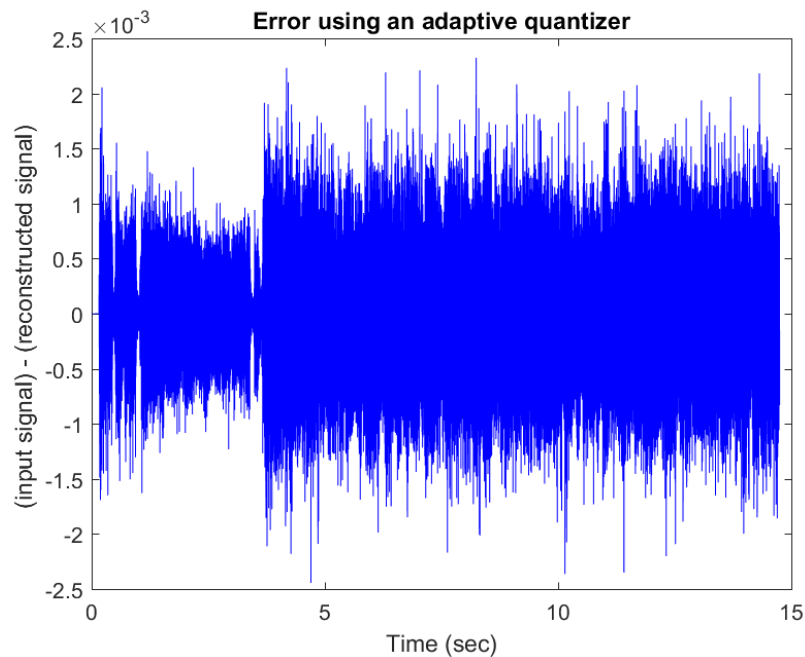


Εικόνα 30. Comparison of signals using an adaptive quantizer

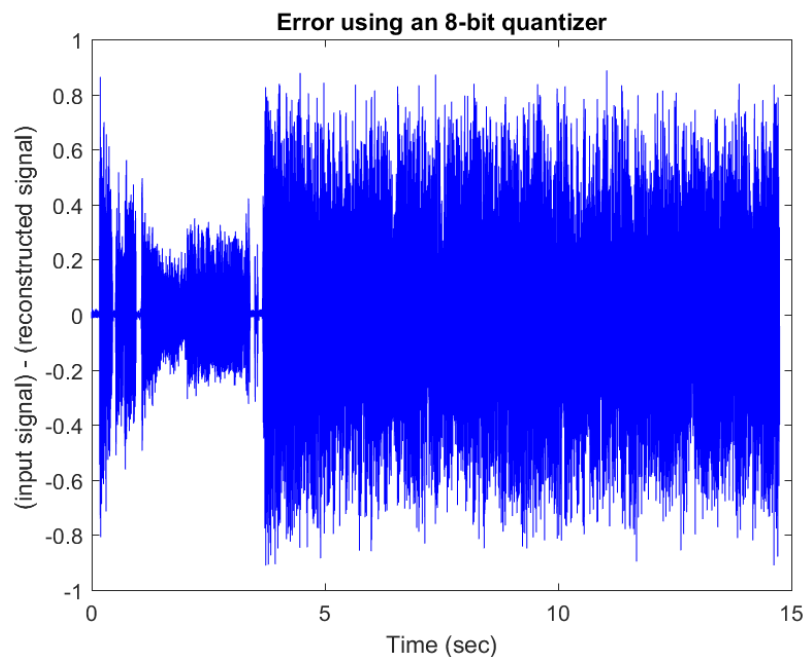


Εικόνα 31. Comparison of signals using an 8-bit quantizer

Τέλος, ακολουθεί και η αναπαράσταση του λάθους (διαφορά των δύο σημάτων) για καθεμία από τις μεθόδους:



Εικόνα 32. Error using an adaptive quantizer



Εικόνα 33. Error using an 8-bit quantizer

Αυτό που βλέπουμε, λοιπόν, είναι ότι με τον προσαρμοζόμενο κβαντιστή έχουμε σχεδόν μηδενικό λάθος, πράγμα που φαίνεται και αν ακούσουμε το ανακατασκευασμένο σήμα, το οποίο ακούγεται ακριβώς ίδιο με τον αρχικό. Αντίθετα, με τον κβαντιστή των 8-bit έχουμε ένα αρκετά σημαντικό λάθος, το οποίο γίνεται αντιληπτό και από τον θόρυβο που ακούμε στο προκύπτον σήμα. Είναι ένας θόρυβος ο οποίος δεν μας εμποδίζει να ακούσουμε το τραγούδι αλλά

## Digital Signal Processing

---

επηρεάζει σε μεγάλο βαθμό την ποιότητα της μουσικής. Όσον αφορά τώρα τον βαθμό συμπίεσης, με τον 8-bit κβαντιστή έχουμε λιγότερα bit εξόδου αλλά πολλά από αυτά που χάθηκαν ήταν χρήσιμη πληροφορία και δημιούργησαν τον θόρυβο που προαναφέραμε. Αντίθετα, με τον προσαρμοζόμενο κβαντιστή πετύχαμε μία σημαντική μείωση στον αριθμό των bits αφαιρώντας ουσιαστικά μόνο τα bits που δεν ήταν χρήσιμα.