

ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΣΗΜΑΤΟΣ

Εργαστηριακό Project 1

Όνοματεπώνυμο: Νίκος Κατσαϊδώνης

Αριθμός Μητρώου: 0 3 1 2 1 8 6 8

Μέρος 1ο - Σύστημα Εντοπισμού Τηλεφωνικών Τόνων (Telephone Touch – Tones)

● Σκοπός της άσκησης αυτής είναι η **κατασκευή των τηλεφωνικών τόνων** και ο **εντοπισμός του πλήκτρου** που έχει πατηθεί. Επιπλέον θα εξοικειωθούμε με το **διακριτό μετασχηματισμό Fourier** (DFT) καθώς και με την **παραθύρωση σημάτων** στο πεδίο του χρόνου.

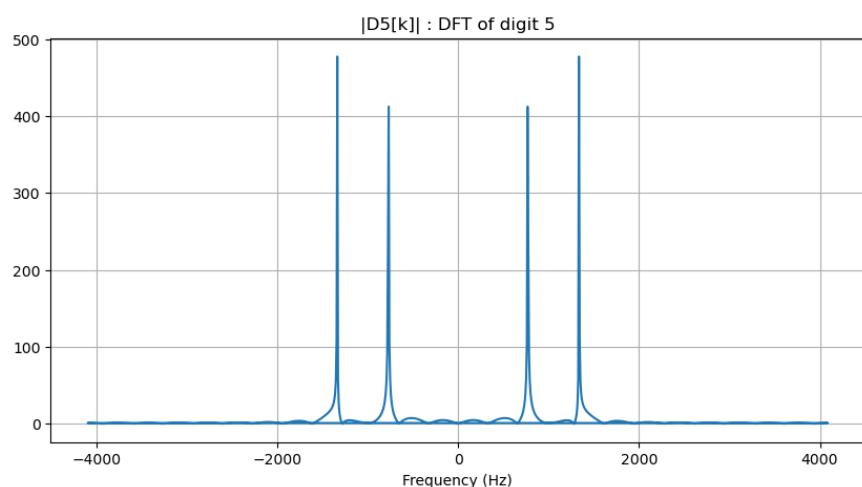
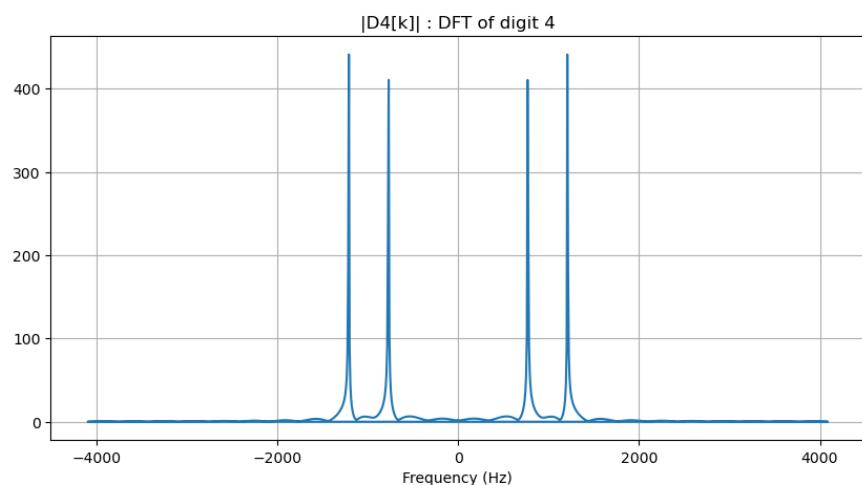
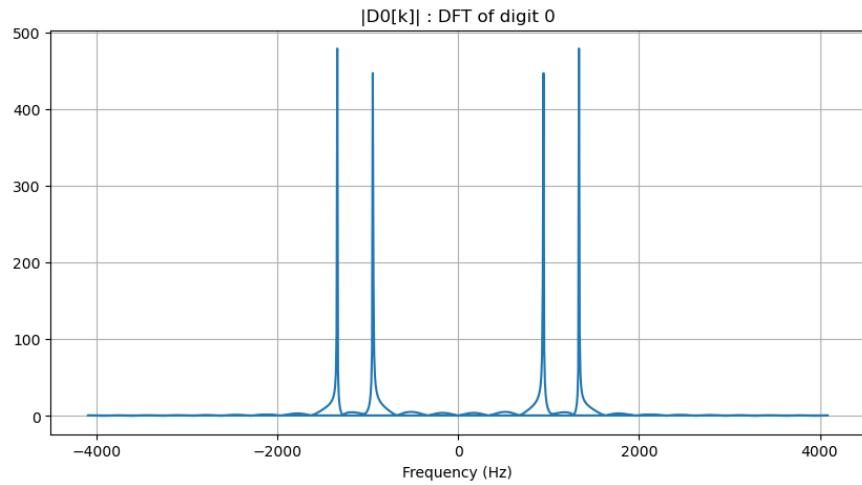
Ερώτημα 1.1

```
d0 = np.sin(0.7217*n) + np.sin(1.0247*n)
d1 = np.sin(0.5346*n) + np.sin(0.9273*n)
d2 = np.sin(0.5346*n) + np.sin(1.0247*n)
d3 = np.sin(0.5346*n) + np.sin(1.1328*n)
d4 = np.sin(0.5906*n) + np.sin(0.9273*n)
d5 = np.sin(0.5906*n) + np.sin(1.0247*n)
d6 = np.sin(0.5906*n) + np.sin(1.1328*n)
d7 = np.sin(0.6535*n) + np.sin(0.9273*n)
d8 = np.sin(0.6535*n) + np.sin(1.0247*n)
d9 = np.sin(0.6535*n) + np.sin(1.1328*n)
```

Κατασκευάζουμε τα παραπάνω σήματα που αντιστοιχούν στους τόνους και με την συνάρτηση **display(Audio())** (συχνότητα δειγματοληψίας **fs=8192Hz**) μπορούμε να ακούσουμε καθέναν από αυτούς.

Ερώτημα 1.2 Υπολογίζουμε τον **DFT 1024** σημείων των **τόνων 0,4,5** με τη χρήση της συνάρτησης **fft.fft()** και **fft.freq()** της numPy.

Παραθέτουμε το αποτέλεσμα:



Όπως αναμέναμε, κάθε DFT αποτελέσται από **4 peaks** συμμετρικά ανά δύο λόγω των **δύο ημιτόνων** που υπάρχουν στον κάθε τόνο. Τα τρία γραφήματα, έχουν αρκετά κοντά τις συχνότητες στις οποίες κάνουν τα peaks τους κάτι το οποίο αιτιολογείται βάση των πολύ κοντινών ω στα αρχικά σήματα.

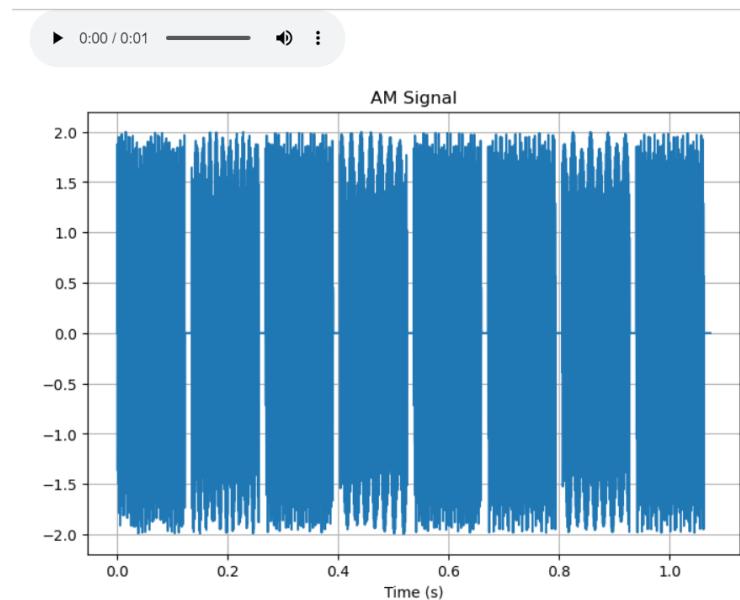
Ερώτημα 1.3

Δεδομένου του αριθμού μητρώου **AM = 0 3 1 2 1 8 6 8**, δημιουργούμε μία λίστα στοιχείων (**AM_signal_list**) που περιέχει τους αντίστοιχους τόνους αφού πρώτα **προεκτείνουμε τον καθέναν από αυτούς με 100 μηδενικά** με την βοήθεια της συνάρτησης **concatenate()** της numpy για να μην υπάρχει δυσκολία διάκρισης των τόνων κατά την ακρόαση και δημιουργούμε μιά νέα λίστα (**extended_signals_list**). Με τη χρήση της της **write()** της `scipy.io.wavfile` γράφουμε το αποτέλεσμα στο αρχείο ήχου **“tone_sequence.wav”** και το κάνουμε **display(Audio())** προκειμένου να το ακούσουμε.

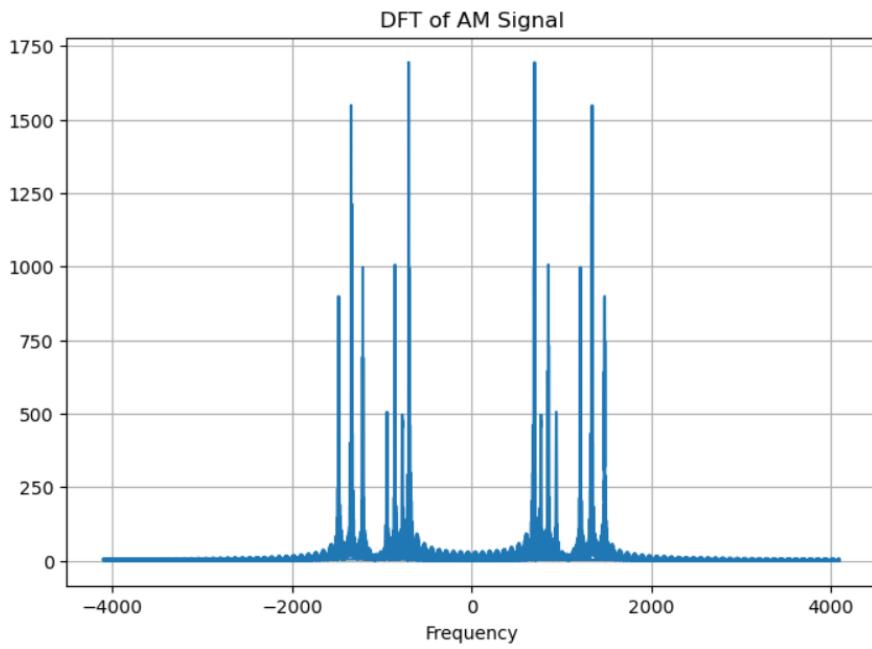
*Σημειώνουμε ότι αντιμετωπίσαμε πρόβλημα με την εγγραφή στο αρχείο και προσθέσαμε την κανονικοποίηση του AM : **normalized_AM = np.int16((AM/AM.max() * 32767))**

Επίσης, στο **tone_sequence.wav φαίνεται να ακούγεται ένας τόνος λιγότερος ενώ στο **display(Audio())** ακούγεται κανονικά.

Παραθέτουμε το **σήμα στο χρόνο** και τον **DFT** αυτού παρακάτω



Τονικό σήμα AM = 0 3 1 2 1 8 6 8

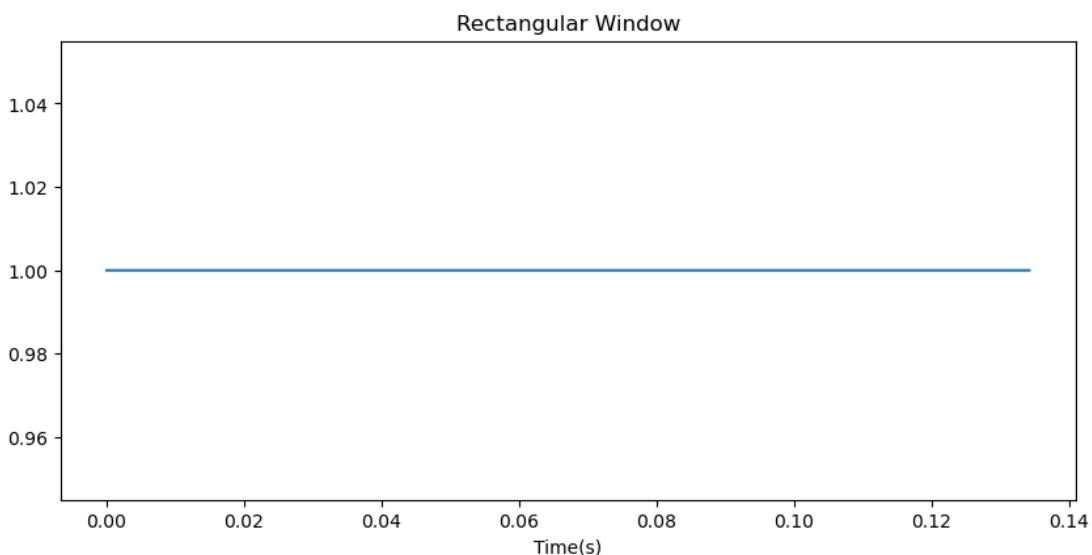


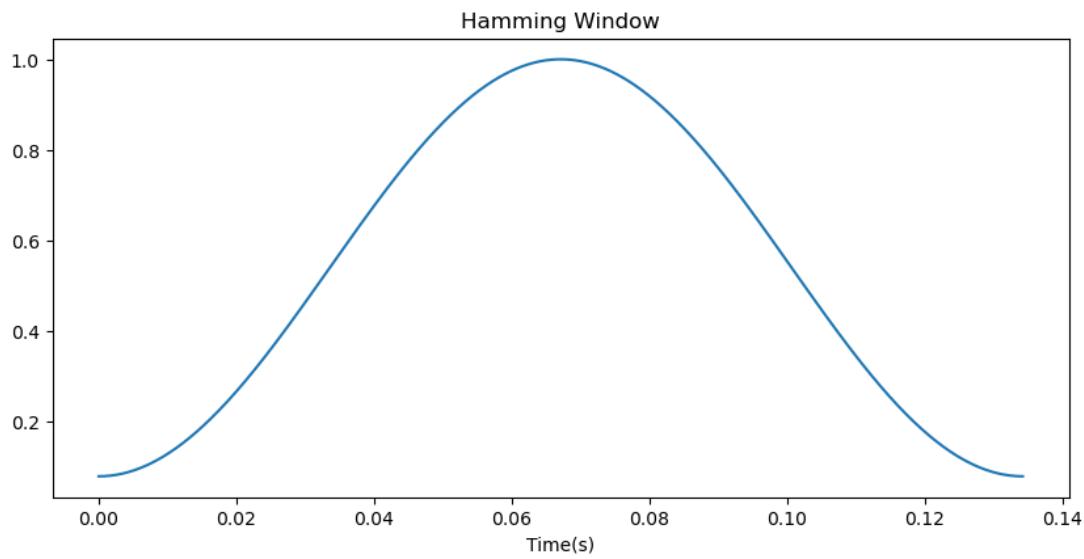
$$DFT \text{ τονικού σήματος AM} = 0 \ 3 \ 1 \ 2 \ 1 \ 8 \ 6 \ 8$$

Παρατηρούμε λιγότερα peaks από ότι αναμέναμε καθώς κάποια ψηφία επαναλαμβάνοντι στο σήμα (1,8) και άρα υπάρχει επικάλυψη των κορυφών καθώς επίσης και για τον λόγο ότι τα peaks είναι πολύ κοντά μεταξύ τους στο φάσμα.

Ερώτημα 1.4

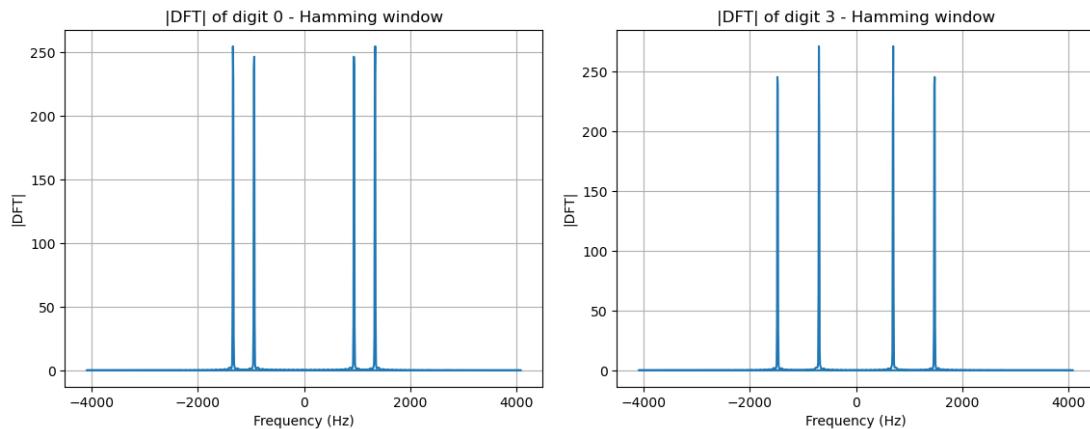
Στο ερώτημα αυτό θα δημιουργήσουμε παράθυρα **hamming** και **τετραγωνικά** μήκους $L=1100$ όσο δηλαδή και των τόνων μας μετά την προσθήκη 100 μηδενικών:

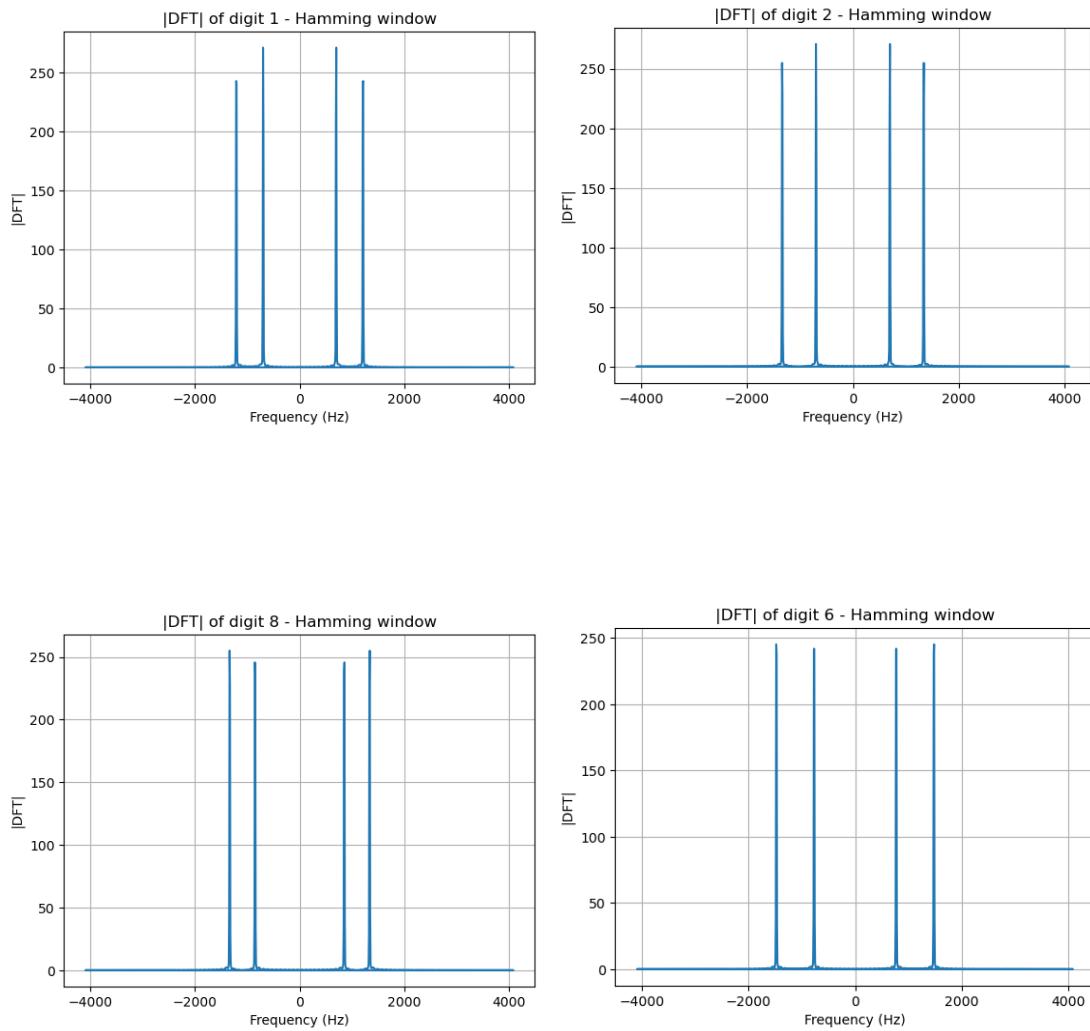




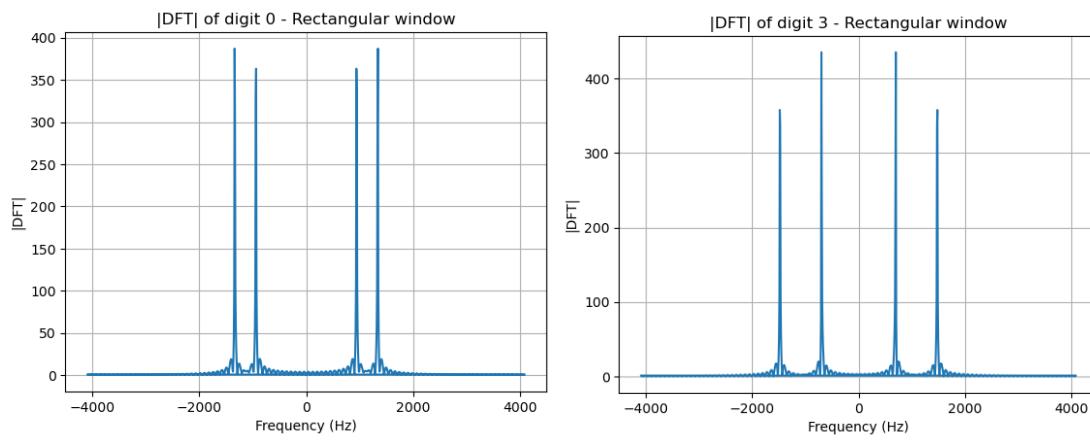
Παραθυρώνουμε τους τόνους [d0,d3,d1,d2,d1,d8,d6,d8] με τα τετραγωνικά και hamming παράθυρα και παριστάνουμε γραφικά τα αποτελέσματα (οι τόνοι 1,8 επαναλαμβάνονται 2 φορές στο σήμα AM):

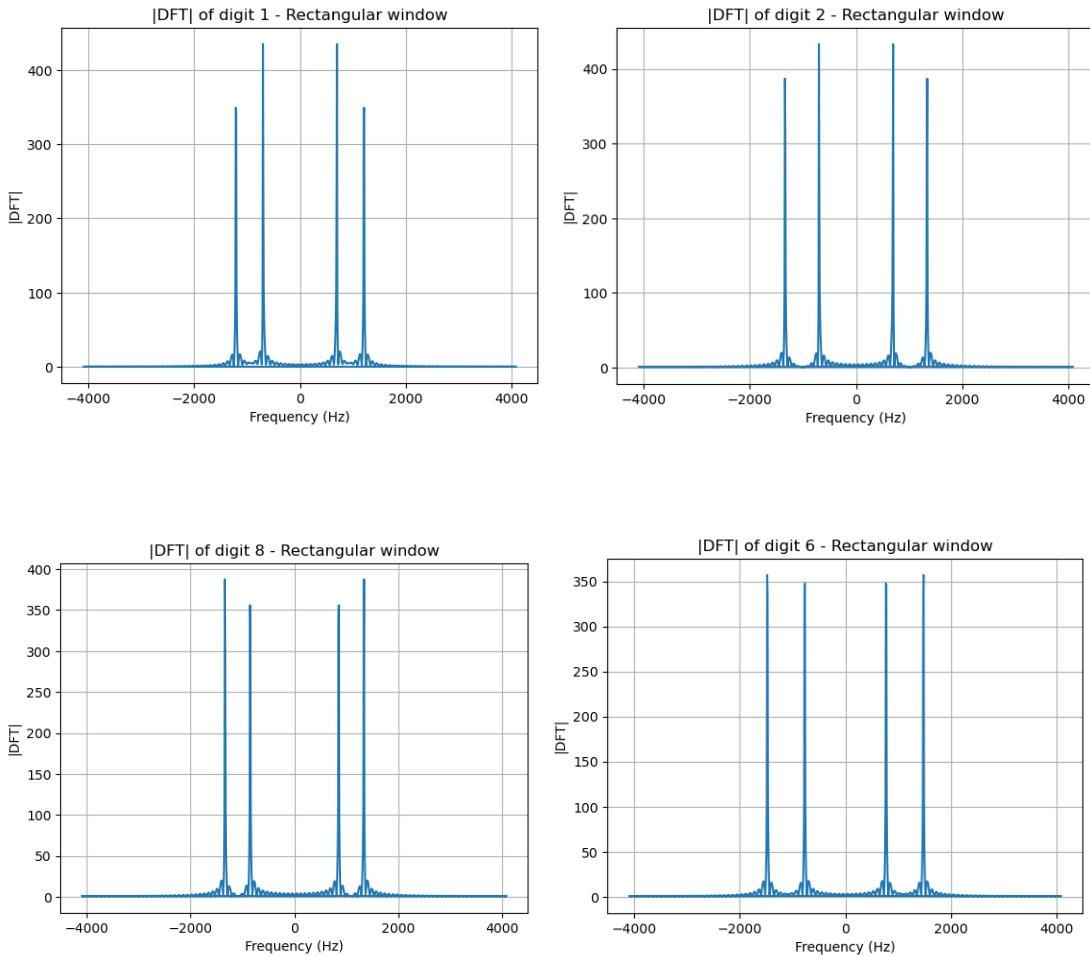
Παράθυρα Hamming:





Τετραγωνικά παράθυρα:





Αυτό που παρατηρούμε συγκρίνοντας τα δύο παράθυρα είναι ότι με το **τετραγωνικό παράθυρο** έχουμε αρκετά **μεγαλύτερο πλάτος των κεντρικών λοβών**. Επιπλέον, παρατηρούμε **υψηλότερο spectral leakage** κάτι το οποίο σημαίνει ότι έχουμε **διασπόρα της ενέργειας** σε μεγαλύτερο εύρος συχνοτήτων (βλέπε μικρά peaks στις χαμηλές συχνότητες).

Ερώτημα 1.5

Κάθε ημιτονικό σήμα, εφόσον έχει συγκεκριμένες τιμές γωνιακής συχνότητας μας παρέχει έναν DFT με peaks σε συγκεκριμένα σημεία. Αφού δημιουργήσουμε μία λίστα με όλους τους DFT των τόνων (**DFTs_of_tones**), υπολογίζουμε την ενέργεια καθενός DFT γύρω από κάθε συχνότητα (από το 0 έως το μήκος του DFT αφού ο DFT είναι συμμετρικός ως προς το μηδέν και άρα αρκεί να βρούμε τις θετικές συχνότητες) και την βάζουμε σε μία λίστα (**energies**). Έπειτα,

διατρέχουμε την λίστα που περιέχει τις ενέργειες γύρω από κάθε συχνότητα (για κάθε τόνο) και για την **μεγαλύτερη και δεύτερη μεγαλύτερη ενέργεια του τόνου**, αποθηκεύουμε την συχνότητά τους στην λίστα k. Τώρα λοιπόν έχουμε τις συχνότητες του DFT στις οποίες το κάθε ψηφίο παρουσιάζει τις δύο μεγαλύτερες τιμές της ενέργειάς του. **Αυτές είναι χαρακτηριστικές για το κάθε ψηφίο.** Παραθέτουμε το μήνυμα που τυπώνεται καθώς και το ζεύγος των δύο μεγαλύτερων ενεργειών του κάθε τόνου που αντιστοιχούν στα peaks.

List of k =	The two greater energies of digit 0 [229076.79430721106, 199350.37704250013]
digit0:[167, 118]	The two greater energies of digit 1 [216220.8361380998, 216220.83613809978]
digit1:[937, 87]	The two greater energies of digit 2 [227455.31999699745, 227455.3199969973]
digit2:[167, 857]	The two greater energies of digit 3 [221957.65273139894, 216948.59833747908]
digit3:[185, 87]	The two greater energies of digit 4 [194312.06347057933, 168323.35386232645]
digit4:[151, 96]	The two greater energies of digit 5 [228014.76713314166, 228014.7671331416]
digit5:[167, 857]	The two greater energies of digit 6 [221573.57742929985, 168704.09347711658]
digit6:[185, 96]	The two greater energies of digit 7 [194296.27568024144, 194296.27568024138]
digit7:[151, 873]	The two greater energies of digit 8 [227823.60072293025, 227823.6007229302]
digit8:[167, 857]	The two greater energies of digit 9 [221808.9511575135, 150805.3735174227]
digit9:[185, 917]	.

Ερώτημα 1.6

Η βασική ιδέα είναι με τη λήψη του σήματος, η συνάρτηση να υπολογίζει τον Fourier, να εντοπίζει ποιο ζεύγος συχνοτήτων από τη λίστα k εμφανίζεται και να αποκωδικοποιεί τα αντίστοιχα ψηφία τόνων που πατήθηκαν.

Ερώτημα 1.7

Διαδοχική εφαρμογή της ttdecode στα σήματα "easySig", "ediumSig" και "hardSig" και εμφάνιση των ψηφίων τους.

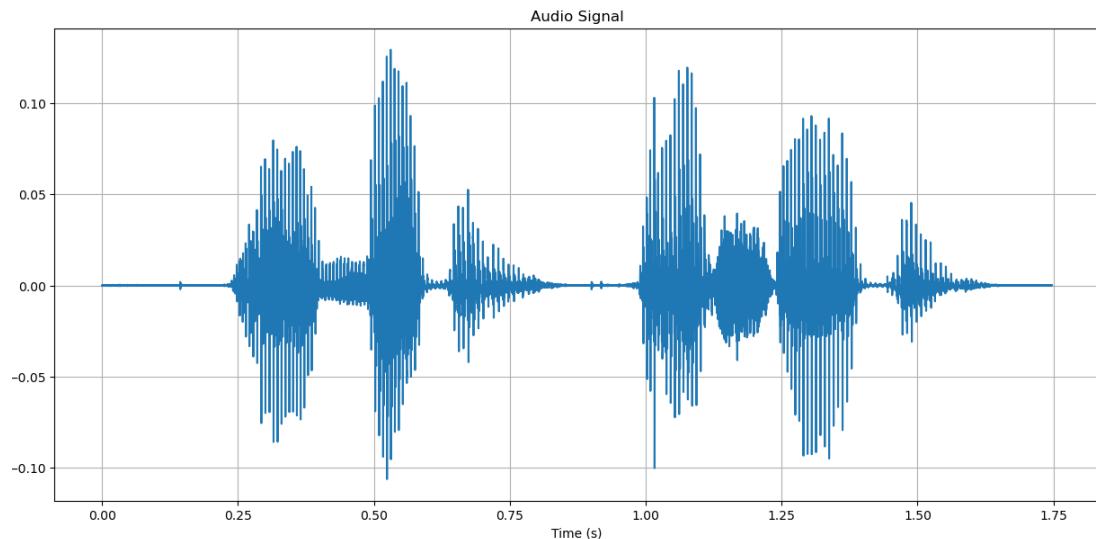
Μέρος 2ο – Χαρακτηριστικά Βραχέος Χρόνου Σημάτων Φωνής και Μουσικής.

- Σκοπός της άσκησης αυτής είναι η εξοικείωση με τις **μετρήσεις βραχέος χρόνου**. Πρόκειται για μετρήσεις που λαμβάνονται είτε στο **πεδίο του χρόνου είτε στο πεδίο της συχνότητας** και τα δύο με τη βοήθεια κάποιου κυλιόμενου παραθύρου (window). Στην πρώτη κατηγορία εντάσσονται τα μεγέθη **Ενέργεια Βραχέος Χρόνου (Short Time Energy)** καθώς και ο **Ρυθμός Εναλλαγής Προσήμου** του σήματος (**Zero Crossing Rate**) ενώ από τη δεύτερη κατηγορία θα ασχοληθούμε την μέτρηση του **Φασματικού Κέντρου (Spectral Centroid)** και την **Φασματική Ροή (Spectral Flux)**.

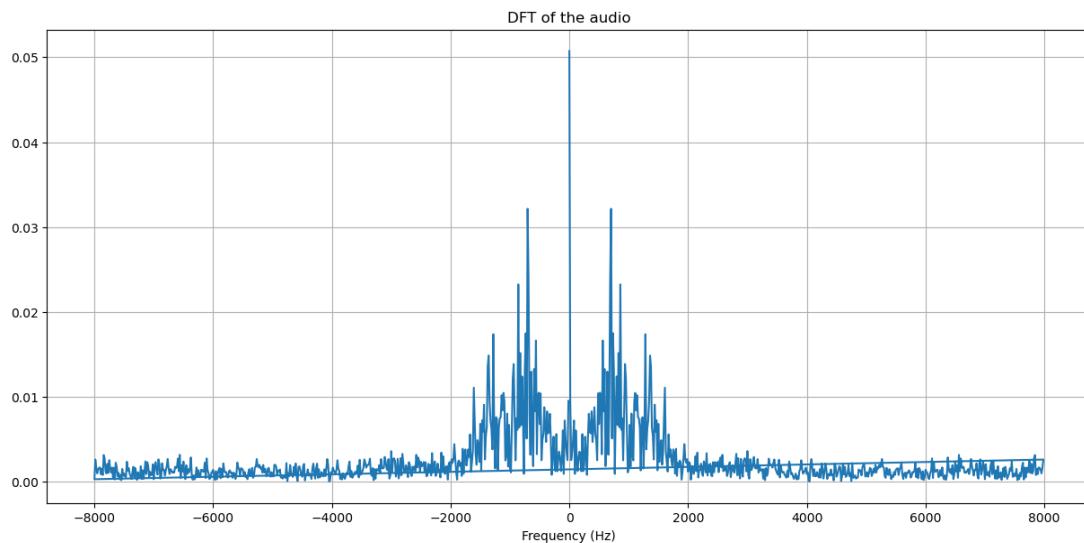
Ερώτημα 2.1

Στο ερώτημα αυτό θα επεξεργαστούμε το **σήμα φωνής speech_utterance.wav** το οποίο παρέχεται το συμπληρωματικό υλικό του Project. Ακούγοντάς το διαπιστώνουμε ότι πρόκειται για το μήνυμα “Oh Mother, I saw them.” Φορτώνουμε το σήμα με την συνάρτηση **load()** της librosa με συχνότητα δειγματοληψίας ίση με **fs = 16kHz**. Απεικονίζουμε το σήμα **διάρκειας 1.75 seconds** στον χρόνο καθώς και τον **DFT 1024 σημείων** του σήματος. Τα αποτελέσματα φαίνονται

παρακάτω:



Σήμα φωνής στο πεδίο του χρόνου



DFT 1024 σημείων σήματος φωνής.

Παρατηρούμε ότι το φάσμα του σήματος εκτείνεται μεταξύ των συχνοτήτων **-8kHz – 8kHz**.

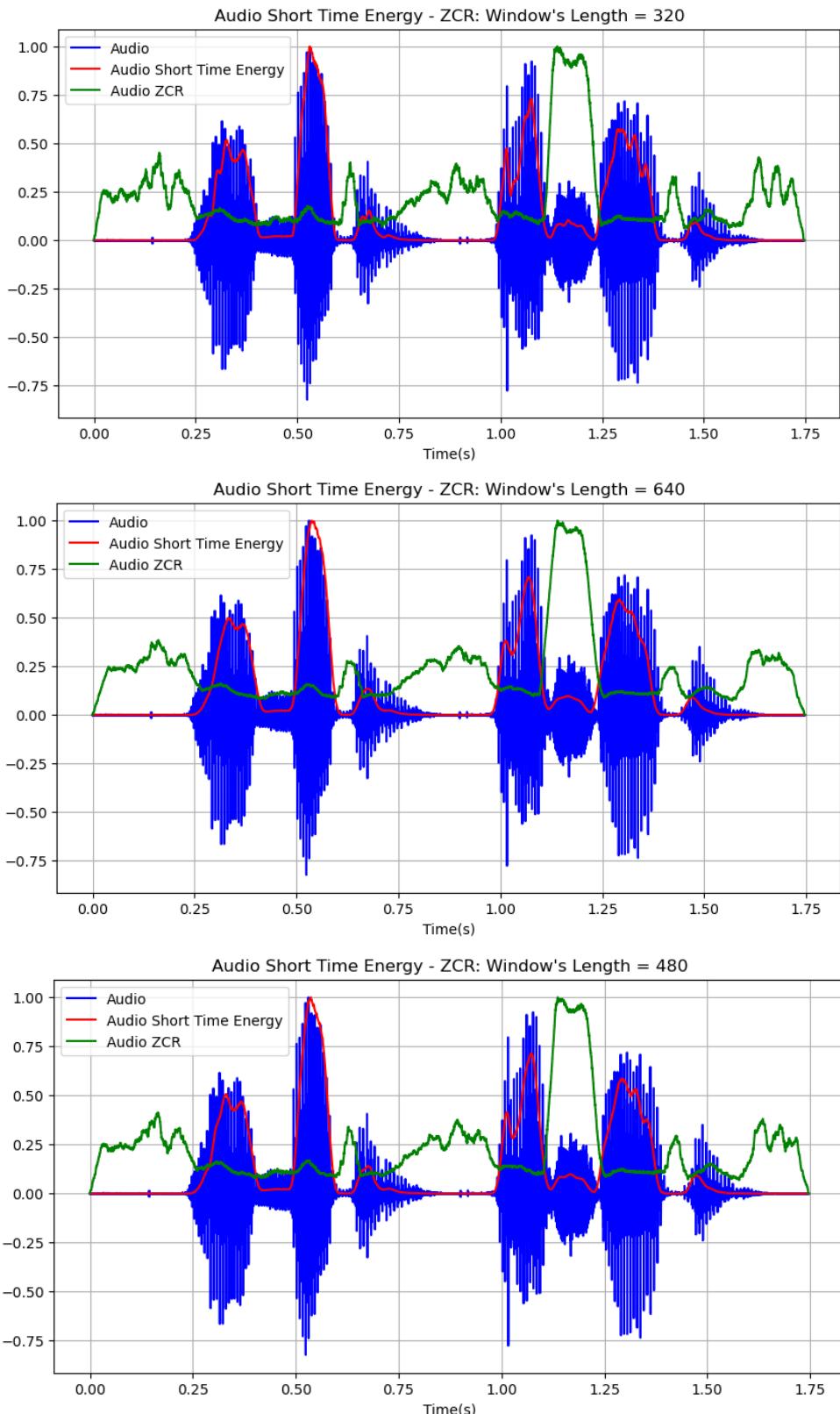
Ερώτημα 2.2

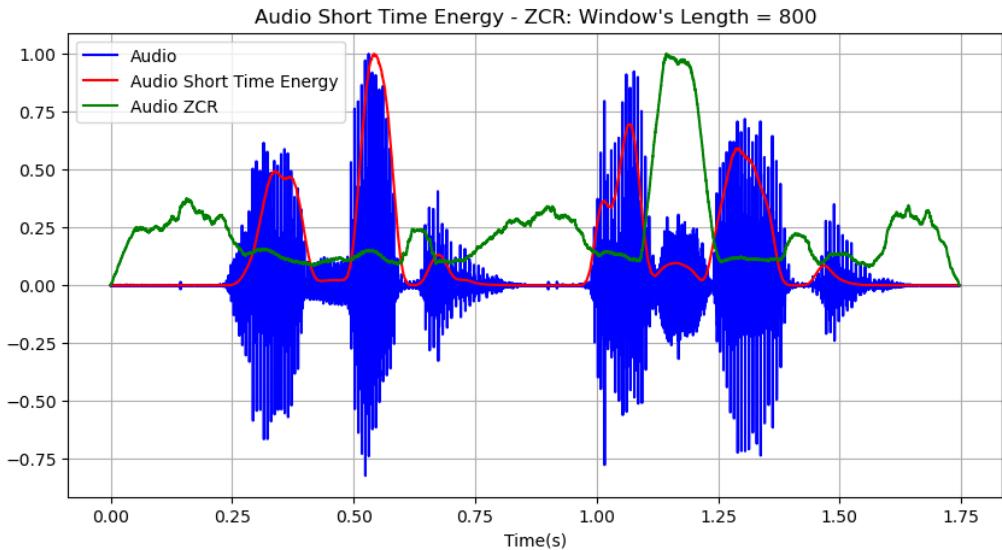
Θα υπολογίσουμε την **ενέργεια βραχέος χρόνου (En)** και τον **ρυθμό εναλλαγής προσήμου (ZCR)** του σήματος για τις τιμές παραθύρων **20-50ms** ανά 10ms άρα μήκους $L = 0.02 * fs = 320$ έως $L = 0.05 * fs = 800$. Η ενέργεια βραχέος χρόνου και το ZCR δίνονται από τους τύπους:

$$En = \sum_{m=-\infty}^{m=\infty} |x[m]|^2 \bullet w[n-m]$$

$$Zn = \sum_{m=-\infty}^{\infty} |sgn(x[m]) - sgn(x[m-1])| \bullet w[n-m]$$

Αφού φτιάξουμε τα παράθυρα **hamming** (για το En) και **rectangular** (για το Zn) δημιουργούμε τα μεγέθη αυτά με τη χρήση της **συνέλιξης** και συγκεκριμένα της συνάρτησης **convolve** της scipy.signal library. Τα μετατοπισμένα σήματα, όπου αυτά χρειάστηκαν, υπολογίστηκαν με χρήση της **roll(signal,n)** της numpy η οποία **μετατοπίζει κυκλικά** κατά n θέσεις δεξιά το σήμα (ή κατά -n αν θέλουμε αριστερή μετατόπιση). Παραθέτουμε τα μεγέθη για τα διάφορα παράθυρα σε κοινό διάγραμμα παρακάτω:





Παρατηρούμε μία **σταδιακή εξομάλυνση των γραφημάτων** Εν και ZCR με την αύξηση του μεγέθους των παραθύρων καθώς εξομαλύνονται πιο γρήγορα οι αλλαγές στο σήμα. Αυτό αφενός μας παρέχει μια ελαφρώς **λιγότερο λεπτομερή εικόνα** της ενέργειας αλλά μας **αποκόπτει τον θόρυβο** παρέχοντας έτσι μια πιο ρεαλιστική εικόνα για την ενέργεια και τον ρυθμό εναλλαγής προσήμου του σήματος μας

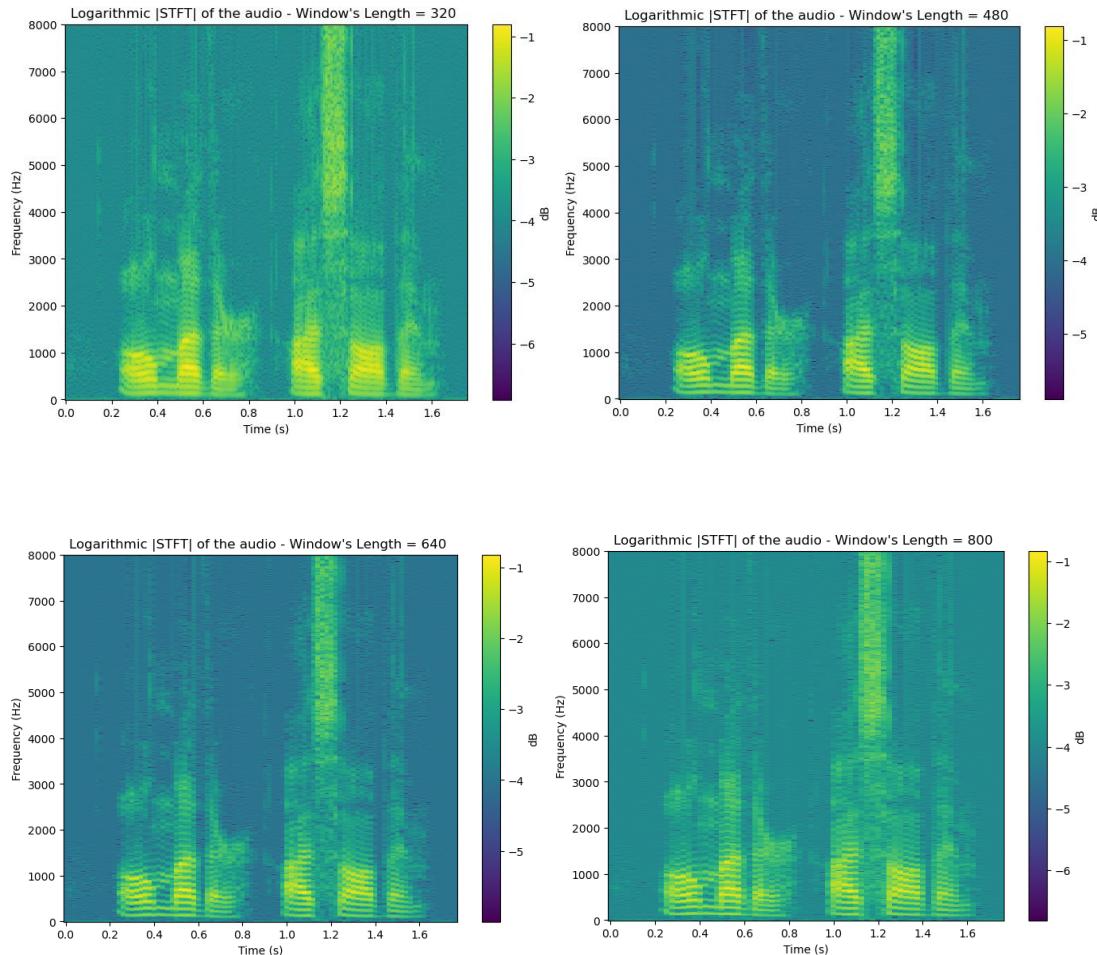
Οι μετρήσεις μας αυτές, θα ήταν χρήσιμες για την περαιτέρω επεξεργασία του σήματος φωνής. Πιο συγκεκριμένα θα μπορούσαμε να διαχωρίσουμε τα χρονικά διαστήματα στα οποία υπάρχει **σιωπή** από αυτά που περιέχουν **ήχο** καθώς τα πρώτα αντιστοιχούν σε **σημεία πολύ μικρής έως μηδενικής ενέργειας** και τα δεύτερα σε **σημεία χαμηλού zero crossing rate**.

Επιπλέον, τα διαγράμματα αυτά μας παρέχουν μία καλή εικόνα για τον εντοπισμό **έμφωνων** (“Oh Mother, I saw them”) ή **άφωνων** ήχων. Οι **έμφωνοι** ήχοι φαίνεται να είναι στις περιοχής όπου η **ενέργεια είναι σε peak και το zero crossing rate είναι χαμηλό** ενώ οι **άφωνοι** σε περιοχές **χαμηλής ενέργειας και σε peaks tou zero crossing rate**.

Ερώτημα 2.3

Για τις διάφορες προαναφερθείσες τιμές μήκους παραθύρου υπολογίζουμε το **Short Time Fourier Transform** που μας δίνει μια εικόνα για την **σχέση χρόνου - συχνότητας** μέσω της συνάρτησης

`sp.signal.stft()` της scipy και απεικονίζουμε το γράφημα του λογάριθμου του μέτρου αυτού (προσθήκη σταθερας 0.0001 για καλύτερη απεικόνηση). Παραθέτουμε τα γραφήματα του μέτρου του STFT τα οποία φτιάξαμε με την χρήση της `pcolormesh()` της matplotlib:



STFT σε λογαριθμική κλίμακα για τις διάφορες τιμές μηκών παραθύρου.

Συμπεράσματα:

Ενέργεια: Περιοχές με **έντονο χρώμα** είναι διαστήματα **υψηλής ενέργειας** και μάλιστα συμπίπτουν με τα αντίστοιχα διαστήματα του γραφήματος Short Time Energy.

Χρόνος: Παρατηρούμε ότι στα πρώτα χρονικά διαστήματα του σήματος υπερισχύουν **χαμηλές συχνότητες** ενώ στη μέση και προς το τέλος συμμετέχουν και οι πιο **υψηλές**.

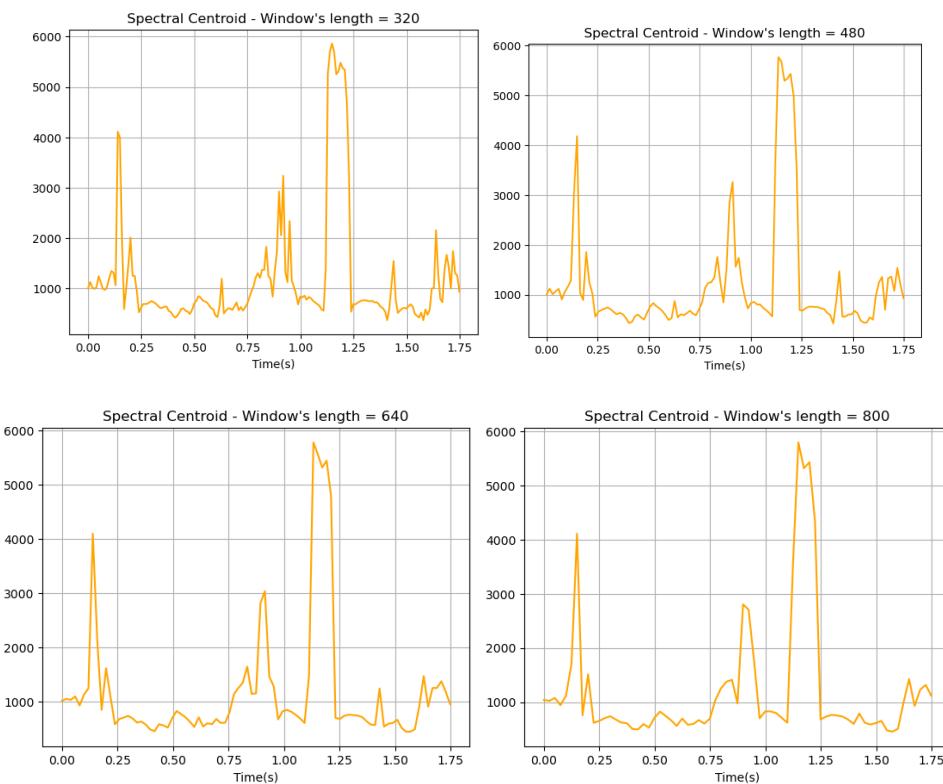
Συχνότητα: Φαίνεται να έχουμε περισσότερη ενέργεια στις **χαμηλές συχνότητες** (πιο έντονο χρώμα)

Παράθυρα: Κατά την σταδιακή αύξηση των μηκών των παραθύρων παρατηρούμε μια ελαφριά **αλλοίωση-ξεδώριασμα** του χρώματος γεγονός που οφείλεται στην **απομάκρυνση του θορύβου** και της μικρής ενέργειας που αυτός προσθέτει στο βασικό σήμα φωνής μας. Έπειτα ασχοληθήκαμε με την αναπαράσταση του **φασματικού κέντρου (SC)** και της **φασματικής ροής (SF)** για τον υπολογισμό των οποίων κατασκευάσαμε συναρτήσεις που “χτίζουν” τους παρακάτω τύπους των παρακάτω τύπων

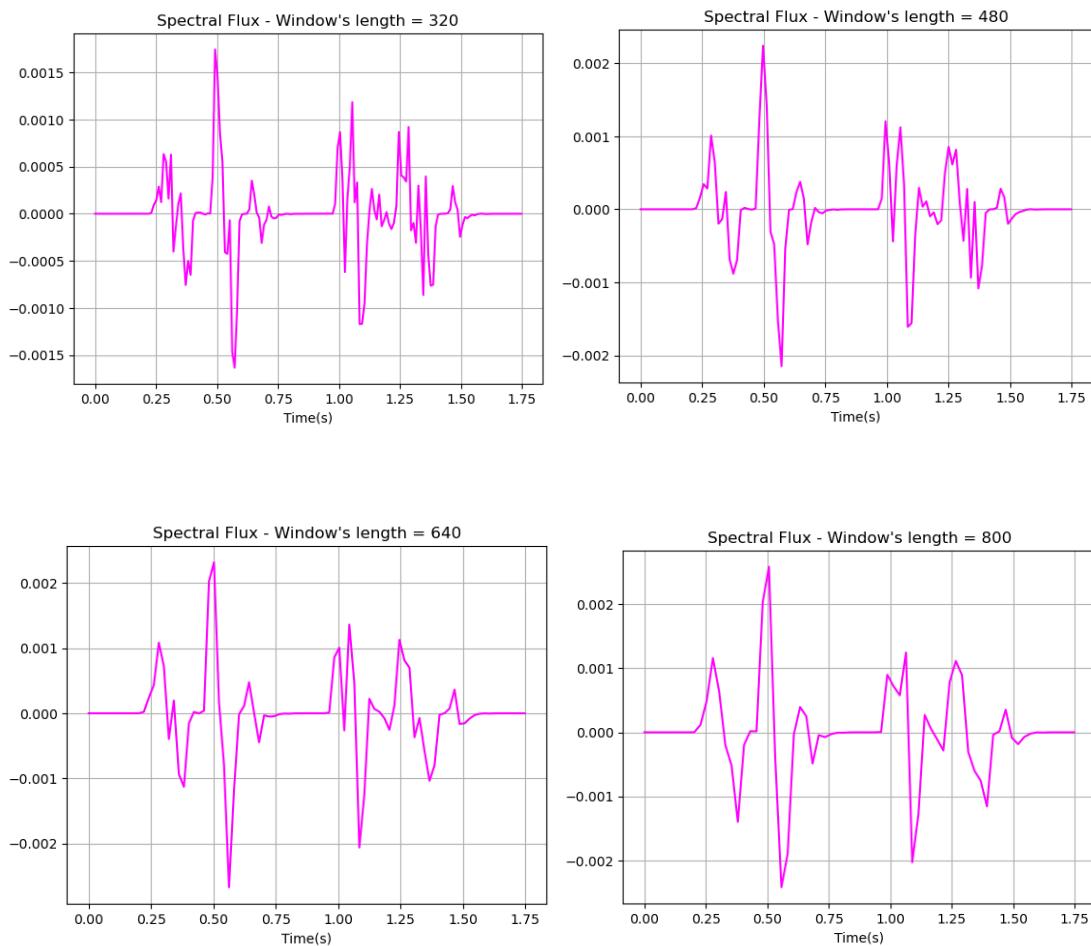
$$SC = \frac{\sum_{k=0}^{\frac{N}{2}} k \cdot |X[k, m]|^2}{\sum_{k=0}^{\frac{N}{2}} |X[k, m]|^2} \cdot \frac{fs}{N}$$

$$SF = \left\| \frac{\frac{|X[k, m+1]|^2}{\sum_{k=0}^{\frac{N}{2}} |X[k, m+1]|^2} - \frac{|X[k, m]|^2}{\sum_{k=0}^{\frac{N}{2}} |X[k, m]|^2}}{\sum_{k=0}^{\frac{N}{2}} |X[k, m+1]|^2 - \sum_{k=0}^{\frac{N}{2}} |X[k, m]|^2} \right\|$$

Παραθέτουμε τα γραφήματα **φασματικού κέντρου** και **φασματικής ροής** για τις διάφορες τιμές μήκους των παραθύρων:



Φασματικό κέντρο για τις διάφορες τιμές μήκων παραθύρου.



Φασματική ροή για τις διάφορες τιμές μηκών παραθύρου.

Το **φασματικό κέντρο** αντιπροσωπεύει το «κέντρο βάρους» του φάσματος ισχύος του σήματος και στην πραγματικότητα μας δίνει μία ιδέα για το που βρίσκεται το **κύριο σημείο ενέργεια του σήματος**.

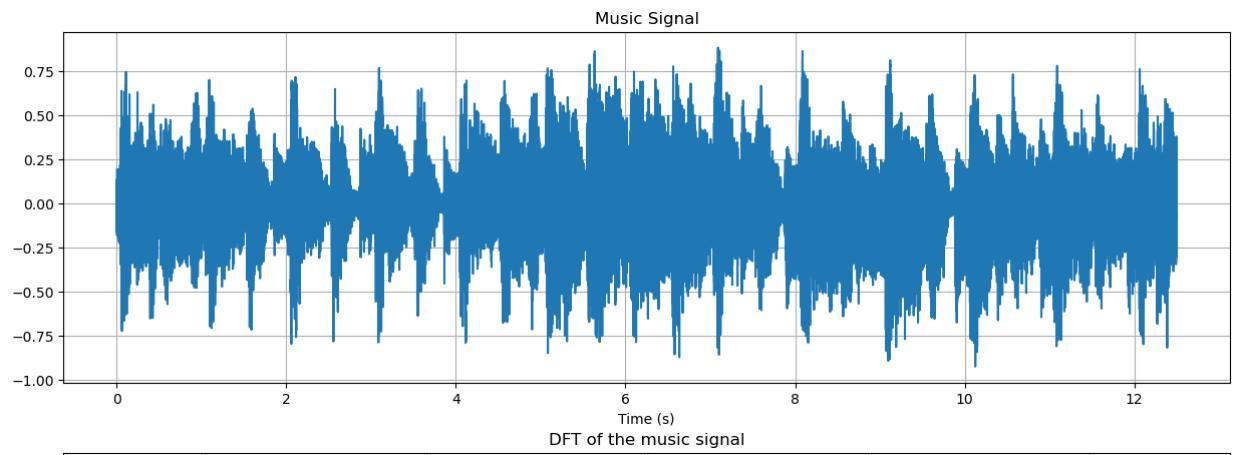
Η **φασματική ροή** μετρά την αλλαγή του φάσματος του ήχου κατά τον χρόνο δηλαδή μπορεί να χρησιμοποιηθεί για την ανίχνευση **αλλαγών στην σύνθεση του ήχου**. Για τον λόγο αυτό έχουμε peaks (εφωνοι-αφωνοι ήχοι) και μηδενική τιμή ανάμεσα (παύσεις ανάμεσα στις λέξεις).

Στα δύο φυσικά μεγέθη αυτά παρατηρούμε **σταδιακή εξομάλυνση** του σήματος με αύξηση του μήκους του παραθύρου καθώς **μειώνεται ο θόρυβος** και **καλύπτονται περισσότερα σημεία στον χρόνο**. Επίσης, παρατηρείται μια **αύξηση της φασματικής ροής**

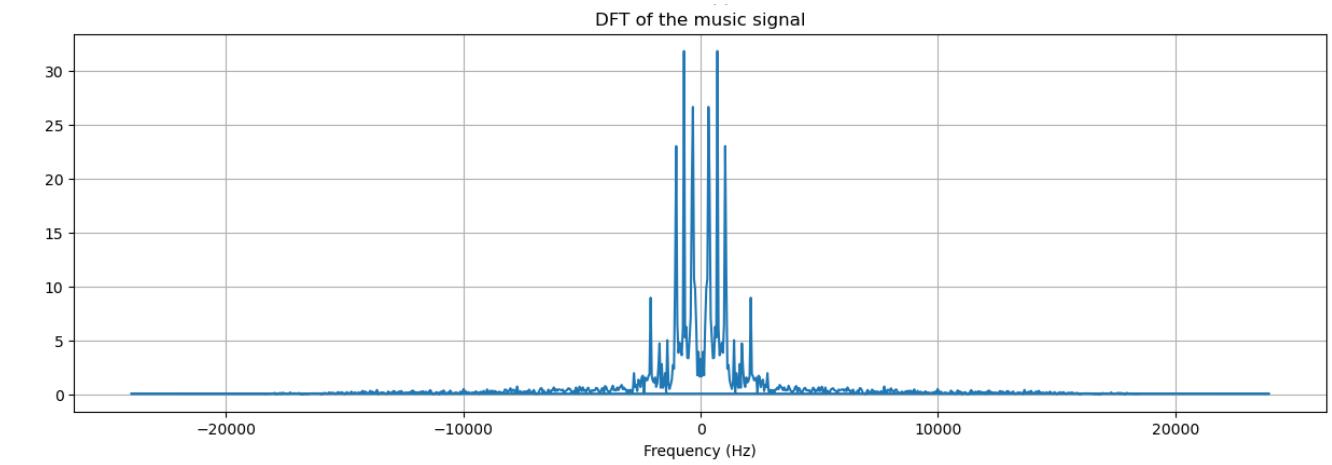
Ερώτημα 2.4

Επαναλαμβάνουμε τα παραπάνω ερωτήματα για το σήμα μουσικής “**music.wav**” συχνότητας δειγματοληψίας $fs2=48kHz$.

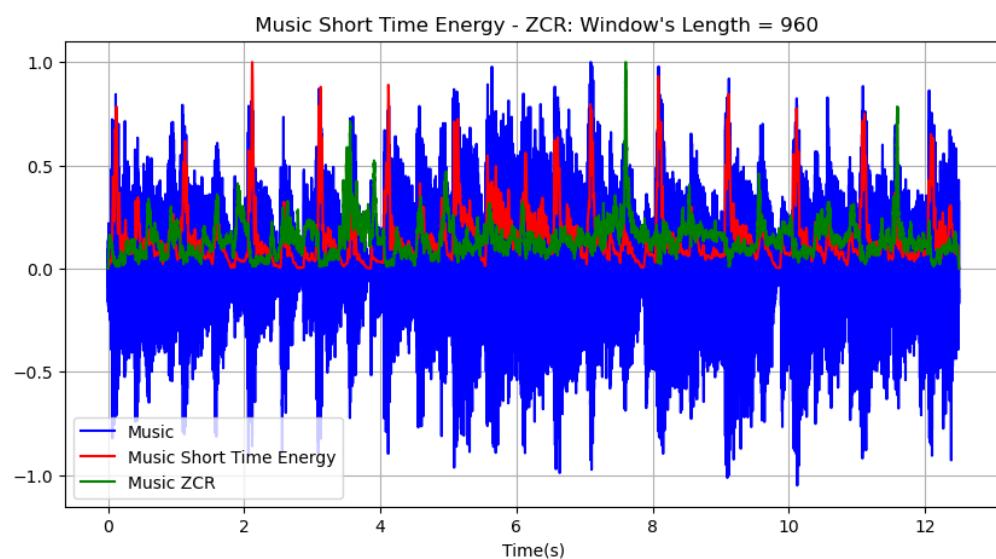
Σήμα μουσικής στον χρόνο:

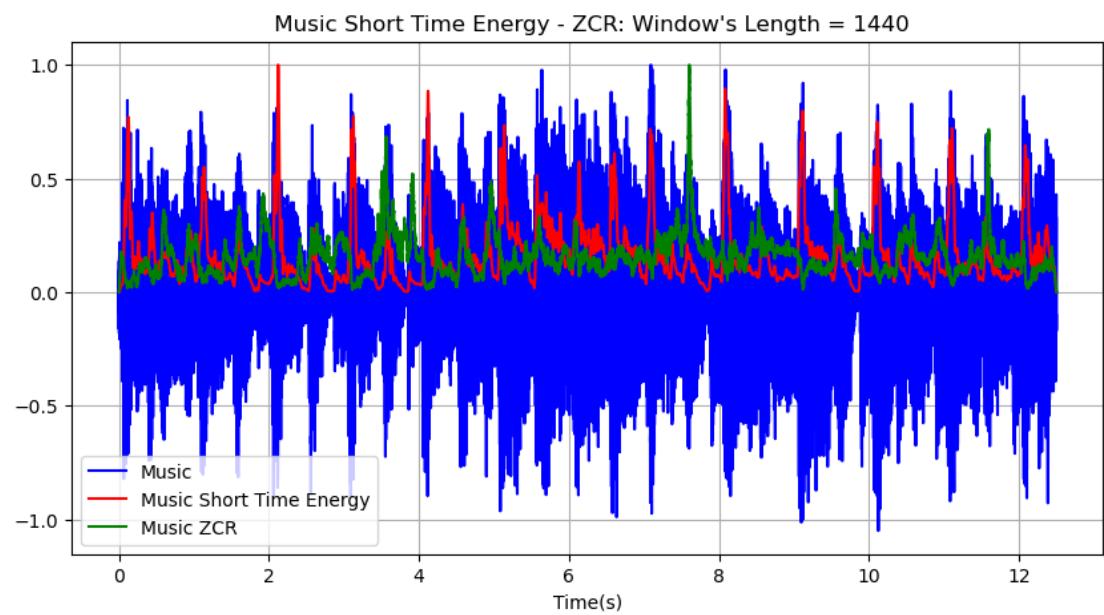
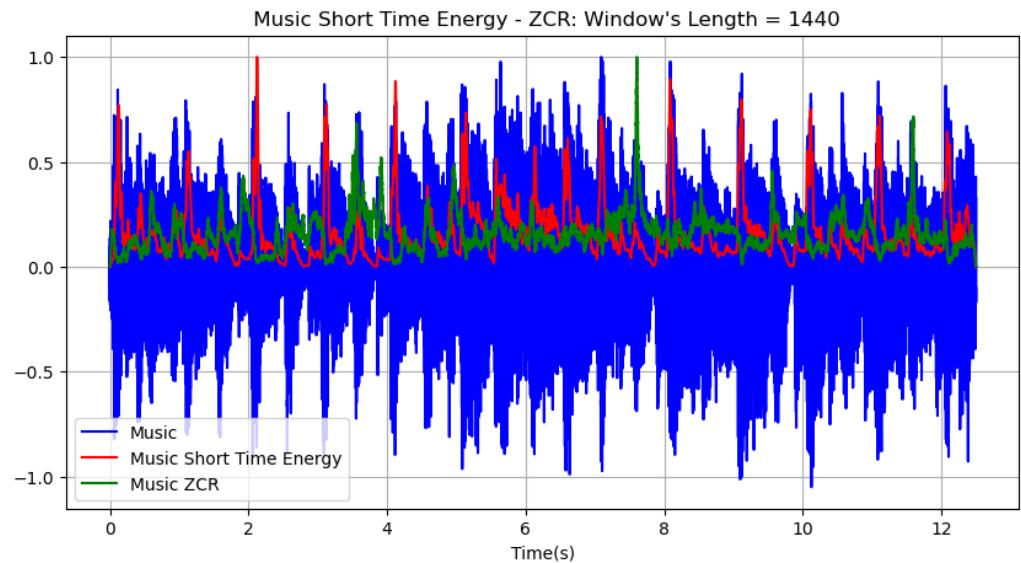


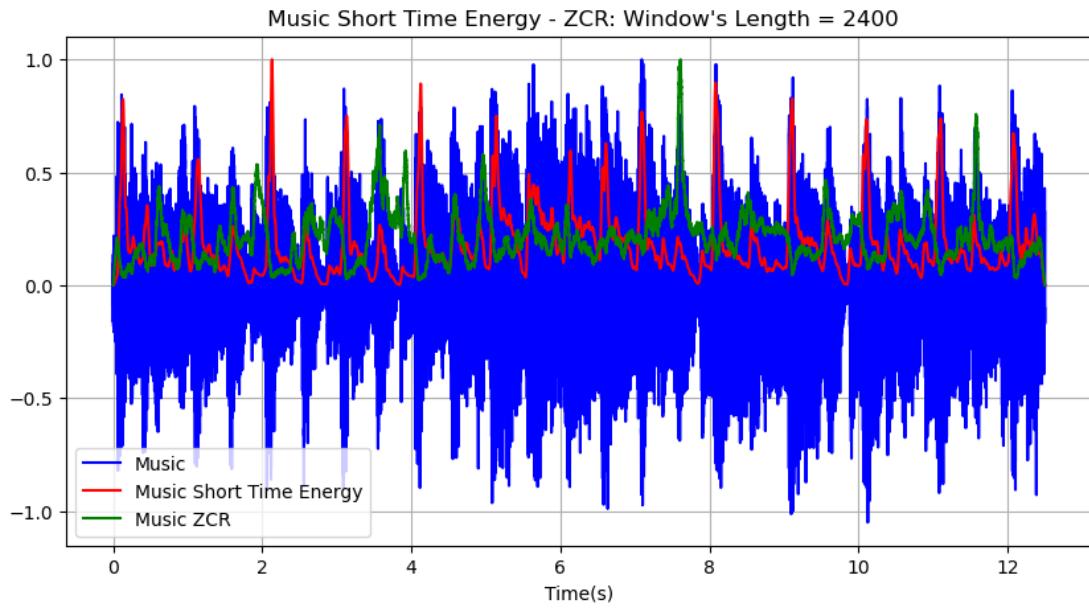
DFT του σήματος:



Ενέργεια βραχέος χρόνου και ZCR του σήματος μουσικής για τα διάφορα παράθυρα:

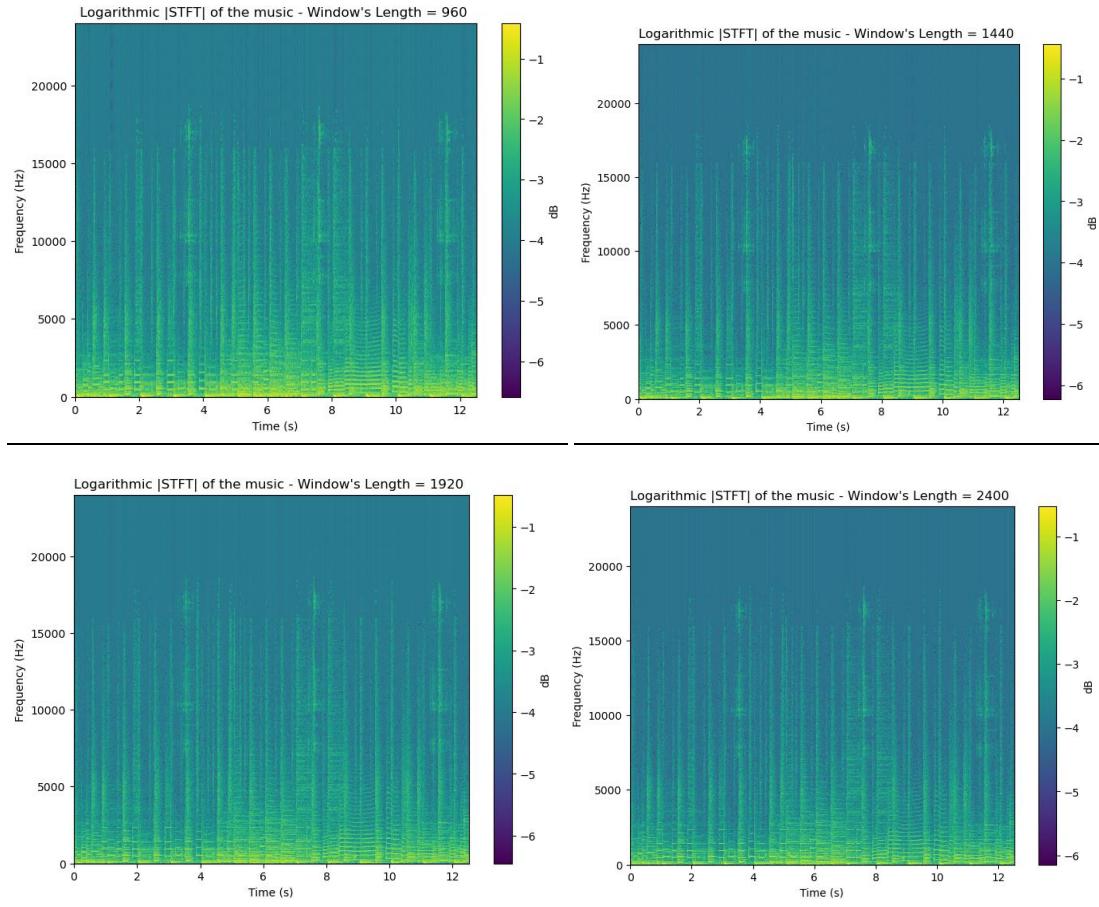






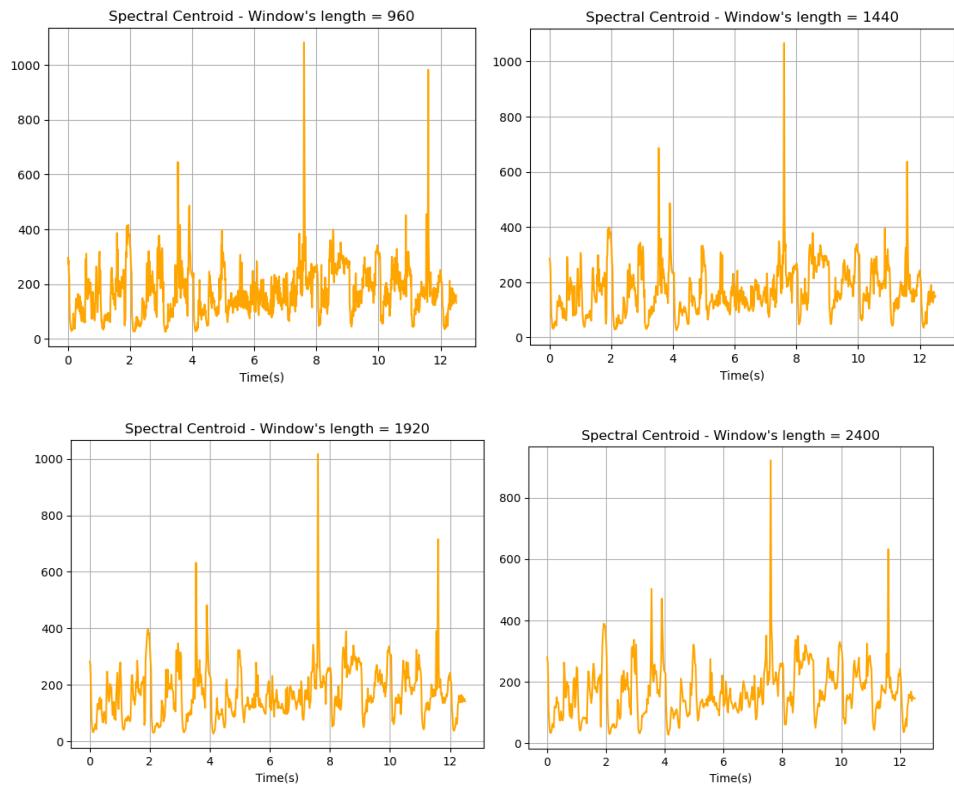
Παρατηρούμε εξομάλυνση με την αύξηση μήκους των παραθύρων.
Επιπλέον, φαίνεται το σήμα να περιέχει αρκετούς ήχους με βάθος
καθώς η ενέργειά του κάνει πολλά peaks (όπως και πριν στα peaks της
ενέργειας έχουμε χαμηλό ZCR ενώ αυτό κάνει peaks στους πιο ρηχούς
ήχους).

STFT σε λογαριθμική κλίμακα για το σήμα μουσικής:

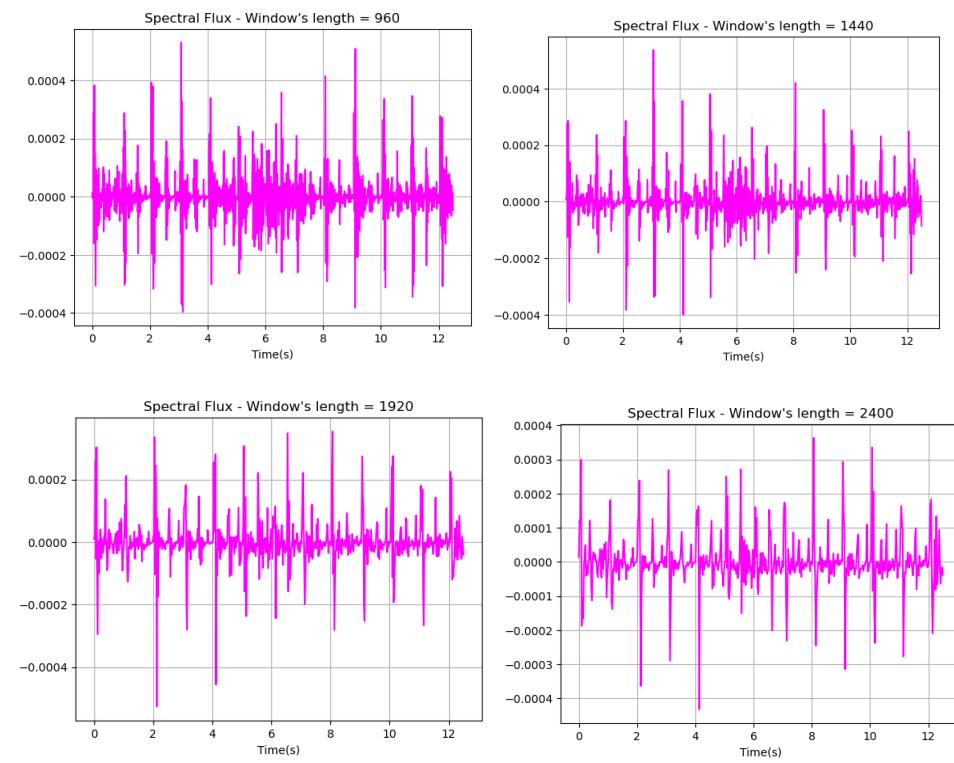


Τα έντονα χρώματα είναι σημεία υψηλής ενέργειας. Φαίνεται το σήμα να έχει μια **ομοιόμορφη κατανομή ενέργειας στο πεδίο του χρόνου** (άξονας x) ενώ περισσότερη ενέργεια φαίνεται να βρίσκεται στις **χαμηλότερες συχνότητες** (χαμηλά στον άξονα y).

Φασματικό κέντρο του σήματος μουσικής για τις διάφορες τιμές του παραθύρου:



Φασματική Ροή σήματος μουσικής για τα διάφορα μήκη παραθύρων:



Είναι εμφανής η εξομάλυνση με την αύξηση μηκών παραθύρου παρέχοντας μας πιο καθαρή εικόνα και για τα δυο μεγέθη SC-SF.

Μέρος 3ο - Ψηφιακή Επεξεργασία και Φασματική Ανάλυση Βιοσημάτων Κητωδών

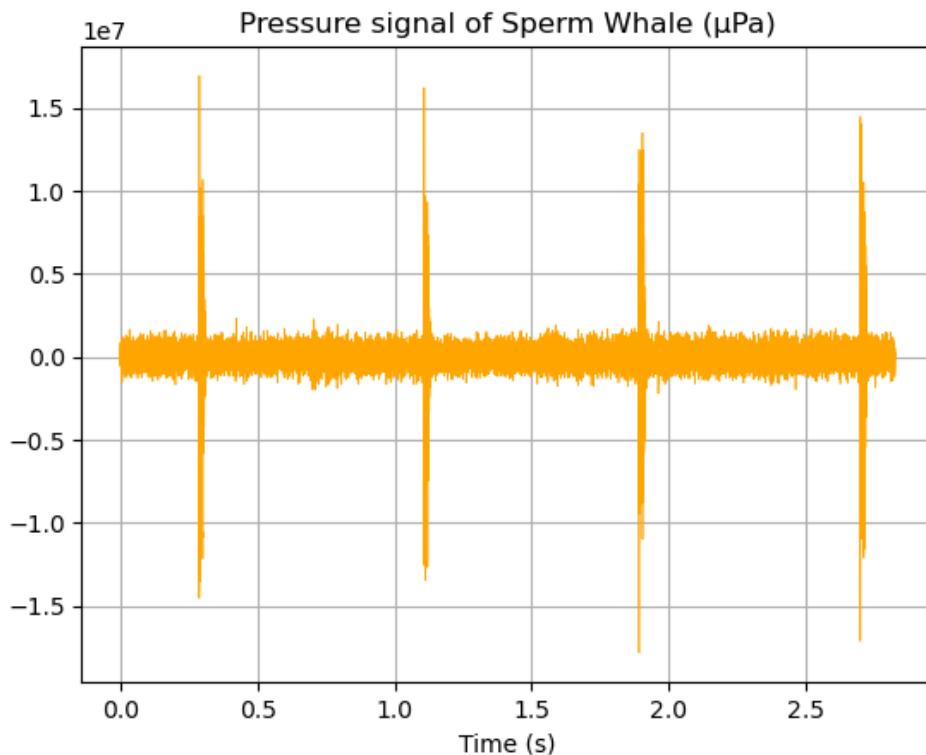
●Σκοπός της άσκησης είναι η **ψηφιακή επεξεργασία δύο βιοσημάτων φαλαινών** τα οποία έχουν συλλεχθεί από υδρόφωνα. Πρόκειται για σήματα από φάλαινες δύο ειδών Sperm whale και Humpback whale. Τέλος μελετάται πως ανθρωπογενής ήχος μπορεί να επηρεάσει τις φάλαινες αυτές.

Ερώτημα 3.1

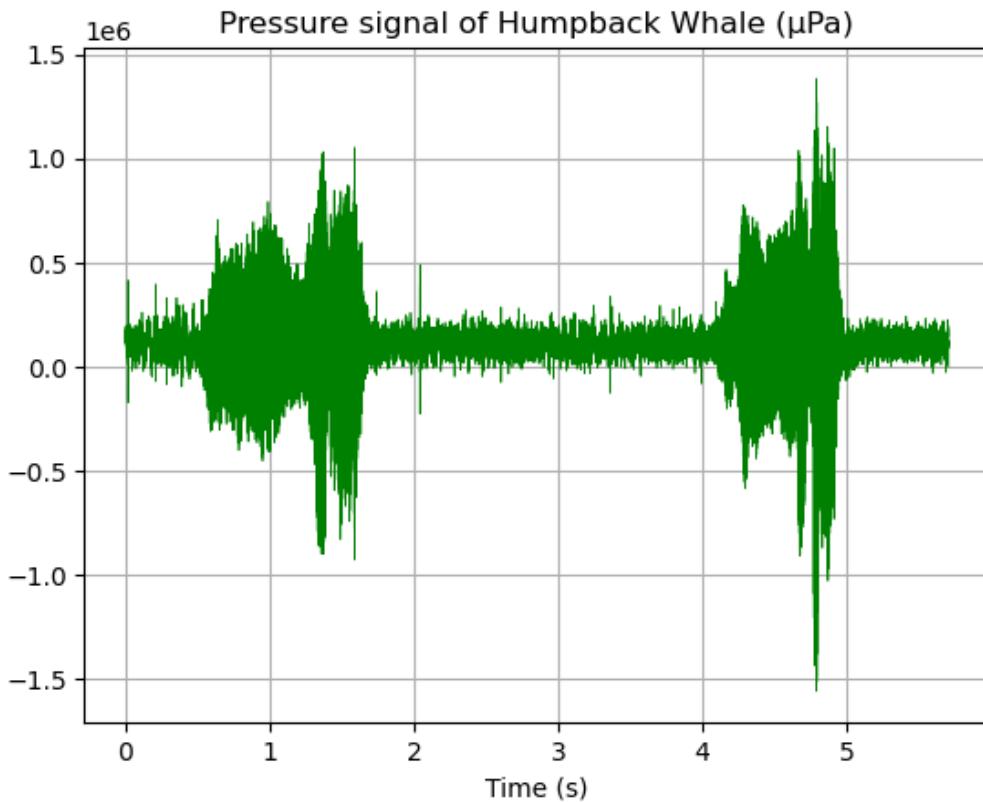
Πρώτα μετατρέπουμε τα βιοσήματα από σήματα τάσης σε σήματα πίεσης σύμφωνα με τον παρακάτω τύπο:

$$pressure = \frac{3.5 \cdot y \cdot GF}{10^{sensitivity/20}}$$

Όπου γι το σήμα μας το οποίο φορτώνουμε με τη βοήθεια της `load()`, GF ο συντελεστής ενίσχυσης και sensitivity η ευαισθησία του υδροφώνου. Με δεδομένες τιμές **GF=0.16** και **sensitivity = -155 dB** relative to 1V/μPa λαμβάνουμε τα παρακάτω διαγράμματα πίεσης-χρόνου:



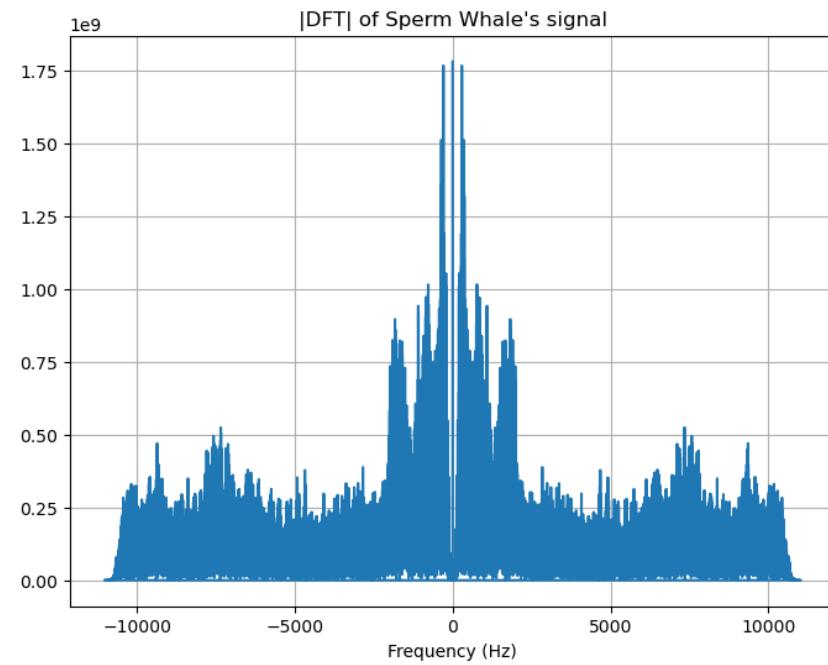
Σήμα πίεσης *Sperm whale*.

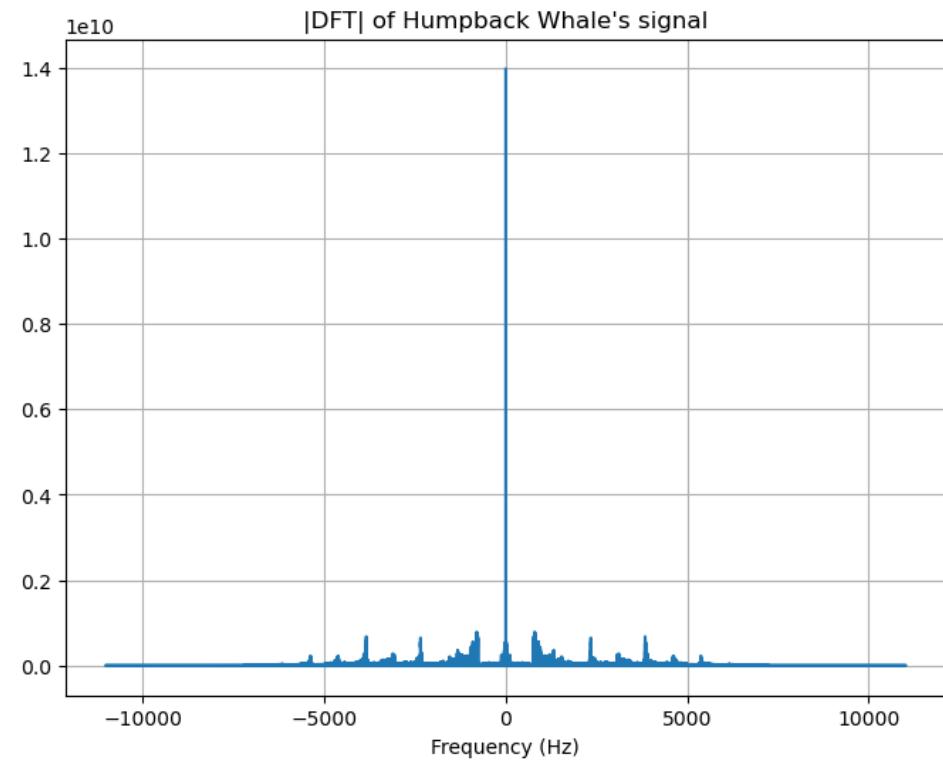


Σήμα πίεσης Humpback whale.

Επίσης διαπιστώσαμε ότι τα δύο σήματα έχουν τον ίδιο **ρυθμό δειγματοληψίας $fs=22050 \text{ Hz}$.**

Επιπλέον παρουσιάζουμε και τους **DFT's** των σημάτων αυτών:





Ερώτημα 3.2

Έπειτα θα υπολογίσουμε την rms τιμή της πίεσης καθώς και το επίπεδο ηχητικής πίεσης (SPL) βάσει των παρακάτω τύπων:

$$P_{rms} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N pressure^2}$$

$$SPL = 20 \log\left(\frac{P_{rms}}{P_{ref}}\right)$$

Όπου N το μήκος του σήματος και $P_{ref} = 1 \mu\text{Pa}$.

Το μήνυμα που τυπώνεται στην οθόνη είναι το εξής:

RMS of Sperm Whale signal is RMS1 = 749840.3376660865 .

RMS of Humpback Whale signal is RMS1 = 195774.37361509659 .

SPL of Sperm Whale signal is SPL1 = 117.49937599177156 dB relative to 1 μPa .

SPL of Humpback Whale signal is SPL2 = 105.83511686211784 dB relative to 1 μPa .

Ερώτημα 3.3

Διαλέγουμε το σήμα της **humpback whale** και επιβεβαιώνουμε το **θεώρημα Parseval** το οποίο αναφέρει ότι η ενέργεια στο χώρο της συχνότητας είναι ίση με τον υπολογισμό αυτή στο πεδίο του χρόνου:

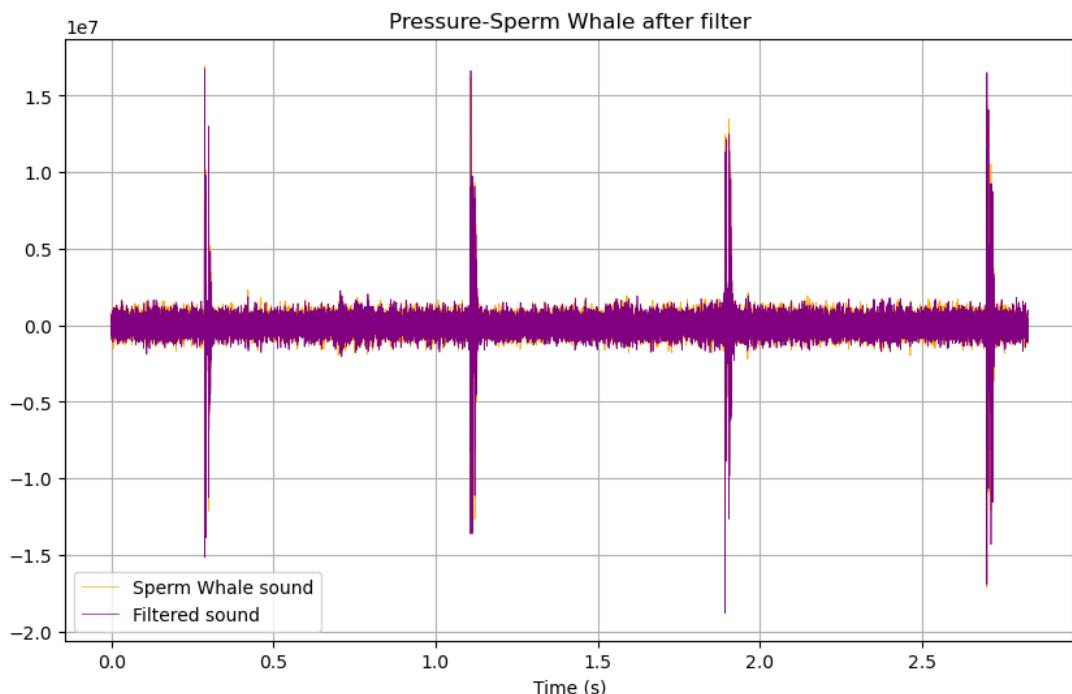
$$En = \sum_{n=0}^{N-1} |p[n]|^2 = \frac{1}{N} \cdot \sum_{k=0}^{N-1} |P[k]|^2$$

Τυπώνεται το παρακάτω μήνυμα γεγονός που σημαίνει ότι το θεώρημα επαληθεύεται (Η μικρή διαφορά οφείλεται σε τυχούσες προσεγγίσεις ή σφάλματα):

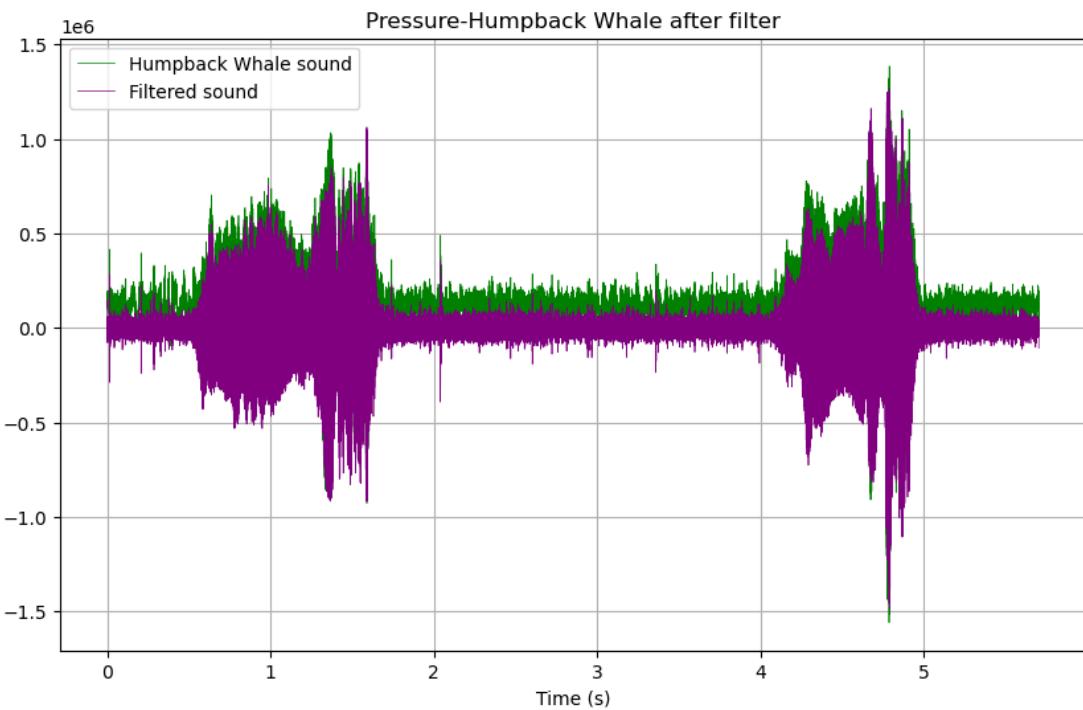
Energy in time domain is 35043394 .
Energy in frequency domain is 35043447 .

Ερώτημα 3.4

Θα **απομονώσουμε** από τα βιοσήματα πίεσης τον **ανθρωπογενή θόρυβο** χαμηλής συχνότητας, κυρίως πλοίων, με την εφαρμογή ενός **υψηπερατού φίλτρου** $Wn=200/(fs/2)Hz$ και **τάξης N=3** το οποίο θα απομονώσει συχνότητες κάτω των 200Hz (**fs/2 συχνότητα Nyquist**). Αυτό το κάνουμε με τη χρήση της **scipy.signal.butter()** και της **scipy.signal.lfilter()**. Τα αποτελέσματα φαίνονται παρακάτω:



Filtered Sperm whale signal.



Filtered Humpback whale signal.

Είναι προφανές ότι το φίλτρο μας έχει **περιορίσει τις χαμηλές συχνότητας** των βιοσημάτων μας ενώ έχει κρατήσει τις συχνότητες άνω των 200Hz **σχεδόν αναλλοίωτες**.

Ερώτημα 3.5

Γνωρίζοντας ότι οι **humpback whales** ακούνε συχνότητες μεταξύ **300Hz-3000Hz** ενώ οι **sperm whales** **όλο το φάσμα** και επιπλέον ότι η συμπεριφορά τους επηρεάζεται για SPL **μεγαλύτερα των 100dB**, θα εξετάσουμε εάν ο ανθρωπογενής ήχος “Pile driving” είναι ικανός να επηρεάσει αρνητικά τις φάλαινες. Πιο συγκεκριμένα, για την sperm whale υπολογίζουμε το SPL του σήματος Pile driving, αφού κι αυτό πρώτα μετατραπεί σε σήμα πίεσης, και υπολογίζουμε το SPL του. Αντίστοιχα, για την Humpback, φιλτράρουμε το ανθρωπογενές σήμα με ένα **bandwidth φίλτρο Wn=[300/(fs3/2),3000/(fs3/2)]** όπου **fs3/2=11025Hz** η **συχνότητα Nyquist**. Το αποτέλεσμα που τυπώνεται στην οθόνη είναι το εξής:

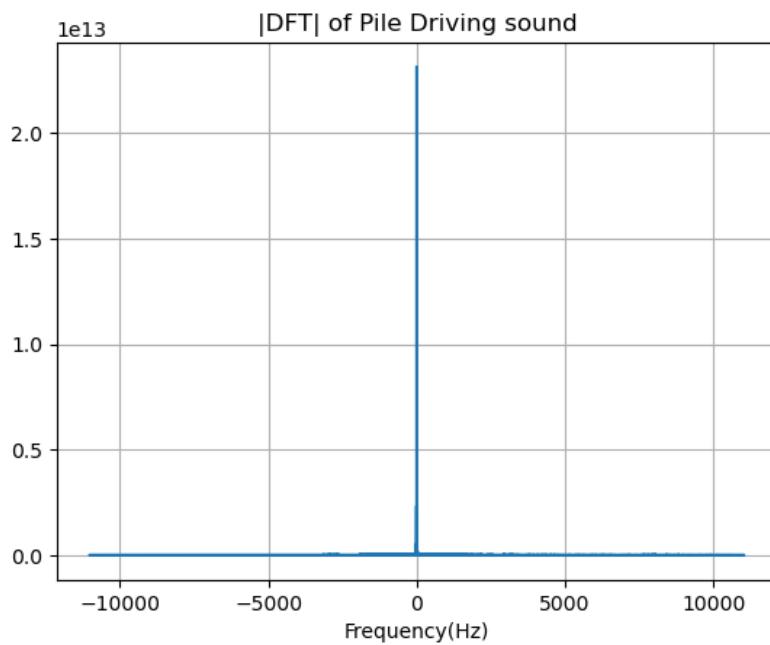
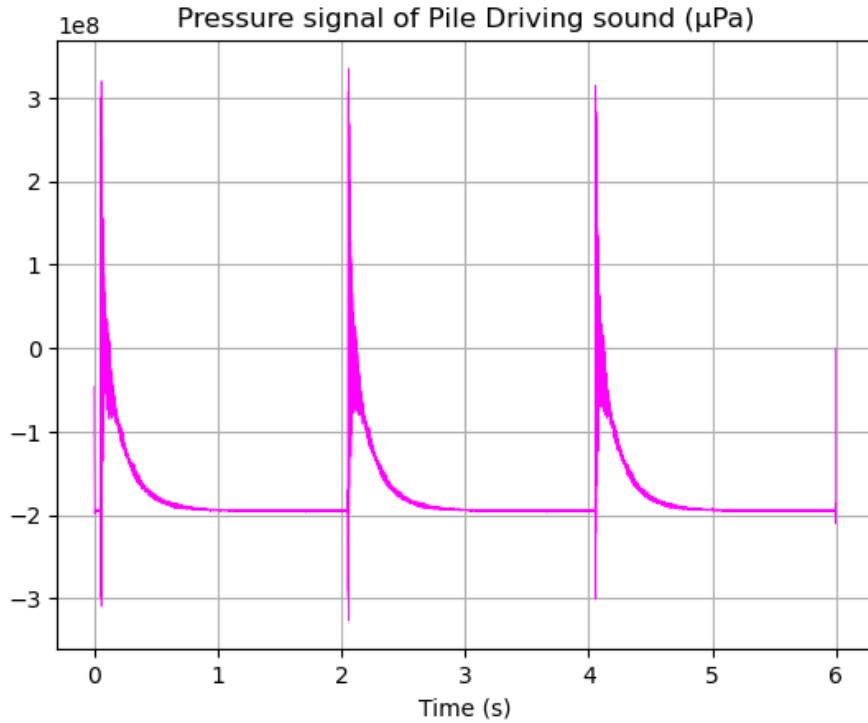
Sampling rate of Pile Driving signal is fs3 = 22050

Sperm Whale: Sound Pressure Level of Pile Driving sound is SPL3_sperm= 165.15296950722166 dB .

Humpback Whale(Filtered Pile Driving): Sound Pressure Level of Pile Driving sound is SPL3_hump = 141.51802420615306 dB .

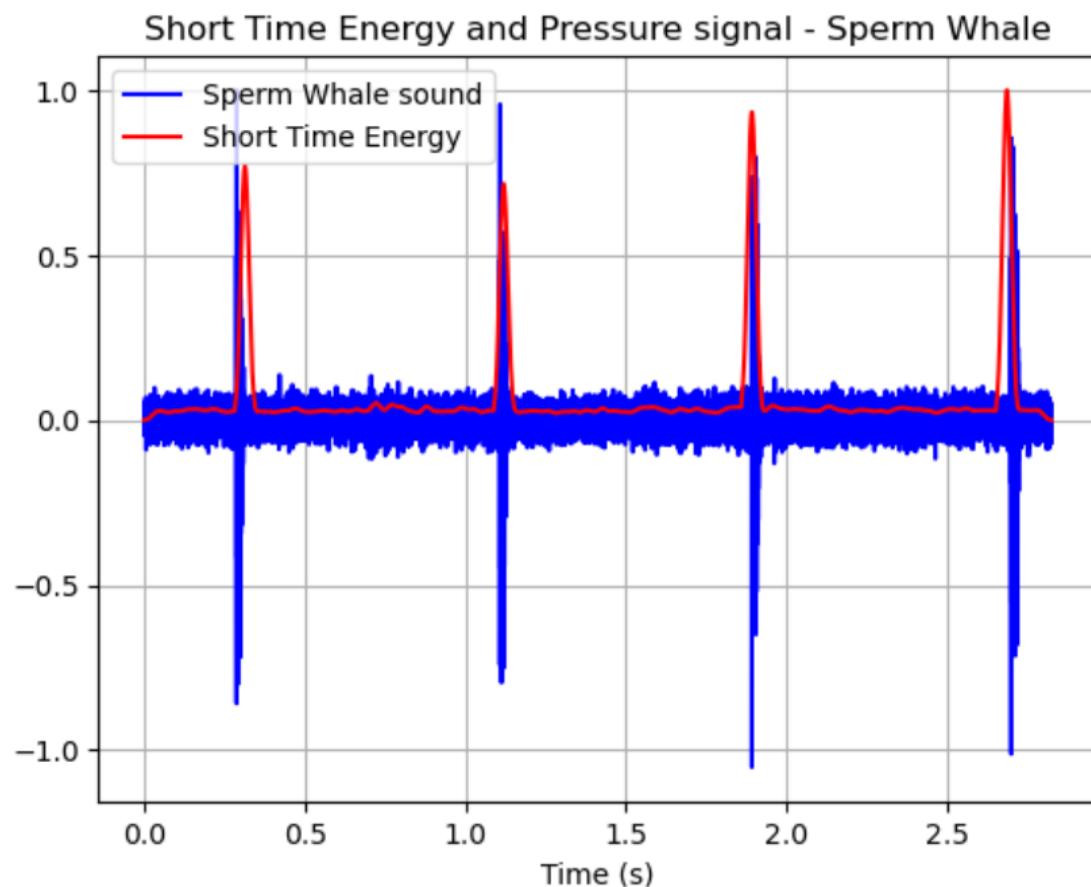
Συνεπώς, το ανθρωπογενές σήμα Pile driving **κρίνεται επιβλαβές και για τα δύο είδη φάλαινας** εφόσον το επίπεδο ηχητικής πίεσης SPL φιλτραρισμένου και αφιλτράριστου σήματος **υπερβαίνει τα 100dB**.

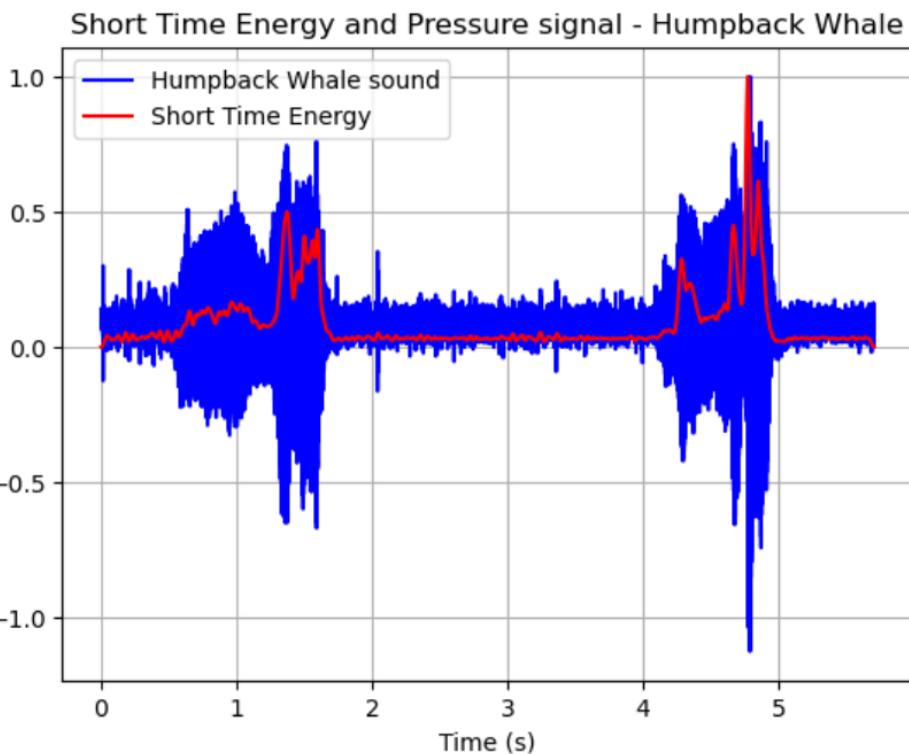
Τέλος παραθέτουμε το **γράφημα πίεσης** του αφιλτράριστου σήματος Pile driving καθώς και τον **DFT** του για μια πλήρη εικόνα του σήματος:



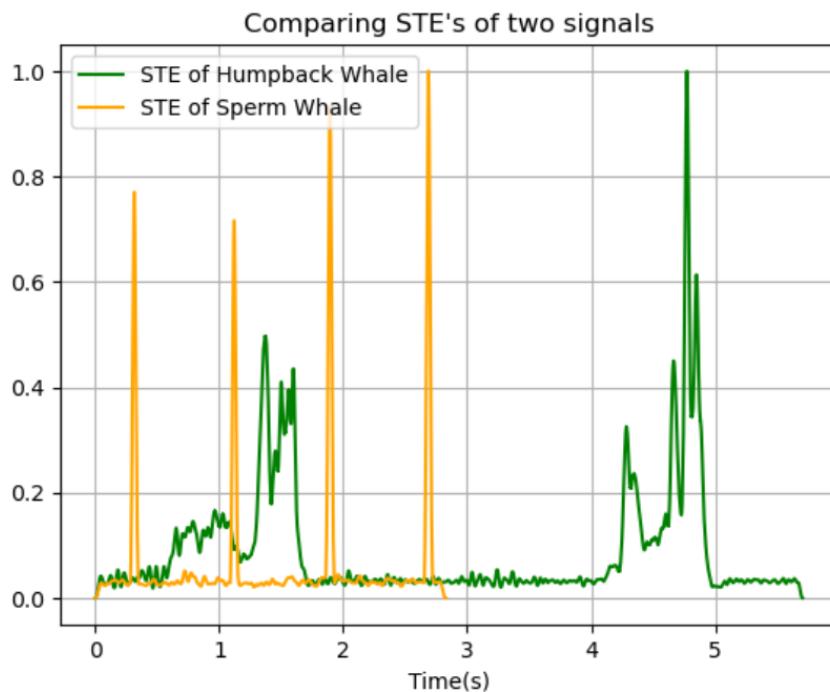
Ερώτημα 3.6

Όπως και σε προηγούμενες ασκήσεις της τρέχουσας αναφοράς, θα υπολογίσουμε την **ενεργεία βραχέος χρόνου** των δύο βιοσημάτων για παράθυρο hamming μήκους $L = 0.05s \cdot 22050Hz = 1102$ ως συνέλιξη του τετραγώνου του κάθε βιοσήματος με το παράθυρο με τη χρήση της **convolve()**. Αφού κανονικοποιήσουμε για καλύτερα οπτικά αποτελέσματα παραθέτουμε τα γραφήματα πίεσης και ενέργειας για κάθε φάλαινα αλλά και τις δύο ενέργειες βραχέος χρόνου μεταξύ τους :





Comparing STE's of two signals



Στα δύο πρώτα γραφήματα, στα σημεία άφωνου ήχου η ενέργεια είναι σχεδόν μηδενική ενώ στα σημεία που ο ήχος είναι έμφωνος έχουμε peaks. Επίσης, παρατηρούμε ότι η sperm whale παράγει ένα βιοσήμα **αρκετά υψηλότερης ενέργειας** από την humpback whale αφού τα peaks

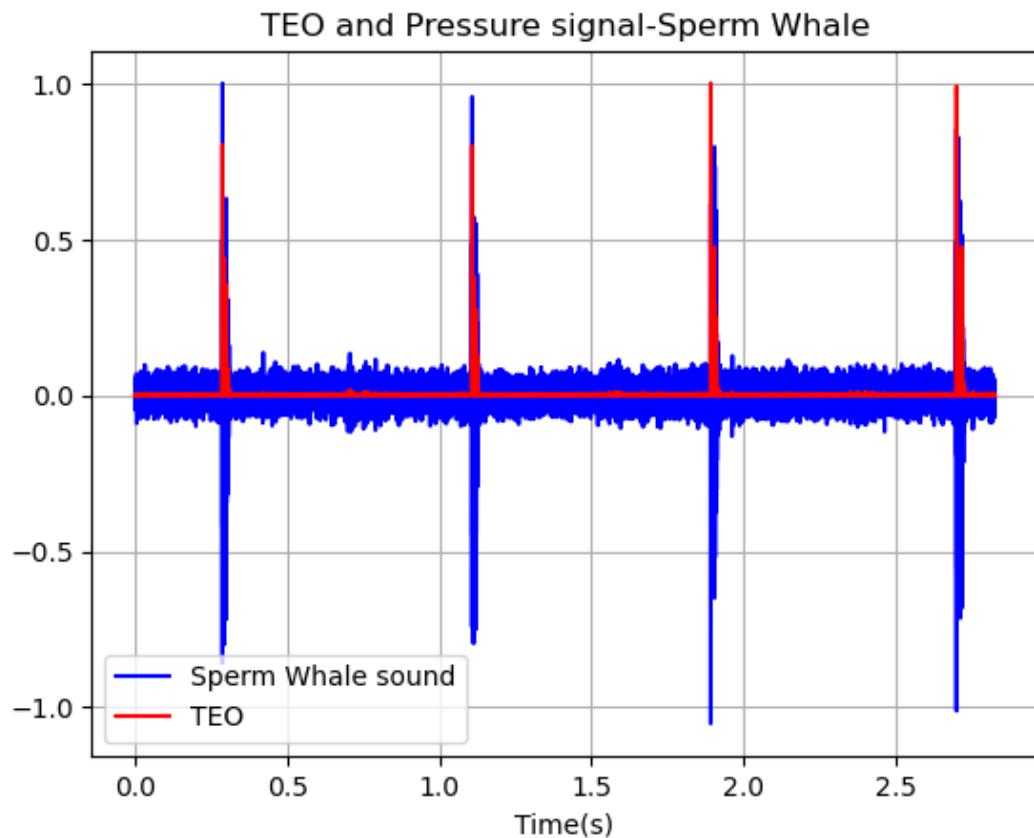
της είναι αρκετά υψηλότερα από τα peaks της humpback (το επιβεβαιώσαμε και από τον τύπο $En = \frac{1}{N} \cdot \sum_{k=0}^{N-1} |P[k]|^2$).

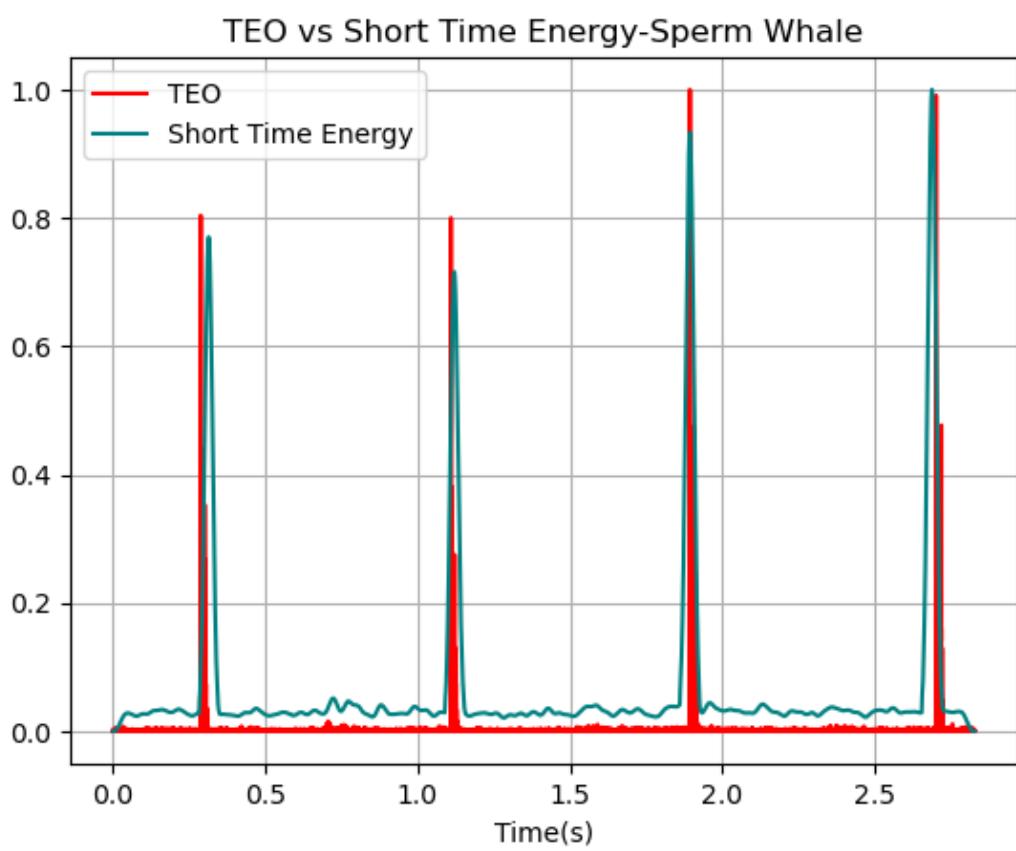
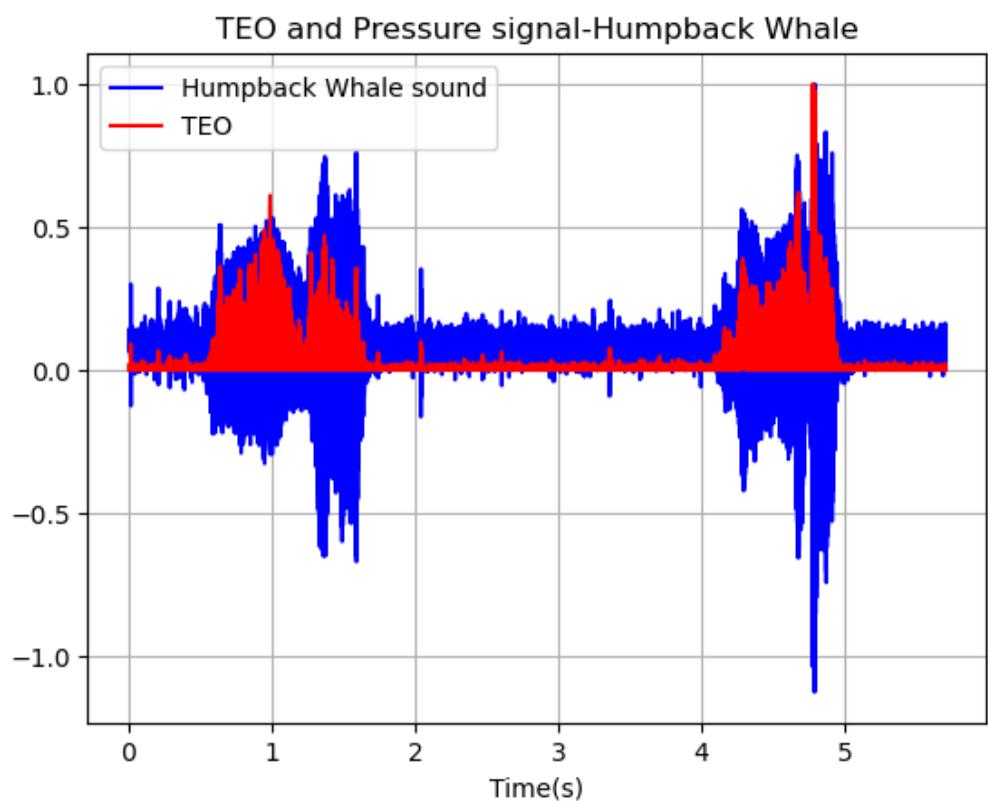
Ερώτημα 3.7

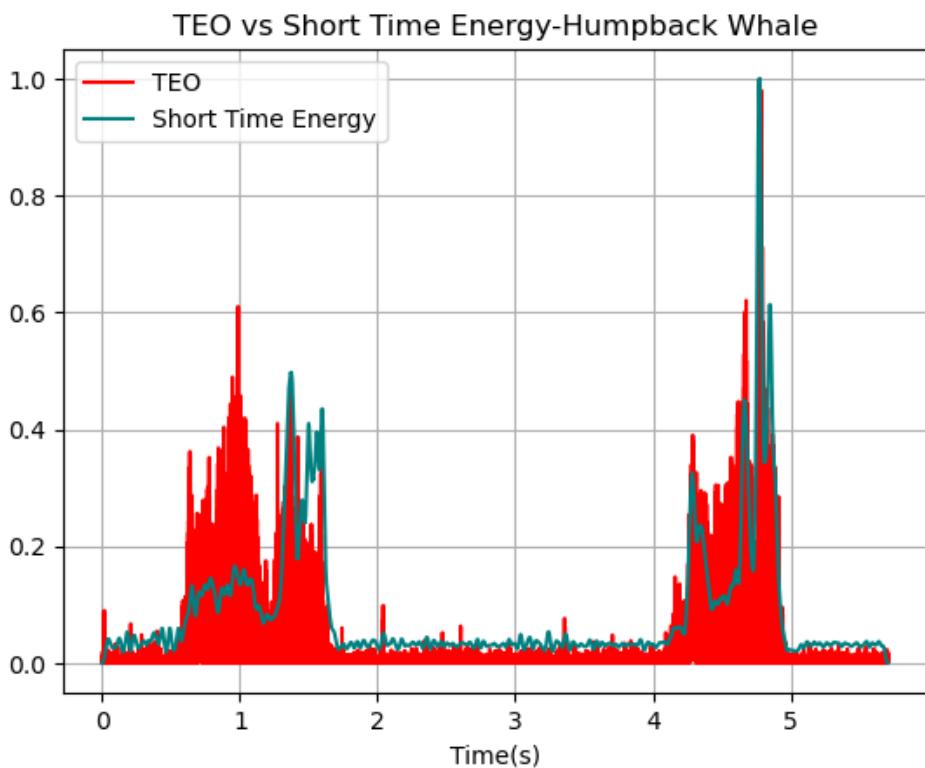
Θα υλοποιήσουμε τον μη γραμμικό τελεστή **Teager Kaiser Energy Operator** (ΤΕΟ) για τα δύο βιοσήματα ο οποίος προκύπτει από τον παρακάτω τύπο:

$$\Psi[p[n]] = p[n]^2 - p[n-1] \cdot p[n+1]$$

Αφού τον κατασκευάσουμε με τη βοήθεια της κυκλικής μετατόπισης `np.roll()` όπου αυτή χρειάζεται ($p[n-1] \rightarrow 1$ θέση δεξιά, $p[n+1] \rightarrow 1$ θέση αριστερά), σχεδιάζουμε την γραφική της παράσταση και την συγκρίνουμε με την ενέργεια βραχέος χρόνου:



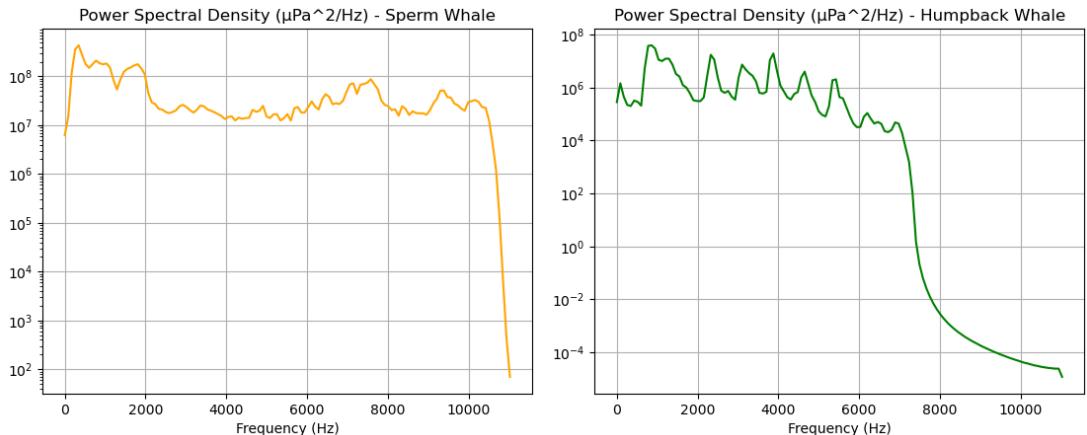




Παρατηρούμε ότι ο ΤΕΟ είναι ελαφρώς υψηλότερο σε σχέση με την ενέργεια βραχέος χρόνου σε ορισμένα σημεία. Σύμφωνα με πληροφορίες που αναζητήσαμε στο διαδίκτυο, γενικά ο ΤΕΟ είναι ένας **δείκτης ανίχνευσης εναλλαγής της ενέργειας και είναι χρήσιμος σε σήματα με αρκετό θόρυβο** όπως τα δικά μας. Παρατηρούμε λοιπόν πως ακολουθεί την μορφή της STE αλλά σε σημεία με πολλές εναλλαγές αυτής, γίνεται ακόμη μεγαλύτερος.

Ερώτημα 3.8

Θα σχεδιάσουμε το **περιοδόγραμμα Welch** σε λογαριθμική κλίμακα με τη χρήση της `.welch()`. Πρόκειται για μία εκτίμηση της φασματικής πυκνότητας ισχύος και θα μας δώσει μια απεικόνιση σχετική με την **ενέργεια στο φάσμα**. Παραθέτουμε το περιοδόγραμμα για τα δύο βιοσήματα παρακάτω:



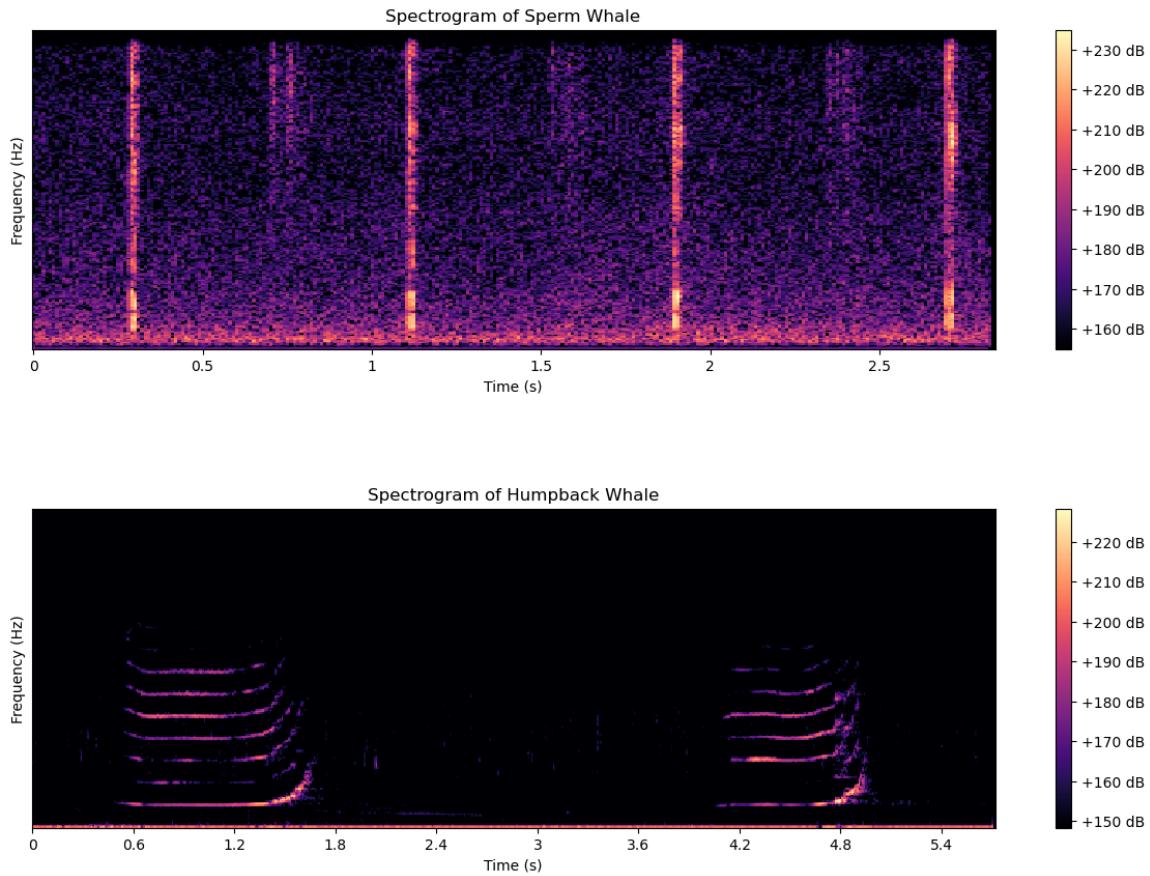
Παρατηρούμε ότι η φασματική πυκνότητα ισχύος και στα δύο σήματα, **εξασθενεί στις μεγάλες συχνότητες** ενώ είναι **ισχυρή στις μικρότερες**.

Επιπλέον, στην sperm whale έχουμε peaks της φασματικής πυκνότητας-άρα και ενέργεια-για μεγαλύτερο εύρος συχνοτήτων από ότι στην humpback. (0 έως 9500Hz έναντι 0 έως 6500Hz)

Τέλος παίρνει μεγαλύτερες τιμές στο βιοσήμα της Sperm whale (έως και πάνω από 10^8) από ότι στης Humpback (έως και κάτι λιγότερο από 10^8). Αυτό βέβαια δεν είναι κάτι που δεν περιμέναμε αφού **η sperm whale δείξαμε πιο πάνω ότι έχει μεγαλύτερη ενέργεια και τα παραπάνω διαγράμματα σχετίζονται άμεσα με την κατανομή της ενέργειας των δύο σημάτων στο φάσμα.**

Ερώτημα 3.9

Θα σχεδιάσουμε το φασματογράφημα των τετραγώνων των δύο βιοσημάτων με τη βοήθεια **της librosa.display.specshow()** αφού υπολογίσουμε τον **STFT** καθενός από αυτά (**sp.signal.stft()**) και μετατρέψουμε το αποτέλεσμα σε λογαριθμική κλίμακα με τη χρήση της **librosa.amplitude_to_db()**.



Όπως είδαμε και στα διαγράμματα φασματικής πυκνότητας ισχύος του **προηγουμένου ερωτήματος**, η humpback whale έχει αρκετά μικρότερη ενέργεια κάτι που υποδεικνύουν τα **λιγότερο έντονα χρώματα στο φασματογράφημα** (βλέπε colorbar).

Πεδίο χρόνου :Παρατηρούμε **υψηλή ενέργεια στα ίδια χρονικά διαστήματα που ήδη είδαμε στα γραφήματα ενέργειας βραχέος χρόνου** και απουσία ενέργειας στα αντίστοιχα διαστήματα, κάτι που περιμέναμε να δούμε.

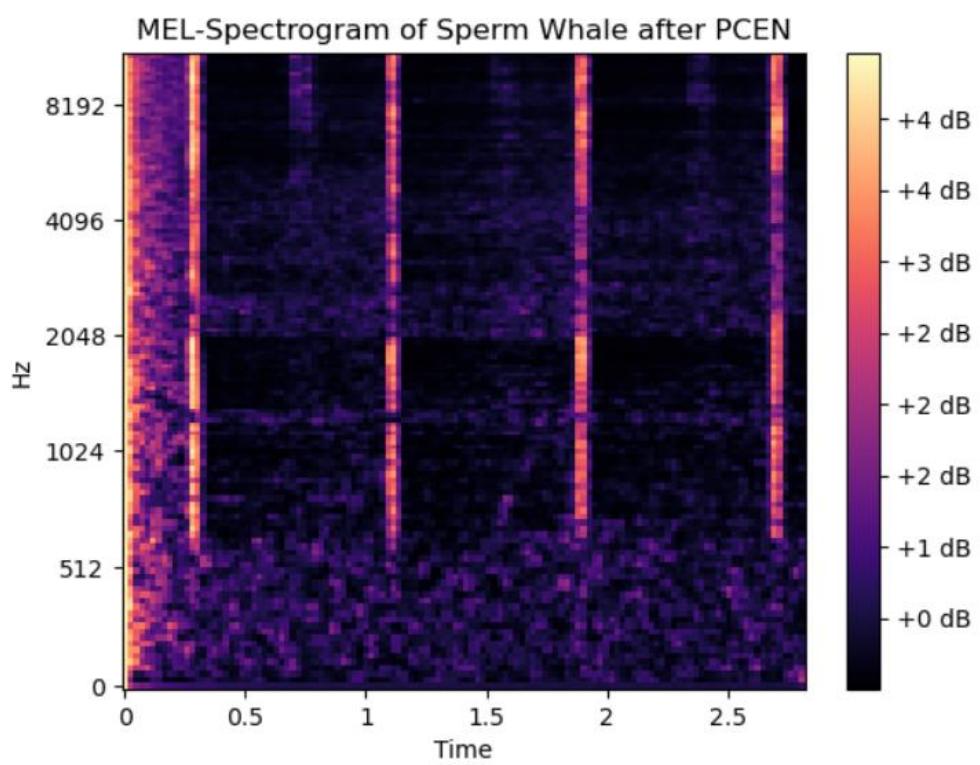
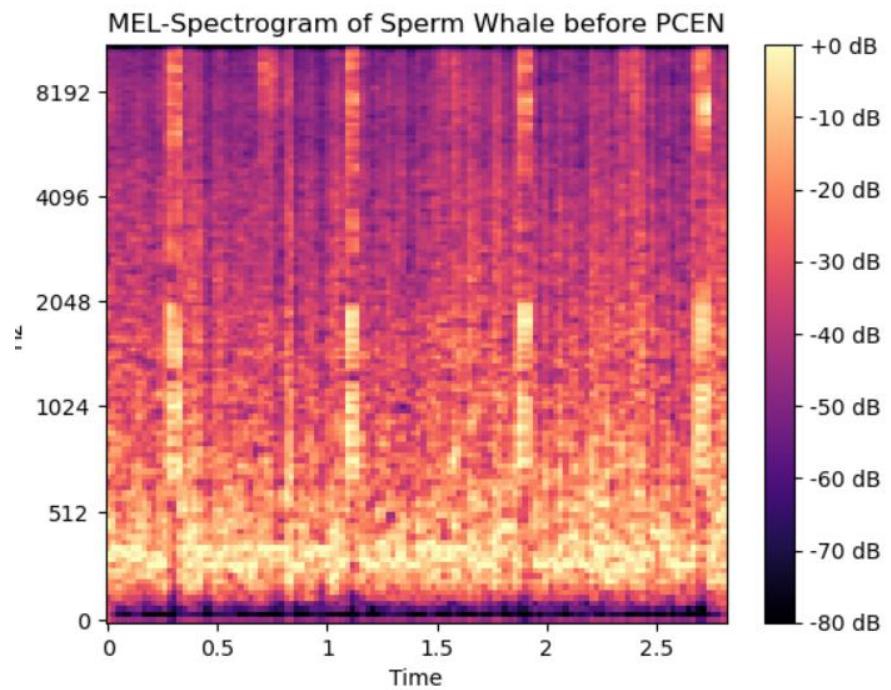
Πεδίο συχνότητας: Όπως ειπώθηκε και στο προηγούμενο ερώτημα, φαίνεται να **έχουμε εξασθένιση της ενέργειας στις υψηλότερες ενέργειες** (ανεβαίνοντας τον άξονα γ). Επίσης, φαίνεται να έχουμε εμφάνιση ενέργειας σε **μικρότερο εύρος συχνοτήτων στην humpback whale** (δεν φτάνει το έντονο χρώμα τόσο ψηλά όσο στην sperm) από ότι στην sperm. Συνεπώς τα φασματογραφήματα **συμφωνούν απόλυτα** με τα διαγράμματα φασματικής πυκνότητας ισχύος.

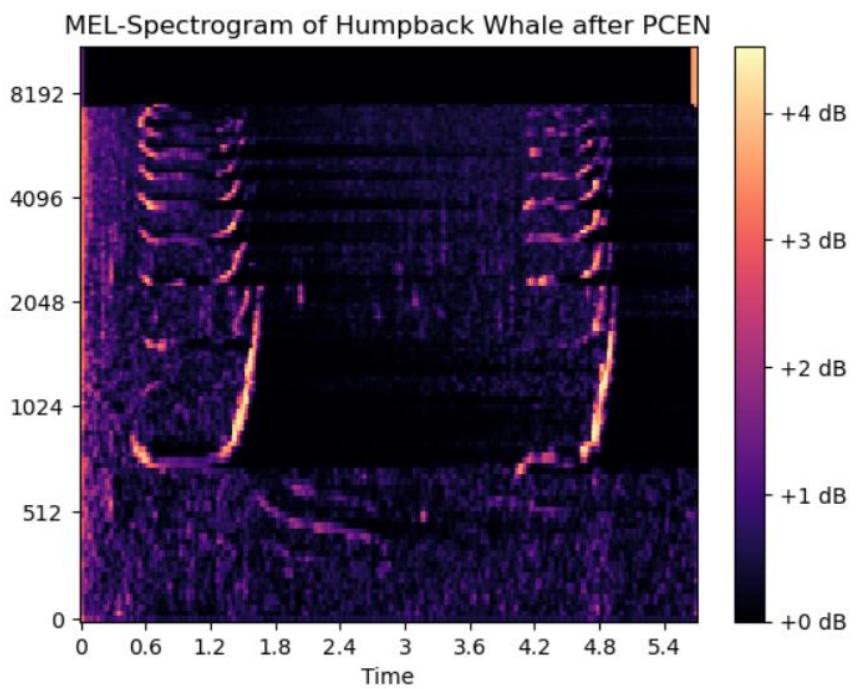
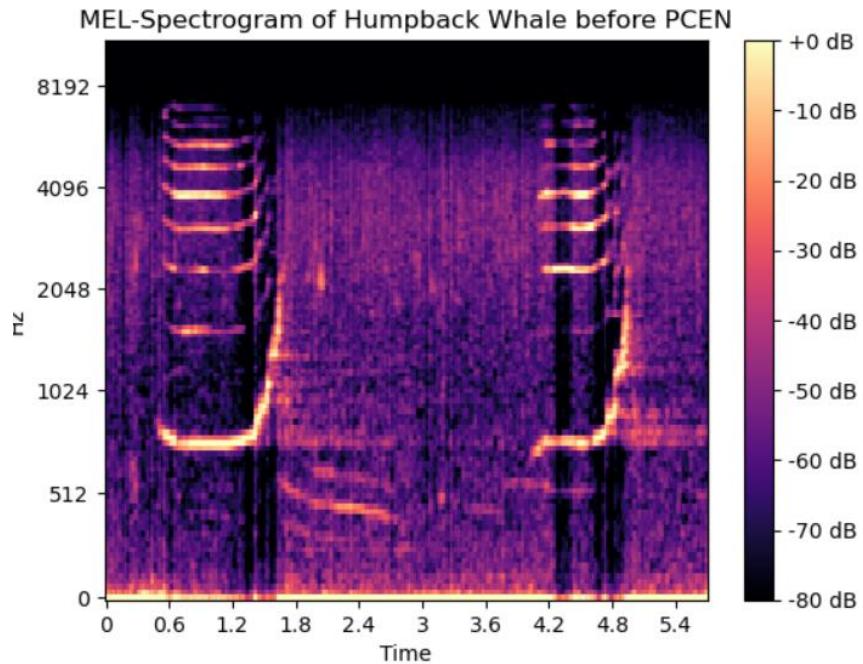
Ερώτημα 3.10

Στο ερώτημα αυτό, θα φτιάξουμε με τη βοήθεια της συνάρτησης ***librosa.feature.melspectrogram()*** το mel -φασματογράφημα των δύο βιοσημάτων πριν και μετά την εφαρμογή της μεθόδου PCEN (***librosa.pcen()***) η οποία αποτελεί τεχνική με την οποία **ενισχύουμε την αντίθεση μεταξύ του θορύβου και των σημάτων μας.**

Όσο για τα mel – φασματογραφήματα, **ύστερα από αναζήτηση στο διαδύκτιο, συμπεράναμε ότι πρόκεινται για φασματογραφήματα προσαρμοσμένα στις ανδρώπινες ακουστικές συχνότητας οι οποίες δεν είναι γραμμικές. Κύρια διαφορά τους από το κοινό φασματογράφημα είναι η παράσταση των συχνοτήτων σε κλίμακα mel.** Αυτά έχουν εφαρμογές στην αναγνώριση φωνής.

Παραθέτουμε τα αποτελέσματα μας:





Σχετικά με την κατανομή της ενέργειας στο χρόνο και το φάσμα δεν παρατηρούμε διαφορά σε σχέση με τα προηγούμενα συμπεράσματα που βγάλαμε **συνεπώς τα φασματογραφήματα και τα διαγράμματα φασματικής πυκνότητας ισχύος συμφωνούν και εδώ**. Ωστόσο αξίζει να σημειωθεί η **έντονη αντίθεση των χρωμάτων** μετά την μέθοδο PCEN

κάτι το οποίο δείχνει ότι όντως **ο θόρυβος φαίνεται να έχει περιοριστεί** από τα γραφήματά μας (**πιο μαύρο χρώμα στα σημεία που υπάρχει θόρυβος και πιο έντονο χρώμα στα σημεία σήματος**).

*Προφανώς υπάρχει κάποιο λάθος με τις κλίμακες, καθώς χωρίς κανονικοποίηση του mel diagram είχαμε πολλά ντεσιμπέλ πριν την εφαρμογή του PCEN **οπότε και αναγκαστήκαμε να κανονικοποιήσουμε.**