

Επεξεργασία Σήματος Φθόγγου με Χρήση Σύγχρονων Τεχνικών

Αθανάσιος Στ. Τσουνάκης
τμ. Ηλεκτρολόγων Μηχανικών &
Τεχνολογίας Υπολογιστών,
Πανεπιστήμιο Πατρών
up1072771@upnet.gr

Περίληψη—Λήψη και επεξεργασία ηχητικού σήματος έναρξης ανθρώπινης φωνής, χρήση της γλώσσας προγραμματισμού ειδικού σκοπού Matlab με σκοπό την αποκοπή υψηλών συχνοτήτων με τη δημιουργία βαθυπερατού φίλτρου, την διαμόρφωση του σήματος και κατόπιν την αποδιαμόρφωσή του.

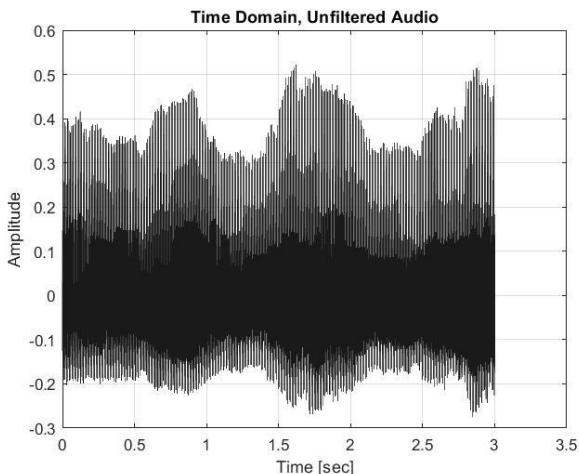
I. ΕΙΣΑΓΩΓΗ

Στην παρούσα εργασία θα αναλυθούν σταδιακά οι διαδικασίες επεξεργασίας ενός ηχητικού σήματος – συγκεκριμένα του φθόγγου /a/ – με χρήση τεχνικών ηλεκτρονικού υπολογιστή. Σε κάθε βήμα, θα παρουσιάζεται το θεωρητικό υπόβαθρο της διαδικασίας, θα παρατίθενται εικόνες, γραφήματα και σχεδιαγράμματα για την επίδειξη των εν λόγω τεχνικών και προβολή των αποτελεσμάτων της επεξεργασίας.

II. ΚΥΡΙΟ ΕΡΓΟ

A. Εγγραφή ηχητικού σήματος

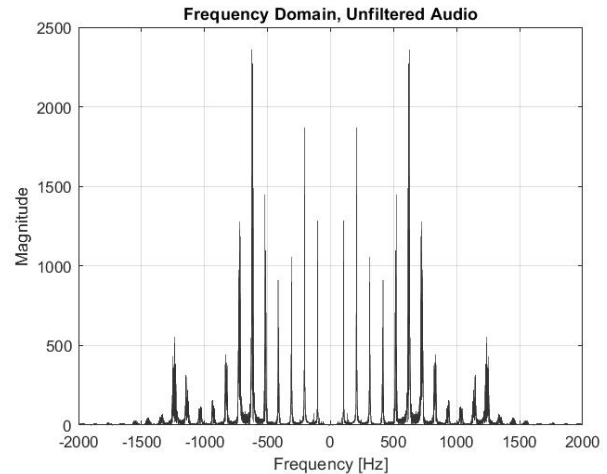
Χρησιμοποιώντας ένα μικρόφωνο, πραγματοποιήθηκε η εγγραφή ηχητικού κομματιού διάρκειας 3 s, σε μορφή μονοφωνικού .wav αρχείου με ρυθμό καταγραφής 48,000, το οποίο στο εξής θα αναφέρεται *x1.wav*. Ο φθόγγος που καταγράφηκε για την ανάγκη της εργασίας ήταν ο /a/. Ακολουθεί η απεικόνιση του σήματος στο πεδίο του χρόνου.



Σχήμα A.1: το ηχητικό σήμα στο πεδίο του χρόνου, δίχως καμία επεξεργασία.

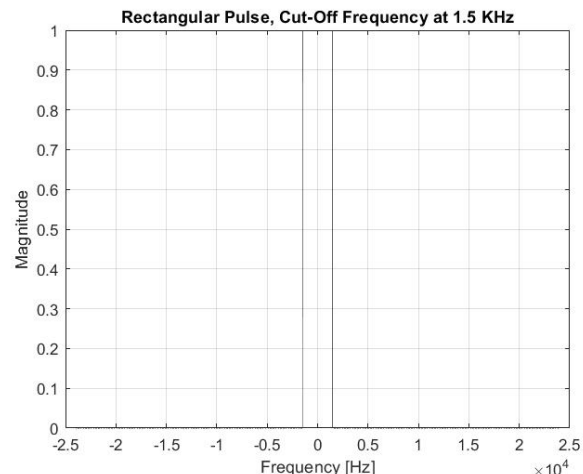
Επίσης, ακολουθεί η απεικόνιση του φάσματος συχνοτήτων του ηχητικού σήματος φθόγγου. Παρατηρώ πως στο πεδίο των συχνοτήτων, το σήμα εμπεριέχει κορυφές ακόμη και σε συχνότητες άνω των 500 Hz. Από τον Traunmueller και τον Eriksson (1991) μπορεί να αντληθεί η πληροφορία πως ελάχιστες πρώτες αρμονικές άνω των 150 Hz ανιχνεύονται γενικώς σε ηχητικά σήματα φωνής άρρενος. Επομένως, η κακή ποιότητα εγγραφής και ο θόρυβος του περιβάλλοντος είναι παράγοντες που συμβάλλουν στην

παρουσία επιπλέον συχνοτήτων στο ηχητικό κομμάτι και ταυτόχρονα δεν προκαλούνται από την προφορά του φθόγγου /a/. Άξιο παρατήρησης είναι πως η πρώτη κορυφή της συχνότητας, δηλαδή η πρώτη αρμονική του σήματος είναι στα 104 Hz, συμβαδίζοντας και με τ



Σχήμα A.2: το ηχητικό σήμα στο πεδίο των συχνοτήτων, δίχως καμία επεξεργασία.

Στη συνέχεια, για την αποκοπή των περιττών συχνοτήτων – άνω των 1.5 KHz –, εφαρμόζω βαθυπερατό φίλτρο στο *x1.wav* σήμα.



Σχήμα A.3: ο τετραγωνικός παλμός της μεταβλητής H, στα πλαίσια της εφαρμογής του φίλτρου.

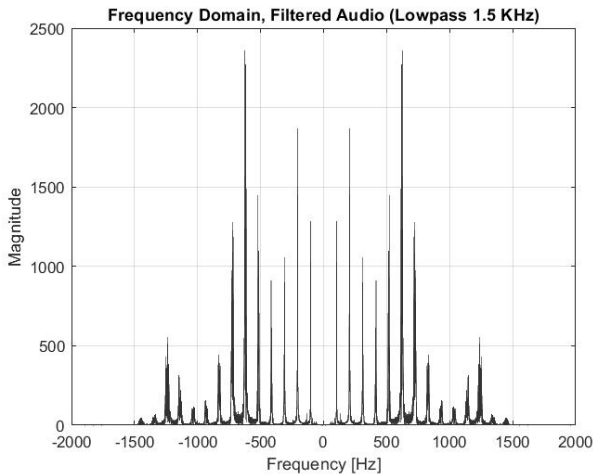
Τα φίλτρα, όπως αναφέρει ο Παπαμάρκος (2011) χρησιμοποιούνται ευρέως προκειμένου να επιτρέπουν το «πέρασμα» σημάτων επιλεγμένων περιοχών συχνοτήτων. Πιο ειδικά, τα βαθυπερατά φίλτρα (low-pass filters) επιτρέπουν τη διέλευση των σημάτων συχνοτήτων μικρότερων από την

συχνότητα αποκοπής ω , χαρακτηριστικό μέγεθος του φίλτρου.

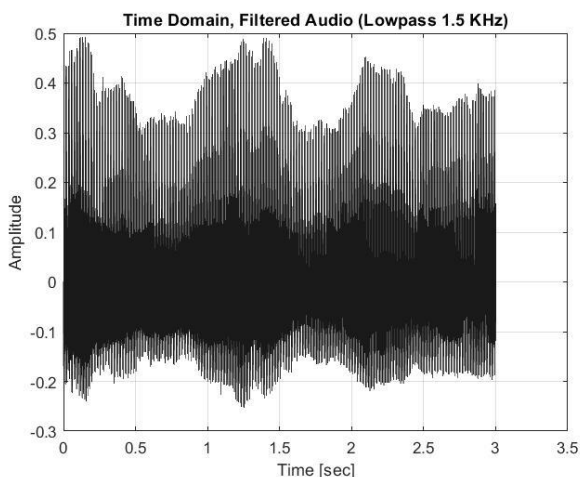
Η συνάρτηση μεταφοράς ενός ιδανικού χαμηλοπερατού φίλτρου ικανοποιεί την εξίσωση της μορφής:

$$H(j\omega) = \begin{cases} \exp(-j\omega t_b), & |\omega| \geq B \\ 0, & |\omega| < B \end{cases}$$

Πολλαπλασιάζοντας στο πεδίο της συχνότητας το σήμα x1.wav με έναν τετράγωνο παλμό από -1.5 KHz έως 1.5 KHz, και επιστρέφοντας με αντίστροφο μετασχηματισμό Fourier στο πεδίο της συχνότητας, μπορώ να επιτύχω το φιλτράρισμα του σήματός μου, αποκόπτοντας τις περιττές συχνότητες – δηλαδή εκείνες που είναι άνω της συχνότητας αποκοπής, όπως δίνεται από την άσκηση. Η διαδικασία αυτή περιγράφεται αναλυτικά στο Παράρτημα της εργασίας, όπου βρίσκεται ο κώδικας, στην συνάρτηση `firstPart()`. Ακολουθεί το γράφημα στο πεδίο της συχνότητας, άμα τη εφαρμογή του φίλτρου, και στο πεδίο του χρόνου, μετά την κατάλληλη εφαρμογή του αντίστροφου μετασχηματισμού Fourier.



Σχήμα A.4: το ηχητικό σήμα στο πεδίο της συχνότητας, έχοντας εφαρμοστεί το φίλτρο.



Σχήμα A.5: το ηχητικό σήμα στο πεδίο του χρόνου, έχοντας εφαρμόσει αντίστροφο μετασχηματισμό Fourier.

Στο εξής, το φιλτραρισμένο σήμα θα καλείται ως «x2».

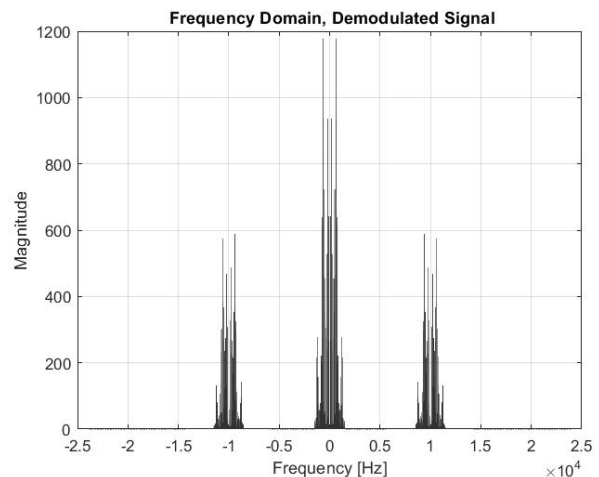
B. Διαμόρφωση και αποδιαμόρφωση του σήματος

Στα πλαίσια των τηλεπικοινωνιών και της επεξεργασίας σημάτων, πολλές φορές είναι κοινή στη βιομηχανία και στις τεχνολογίες του τελευταίου αιώνα η τεχνική της διαμόρφωσης πλάτους. Σύμφωνα με αυτήν την τεχνική, το σήμα πολλαπλασιάζεται με ένα υψίσυγχο σήμα, που καλείται φέρον, έτσι ώστε να τεθεί εφικτή η διέλευση του σήματος μέσα από ένα τηλεπικοινωνιακό κανάλι. Με την διαμόρφωση ενός εκπεμπτέου σήματος από έναν ραδιοφωνικό σταθμό, επιτυγχάνεται η συρρίκνωση της κεραίας του δέκτη, που εκ των πραγμάτων πρέπει να είναι μικρή για να εξοικονομείται χώρος στις συσκευές των δεκτών, αφού το μέγεθος της κεραίας είναι αντιστρόφως ανάλογο της συχνότητας του σήματος, όπως επιβεβαιώνει κι ο Wright (1987).

Όπως φαίνεται στο Παράρτημα του κώδικα, η τρέχουσα διαδικασία διαμόρφωσης και αποδιαμόρφωσης πλάτους που εφαρμόζεται στο σήμα x2, εξάγοντας τα σήματα x3 και x4 αντίστοιχα, παρουσιάζονται και αναλύονται στις συναρτήσεις `secondPart()` και `thirdPart()`.

Παράγοντας και πολλαπλασιάζοντας ένα φέρον σήμα με πολύ υψηλή συχνότητα – της τάξης των 5 KHz – με το σήμα x2 στον χρόνο, προκύπτει το διαμορφωμένο πλέον σήμα x3. Η σημασία των γραφικών του x3 δεν έχει ιδιαίτερο ενδιαφέρον καθώς λίγη πληροφορία για το αρχικό σήμα μας περιέχει, σε σχέση πάντα με τις πολύ ψηλές συχνότητες του φέροντος σήματος.

Επαναπολλαπλασιάζοντας στον χρόνο με το φέρον σήμα, προκύπτει το x4 σήμα, το αποδιαμορφωμένο πλέον σήμα. Ενδιαφέρον έχει το φάσμα του, όπως παρουσιάζεται ακολούθως. Παρατηρούνται κορυφές μακριά από τα λογικά πλαίσια των <0.2 KHz που αναφέρθηκαν ως αναμενόμενες συχνότητες για την αντρική φωνή. Αυτό προκύπτει ως επακόλουθο της συνέλιξης στην συχνότητα, του φέροντος και του x3 κατά τη διαδικασία αποδιαμόρφωσης.

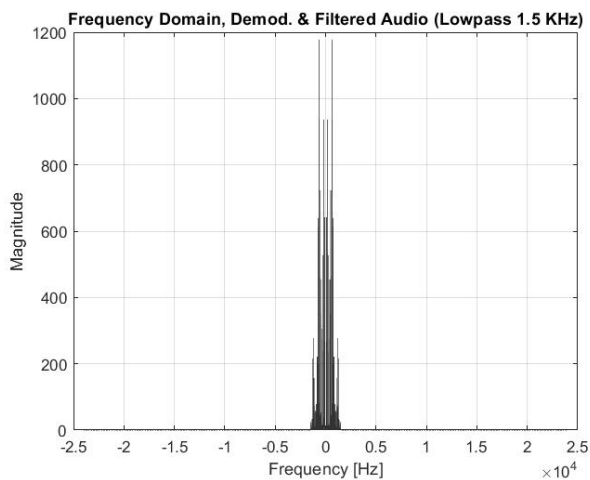


Σχήμα B.1: το αποδιαμορφωμένο σήμα, δίχως καμία επεξεργασία ή φιλτράρισμα.

Εφαρμόζοντας ένα φίλτρο, ουσιαστικά για να αποκόψουμε τις υψηλές συχνότητες που καθιστούν το εξαχθέν x4.wav αρχείο ακατάλληλο για ακρόαση, εξαιτίας της χαρακτηριστικού υψηλού τονισμού (συχνότητας) που υπερκαλύπτει την χρήσιμη πληροφορία, μπορεί να πραγματοποιηθεί η επαναφορά του ηχητικού σήματος.

Ακολουθώντας παρόμοια διαδικασία με το φιλτράρισμα στην αρχή της εργασίας, συνελίσσω στην συχνότητα το σήμα x4 με έναν τετραγωνικό παλμό ούτως ώστε να αποκόψω τις περιθωριακές συχνότητες.

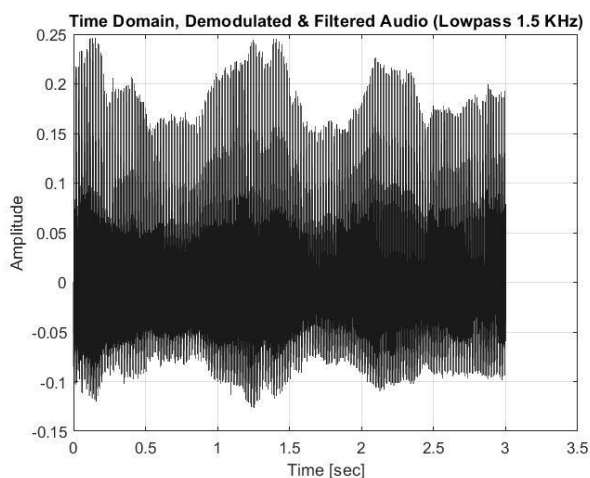
Ως αποτέλεσμα έχει:



Σχήμα B.2: το φιλτραρισμένο, πλέον, αποδιαμορφωμένο σήμα.

Για το πέρας της διαδικασίας, λαμβάνει χώρα αντίστροφος μετασχηματισμός Fourier στο σήμα x4 και προκύπτει το τελικό σήμα x5, που είναι πανομοιότυπο ακουστικά και γραφικά με το σήμα x2, δηλαδή το φιλτραρισμένο έως 1.5 KHz σήμα.

Ακολουθεί η γραφική απεικόνιση στο πεδίο του χρόνου:

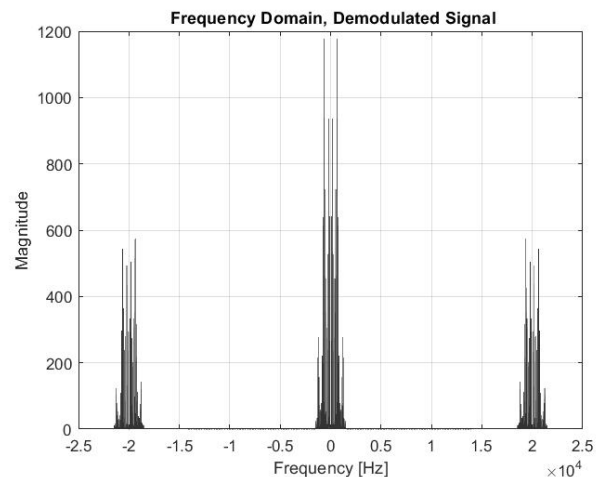


Σχήμα B.3: γραφική απεικόνιση στον χρόνο του τελικού αποδιαμορφωμένου και φιλτραρισμένου σήματος.

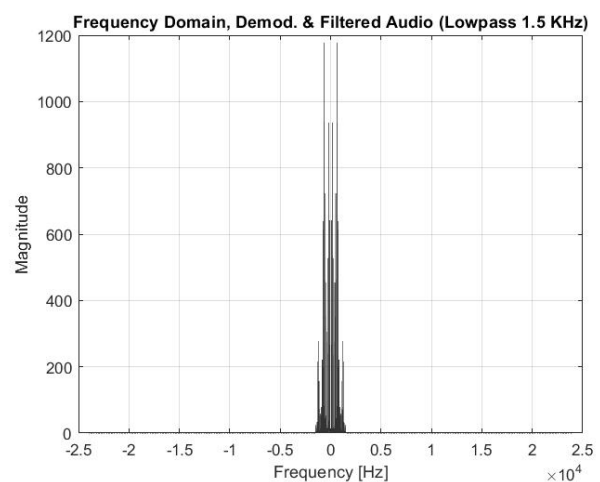
C. Επανάληψη διαδικασιών με φέρον σήμα 10 KHz

Η διαδικασία είναι ακριβώς η ίδια, παρατηρώντας τα εξής αποτελέσματα, αναφορικά με τα φάσματα και τις απεικονίσεις στο πεδίο του χρόνου.

Λαμβάνουμε υπόψη πως σε αυτή τη φάση της εργασίας, το υψίσυχνο σήμα με το οποίο συνελίσσουμε στη συχνότητα το σήμα ήχου, είναι της τάξης των 10 KHz, σε αντίθεση με τα 5 KHz του προηγούμενου μέρους της ασκήσης.



Σχήμα C.1: το αποδιαμορφωμένο σήμα, χωρίς κανένα φιλτράρισμα.



Σχήμα C.2: το φιλτραρισμένο αποδιαμορφωμένο σήμα.

Παρόμοια, το διαμορφωμένο αρχείο ήχου είναι ακατάληπτο λόγω του υψίσυχνου σήματος, ωστόσο μετά την αποδιαμόρφωση και το φιλτράρισμα των επιπλέον συχνοτήτων, το τελικό x5 σήμα, εξαχθέν ως .wav αρχείο είναι πανομοιότυπο ακουστικά, γραφικά και μαθηματικά με το αρχείο x2.

III. ΣΥΜΠΕΡΑΣΜΑΤΑ

Αναλύοντας θεωρητικά και λαμβάνοντας πειραματικά τα αναμενόμενα αποτελέσματα, διεξήχθη μία προσεκτική επεξεργασία του σήματος φθόγγου /a/ από άντρα, έλαβε χώρα μια μελετημένη και σχολαστική χρήση και οικειοποίηση των σύγχρονων τεχνικών διαμόρφωσης σημάτων, των εφαρμογών της θεωρίας του μαθήματος *Σήματα και Συστήματα* και επίσης της γλώσσας προγραμματισμού ειδικού σκοπού *MATLAB*.

ΑΝΑΦΟΡΕΣ

- Μάργαρης, Α. (2021). *ΣΗΜΑΤΑ ΚΑΙ ΣΥΣΤΗΜΑΤΑ, Συνεχούς και Διακριτού Χρόνου*. ΕΚΔΟΣΕΙΣ ΤΖΙΟΛΑ, Θεσσαλονίκη
- Παπαμάρκος, Ν. (2011) *ΗΛΕΚΤΡΙΚΑ ΚΥΚΛΩΜΑΤΑ ΤΟΜΟΣ ΑΒ*. ΑΦΟΙ ΠΑΠΑΜΑΡΚΟΥ ΟΕ, Ξάνθη
- Eriksson, A., Traunmueller, H.(1991) *The frequency range of the voice fundamental in the speech of male and female adults*. Institutionen foer lingvistik, Stockholms universitet. Στοκχόλμη, Σουηδία
- Wright, H. C. (1987) *An Introduction to Antenna Theory*. BERNARD BABANI (publishing) LTD, Λονδίνο, Ηνωμένο Βασίλειο

IV. ΠΑΡΑΡΤΗΜΑ

A. Κώδικας

```
function [] = signalEditing()
    [x1, rate] = audioread("x1.wav");

    x2 = firstPart(x1, rate, 1500);
    x3 = secondPart(x2, rate);
    x4 = thirdPart(x3, rate);
    x5 = lastPart(x4, rate, 1500);
end

function filteredSignal = lastPart (signal, rate, cut)
    f = (-length(signal)/2 : length(signal)/2 -1)*rate/length(signal);
    fourierSignal = fftshift(fft(signal));

    H = 1 - heaviside(f - cut) - heaviside(- f - cut);
    filteredFourierSignal = fftshift(H .* fourierSignal);

    figure
    plotFreq(fftshift(filteredFourierSignal), f, 'Frequency Domain, Demod. & Filtered Audio
    (Lowpass 1.5 KHz)');

    figure
    filteredSignal = 2*ifft(filteredFourierSignal);
    plotTime(filteredSignal, rate, 'Time Domain, Demodulated & Filtered Audio (Lowpass 1.5
    KHz)');

    audiowrite('x5.wav', filteredSignal, rate);
end

function demodulatedSignal = thirdPart (signal, rate)
    t = linspace(0, length(signal)/rate, length(signal));
    f = (-length(signal)/2 : length(signal)/2 -1)*rate/length(signal);

    cosine = cos(2*pi*10000*t);
    demodulatedSignal = signal .* cosine;
    figure
    plotTime(demodulatedSignal, rate, 'Time Domain, Demodulated Signal');

    figure
    plotFreq(fftshift(fft(demodulatedSignal)), f, ' Frequency Domain, Demodulated Signal');

    audiowrite('x4.wav', demodulatedSignal, rate);
end

function modSignal = secondPart (signal, rate)
    t = linspace(0, length(signal)/rate, length(signal));
    f = (-length(signal)/2 : length(signal)/2 -1)*rate/length(signal);

    cosine = cos(2*pi*10000*t);
    modSignal = signal .* cosine;
    figure
    plotTime(modSignal, rate, 'Time Domain, Modulated Signal');

    figure
    plotFreq(fftshift(fft(modSignal)), f, 'Frequency Domain, Modulated Signal');

    audiowrite('x3.wav', modSignal, rate);
```

```
end
```

```
function filteredSignal = firstPart (signal, rate, cut)
    figure
    plotTime(signal, rate, 'Time Domain, Unfiltered Audio');

    figure
    f = (-length(signal)/2 : length(signal)/2 -1)*rate/length(signal);
    fourierSignal = fftshift(fft(signal));
    plotFreq(fourierSignal, f, 'Frequency Domain, Unfiltered Audio');

    figure
    H = 1 - heaviside(f - cut) - heaviside(- f - cut);
    plotFreq(H, f, 'Rectangular Pulse, Cut-Off Frequency at 1.5 KHz');
    figure
    filteredFourierSignal = H .* fourierSignal';
    plotFreq(filteredFourierSignal, f, 'Freq Domain, Filtered Audio (Lowpass 1.5 KHz)');
    filteredFourierSignal = fftshift(filteredFourierSignal);

    figure
    filteredSignal = ifft(filteredFourierSignal);
    plotTime(filteredSignal, rate, 'Time Domain, Filtered Audio (Lowpass 1.5 KHz)');

    audiowrite('x2.wav', filteredSignal, rate);
end
```

```
function [] = plotFreq (data, f, ntitle)
    plot(f, abs(data),'color', [0 0 0] + 0.2);
    grid on;
    xlabel('Frequency [Hz]');
    ylabel('Magnitude');
    title(ntitle);
end
```

```
function [] = plotTime (data, Fs, ntitle)
    t = linspace(0, length(data)/Fs, length(data));
    plot(t, data, 'color', [0 0 0] + 0.1);
    grid on;
    xlabel('Time [sec]');
    ylabel('Amplitude');
    title(ntitle);
end
```