



המרכז האקדמי רופין  
Ruppin Academic Center

# דו"ח התקדמות - פרוייקט גמר תשפ"ד עיבוד דיגיטלי של אותות אודיו

מגיש: רותם צלישר 203773601

מנחה: ד"ר בני גור סולומון



## 1. מטרת הדו"ח

מטרת דו"ח ההתקדמות תהיה להציג את שלבי הפרוייקט שבוצעו עד כה, ולהציג את המשך התהליך. בדו"ח זה, אציג את חלקיו השונים של הפרוייקט, אתייחס למה בוצע עד כה, אסביר את התיאוריה והמימושים של מה שבוצע עד כה, אציג את מה עתיד להיות מבוצע ואסכם.

בתצוגת החומר איעזר, במידת הצורך, בגרפים, הוכחות מתמטיות, נוסחאות רלוונטיות ודיאגרמות בלוקים רלוונטיות.

להמחשת החומר (במידת הצורך בלבד) יצורפו להגשה קבצי מילב המיישמים את המובא בטקסט.

## 2. תזכורת לרקע וחלקי הפרוייקט

### 2.1 רקע:

בעולם המודרני שלנו, בו כל כך הרבה מכשירים משדרים מולטימדיה רוב שעות היממה, עולם עיבוד אותות האודיו תופס תאוצה. הצורך בשידור, שמיעה והעברת אודיו יוצר צורך בשיטות עיבוד טובות ויעילות, על מנת שנוכל להשתמש בשימושים השונים של המכשירים שלנו.

אותות אודיו, וביניהם נכללים גם אותות הדיבור, ניתנים לניתוח ועיבוד גם במישור הזמן וגם מתוך הסתכלות על התכולה הספקטרלית שלהם. האותות השונים, מורכבים מתכולת תדרים שניתן לנתח, לאפיין וגם לייצר מערכות שיגבירו (או ינחיתו) נוכחות של תדרים מסויימים. לכל כלי נגינה, כולל מיתרי הקול של האדם, יש תכולה ספקטרלית אופיינית לאקוסטיקה של המערכת המייצרת את הצליל. עובדה זו תעזור לנתח ולעבד את אותות האודיו.

יחד עם הצורך בשיטות העיבוד, ועם התפתחות ה AI בעולם, נולדו גם שימושים חדשים המקשרים את תחום למידת המכונה לעולם עיבוד האודיו. למשל - סירי ואלקסה, זאת של אפל וזאת של אמאזון, דורשות רמה גבוהה של speech recognition. בפרוייקט זה אנסה למדל יחידה שתדע לבצע Musical Genre Classification. אשתמש ב data set גדול על מנת לנסות ללמד את המודל להצליח להבדיל בין סגנונות מוסיקה (רוק, היפ הופ, ג'ז וכו'), תוך התחשבות בהיררכיית סגנונות נתונה, כאשר המטרה תהיה להגיע לדיוק של לפחות 85 אחוז בסיווג בין שני סוגי מוסיקה, ולפחות 60 אחוז דיוק בסיווג בין כל סגנונות המוסיקה. אנסה לנתח את תוכן האודיו ברמת הפיצ'רים, לרבות התוכן הריתמי, גבהי צליל ופיצ'רים נוספים על מנת להגיע ליכולת לסווג את קטע האודיו בדיוק מוצלח מספיק.

## 2.2 חלקי הפרוייקט:

הפרוייקט יחולק לשלושה חלקים.

- **משוונים:** בפרק זה, אצלול לתוך השיטות השונות לאיזון ספקטרלי של אותות אודיו. כחלק מפרק זה, איישם מימוש של משוון 7-Band (בעל חלוקה ל 7 יחידות עיבוד מקבילות). אתחיל מלהסביר ולפתח את הרכיבים ממנו בנוי המשוון – המסננים השונים.  
  
משוונים שונים ניתנים למימוש ע"י אסופה של מסננים שונים. משפחות המסננים מתחלקות למסנני peak, shelving, notch, וניתן לבצע סינונים ע"י אלגוריתמים מתקדמים יותר. אם נרצה, לצורך המחשה, להגביר תכולת תדר סביב  $f_c$  כלשהו, נוכל לממש מערכת המורכבת מ all pass במקביל עם BP פילטר, והמערכת המשותפת תיקרא peak filter. המערכת תעביר את כל הסיגנל במלואו, ובנוסף תגביר (או תנחית) את תחום התדר הרצוי. בעזרת פילטר זה, ובשיטות של חישוב מקדמי המסננים כתלות בפרמטרים של המערכת (תדר קטעון, רוחב סרט לתחום ההגבר \ הנחת וכו') נוכל לממש את המשוון המלא (למשל: פונקציה שתקבל את ההגבר הרצוי DB, תדר FC סביבו מגבירים ורוחב סרט BW, ותדע לייצר את מקדמי המסנן PEAK\SHELVING הרצוי).  
  
לאחר פיתוח סוגי המסננים השונים, אדגים שילוב שלהם לכדי יחידת EQ אחת גדולה.
- **אם יותיר הזמן, לצורך המחשה להצגת הפרוייקט המלא (ההצגה הסופית) –**  
**אבנה יחידת GUI אשר תיתן למשתמש ממשק, דרכו המשתמש ייגש לפרמטרים השונים של המערכת ויוכל לערוך את האודיו דרכם.**
- **סימולציית חדר:** בפרק זה, אדגים שיטה לעיבוד אודיו, המאפשרת לנו לערוך קטע אודיו "יבש" (מוקלט בתנאים אופטימליים) לכדי אודיו מעובד, שמטרתו תהיה להישמע כאילו הקטע הוקלט (או נוגן) בחדר בעל הד.  
בשיטה זו, ניתן לקחת את האות היבש ולהעביר אותו בקונבולוציה יחד עם ה"תגובה לתדר" של החדר שאותו אנו מנסים לתאר. כל חדר מתאפיין ביכולת מסויימת להחזיר (או לבלוע) את התדרים המתפשטים בו, כתלות בחומר ממנו עשויים כל ה"מכשולים" בחדר שגורמים לגלים חוזרים. ע"פ זאת, אם נצליח לערוך באופן דיגיטלי את האות, כך שיייתקבל ייצוג לא רק של האות המוקלט, אלא של האות המוקלט הכולל בתוכו את החזרי הגלים הפוטנציאליים מהחדר, נוכל לייצר את

האשליה שהאות הוקלט (או הושמע) בחדר הנ"ל. התיאור המתמטי של החזרי הגלים המתקבלים מהחדר ייקרא "התגובה לתדר" של החדר. בפרק זה אדגים שני אלגוריתמים, בשם SCHRODER ו-MOORER, המיישמים את הלוגיקה הזו לעיבוד הסאונד.

- למידת מכונה: בפרק זה, אנסה להעמיק בשיטות השונות לערב למידת מכונה עבור עיבוד אודיו. למידת מכונה תבוא לידי ביטוי בניתוחי ספקטוגרמות, אלגוריתמים שונים לזיהוי פיצ'רים ונושאים מתקדמים נוספים. בפרוייקט זה, אתמקד בשיטות לסיווג סגנון מוסיקלי.

לצורך פרק זה, ישנו data set גדול, המורכב מקטעי אודיו המסווגים לסגנונות השונים, ביניהם, קלאסי, רוק, ג'ז וכולי, כמו גם תמונות של ספקטוגרמות התואמות את הסגנונות השונים, מהן יהיה ניתן לנתח את התמונה האופיינית המתקבלת עבור סגנון ספציפי. ישנם גם קבצי EXCEL המכילים תיוגים על מידע אופייני לסוגי המוזיקה השונים (טמפו, תכולות תדר, RMS וכו'). אנסה לייצר מערכת שתגיע לכל הפחות ל60 אחוז דיוק בסיווג בין כלל הסגנונות השונים, ו85 אחוז דיוק בקביעת סיווג בין שני סגנונות אפשריים.

**פרק זה טרם יושם ולכן מובא כמבוא בלבד, המימושים של הפרק יוצגו בהצגת הפרוייקט הסופי.**

### 3. פרק 1 - משוונים:

איזון ספקטרלי של אותות אודיו הינה מתודה חשובה ושימושית מאוד להמון אפליקציות מודרניות. החל מהרדיו שיושב ברכב, ועד לאולפני הפקת סאונד (מוסיקה, פודקאסטים וכו'..). – בכולם ניתן למצוא שימוש במשווני אודיו.

בפרק זה אציג את המהלך השלם למימוש משוון פרמטרי בעל 7 יחידות עיבוד (7-Band Parametric EQ), החל מאבני הבניין הקטנה ביותר במערכת – המסנן (פילטר).

#### הגדרות ומושגי יסוד:

סינון אותות אודיו מתבצע ע"י אחד (או כמה) מסוגי המשוונים הבאים:

1. LP/HP – מסננים אלו מוגדרים ע"י "תדר קטעון" (מסומן:  $f_c$ ). תדר זה מסמל ירידה במגניטודת האות ב-3 db ואחריו (או לפניו) מגיע תחום הקטעון (LP/HP בהתאמה).

2. BP/BS – מסנני band pass/stop אלו מסננים המקיימים מגוון תחומי מעבר וקטעון לאות בודד. מוגדים ע"י שני תדרי קטעון  $f_l, f_u$  אשר מסמלים את נק' ההנחת ב-3 db בין תחום תדרים זה מתקיים פס ה"מעבר" (בBS: קטעון). ניתן לתאר את תחום המעבר (או קטעון):  $fb = f_u - f_l$ .

3. Shelving Filters – פילטרים דומים בתכונותיהם לLP/HP, אך ההבדל הוא שלפילטר אין "תחום קטעון". במקום – הפילטר יודע לייצר הגבר \ הנחת לתחום תדר מסויים, בעוד ששאר התחומים נשארים ללא עיבוד.

4. Peak / Notch Filters – פילטרים דומים בתכונותיהם לBP/BS, אך, כמו במקרה הקודם, ההבדל הוא שלפילטר אין "תחום קטעון". במקום – הפילטר יודע לייצר הגבר \ הנחת לתחום תדר מסויים, בעוד ששאר התחומים נשארים ללא עיבוד.

#### פילטרים רקורסיביים:

##### הקדמה:

למימוש פילטרים ישנן שתי גישות. האחת: מימוש הפילטר תוך שימוש ב"אפסים" בלבד, במישור s (או z) – פילטר חסר משוב (לא רקורסיבי). השניה: מימוש הפילטר תוך שימוש ב"אפסים" וב"קטבים" (המעידים על היותו פילטר רקורסיבי, בפרספקטיבה כללית). לשתי הגישות ישנן משמעויות ותכונות שונות, בעוד שהנקודה המרכזית שניתן לגעת בה כרגע היא זמן החישוב של המערכת. בעזרת פילטר רקורסיבי (בעל קטבים), נוכל לממש מסנן מסדר נמוך בהרבה יותר מכך של מסנן בעל אפסים בלבד, על מנת לקרב פתרון של בעיה זוהה. מכך, נסיק שפילטר רקורסיבי יוכל לבצע את החישוב בזמן קצר יותר (סדר נמוך יותר של מערכת).

את המסננים שלנו נתכנן במישור S (לפלט), המתאים לצורך המחשת הדיון לכרגע. נזכור, כי בפועל, התמרת לפלט תתאר משוואות דיפרנציאליות של פונקציות רציפות, בעוד, הייצוג

הדיגיטלי של המידע מכריח אותנו לחשוב על עיבוד המידע שיש לנו כעיבוד של מידע **בדיד**. בהמשך, נדון בטרנספורמציה למישור Z המטפל בדיוק בנושא זה.

HP/LP/BP:

על מנת לעבד מידע המיוצג בתחום התדר, נוכל לעשות מימוש במסנני LP, BP ו-HP. נוכל לייצר תחומי הגבר וקיטעון לכל תדר ורוחב סרט במרחב האודיו שלנו, שבאופן טיפוסי נע בין 0-20 קילו הרץ (האוזן האנושית לא כך כך שומעת מעבר ל-20 קילו הרץ).

ייצוג במישור לפלס של מסננים אלו (כתלות בתדר הקטעון):

$$H_{HP}(s) = \frac{s^2}{s^2 + \frac{\omega_c}{Q_\infty}s + \omega_c^2} \quad H_{LP}(s) = \frac{\omega_c^2}{s^2 + \frac{\omega_c}{Q_\infty}s + \omega_c^2},$$

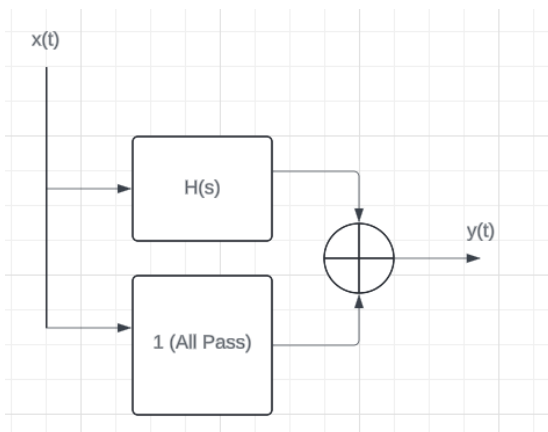
$$H_{BS}(s) = \frac{s^2 - \omega_c^2}{s^2 + \frac{\omega_c}{Q_\infty}s + \omega_c^2} \quad H_{BP}(s) = \frac{\frac{\omega_c}{Q_\infty}s}{s^2 + \frac{\omega_c}{Q_\infty}s + \omega_c^2},$$

כאשר רוחב הסרט של המסננים ייקבע ע"פ גורם הטיב (Q), אשר מייצג את היחס בין רוחב הסרט לתדר הקטעון.

Shelving Filters:

בעיה מסוימת צצה כאשר אנחנו רוצים לעבד תחום תדר, ולהשאיר את שאר התחומים בלי עיבוד. המסננים שהצגתי עד כה אמנם מעבדים את התחום המתבקש אך מייצרים קטעון לשאר תחום התדר.

לשם כך, נתאר את העיבוד הרצוי כמערכת שמשלבת העברת אות ללא עיבוד, ובמקביל העברת אות ליחידת עיבוד. כלומר:



במוצא נקבל את האות המעובד בתחום הרצוי, בעוד בשאר התחומים האות יעבור כמו שהוא ללא עיבוד.

תיאור מתמטי של מסננים אלו בעבור LP\HP:

$$H(s) = 1 + H_{HP}(s) \quad H(s) = 1 + H_{LP}(s)$$

ופונקציית התמסורת למערכת מסדר שני התלויה בפרמטר הגבר כלשהו  $V_0$ :

עבור LP:

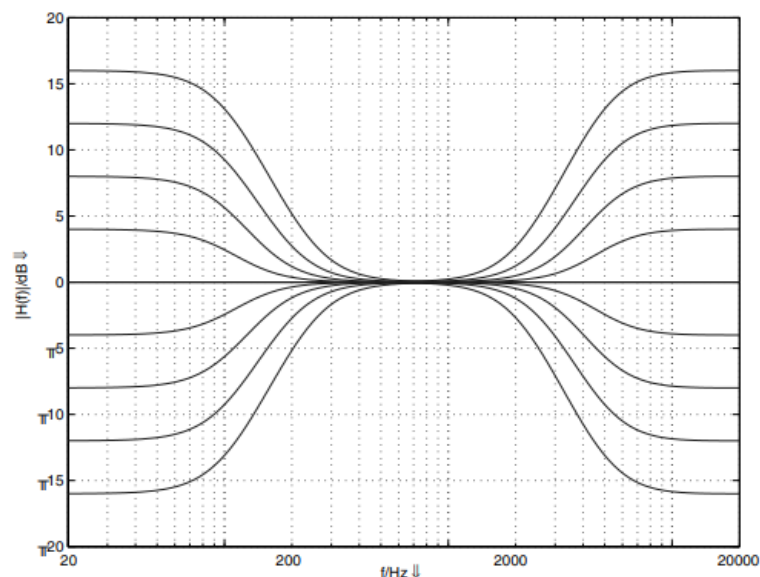
$$H(s) = \frac{s^2 + \sqrt{2V_0}s + V_0}{s^2 + \sqrt{2}s + 1}$$

עבור HP:

$$H(s) = \frac{V_0s^2 + \sqrt{2V_0}s + 1}{s^2 + \sqrt{2}s + 1}$$

(ניתן לשים לב לאופי התגובה כאשר לוקחים את  $s$  לאינסוף ולאפס).

דוגמא לתגובות התדר של מסננים אלו כתלות בערכי  $V_0$  שונים:



**Figure 5.14** Frequency responses of second-order low-/high-frequency shelving filters – low-frequency shelving filter  $f_c = 100$  Hz (parameter  $V_0$ ), high-frequency shelving filter  $f_c = 5000$  Hz (parameter  $V_0$ ).

:Peak Filters

סוג נוסף של מסנן המתוכנן לתת הגבר או הנחת לתחומי תדר ספציפיים, בעוד שאר התדרים נשארים בלא עיבוד. בניגוד למסננים בסעיף הקודם, הנותנים מענה לתחום "כללי" התואם לאיזורי תדר "גבוהים" או "נמוכים", מסנן זה יותר דומה לBP, ולכן BP יהיה התשתית שעליה נעבוד. בדומה למסננים מסעיף קודם, גם את מסנן זה ניישם ע"י שליחה של האות דרך נתיב ALL PASS, ובמקביל שליחה דרך נתיב BP, ונסכום את המוצא. כלומר:

$$H(s) = 1 + H_{BP}(s).$$



דוגמא לתגובות התדר של מסננים אלו בעבור ערכי  $v_0$  שונים:

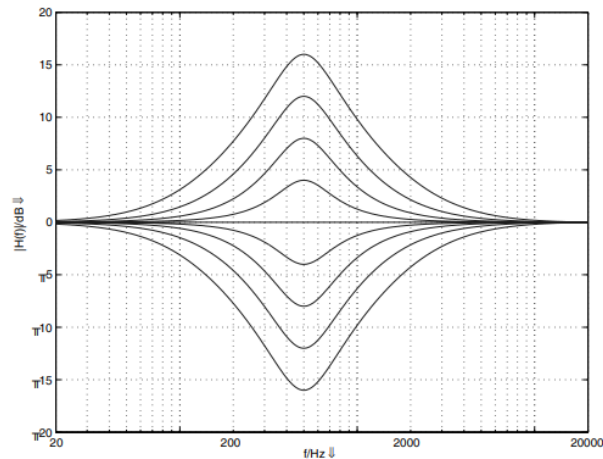


Figure 5.16 Frequency response of a peak filter –  $f_c = 500$  Hz,  $Q_\infty = 1.25$ , cut parameter  $V_0$ .

### טרנספורמצית Z (מעבר מתכנון אנלוגי לספרתי):

כאמור, לאחר שדנים בתכנון מערכת כלשהי במישור לפלס, יש לזכור כי במישור זה אנו משערכים פתרונות למשוואות דיפרנציאליות ואינטגרליות של פונקציות רציפות. כאשר אנו דנים בעיבוד ספרתי של אותות, יש לייצר את השערוך הנ"ל בעולם הבדיד (לדוג' – במחשב לא קיימת שום סוג של "רציפות". המידע הוא אותות בדידים המיוצגים על מרחב בדיד).

על מנת ליישם את העיבוד בעולם הבדיד, ניתן לבצע מעבר ממישור לפלס למישור Z, אשר מייצג לנו את התכנון של מערכת בדידה, עבור אותות בדידים.

המטרה תהיה, למפות את הקטבים והאפסים, אשר מגדירים את המערכת במישור לפלס, למישור Z.

מיפוי זה הולך להתקבל ע"פ הטרנספורמציה הבאה:

$$w_d = \tan\left(w_c * \frac{T}{2}\right) \quad s = \frac{2}{T} \frac{z-1}{z+1}.$$

תתקבל לנו פונקציית תמסורת במישור Z. נתאר אותה באופן כללי בעזרת המקדמים שלה:

$$H(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}},$$

ניתן להראות, בהצבה של משתנה עזר  $K = \tan\left(w_c * \frac{T}{2}\right)$ , שאפשר לקבל טבלה המייצגת את כל ערכי המקדמים הנדרשים, לכל פונקציית תמסורת מהפונקציות שנדרשות לנו למימוש הפילטרים שצינו למעלה (טבלה בעמוד הבא).

טבלת המקדמים כתלות בפרמטרים  $v_0, \omega_c$ :

**Table 5.3** Peak filter design.

Peak (boost $V_0 = 10^{G/20}$ )				
$a_0$	$a_1$	$a_2$	$b_1$	$b_2$
$\frac{1 + \frac{V_0}{Q_\infty} K + K^2}{1 + \frac{1}{Q_\infty} K + K^2}$	$\frac{2(K^2 - 1)}{1 + \frac{1}{Q_\infty} K + K^2}$	$\frac{1 - \frac{V_0}{Q_\infty} K + K^2}{1 + \frac{1}{Q_\infty} K + K^2}$	$\frac{2(K^2 - 1)}{1 + \frac{1}{Q_\infty} K + K^2}$	$\frac{1 - \frac{1}{Q_\infty} K + K^2}{1 + \frac{1}{Q_\infty} K + K^2}$
Peak (cut $V_0 = 10^{-G/20}$ )				
$a_0$	$a_1$	$a_2$	$b_1$	$b_2$
$\frac{1 + \frac{1}{Q_\infty} K + K^2}{1 + \frac{V_0}{Q_\infty} K + K^2}$	$\frac{2(K^2 - 1)}{1 + \frac{V_0}{Q_\infty} K + K^2}$	$\frac{1 - \frac{1}{Q_\infty} K + K^2}{1 + \frac{V_0}{Q_\infty} K + K^2}$	$\frac{2(K^2 - 1)}{1 + \frac{V_0}{Q_\infty} K + K^2}$	$\frac{1 - \frac{V_0}{Q_\infty} K + K^2}{1 + \frac{V_0}{Q_\infty} K + K^2}$

**Table 5.4** Low-frequency shelving filter design.

Low-frequency shelving (boost $V_0 = 10^{G/20}$ )				
$a_0$	$a_1$	$a_2$	$b_1$	$b_2$
$\frac{1 + \sqrt{2V_0} K + V_0 K^2}{1 + \sqrt{2} K + K^2}$	$\frac{2(V_0 K^2 - 1)}{1 + \sqrt{2} K + K^2}$	$\frac{1 - \sqrt{2V_0} K + V_0 K^2}{1 + \sqrt{2} K + K^2}$	$\frac{2(K^2 - 1)}{1 + \sqrt{2} K + K^2}$	$\frac{1 - \sqrt{2} K + K^2}{1 + \sqrt{2} K + K^2}$
Low-frequency shelving (cut $V_0 = 10^{-G/20}$ )				
$a_0$	$a_1$	$a_2$	$b_1$	$b_2$
$\frac{1 + \sqrt{2} K + K^2}{1 + \sqrt{2V_0} K + V_0 K^2}$	$\frac{2(K^2 - 1)}{1 + \sqrt{2V_0} K + V_0 K^2}$	$\frac{1 - \sqrt{2} K + K^2}{1 + \sqrt{2V_0} K + V_0 K^2}$	$\frac{2(V_0 K^2 - 1)}{1 + \sqrt{2V_0} K + V_0 K^2}$	$\frac{1 - \sqrt{2V_0} K + V_0 K^2}{1 + \sqrt{2V_0} K + V_0 K^2}$

**Table 5.5** High-frequency shelving filter design.

High-frequency shelving (boost $V_0 = 10^{G/20}$ )				
$a_0$	$a_1$	$a_2$	$b_1$	$b_2$
$\frac{V_0 + \sqrt{2V_0} K + K^2}{1 + \sqrt{2} K + K^2}$	$\frac{2(K^2 - V_0)}{1 + \sqrt{2} K + K^2}$	$\frac{V_0 - \sqrt{2V_0} K + K^2}{1 + \sqrt{2} K + K^2}$	$\frac{2(K^2 - 1)}{1 + \sqrt{2} K + K^2}$	$\frac{1 - \sqrt{2} K + K^2}{1 + \sqrt{2} K + K^2}$
High-frequency shelving (cut $V_0 = 10^{-G/20}$ )				
$a_0$	$a_1$	$a_2$	$b_1$	$b_2$
$\frac{1 + \sqrt{2} K + K^2}{V_0 + \sqrt{2V_0} K + K^2}$	$\frac{2(K^2 - 1)}{V_0 + \sqrt{2V_0} K + K^2}$	$\frac{1 - \sqrt{2} K + K^2}{V_0 + \sqrt{2V_0} K + K^2}$	$\frac{2(K^2/V_0 - 1)}{1 + \sqrt{2/V_0} K + K^2/V_0}$	$\frac{1 - \sqrt{2/V_0} K + K^2/V_0}{1 + \sqrt{2/V_0} K + K^2/V_0}$

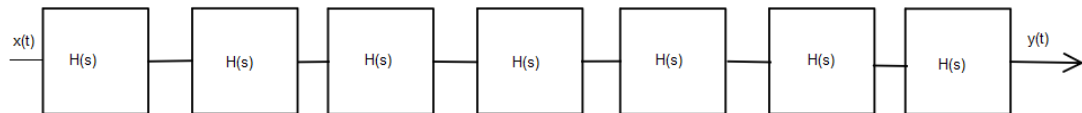
בטבלה זו נשתמש בפועל, על מנת לחשב את המקדמים לכל אחד מהפילטרים הנדרשים למימוש המשוון.

### משוון פרמטרי בעל 7 יחידות עיבוד (7 Band Parametric EQ):

כעת אוכל לממש את המערכת הנדרשת לעיבוד האות הדיגיטלי. לבחירה לממש את המשוון בעזרת 7 יחידות עיבוד יש משמעויות נרחבות, שמפעת גבולותיו הסופיים של הפרוייקט לא כנסתי לחקור אותם (לעת עתה). הבחירה ב-7 יחידות הינה (פסאדו) שרירותית, מכיוון שמהיכרות עם מוצרים שקיימים בשוק, רוב המשוונים שראיתי עד כה היו בעלי 7 יחידות.

אחלק את המשוון ל 2 יחידות בקצוות הספקטרום, Low Shelving – High Shelving,  $\omega - 5$  יחידות במרכז הספקטרום: Peak Filters. לכל יחידה יש שליטה בפרמטרים  $f_c, G$ , כאשר  $f_c$  הוא תדר הקיטעון בהרצים,  $G$  הוא ההגבר הנדרש ב dB.

בעזרת הטבלאות הפרמטריות, נחשב את המקדמים של כל אחת משבעת פונקציות התמסורת, לאחר מכן נעביר את האות עיבוד בקו הכולל את שבעת המסננים (כאשר בעבור מסנן במצב BYPASS, ללא עיבוד, נקבל במוצא את אות הכניסה).



#### 4. משוונים - מימוש:

את המימוש בפרק זה אראה בעזרת פונקציות מטלב, ואיעזר בגרפים להמחשת העיבוד.

ישנן 3 פונקציות מטלב (מצורפות להגשה) שמטרתן לחשב, ולהחזיר את מקדמי המסננים. שלושת הפונקציות מקבלות כקלט את התדר המרכזי  $f_c$ , ההגבר  $G_{db}$  ותדר הדגימה של המערכת  $f_s$ . חישוב מקדמי ה PEAK דורש גם את גורם הטיב ולכן הפונקציה מקבלת גם את רוחב הסרט כפרמטר.

חתימות הפונקציות:

```

function [b,a] = calc_bp_coeffs(fc,Gdb,BW, fs)
function [b,a] = calc_hp_coeffs(fc,Gdb, fs)
function [b,a] = calc_lp_coeffs(fc,Gdb, fs)
  
```

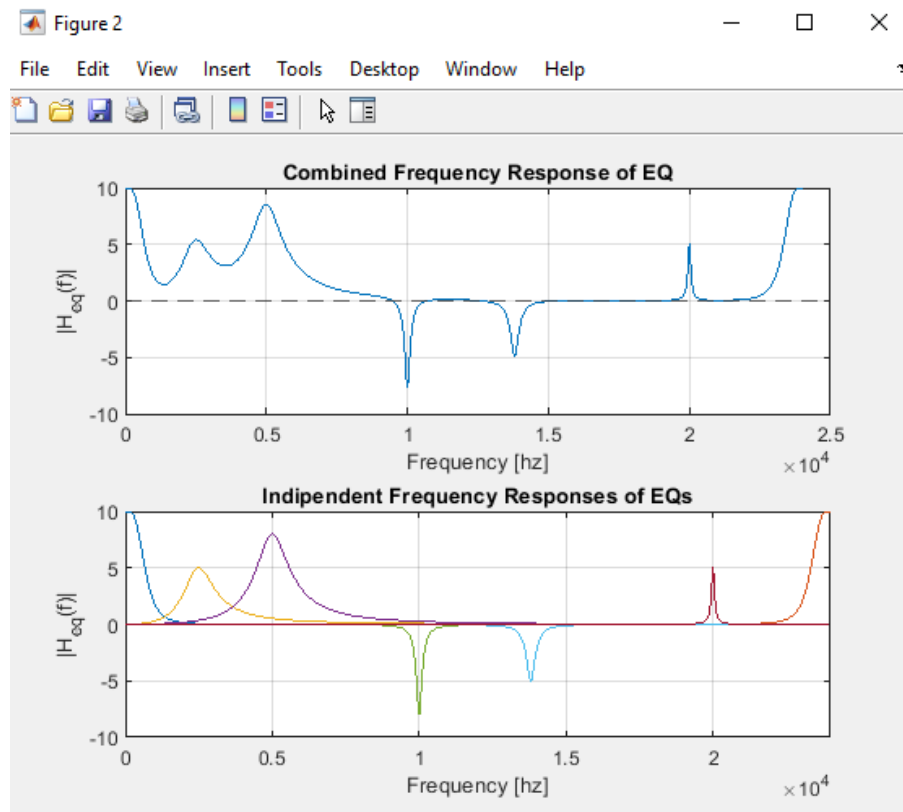
כל אחת משבעת יחידות העיבוד מקבלת מבנה לוגי המכיל את המקדמים שלה, התדר המרכזי ביחס אליו היא מעבדת, ההגבר ותדר הדגימה (כאשר הגבר 0 אומר שהמערכת לא מעבדת והיא ב BYPASS). לדוגמא:

```

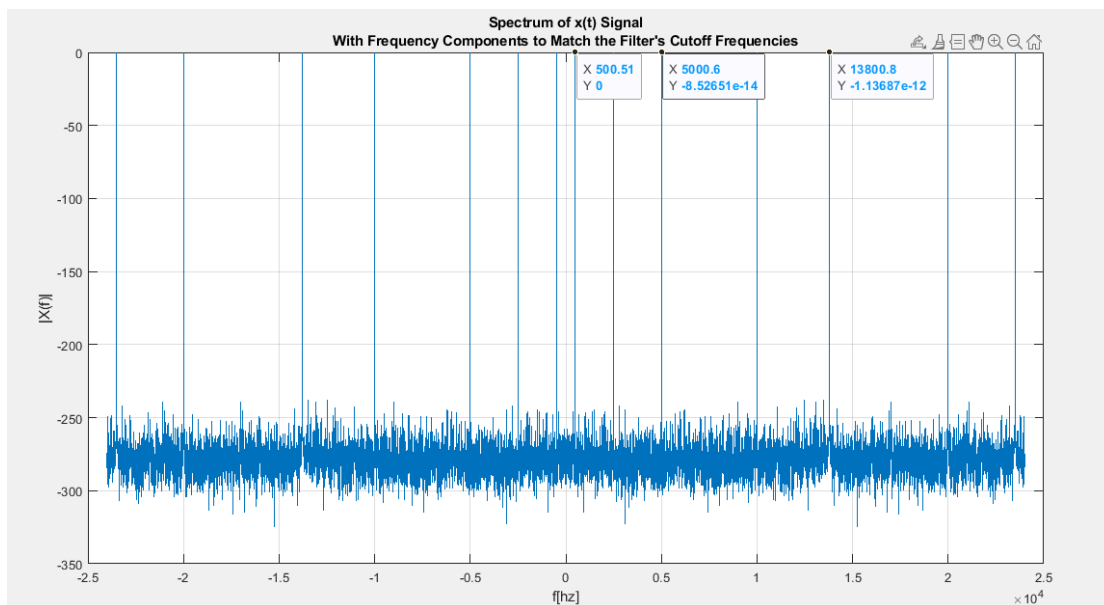
% calc coeffs
hp1.Gdb = 10; hp1.fc = 500;
[hp1.b, hp1.a] = calc_lp_coeffs(hp1.fc, hp1.Gdb, fs);
  
```

השימוש ב struct נובע מהצורך לשמור את המקדמים במערכת, כל עוד לא התבצע שום "שינוי" מאז החישוב האחרון. כך, נוכל להקטין את כמות החישובים בכך שנבצע חישוב מקדמים בכל פעם שהמשתמש משנה את הפרמטרים בלבד, ולא בכל פעם שנכנסת סדרה לעיבוד.

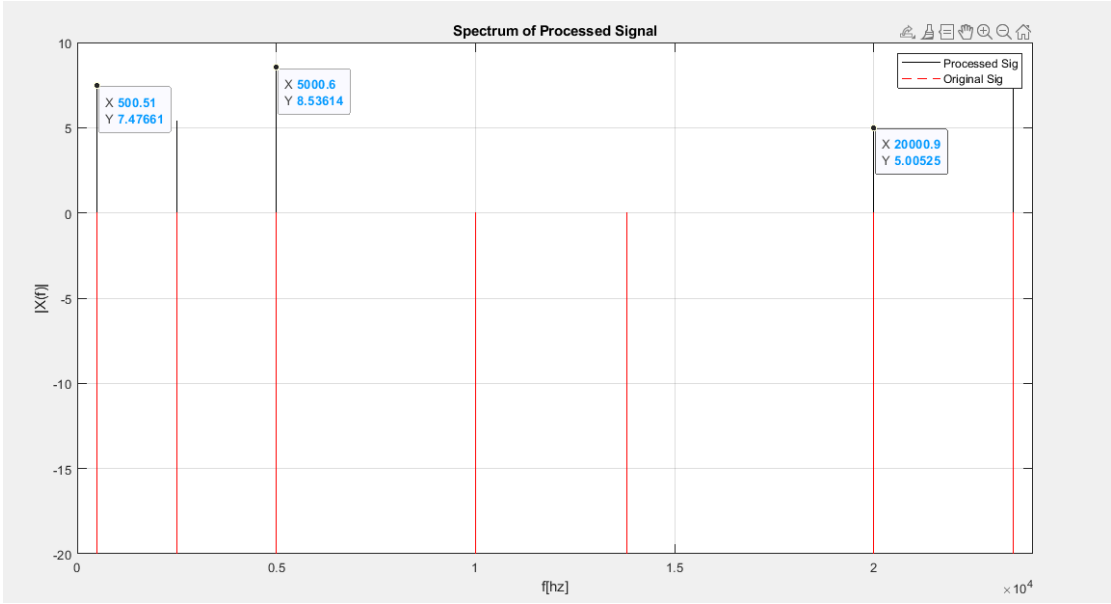
הדגמה של תגובת התדר הכוללת של המשוון המורכב מ-7 יחידות עיבוד:



בעבור האות הבא:



כשנעביר את האות ביחידת העיבוד בעלת תגובת התדר שצורפה בגרף למעלה, נקבל  
במוצא את התמונה:



## 5. פרק 2 – סימולציית חדר:

גלי קול הינם גלי לחץ אוויר הנעים ומתפשטים במרחב בו הם חיים. כמו כל גל המתפשט, הגל יכול לעבור תופעות שונות ומשונות במפגש עם תכונות תווך מסוימות.

גל קול אשר נולד בחדר סגור, יתפשט עד שייפגע בעצמים שונים בתווך, או בקירות התווך. פגיעה זו, תייצר אפקטים של החזרה ובליעה של הגל. ניתן להבין, שהגל שמגיע לנו לאוזן בחדר כזה הוא תרכובת של גלי "מקור" וגלים "חוזרים".

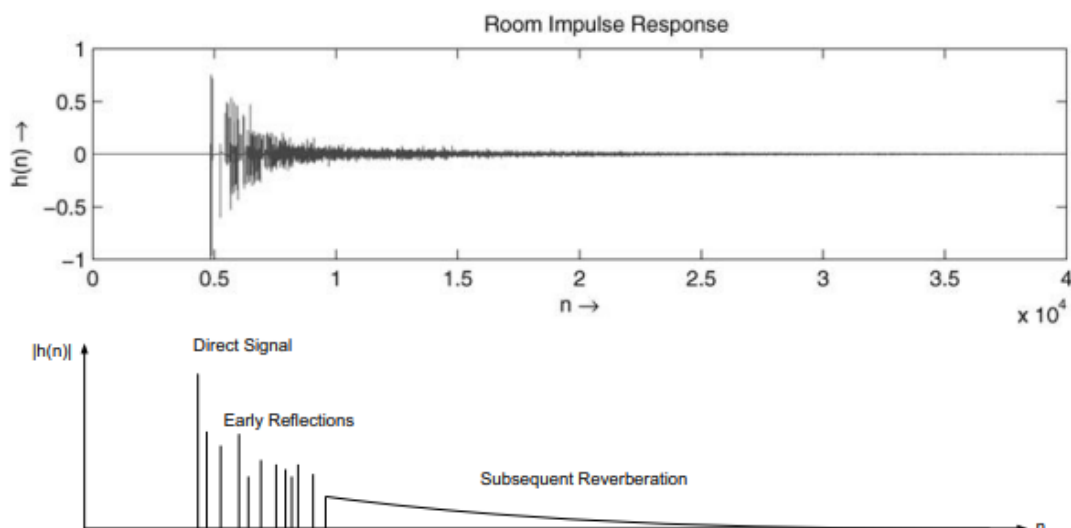
בעולם הדיגיטלי, ניתן לעשות שימוש בשיטות עיבוד על מנת לדמות תופעות אשר קורות מהחזרים שונים שלא קיימים בסיגנל המקורי, ה"יבש", ביניהם לדוגמא נמצא REVERB (הדהוד). רוורב הוא אפקט אשר מדמה את החזרי גלי הקול ע"פ אופי החדר אותו מנסים לדמות. בפרק זה אכנס לשיטות העיבוד בהן משתמשים על מנת לייצר את האפקט.

הרעיון עובד באופן הבא: אם נוכל לשערך איך נראית תגובת החדר להלם, ידוע לנו כי כעת נוכל לקחת כל סיגנל, להעביר אותו קונבולוציה עם התגובה להלם של החדר ובכך נקבל את הסיגנל המקורי "מעובה" באפקט ההדהוד של החדר. אראה אלגוריתמים שונים המתעסקים בסימולציות חדר, כמו MOORER, SCHROEDER. בעזרת אלגוריתמים אלו נוכל להגיע לאפקט החדר הרצוי.

### אקוסטיקת חדר:

ניתן להרחיב המון על התפשטות של גלים בחדר כתלות בגיאומטריית החדר, מקדמי החזר ובליעה של חומרים בחדר ואלמנטים נוספים. אני אסתפק בלהציג את התהליך של התפשטות גל לחץ אוויר בחדר באופן פשוט על מנת לפתח את השפה המשותפת עליה ייבנה האלגוריתם הדיגיטלי שממדל את האפקט.

תגובת ההלם של נק' בחדר יכולה להיות ממודלת באופן הבא: האות המקורי, מלווה בהחזרים "מידיים" יותר (מהקירות וכדומה), כאשר מספר ההחזרים הולך וגדל עם הזמן, סכום ההחזרים מתווסף לכדי אות עם דעיכה אקספוננציאלית הנקרא "רוורב עקיבה" (subsequent reverb).



נגדיר את "זמן ההדהוד" (reverberation time) להיות הזמן בו לוקח לגלי לחץ האוויר לרדת ב-60 dB.

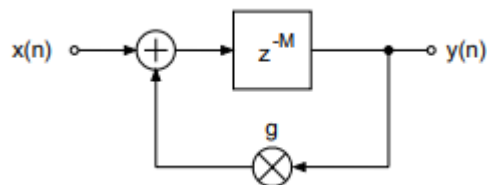
ניתן להבחין בשני היבטים חשובים:

- האחד, צפיפות ההחזרים, ההולכת וגדלה עם הזמן. ככל שהזמן עובר, יותר ויותר החזרים יגיעו מהנתיבים הפחות "ישירים" בחדר וייצרו צפיפות גדולה יותר של גלים חוזרים.
- השני, המרווחים בין ההחזרים, שלא מפולגים בצורה אחידה אלא בצורה יותר "אקראית". הדבר נובע גם מההבדל בין גלים בעלי נתיב ישיר להחזר, לבין גלים בעלי נתיבים יותר מורכבים.

שתי אבחנות אלו יעזרו לנו לבנות את האלגוריתם שיממש את הדימוי של החזרי החדר, אלגוריתם Schroder.

## 6. אלגוריתם Schroeder - מימוש:

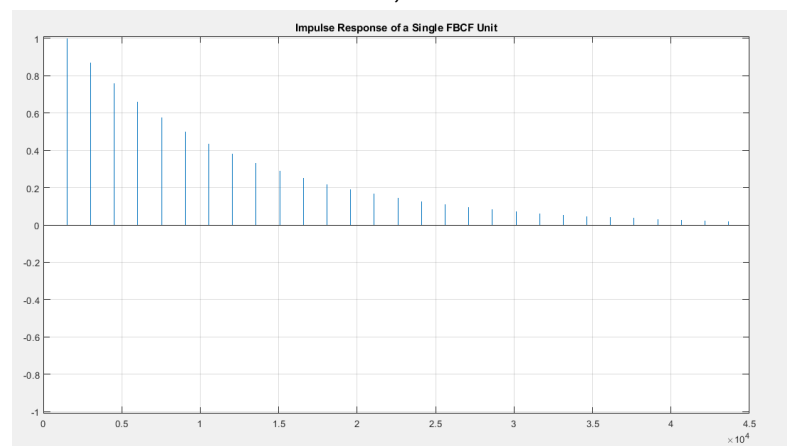
אלגוריתם Schroeder ממדל את התופעות האקוסטיות שהובאו בפרק הקודם. ליבת האלגוריתם – הינו פילטר רקורסיבי בעל הגבר משוב קטן מאחד, כך שתיווצר "סדרה" שדועכת כתלות בהגבר. לפילטר יש יחידת השהייה בנתיב הלוך (Forward Path) כך שגם המשוב המונחת וגם האות הנקי שמסתכם איתו, יעברו השהייה של M דגימות:



יחידה אחת כזו תמדל סדרה של החזרים בדעיכה אקספוננציאלית ( $g < 1$ ). לפילטר בודד ישנה חתימת הפונקציה:

```
function [y, buff] = fbcomb(x, buff, n, d, G_linear)
```

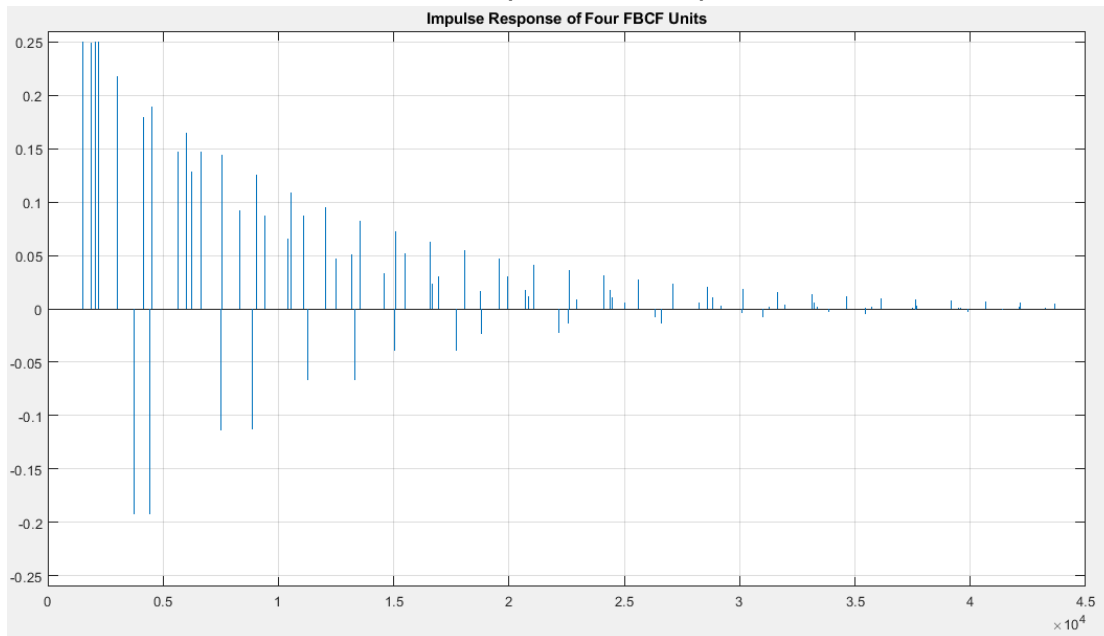
כאשר, על מנת לעזור לבנות את המערכת השלמה, כל יחידה כזו מקבלת גם את הבאפר שהיא תמלא, גם את הדגימה שהיא מעבדת וגם את ההשהייה בזמן d. בעבור פרמטרי התחלה מסויימים, תגובת ההלם של יחידה בודדת כזו תיראה כך:



ניתן לראות כי ההחזרים דועכים כמו שציפינו, אך על מנת למדל באופן מלא את החדר, נצטרך לטפל בפילוג ההחזרים, שלא יהיה אחיד. לצורך כך, נבנה יחידה המורכבת מארבעה פילטרים כאלו, בעלי זמני  $d$  שונים והגברים שונים, כך שהפילוג בין מופעים של גלים חוזרים יקבל אלמנט יותר "אקראי" (לא באמת אקראי במלו מובן המילה, אבל מפולג באופן יותר טוב מהפילוג האחיד של יחידה בודדת).

```
[w1,buffer1] = fbcomb(x(n,1),buffer1,n,d1,g1);
[w2,buffer2] = fbcomb(x(n,1),buffer2,n,d2,g2);
[w3,buffer3] = fbcomb(x(n,1),buffer3,n,d3,g3);
[w4,buffer4] = fbcomb(x(n,1),buffer4,n,d4,g4);
combPar = 0.25*(w1 + w2 + w3 + w4);
```

תגובת ההלם של ארבע יחידות מקבילות תיראה כך:

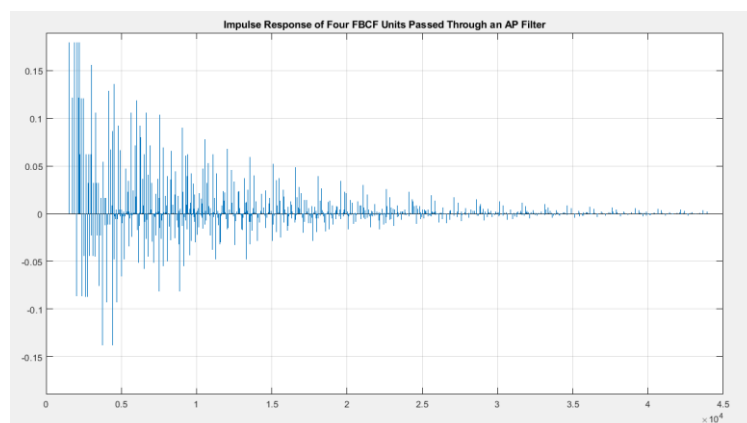


וקיבלנו תגובה מפולגת באופן קצת יותר מוצלח מהתגובה הקודמת.

עוד טכניקת עיבוד שתמומש כאן, היא מעבר ביחידת All Pass של האות המורכב מארבעת יחידות העיבוד שמומשו עד כה. המעבר במסנן AP "מעבה" את צפיפות ההחזרים, מה שעוזר לנו לייצר אפקט השהייה פסאדו אקראי נוסף. חתימת AP פילטר בתכנית:

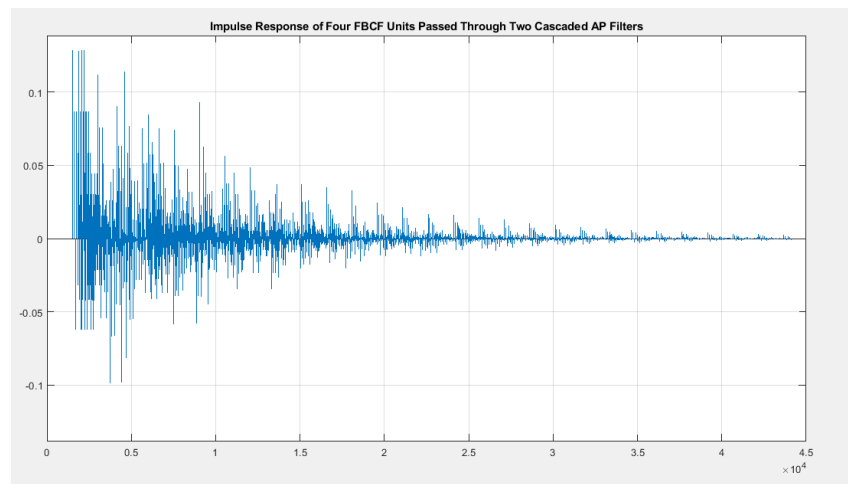
```
function [y,buffer] = apfilt(x,buffer,n,d,G_linear)
```

מעבר של מוצא היחידה המורכבת מארבעה FBCF ב AP פילטר:

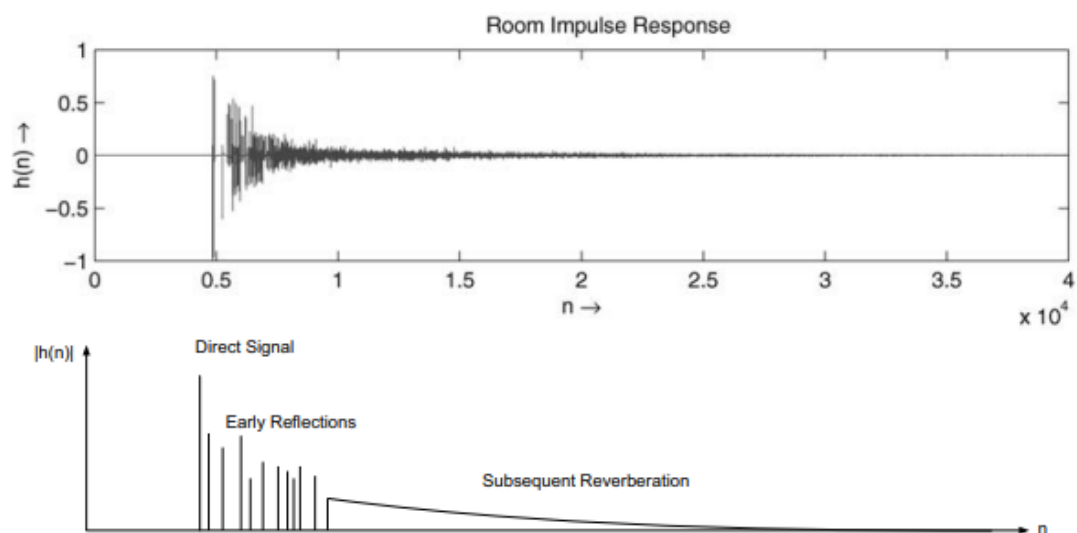




לאחר מעבר בAP נוסף (שני):



ולמעשה הצלחנו למדל באופן יחסית טוב את הגרף שתיאר את התנהגות החדר, שממנו התחיל הפרק:



- להגשה מצורף קובץ SchroederAlgoAudioFileDemonstration.m המדגים את ביצועי האלגוריתם על קטע דיבור קצר.

## **7. מה נשאר לממש? (פרק שלישי: למידת מכונה)**

בפרק זה, אנסה להעמיק בשיטות השונות לערב למידת מכונה עבור עיבוד אודיו. למידת מכונה תבוא לידי ביטוי בניתוחי ספקטוגרמות, אלגוריתמים שונים לזיהוי פיצ'רים ונושאים מתקדמים נוספים. בפרוייקט זה, אתמקד בשיטות לסיווג סגנון מוסיקלי.

לצורך פרק זה, ישנו data set גדול, המורכב מקטעי אודיו המסווגים לסגנונות השונים, ביניהם, קלאסי, רוק, ג'ז וכו', כמו גם תמונות של ספקטוגרמות התואמות את הסגנונות השונים, מהן יהיה ניתן לנתח את התמונה האופיינית המתקבלת עבור סגנון ספציפי. ישנם גם קבצי EXCEL המכילים תיוגים על מידע אופייני לסוגי המוזיקה השונים (טמפו, תכולות תדר, RMS וכו'). אנסה לייצר מערכת שתגיע לכל הפחות ל60 אחוז דיוק בסיווג בין כלל הסגנונות השונים, ו85 אחוז דיוק בקביעת סיווג בין שני סגנונות אפשריים.

אחלק את הגישה לפרק זה לשני חלקים:

### **למידה סטטיסטית:**

בחלק זה אתעסק בחקר של גישות יותר "קלאסיות" ללמידת מכונה. אחקור את הביצועים של מסווגים מבוססי Discriminant Analysis ועצי החלטה על הדאטא סט הנתון, כאשר המדדים המרכזיים הם ביצועי המסווג על סטים שיוקצו להיות "סט מבחן", ומטריצות הערבול (יורחבו בפירוט התיאוריה והפיתוחים כאשר אממש את הפרק במלואו).

### **למידה עמוקה:**

בחלק זה אשתמש בגישה מבוססת רשתות עצביות ללמידת הדאטא סט, וסיווג הסגנונות והקטעים השונים. גישה זו חוקרת יותר לעומק מערכות מרובות משתנים, ושיטות מורכבות יותר לסיווג וקבלת החלטה של הרשת. המדד העיקרי בחלק זה יהיה ביצועי הרשתות על סטים שיוקצו להיות "סט מבחן".

### **מימושים נוספים:**

עם הצגת הפרוייקט הסופי, ארצה לחזור לשני הפרקים הקודמים (משוונים, סימולציות חדר) ולממש להם ממשק משתמש במטלב. המטרה שלי היא לאפשר

למשתמש גישה לכל הפרמטרים שדנו בהם בפרקים אלו, כך שיוכל להפעיל אותם ולחוות את התוצאה בידים ובאוזניים.

ארצה למטרה זו גם לממש סוג של מערכת שתוכל לבצע עיבוד "בזמן אמת", כמה שניתן (לא מערכת זמן אמת מבוססת מיקרו בקר מלאה, מה שיהיה ניתן לעשות במסגרת האפשרויות של ממשקי משתמש במטלב).

## **9. רשימת מקורות:**

- **Digital Audio Signal Processing**, Udo Zolzer [1]
- **Audio & Speech Processing with Matlab**, Paul R. Hill. [2]
- **Hack Audio, An Introduction to Computer Programming and Digital Signal Processing in MATLAB**, Eric Tarr
- **Theory and Applications of Digital Speech Processing**, Lawrence R Rabiner, Ronal W. Schafer. [3]
- **Introduction to Digital Speech Processing**, Lawrence R. Rabiner, Ronal W. Schafer. [4]
- **Speech and Audio Processing - a MATLAB Based Approach**, Ian Vince McLoughlin [5]
- **All About Audio Equalization**, Vesa Välimäki and Joshua D. Reiss, Applied Science [6]
- **Speech Recognition Using Articula** [7]