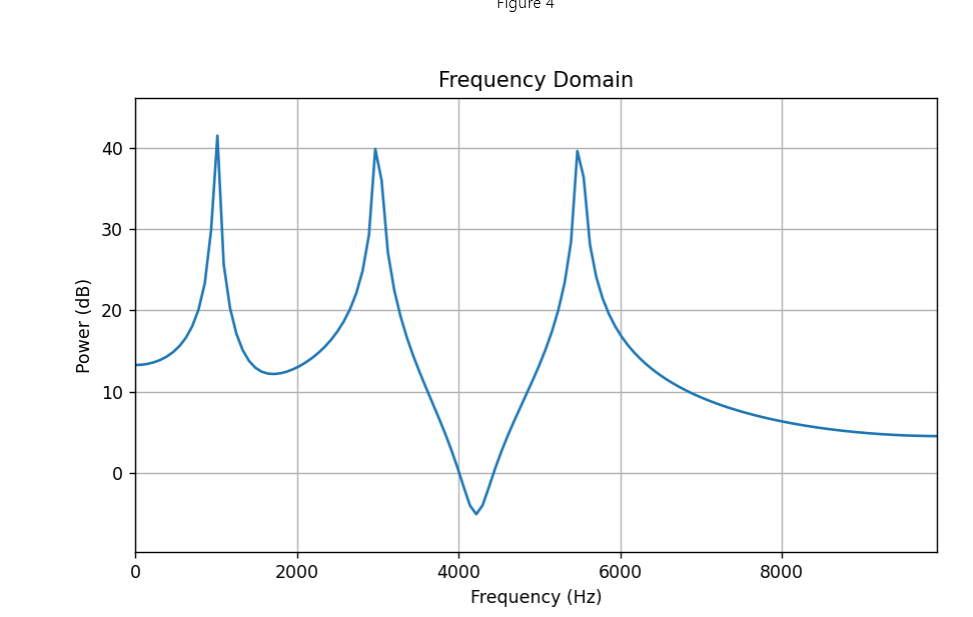
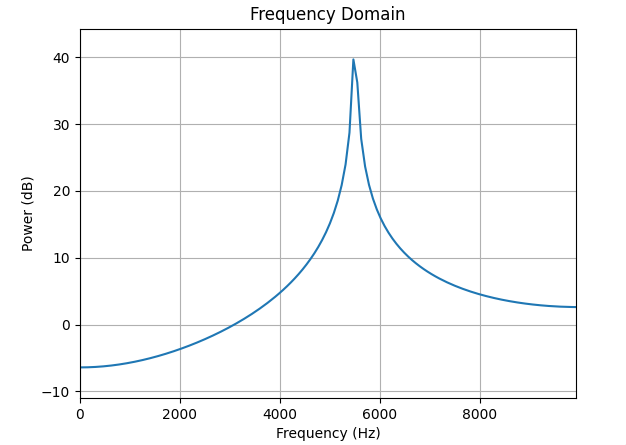
תרגיל בית 2 – למידה עמוקה באותות דיבור

# שאלה 1:

1. בקוד
2. A graph of a frequency

   Description automatically generatedמחשבים log10 על העוצמה של כל תדר. ה ולכן גם גודל הFFT. תדירות הדגימה היא 20,000 אלף הרץ ולכן רזולוציית התדר היא .  
   תמונת התדר שהתקבלה מייצגת את האמפליטודה של הסינוס שמרכיב את התדר, לכל אחד מהתדרים המוצגים. ניתן לראות כי בערכים שהיינו רוצים לקבל – התדרים של הסינוסים שבאמת מרכיבים את הסיגנל, יש קפיצה משמעותית.
3.   
     
   השיאים מתקבלים ב1000, 3000 ו5500 הרץ. כעת התדריים הם בין 0 ל10,000 אך עכשיו יש 256 דגימות (אזי גודל הFFT הוא 256) ולכן, רזולוציית התדר היא תמונת התדר שמתקבלת היא כמה "פיקים" ב1000, 3000 ו5500 הרץ – שאלו אכן התדרים שהיינו אמורים לקבל – אלה התדרים של הסינוסים השונים אותם סכמנו.

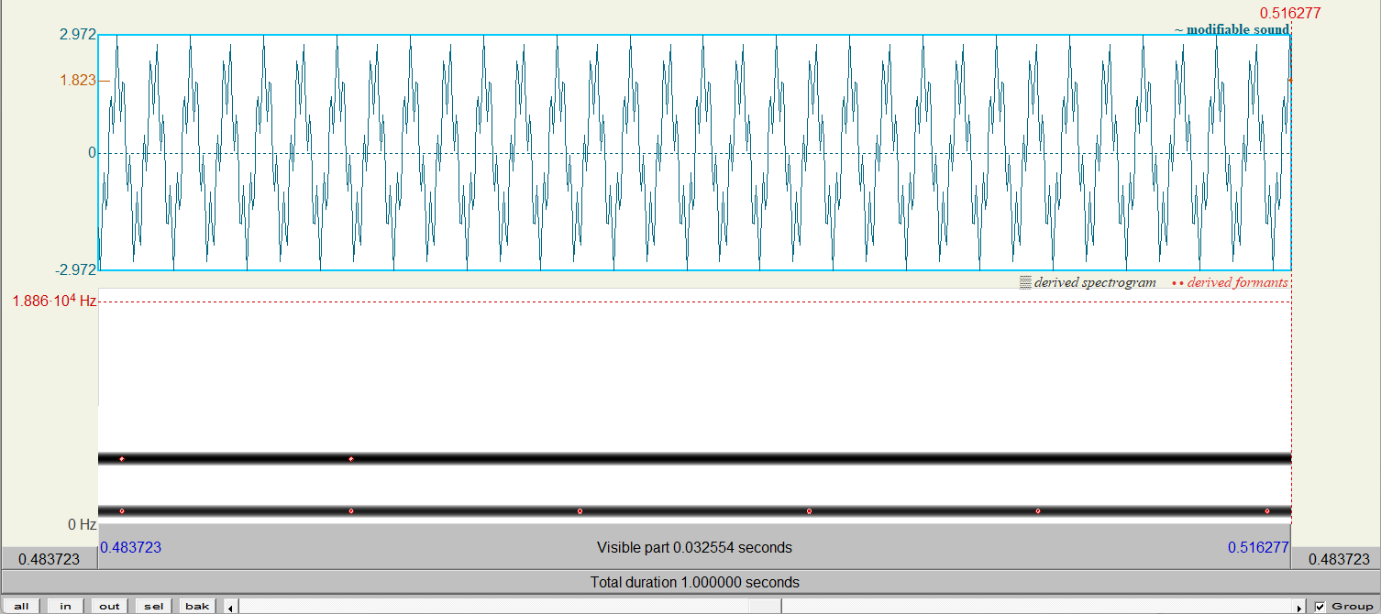


פה ניתן לראות שהשיא ב19K לא קיים (למרות שיש סינוס בתדר הזה) והשיא ב1000 נעלם - כלומר רק מה שב5.5 נשאר.  
האות נדגם ב20 אלף הרץ ולכן, כל דגימה ניתן לתאר על ידי . אם נציב ב

כלומר, האותות מבטלים זה את זה.

באופן סכמתי, מאחר ואנו דוגמים ב20 אלף הרץ ויש תדר בתדירות 19 אלף הרץ שגדולה ב – קורית תופעה של aliasing. כלומר, הסיגנל בתדירות 19 אלף מבטל את הסיגנל בתדר של 1000 הרץ.

# שאלה 2:



# שאלה 3

1. לפי חומר העזר שניתן לתרגיל, תחום התדרים של הpitch של דיבור אנושי נע בתחום
2. מאחר ותדירות היא הופכית לזמן המחזור, משך הזמן האופייני של מחזור pitch הוא בין  
    .
3. האות נדגם בתדירות של ויש p דגימות בכל מחזור של pitch ולכן, הזמן שייקח לכל מחזור הוא ומכאן התדירות היא .
4. כפי שהסברנו לעיל, זמן המחזור הוא ולכן התדר המתקבל הוא ואכן ולכן הוא בתחום.

# שאלה 4

ראשית, נחשב את הפורייה כמו שעשינו בתרגיל בית הראשון

# שאלה 5

## א:

## ב:

1. האוטוקורלציה היא מכפלה פנימית של וקטור עם עצמו (עם הזחה של k וריפוד עם אפסים בצדדים). עבור k=0, זוהי פשוט מכפלה פנימית בין הוקטור לעצמו ואנחנו יודעים שהמקסימום האפשרי של מכפלה פנימית של וקטור עם וקטור אחר היא עם עצמו. המכפלה הפנימית עם עצמו היא הנורמה בריבוע וזו אנרגיית האות.
2. אם האות מחזורי במחזור של P, אז בהזחה של כאשר המכפלה שבסכום היא מכפלה של איבר בעצמו. אם נסתכל על הזחה קטנה באחד או גדולה באחד מ, נקבל שכמות האיברים היא אותה כמות אך במקום מכפלה פנימית עם עצמו, מכפילים את הוקטור עם וקטור אחר.
3. כאשר אנו מתרחקים מ, ככה יש יותר ריפוד של אפסים שמאפס את המכפלה וכך עבור כל מחזור שנתרחק מ0 נאבד אנרגיה של מחזור שלם.

# שאלה 6

## א

A graph with a red dot

Description automatically generatedנסתכל על שתי נקודות בגרף אשר נוגעות בgrid שקל לחשב בעזרתן את זמן המחזור.  
נשים לב כי ההפרש בין שתי הנקודות ה זמן מחזור. ומבחינת זמן הוא שווה ל0.010 sec.

ולכן נציב במשוואה ונפתור. . מכאן התדר הוא – כמו שציפינו.

## ב

A screen shot of a graph

Description automatically generated

*לפי התמונה ניתן לראות (באמצע למעלה) שהתדר של הpitch הוא 125.5 הרץ. זה לא בדיוק אותו מספר אך מאוד קרוב וניתן להסביר זאת בגלל חוסר דיוק שלנו בסימון הpeaks של הגל.*

## ג

A screen shot of a graph

Description automatically generated

סעיפים ד' ה' נמצאים בקוד.

# שאלה 7

## 1

1. בקוד
2. בקוד
3. 2 blue fish swam in the tank

## 2

1. A graph of a sound wave

   Description automatically generated

## A diagram of a sound wave Description automatically generatedA colorful graph of sound waves Description automatically generated with medium confidence3

Voiced, are parts of the speech where the vocal cords are used, it can be seen by noting a pitch in the wave form, unlike areas where the is no prominent pitch, but just a “broad band” noise like sound, it is when the vocal cords aren’t used.

## 4

1. בקוד
2. אינדקס המסגרת הוא 34 וההבהרה היא A מתוך המילה swam.
3. אינדקס המסגרת הוא 2 וההבהרה היא T מתוך המילה two.

# שאלה 8

## 1

1. A graph showing a number of blue lines

   Description automatically generated  
   ניתן לראות את הפיקים של רכבת ההלמים בערך ב40 וב115 ולכן, אורך המחזור בדגימות הוא 75.  
   האות נדגם בתדירות של ולכן, כמו שעשינו בתחילת התרגיל, התדר של הpitch הוא ובשניות הוא .
2. A graph showing a waveform

   Description automatically generated with medium confidenceיש 256 דגימות ומכאן גודל הFFT הוא 256. תדירות הדגימה היא 8000 ולכן רזולוציית התדר היא .   
   ניתן לראות כמה פורמנטים שונים (העוצמה שלהם גבוהה יותר משאר התדרים) והם בערך 105, 205, 515. אנחנו יודעים שהpitch הוא הפורמנט הקטן ביותר ולכן, מהתמונה אנחנו יכולים להעריך שהpitch הוא 105 וזה אכן תואם את הסעיף הקודם.
3. ניזכר כי ההרמוניה הK היא תהיה בעלת תדירות K כפול התדר המרכזי, כאשר התדר המרכזי הוא הpitch, ולכן אם נסתכל על גרף ונבדוק עבור כל peak שנראה, אם הוא נמצא בכפולה טבעית של התדר המרכזי, נדע לזהות את ההרמויות. עשינו plot בו סימנו כפולות טבעיות של הpitch ונספור רק את הדוגמאות בהן הפסגה מסתדרת עם הקו האדום שסימנו.   
   A graph with red and blue lines

   Description automatically generatedהצלחנו לספור 13 הרמוניות.
4. המעטפת מסומנת באדום A graph with a red line

   Description automatically generated

## 2

1. A graph with blue lines

   Description automatically generatedניתן לראות כי המחזור של הpitch הוא אחרי 76 דגימות (מאחר והפיקים הקטנים יותר הם הרמוניות ולא הpitch) ולכן, הכמות זמן למחזור היא שניות והתדר הוא .
2. בעזרת הפונקציה pitch\_detect\_corr, מצאנו כי הכמות דגימות במחזור של הpitch הוא 77 דגימות והתדר הוא 103.9 הרץ.

## 3

1. לא, לאחר הסתכלות בהתמרת הפוריה, לא ניתן לראות מחזור pitch, האות נראה רועש וההפרש בין הpeaks אינו קבוע ואין שום חוקיות שניתנת לזיהוי בעין.
2. ההבדל העיקרי שניתן לשים לב אליו בהסתכלות על שני הגרפים, והוא שבגרף הקולי, פונקציית האוטוקורלציה נראית "חלקה", לעומת הפונקציה של הגרף הלא קולי, שנראית מחוספסת ומלאה ברעש, בלי שום הרמוניות או קפיצות קבועות.

# שאלה 9

## 1

1. מומש בקוד
2. ניתן לראות שכמות הzero crossings משמעותית יותר גדולה בעבור הרעש מאשר הסיגנל של הסינוס.

## 2

מומש

## 3

אכן קיבלנו שהפונקציה חזתה שהvoiced הוא voiced והunvoiced הוא unvoiced.

## 4

1. A screen shot of a computer program

   Description automatically generated

ניתן לראות שהאנרגיה הממוצעת של הvoiced היא יותר גדולה.

1. A screenshot of a computer program

   Description automatically generated

קצב חציית ה0 של הunvoiced יותר גדולה.

1. A computer screen shot of a black screen

   Description automatically generatedהפונקציה אכן סיווגה נכון את הדוגמאות.