



M Ű E G Y E T E M 1 7 8 2

## DIPLOMATERVEZÉSI FELADAT

**Gera Nándor (HNEDSH)**

Villamosmérnök hallgató részére

### Irányított karakterisztika kialakítása mikrofontömb segítségével

Több esetben is felmerülhet az igény olyan hangérzékelő rendszerekre, amelyben egy térrész adott pontjára lehet irányítani a felvételt készítő mikrofont. Előfordulhat ilyen eset például stúdiófelvételek, konferenciabeszélgetések, különböző multimédiás alkalmazások esetén, amikor több szereplő közül csupán egy valakinek a beszédére kell koncentrálni, de az is egy előnyös megoldás lehet, ha az irányított mikrofon segítségével valamilyen zajforrást lehet szelektíven kizárni a felvételből.

Irányított mikrofonok készen is kaphatók, ezek iránykarakterisztikái általában fixek, így a kiválasztott hangforrásra való irányítást fizikai mozzgatással kell megoldani. Egy másik módszer az úgynevezett irányított mikrofontömb alkalmazása, amelyben több, általában nem irányított karakterisztikájú mikrofon jelének együttes feldolgozásával lehet irányított karakterisztikát kialakítani különböző jelfeldolgozási algoritmusok segítségével.

A dolgozat célja, hogy megvizsgálja a fent vázolt, irányított mikrofontömb kialakításához szükséges jelfeldolgozási eljárásokat, és a gyakorlatban is tesztelje azokat.

A hallgató feladatának a következőkre kell kiterjednie:

- Állítson össze megfelelő mérési elrendezést, végezze el hangfelvételek gyűjtését különböző körülmények között!
- Ismerjen meg különböző módszereket, amelyek segítségével mikrofontömbök karakterisztikája irányítottá tehető!
- Vizsgálja meg az egyes módszerek tulajdonságait, pontosságát!
- Vizsgálja meg azon körülményeket, amelyek befolyásolják az irányított karakterisztika kialakításának minőségét!

**Tanszéki konzulens:** Dr. Orosz György, docens

Budapest, 2020.02.28.

.....  
.....  
Dr. Dabóczi Tamás  
tanszékvezető  
egyetemi tanár, DSc





M Ű E G Y E T E M 1 7 8 2

**Budapesti Műszaki és Gazdaságtudományi Egyetem**

Villamosmérnöki és Informatikai Kar

Méréstechnikai és Információs Tanszék

# **Irányított karakterisztika kialakítása mikrofontömb segítségével**

*Készítette*

Gera Nándor

*Konzulens*

Dr. Orosz György

2020

# TARTALOMJEGYZÉK

Kivonat .....	7
Abstract.....	8
1. Bevezetés .....	9
1.1. Akusztikus iránykarakterisztika bemutatása .....	10
2. Célok.....	12
2.1. Követelmények.....	13
2.2. Rendszerterv.....	13
3. Jelrekonstrukció okai, zavarhatások.....	15
3.1. Zavarok típusai .....	15
3.1.1. Reflexió .....	17
3.1.2. Interferencia.....	17
4. Mikrofontömb .....	19
4.1. Mikrofontömb iránykarakterisztikája .....	19
4.2. Mikrofontömb elrendezések.....	31
5. Mérések.....	33
6. Nyalábformálás .....	39
6.1. Jelmodell .....	39
6.2. Szűrőmodell .....	41
6.2.1. Keskenysávú szűrő modellje .....	42
6.2.2. Szélessávú szűrő modellje.....	48
6.2.3. Valós bemenet .....	50
7. Eredmények .....	52
7.1. Keskenysávú szűrő .....	52
7.1.1. Modellezett környezet.....	54
7.1.2. Valós környezet .....	59
7.2. Szélessávú szűrő.....	62
7.2.1. Modellezett környezet.....	63
7.2.2. Valós környezet .....	76
8. Összefoglalás .....	84
9. Köszönetnyilvánítás .....	85

10. Irodalomjegyzék.....86

## HALLGATÓI NYILATKOZAT

Alulírott Gera Nándor, szigorló hallgató kijelentem, hogy ezt a szakdolgozatot meg nem engedett segítség nélkül, saját magam készítettem, csak a megadott forrásokat (szakirodalom, eszközök stb.) használtam fel. Minden olyan részt, melyet szó szerint, vagy azonos értelemben, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Hozzájárulok, hogy a jelen munkám alapadatait (szerző(k), cím, angol és magyar nyelvű tartalmi kivonat, készítés éve, konzulens(ek) neve) a BME VIK nyilvánosan hozzáférhető elektronikus formában, a munka teljes szövegét pedig az egyetem belső hálózatán keresztül (vagy hitelesített felhasználók számára) közzétegye. Kijelentem, hogy a benyújtott munka és annak elektronikus verziója megegyezik. Dékáni engedéllyel titkosított diplomatervek esetén a dolgozat szövege csak 3 év eltelte után válik hozzáférhetővé.

Kelt: Budapest, 2020. 12. 18.

.....  
Gera Nándor

## **Kivonat**

Egyes alkalmazásokban fontos lehet, hogy egy adott akusztikus hangforrás jelét minél pontosabban, tisztábban tudjuk felvenni. Ilyen alkalmazás lehet, amikor egy autó, lakás, vagy egyéb eszköz funkcióit hang segítségével szeretnénk vezérelni. Másik gyakori lehetőség, amikor egy konferenciáról, előadásról akarunk felvételt készíteni, és szükséges, hogy megfelelő hangminőséget érjünk el.

Diplomatervem során prezentálni fogom, hogyan készítettem egy olyan alkalmazást, amely beltérben egy meghatározott hangforrás jelét vételezi, majd rekonstruálja, ezzel előállítva a becsült eredeti jelet. A beltér tulajdonságai miatt a jó hangminőség elérése érdekében szükség van egy ilyen, csak a forrás jelét átteresztő szűrőre, hiszen a jel-zaj viszony és a jel-interferencia viszony zavaróan alacsony tud lenni. Munkám során, két fő csoportra osztottam ezeket az alkalmazásokat, keskenysávú és szélessávú esetekre, melyeket prezentálni is fogok. Az időtartományban végzett digitális jelfeldolgozás során a használt iránykarakterisztikus szűrők térbeli és/vagy időbeli szűrést hajtanak végre a min-tavételezett jelen.

Beszámolóim során bemutatom a fent felsorolt algoritmusok működését és jellemzőit, az általam végzett méréseket, illetve az elkészült rendszer működését modellezett és valós környezetben is.

## **Abstract**

In some applications might be important to record the sound of an acoustic source as clear as possible. These applications can be found in cars, smart houses and other electronic devices which can be actuated by the sound of a person. Another possible case is when we need to record a conference or a presentation and the quality of the sound is crucial.

In my thesis I will present an application, which samples and reconstructs the sound of a given source, mainly speech, in a room. Because of the attributes of a room, the signal-to-noise and the signal-to-interference ratios are low, so we have to consider some kind of algorithm to make it better, to make the speech of a person more understandable by suppressing the noise. In my thesis I dealt with two different types of these algorithms, these were the narrowband and wideband filters, which I will present here. Generally, these programs work in time domain and can be applied as spatial and/or frequency filters.

In my thesis I will show how these algorithms work, and how they can be implemented, and used in my model and in real environments.



## 1. Bevezetés

Az akusztikus iránykarakterisztikának, mint módszernek számos alkalmazása lehetséges. Sok olyan eset van, amikor egy eszközt, például egy autó multimédiás rendszerét, okos otthont, vagy bármilyen más okos eszközt hanggal szeretnénk vezérelni. Egy másik lehetőség, amikor konferenciákról, előadásokról szeretnénk jó minőségű hangfelvételt készíteni, és esetleg automatikusan feliratozni azt. Ezekben az esetekben az a közös, hogy az éppen beszélő ember szavait fel kell ismerni, ami alapján az adott digitális eszköz vezérelheti valamelyik perifériáját, például egy autóban a központi számítógép utasításra emel a rádió hangerőjén. Továbbá az iránykarakterisztikus szűrő felhasználható olyan környezetben, ahol a beszélő ember - előadó, riporter - szavait elnyomja egy kívülről beszűrődő motor, kompresszor vagy más eszköz hangja, ezáltal akusztikus szűrés nélkül megnehezítve a beszéd felismerését a hallgatók, nézők számára.



1-1. ábra, Hangvezérlés [1]

Mivel a fenti alkalmazások általában beltérek, ezért számos problémával szembesülhetünk a használatuk során megfelelő jelfeldolgozás hiányában. A terem falairól és az egyéb tárgyakra visszaverődő hangok nehezítik az eredeti jel rekonstrukcióját. A visszavert hanghullámok elrontják a jel tisztaságát, kevésbé lesz érthető a vett hang. Ezen kívül, beltérekben gyakran előfordulnak különböző zajokat kibocsátó készülékek, például klímaberendezések, számítógépek, szellőztetők, illetve a város zajai is bejutnak a terembe,

ezáltal rontva a hangminőséget. Ugyanakkor előfordulhat az is, hogy egyszerre több dedikált akusztikus kibocsátó van jelen a vizsgált térben, például egyszerre ketten beszélnek egyidejűleg egy konferencia közben, egymás szavába vágva. Ilyen esetekben fontos lehet, hogy valamilyen jelfeldolgozási módszerrel elkülönítsük a dedikált forrás jelét más akusztikus forrásokétól, vagy teljesen elnyomjuk az egyiket, és átengedjük a másikat. A következő alfejezetben bemutatom, hogy ezt milyen módszerek segítségével lehet megoldani.

### **1.1. Akusztikus iránykarakterisztika bemutatása**

Akusztikus jelfeldolgozási feladatok során ritkán adatik meg az a lehetőség, hogy nincsen szükség a vett jel semmilyen szűrésére. Mind kültéri, de még inkább beltéri alkalmazások esetén a felvétel nem csak az általunk kívánt forrás jelét tartalmazza, hanem egyéb zajokat, interferenciát, illetve esetlegesen más források jeleit is. A jelfeldolgozó algoritmus egyik elsődleges feladata, hogy a felvételtől kiszűrje az alkalmazás szempontjából hasznos információt, melyre több különböző lehetőségünk is van.

Az első lehetőség, hogy térben próbáljuk meg megkülönböztetni a mikrofonok által felvett hangot. Igazán hasznos lehet ez azokban a helyzetekben, amikor jól elkülöníthetők a dedikált források egymástól térben, illetve számuk is viszonylag alacsony a térbeli felbontási képességünkhöz képest. A térbeli szűrés felbontásának finomságára egy felső korlátot határoz meg a mikrofonok száma (lásd 4. fejezet), ezért nem tudunk tetszőlegesen kis felbontású szűrőt tervezni a valóságban.

Egy másik lehetőség, amikor időben szűrjük a vett jelet. Ekkor időtartományban hajtunk végre szűrést oly módon, hogy bizonyos frekvenciasávban lévő jeleket átengedünk, bizonyosakat pedig kiszűrünk a vett jeltől. Az előző esettel ellentétben itt nem korlátozza a szűrést a mikrofonok száma, akár egyetlen mikrofon használatával is végezhetjük a jelfeldolgozást, csupán a számítási kapacitásunk nagysága határoz meg limitet.

Az előző kettő megoldástól egy lényegesen különböző, de manapság egyre gyakrabban használt lehetőség az, amikor valamilyen mesterséges intelligenciával dolgozunk. Mivel az emberi agy is előfeldolgozatlan jeleket dolgoz fel (iránymeghatározás, hangfelismerés, stb.), ezért kiindulhatunk abból az ötletből, hogy valamilyen mesterséges intelligenciával végezzük el a jelfeldolgozást. Ahogyan az agy is folyamatos tanulás során sajátítja el ezt a képességét, úgy egy mesterséges intelligencia is megtanítható valamilyen fokig erre. Ennek hátránya az általában jóval bonyolultabb algoritmus, aminek a tanítása, illetve a

tanító minták előállítására tapasztalatot igényel. A tanítás sokszor nehézkes, rengeteg minta gyűjtése szükséges hozzá, amelyre nem mindig van lehetőségünk. Ugyanakkor bizonyos esetekben jobb eredményeket kaphatunk vele, mint a hagyományos, analitikus módszerek használatával.

A fenti három gyakori módszer közül a térbeli és időbeli szűrést alkalmaztam a munkám során. Ennek oka, hogy irodalomkutatásom folyamán arra jutottam, hogy az általam használt esetben analitikus jelfeldolgozás segítségével megfelelő minőségű eredményeket lehet elérni.

A következő fejezetben a félév elején felállított célokról, az általam készített alkalmazással szemben felállított követelményekről, illetve a felhasznált eszközökről írok.

## 2. Célok

Munkám során az elsődleges célom egy dedikált hangforrás jelének minél jobb becslőjének előállítására. A vizsgált térben egyéb források is jelen vannak, melyek lehetnek más, a megfigyelt forráshoz hasonló akusztikus források, különböző zavarok, illetve az interferencia és reflexió jelenségei is megtalálhatóak a térben. A cél, hogy a zavaró hatások teljesítményét a lehető legjobb hatásfokkal elnyomjam anélkül, hogy a dedikált forrás jelét zavaró mértékben torzítanám.

A rendszer lineáris mikrofontömb segítségével veszi fel a hangot. A lineáris mikrofonelrendezés előnye, hogy viszonylag kevés darabszám esetén is megfelelő felvételt lehet vele készíteni, ellentétben a kettő-, vagy háromdimenziós elrendezésekkel (lásd 4. fejezet). Ezáltal az alkalmazás működéséhez szükségem lesz a források irányára a priori ismeretként, amik a mikrofontömb körül, egy körvonal mentén helyezkednek el a modell szerint.

Az alkalmazás működése passzív, azaz nem bocsát ki semmilyen jelet működése közben. Bemenetként kapja meg az elkészült felvételt, a dedikált forrás és a jelentősebb zavarforrás(ok) irányát, továbbá az iránykarakterisztikus szűrő konfigurálásához előre ismert ezek frekvenciaspektruma is. A jelfeldolgozás végeztével a kimenetén megjeleníti a dedikált hangforrás becslőjét, illetve a becslő előállításánál használt szűrő iránykarakterisztikáját. Ezen kívül fontos megjegyezni, hogy az alkalmazás működése során a források iránya nem változik, azok meghatározott irányban vannak a mikrofontömbhöz képest.

Diplomatervem elkészítése során az alábbi célok kerültek megvalósításra:

- mérések végzése különböző elrendezések esetén különös hangsúlyt fektetve
  - a terem méretére, a benne lévő tárgyak szerepére akusztikai szempontból
  - a hangforrások számára a zárt térben
  - a zavarforrások hangjának erősségére
  - a mikrofonok számára a lineáris mikrofontömbben
- Matlab segítségével a szűrő iránykarakterisztikájának megtervezése modellszinten, a hangforrások helyének ismeretében
- A szűrő modelljének tesztelése, és esetleges javítása az elvégzett mérések segítségével végrehajtott tesztek során
- Eredmények kiértékelése, következmények levonása, az alkalmazás használhatóságának, korlátjainak megismerése

## 2.1. Követelmények

A diplomatervem eredményeként létrejövő alkalmazáshoz a következő követelményeket támasztottam a munkám elején:

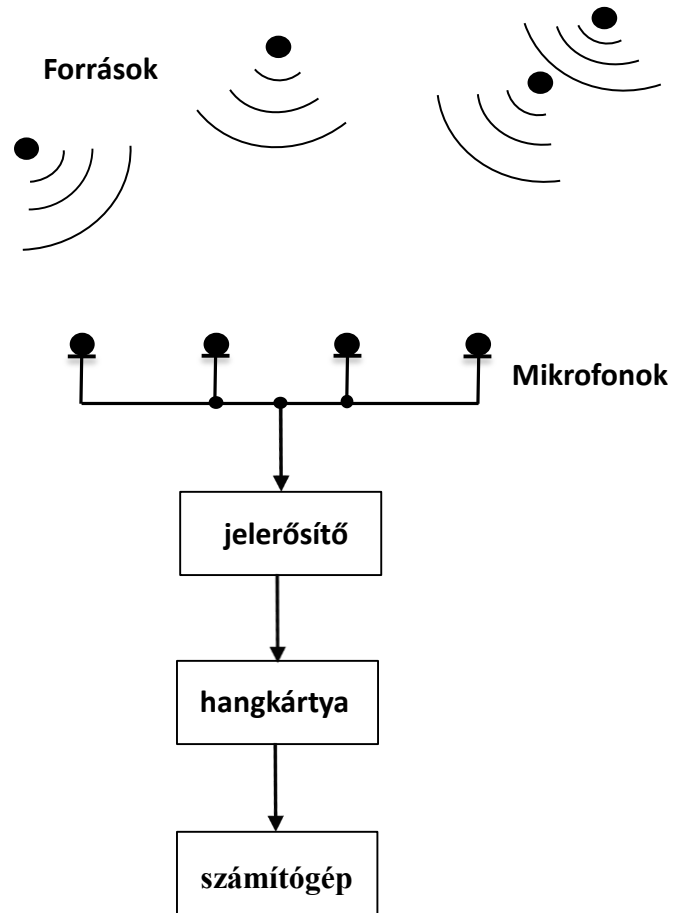
- passzív működés, offline jelfeldolgozás
- lineáris mikrofontömb által vételezett jel feldolgozása
- dedikált jel rekonstrukciója keskenysávú és szélessávú iránykarakterisztikus szűrés segítségével
- dedikált hangforrás jelének rekonstrukciója a vett jelből minél kisebb torzítással
- zavarforrások teljesítményének elnyomása azok irányának és frekvenciaspektrumának ismeretében
- a használt iránykarakterisztikus szűrő megjelenítése alkalmazástól függően kettő- vagy háromdimenziós grafikonon.

## 2.2. Rendszerterv

A munkám során végzett mérésekhez a Budapesti Műszaki és Gazdaságtudományi Egyetem, Méréstechnikai és Információs Tanszéke által számomra elérhető eszközöket tudtam használni. Ez magában foglalta azt is, hogy a mérések elvégzéséhez maximum 8 darab mikrofont használhattam, mivel sem több mikrofon nem állt rendelkezésemre, sem olyan hangkártya, amellyel párhuzamosan ennél több mikrofon jelét egyszerre tudnám mintavételezni. A jelfeldolgozást személyi számítógépen végeztem Matlab segítségével, a bemenetként használt hangfájlok elkészítéséhez pedig az Audacity programot választottam. Az általam használt eszközök listája:

- 8 darab mikrofon (Behringer ECM-8000), ismert iránykarakterisztikával [2]
- OneWay RTO3 8 csatornás mikrofonerősítő
- Cakewalk UA101 külső hangkártya, mely a mikrofonok által vett jelet pontos időbélyeggel, azaz párhuzamosan továbbítja [3]
- személyi számítógép, amelyen a jelfeldolgozó algoritmus és a mérésekhez szükséges hangrögzítő program (Audacity 1.3.13 verzió) [4] futtat
- 2 db Genius SP-HF1250x hangszóró [5]

A fent említett eszközök segítségével az 2.1-es ábrán látható blokkdiagram alapján végeztem el a méréseket.



2-1. ábra, blokkdiagram

### **3. Jelrekonstrukció okai, zavarhatások**

Jelrekonstrukció során a cél az eredeti jel egy minél jobb becslőjének előállítás. Ennek oka, hogy akár zárt, akár nyílt térről van szó, mindig szenved valamilyen torzítást az eredeti jel, illetve jelen van valamilyen típusú zaj is. Ahogyan azt már a korábbi fejezetekben említettem, belterekben az egyik legfontosabb befolyásoló tényezők a reflexió és az interferencia. Ha a hangforrás teljesítménye megfelelő jel-zaj viszonyt biztosít, akkor az általános alapzaj értéke nem zavaróan magas, ugyanakkor szűrése továbbra is szükséges, hiszen javítja a jel rekonstrukciójának minőségét. Ezen kívül vannak egyéb, viszonylag nagy teljesítményű zavarforrások is, például szellőztető rendszerek, klímaberendezések hangja működés közben, vagy kívülről, az utcáról beszűrődő, közlekedésből származó zajok. A következő alfejezetekben a különböző zavarokat, torzításokat fogom részletesen bemutatni, melyek megismerése nélkülözhetetlen az iránykarakterisztikus szűrő megtervezéséhez. Ezen jelenségekkel az egyetemen végzett korábbi munkáim [6] során sikerült viszonylag mélyen megismerkednem, ami segítségemre volt a jelenlegi munkám kezdeti szakaszában.

#### **3.1. Zavarok típusai**

Akusztikus zavarforrások csoportosítására számos lehetőség áll rendelkezésre mind teljesítmény, frekvencia, irány, időtartam, periodicitás, lecsengés, stb. szempontok alapján. Diplomatervem során a csoportosításom fő szempontjai a szűrő tervezésében voltak segítségemre, illetve a későbbiekben működésének tesztelésében is hasznosnak bizonyultak.

Az első csoportosítás a dedikált hangforrástól való függés szempontjából készült. A dedikált hangforrástól függő zajokról beszélhetünk, amikor az emberi beszéd a hangforrásnak tekinthető szájtól nem csak egyenes úton jut el a mikrofonokig, hanem más irányokból is (lásd 3.1.1.-es fejezet). Ezen kívül ide sorolandó a jelterjedés során bekövetkező teljesítménycsökkenés is, aminek hatása nem számottevő kisméretű termekben, viszont jelentős tud lenni például előadótermekben. Minden más előforduló zavarforrás a vizsgált hangforrástól függetlenül fejt ki hatását csoportosításom szerint. Ide sorolhatók többek között a dedikált forráson kívüli egyéb, tulajdonságaiban hasonló források, például amikor nem csak az általunk vizsgált ember beszél, hanem vele egyidőben valaki más is. Az ilyen jellegű jelenségek interferenciát okoznak, melynek szűrése megfelelő iránykarakte-

risztikával rendkívül fontos, hiszen sokszor nagyjából azonos teljesítményszinttel rendelkeznek ezek a források, emiatt nem hagyatkozhatunk a jel-zaj viszony megfelelő mértékére. Ezen kívül ebbe a csoportba sorolom a viszonylag rövid ideig tartó, általában kis frekvenciasávban sugárzó zavarokat. Ilyen lehet például a terem padlóján elmozduló szék csikorgása, az elinduló klímaberendezés nagyobb teljesítményű tranziens hangja, vagy éppen egy, az utcán elhaladó hangosabb jármű. Ezen kívül ide tartozik a belső tér, és a benne található objektumok jellegéből adódó állandó alapzaj. Ez nem csak a természetben megtalálható természetes zajt, hanem egyéb, a felvétel során állandóan jelen levő zajokat is tartalmazza. Többek közt lehet ez egy háttérben zúgó számítógép, vagy egy zárt terem állandóan működő szellőztető rendszere.

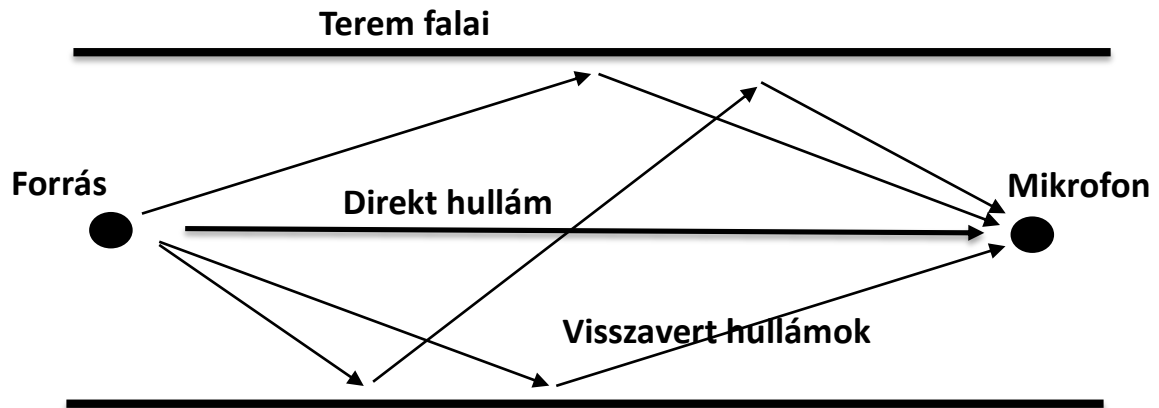
Egy másik általam használt felbontás, amikor periodicitás és frekvenciatartomány szerint csoportosítom a zavarokat. Könnyen belátható, hogy szűrés szempontjából minél keskenyebb frekvenciatartományban, illetve minél inkább periodikus hangot bocsát ki egy zavarforrást, annál könnyebb annak szűrése a vett jelből, tehát annál jobb minőségben tudjuk rekonstruálni a kívánt beszédet. Ezen szempont alapján 4 különböző csoportot hoztam létre munkám során. Az első, amikor egyetlen, tiszta szinuszos zavar van jelen a vett jelben. A második, amikor egyszerre több, különböző frekvenciájú szinusz is jelen van, például amikor egy ideális hangszert szólaltatunk meg. Ebben az esetben tehát még nem foglalkozok nem ideális hatásokkal, a zavarjel csak a dedikált frekvenciákon rendelkezik teljesítménnyel. A következő eset az előzőnek a valóságbéli megfelelője, például egy háttérben hallatszódó elektromos csavarhúzó. Ebben az esetben az elektromos motor, illetve a forgó alkatrészek a viszonylagosan állandó fordulatszáma miatt néhány frekvencián jelentős teljesítmény van jelen, ugyanakkor széles frekvenciatartományban található valamekkora zavarteljesítmény. Az utolsó eset, amikor például egy másik ember is beszél a vizsgált személy mellett. Az emberi beszéd széles frekvenciatartományban tartalmaz meghatározó teljesítményt, ezért ennek szűrése a korábbi zavarásokhoz képest jóval nehezebb.

Mivel a dedikált forráshoz kötődő zavarjelenség főleg a reflexióhoz kapcsolódik, ezért erről bővebben értekezek a következő alfejezetben. Arra vonatkozóan, hogy ezeket a jelenséget hogyan, és milyen elrendezés esetén lehet kompenzálni, a 6. fejezetben mutatom be [7].



### 3.1.1. Reflexió

Mivel a vizsgált tér, ahol a felvételek készülnek, illetve ahol a legtöbb akusztikus alkalmazás működik, egy zárt tér, ezért gondolni kell arra, hogy a terem falain, illetve a benne található objektumokon reflektálódhat, azaz visszaverődhet a hang. Ennek egyszerű modellje a 3-1.-es ábrán látható.



3-1. ábra Hanghullám reflexiója

Mivel a visszavert hullám teljesítménye nem csökken nagyságrendileg a direkt hulláméhoz képest, ezért nem csak a forrás irányából érkezik számottevő teljesítmény a mikrofonokhoz. Miután a reflektált hullám nagyobb utat jár be a mikrofonokig, mint a direkt hullám, ezért időben késik ahhoz képest. Szélsőséges esetben akár visszhangos is lehet a zárt tér, ekkor olyan nagy a jelütkülönbség a direkt és a visszavert hullám megtett útja között, hogy a hangforrás, általában az ember által kimondott szó időtartama egy nagyságrendbe esik, vagy kisebb, mint a direkt hullám és a visszavert hullám mikrofonokhoz való érkezésének időkülönbsége.

A terem falain reflektálódó hullámok, illetve egyéb források zavarjelei interferenciát okozhatnak a mikrofonok közelében, melyet a következő alfejezetben fejtek ki részletesen.

### 3.1.2. Interferencia

Az interferencia jelensége előfordulhat általános zavarforrások, illetve reflexió miatt is. Interferencia során az azonos frekvenciájú, de különböző forrású és/vagy különböző terjedési irányú hangok a tér egy pontjában azonos időpontban vannak jelen. Ekkor a szuperpozíció elvének megfelelően az egyes jelek összetartozó frekvenciakomponenseinek összege fogja megadni a tér adott pontjában a jelerősséget. Tulajdonképpen minden egyes zavarforrás, amely legalább részben megegyező frekvenciatartományban sugároz, mint a

dedikált forrás, interferencia jelenséget okoz a mikrofonok közelében. A belső tér falai csak felerősítik ezt a jelenséget, hiszen a jelnek általában nem kell nagyságrendekkel nagyobb utat bejárni a terem falain reflektálódva a mikrofonokig, mint a direkt úton. A jelenség hatását leginkább térbeli szűrés segítségével tudjuk minimalizálni, hiszen ha elég sok pontra felbontjuk a teret, akkor viszonylag kicsi az esélye, hogy a tér egyazon pontjában legyen a dedikált forrás és a zavarforrás is, illetőleg a reflektált hullám is ritkán érkezik a direkt hullámmal azonos irányból.

Az interferencia és a reflexió együttes jelenlétének egy kiváló példája az, amikor egy szobában egy hangszóró segítségével adott frekvenciájú szinuszos hangot szólaltatunk meg. Ekkor a szoba méretétől függően csupán néhány 10 cm-es távolság alatt látványosan hallható mértékben megváltozhat a jel erőssége a szoba egyes pontjaiban. Ezen egyszerű kísérlet elvégzése is bizonyítja, hogy mennyire fontos a jel rekonstrukciója egy olyan komplex jel esetén, mint az emberi beszéd, hogy visszakapjuk az eredeti jelminőséget zajos esetekben is.

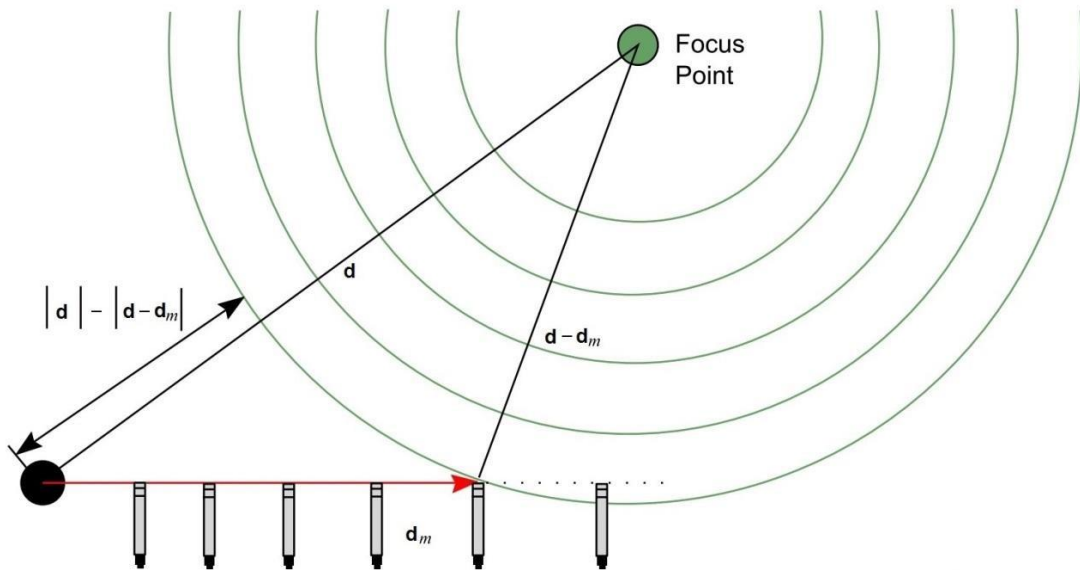
## 4. Mikrofontömb

A felvételek elkészítéséhez használt mikrofontömb típusának választásakor számos tételt kell szem előtt tartanunk. Elsősorban definiálnunk kell, hogy hány dimenzióban szeretnénk végezni a kísérleteket, méréseket, tehát a későbbiekben az elkészült alkalmazásunk hány dimenzióban legyen képes iránykarakterisztikus szűrést megvalósítani. Ahány dimenzióban szeretnénk szűrést létrehozni, legalább annyi dimenzióban kell elhelyezkednie a mikrofontömbnek. Munkám során amellettt döntöttem, hogy csupán irányokat határozok meg, nem foglalkozok a tér mélységével, illetve azzal, hogy az adott irányban milyen magasan helyezkednek el a források, zavarok. Ennek oka, hogy alkalmazásom zárt, rendszerint viszonylag kicsi, közepes méretű beltérekhez készült (például kis előadóterem, lakásban egy szoba, autók beltére, stb.), ezért elhanyagolom annak lehetőségét, hogy a mikrofontömbhöz képest egy irányban egymás mögött, vagy egymás felett több forrás helyezkedik el. A következő fontos kényszer az, hogy hány mikrofont tudunk felhasználni, és mekkora azok mérete, milyen sűrűn, vagy adott esetben milyen távol tudjuk őket egymástól elhelyezni a térben. Diplomatervem során a Budapesti Műszaki és Gazdaságtudományi Egyetem Méréstechnikai és Információs Rendszerek Tanszék laborjában található eszközöket volt lehetőségem felhasználni. A 2.2-es fejezetben ismertettem a felhasznált hardverek között a mikrofonokat is, amelyekből maximum 8 darab felhasználására volt lehetőségem munkám során.

### 4.1. Mikrofontömb iránykarakterisztikája

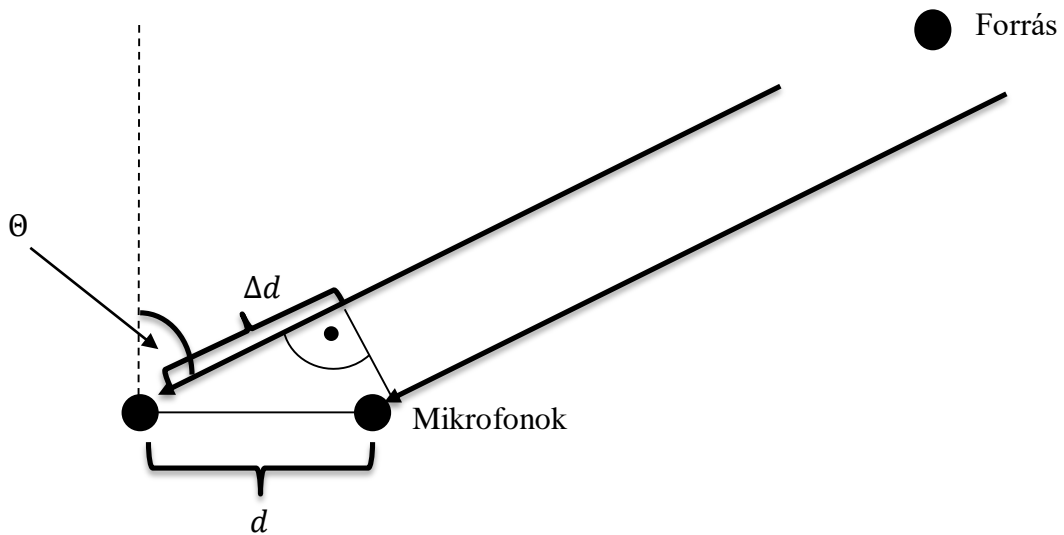
Mielőtt bármiféle iránykarakterisztika formálásába kezdtem volna bele munkám során, előbb szükséges volt megismerkednem a mikrofontömb tulajdonságaiból adódó saját iránykarakterisztikájával. Ennek megismeréséhez a kezdetekben segítségemre voltak többek közt Peter Delos, Bob Broughton és Jon Kraft két cikkei [8] [9], melyekben bemutatták, hogy milyen fizikai tulajdonságoktól függ egy egyszerű vonalbeli elrendezés során a mikrofontömb akusztikus átvitele.

A hang a forrásától számítva koncentrikus körök/gömbök mentén terjed a tér minden irányába (lásd 4-1. ábra).



4-1. ábra, hang koncentrikus körök menti terjedése [10]

Ha elegendően távolról vizsgáljuk a hangforrást, akkor a mikrofonok közelében a koncentrikus gömbök sugara nagyságrenddel nagyobb lesz a mikrofontömbünk fizikai méreteinél, ezért a hanghullámot síkhullámmá egyszerűsíthetjük, ezt láthatjuk sztereó mikrofonok esetén a következő ábrán.



4-2. ábra síkhullám terjedése

Ahogy az ábrán látható, a forrás nem a mikrofonok előtt, hanem attól  $\theta$  szögben helyezkedik el. Ha célunk, hogy a mikrofonok által vett jeleket azonos fázisba forgatva összegezzük, akkor a következő műveletet kell elvégeznünk.

$$\sin \theta = \frac{\Delta d}{d} \quad (4.1)$$

$$\Delta \varphi = 2 * \pi * \frac{\Delta d}{\lambda} \quad (4.2)$$

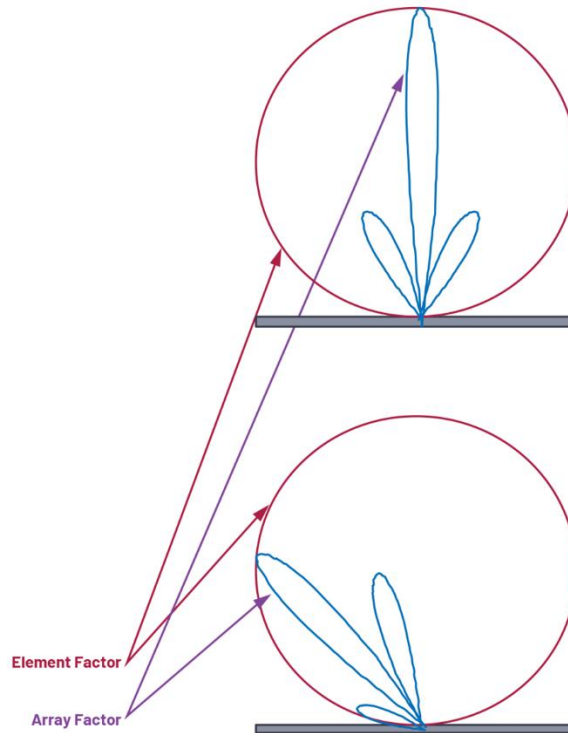
$$\Delta \varphi = 2 * \pi * \frac{d}{\lambda} * \sin \theta \quad (4.3)$$

ahol:

$\lambda$  – a vizsgált frekvencia hullámhossza

A (4.3)-as egyenlet szerint tehát ahhoz, hogy azonos fázisba forgassuk a mikrofonok által vett jeleket  $\Delta \varphi$  fázisforgatást kell elvégeznünk egy  $\lambda$  hullámhosszú jelen. Nagyobb mikrofonszám esetén is egyszerű dolgunk van, hiszen ha a mikrofonokat egyenlő távolságra helyezzük el egymástól, akkor a szomszédosok között a jelútkülönbség, azaz a  $\Delta d$  távolság mindig azonos lesz.

Amikor arról beszélünk, hogy mi az iránykarakterisztikája egy jelen esetben lineáris, egyenletesen elrendezett mikrofontömbnek, akkor két faktort kell figyelembe vennünk. Egyrészt az egyes mikrofonok iránykarakterisztikáját, másrészt pedig a mikrofontömb, mint tömb hatását. Az eredő iránykarakterisztikát az egyes mikrofonok és a mikrofontömb karakterisztikájának összegzésével kaphatjuk meg. A munkám során használt mikrofonok karakterisztikája „omnidirectional”, azaz minden irányból egységes átvitele van az eszköznek [2].



4-3. ábra, egyes mikrofon, illetve lineáris mikrofontömb iránykarakterisztikája

A teljes átvitel tehát leírható a következő egyenlettel (dB-ben):

$$G = G_E + G_A \quad (4.4)$$

$G_E$  jelenti egy darab mikrofon átvitelét, ami mivel minden irányból egységes, ezért a teljes tartományon belül zéró, dB-ben számolva.  $G_A$ , a lineáris mikrofontömb átvitele a következőképpen számolható [8], adott frekvencián:

$$AF(\Theta) = \frac{\sin(N * \pi * \frac{d}{\lambda} * [\sin \Theta - \sin \Theta_0])}{N * \sin(\pi * \frac{d}{\lambda} * [\sin \Theta - \sin \Theta_0])} \quad (4.5)$$

ahol:

$N$  – mikrofonok száma

$AF(\Theta)$  – mikrofontömb átvitele egy adott irányban

$\Theta_0$  – az az irány, amerre a fejezet korábbi részében meghatározott  $\Delta\varphi$  fázisforgatással tudjuk forgatni a mikrofontömböt

Az átvitelnek az egyenletben szereplő  $\frac{d}{\lambda}$  aránytól való függésének vizsgálatához rendezzük át a (4.3)-as egyenletet:

$$\Theta = \sin^{-1}\left(\frac{\Delta\varphi}{2*\pi} * \frac{\lambda}{d}\right) \quad (4.6)$$

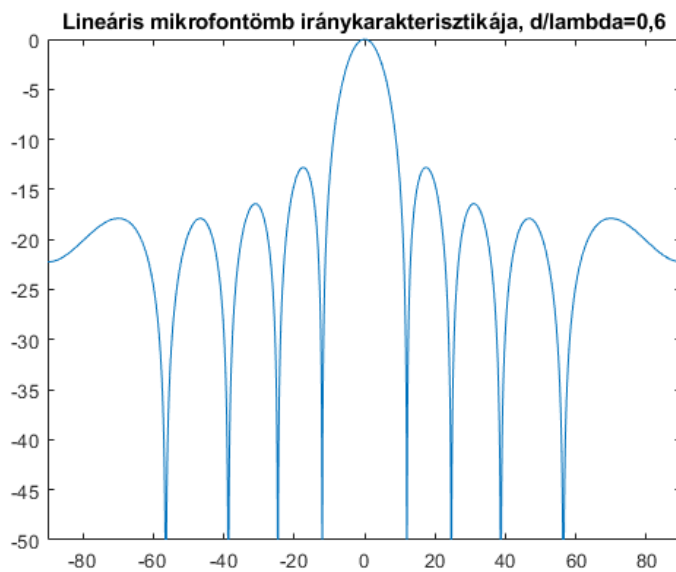
Az *arcsin* függvény periodikus, valamint valós megoldást csak akkor ad, ha -1 és 1 közötti értéket kap argumentumaként. A  $2\pi$ -szerinti periodikusságot, valamint az argumentum határértékeit figyelembe véve az *arcsin* függvény a következő esetekben ad valós megoldást:

$$\left| \frac{m * 2\pi + \Delta\varphi}{2 * \pi} * \frac{\lambda}{d} \right| \leq 1 \quad (4.7)$$

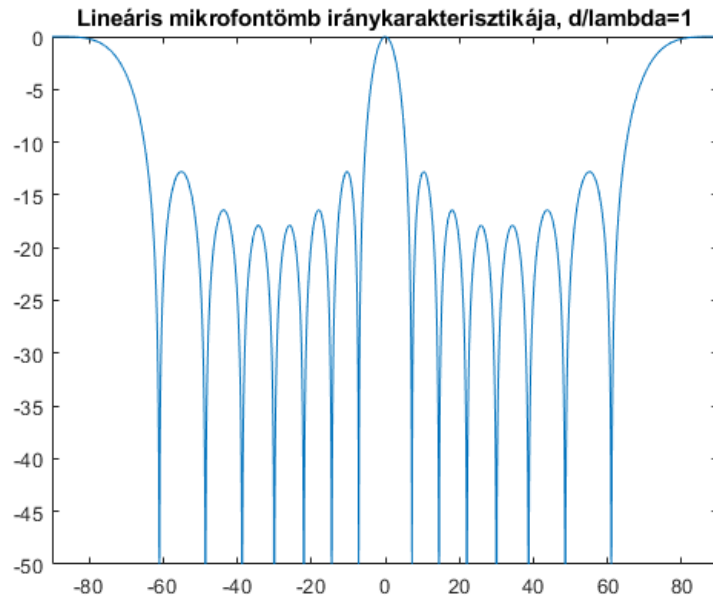
ahol:

$$m = 0, \pm 1, \pm 2, \dots$$

Ha  $\frac{d}{\lambda} > 1$ , akkor az  $m = 0$ -án kívül más megoldása is lehet a (4.7)-es egyenletnek, ami azt jelenti, hogy több irányból is 0 dB-es lesz az átvitel (lásd 4-5. ábra). Ha  $\frac{d}{\lambda} < 1$ , akkor csak  $m = 0$  esetén van megoldása az egyenletnek (lásd 4-4. ábra).

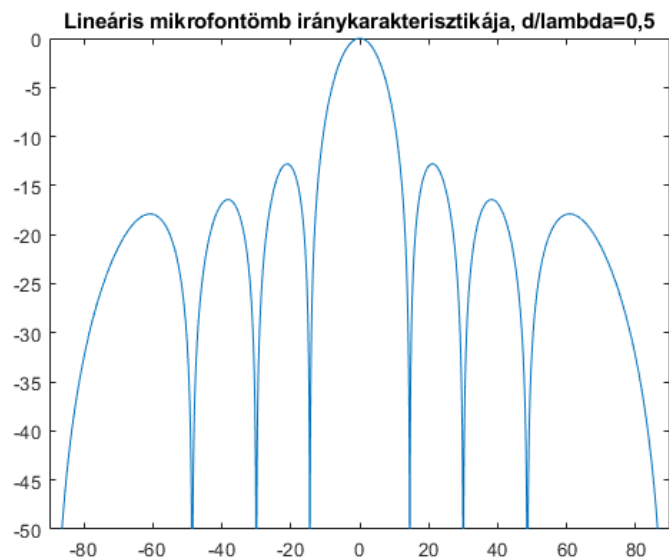


4-4. ábra, lineáris mikrofontömb iránykarakteristikája,  $\frac{d}{\lambda} = 0,6$



4-5. ábra, lineáris mikrofontömb iránykarakterisztikája,  $\frac{d}{\lambda} = 1$

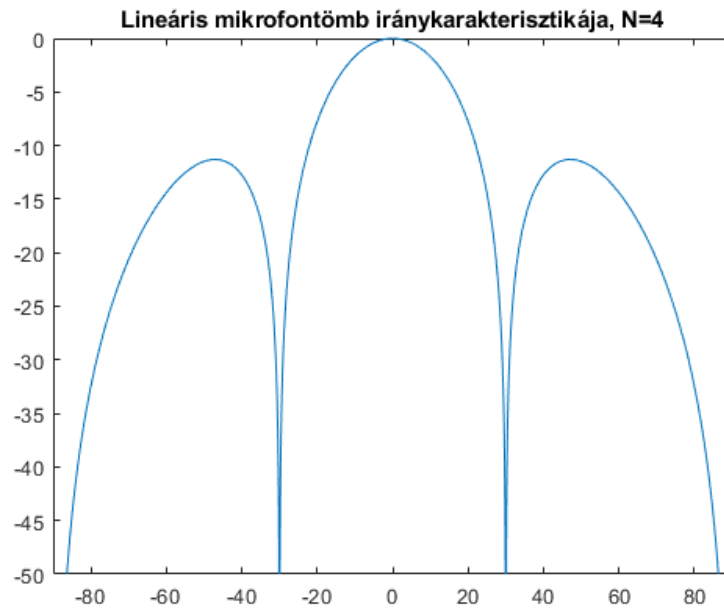
Tovább vizsgálva a  $\frac{d}{\lambda} < 1$  esetet a (4.7)-es egyenletben azt kapjuk, hogy  $\frac{d}{\lambda} = 0,5$  esetén igaz, hogy:  $-\pi \leq \Delta\varphi \leq \pi$ , azaz pontosan félperiódusnyit tudjuk maximum késleltetni a mikrofonok által vett jelek fázisát. Ebben az esetben érhetjük el a széleken, azaz  $\Theta = -90^\circ$  és  $\Theta = 90^\circ$ -os irányokban a maximális elnyomást, lásd a következő ábrán.



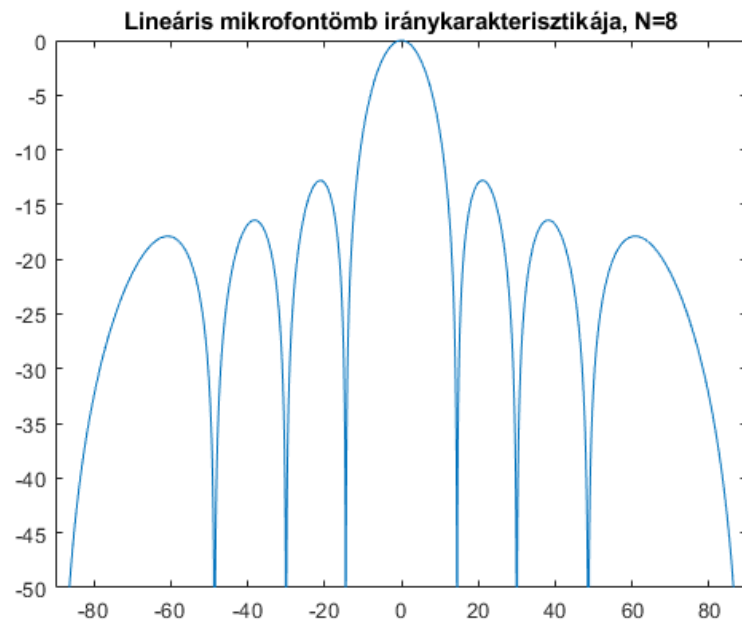
4-6. ábra, lineáris mikrofontömb iránykarakterisztikája,  $\frac{d}{\lambda} = 0,5$



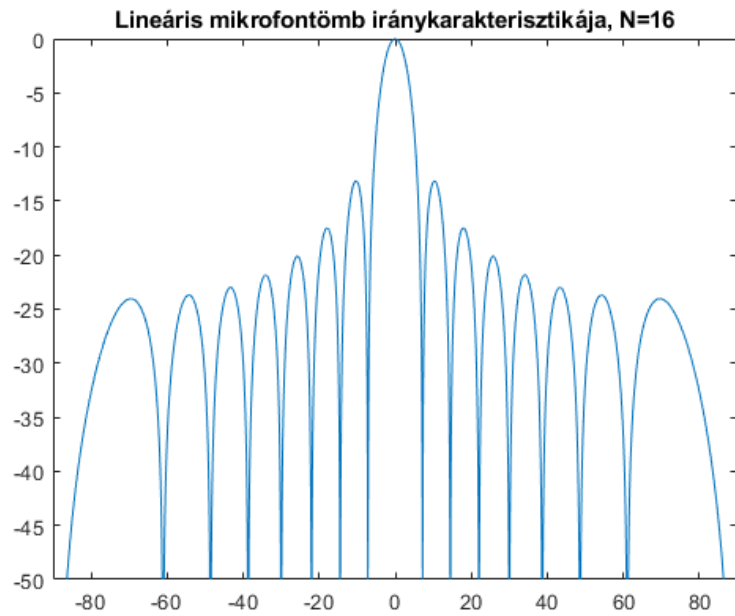
A (4.5)-ös egyenlet a  $\frac{d}{\lambda}$  arányon kívül függ még a mikrofonok számától, azaz  $N$ -től. Minél több mikrofont használunk fel a lineáris tömbben, annál több zéró erősítésű pont lesz az átvitelben, illetve annál nagyobb lesz az elnyomás a főnyalábtól távolodva. A következő 4 ábrán  $N = 4$ ,  $N = 8$ ,  $N = 16$  és  $N = 64$  esetekben látható a lineáris, egymástól egyenletes távolságban elhelyezett mikrofonok alkotta tömb iránykarakterisztikája ( $\frac{d}{\lambda} = 0,5$ ).



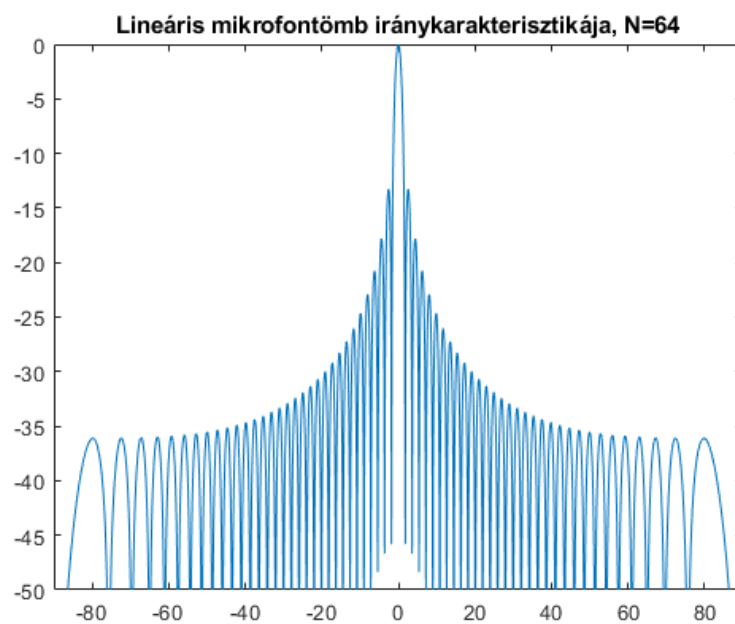
4-7. ábra, lineáris mikrofontömb iránykarakterisztikája,  $N = 4$



4-8. ábra, lineáris mikrofontömb iránykarakterisztikája,  $N = 8$



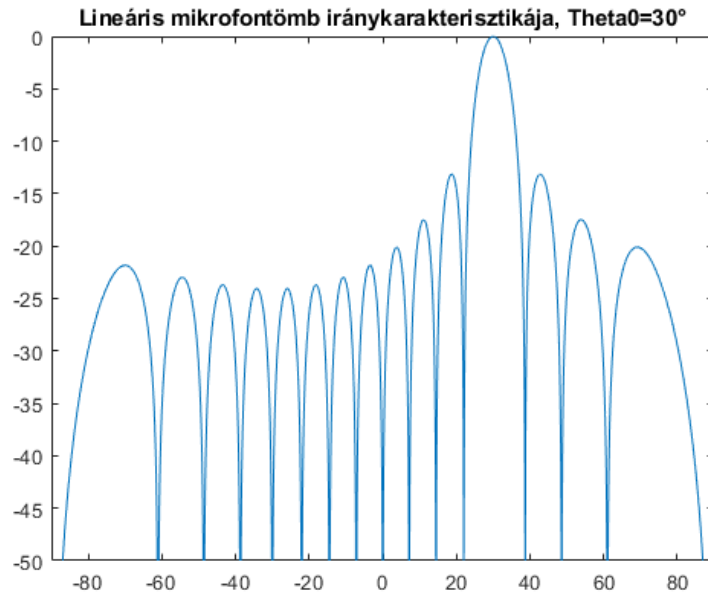
4-9. ábra, lineáris mikrofontömb iránykarakterisztikája,  $N = 16$



4-10. ábra, lineáris mikrofontömb iránykarakterisztikája,  $N = 64$

Leolvasható, hogy míg például 8 mikrofon esetében a főnyalábot és melléknyalábokat burkoló görbe alig -20 dB-es elnyomással rendelkezik  $\theta = -90^\circ$  és  $\theta = 90^\circ$ -os irányokban, addig mindez 64 mikrofon esetében már -35 dB. Ezen kívül egy másik észrevétel, hogy az átviteli iránykarakteristikák pontosan annyi teljes mértékben elnyomó irányt tartalmaznak, ahány mikrofonból állt a mikrofontömb.

Az eddigi esetekben a minden alkalommal a  $\Theta_0 = 0^\circ$ -os irányt választottuk a főnyaláb irányának. Ezáltal az átvitelekből készített grafikonok szimmetrikusak voltak, azaz  $\Theta = 30^\circ$ -os irányban ugyanakkora volt az erősítés, mint  $\Theta = -30^\circ$ -os irányban. A következő ábrákon bemutatok néhány esetet  $N = 16$  mikrofonnal, amikor a főnyaláb iránya  $\Theta_0 = 0^\circ$ -tól eltér, ugyanakkor a  $\frac{d}{\lambda}$  arány továbbra is 0,5 lesz.

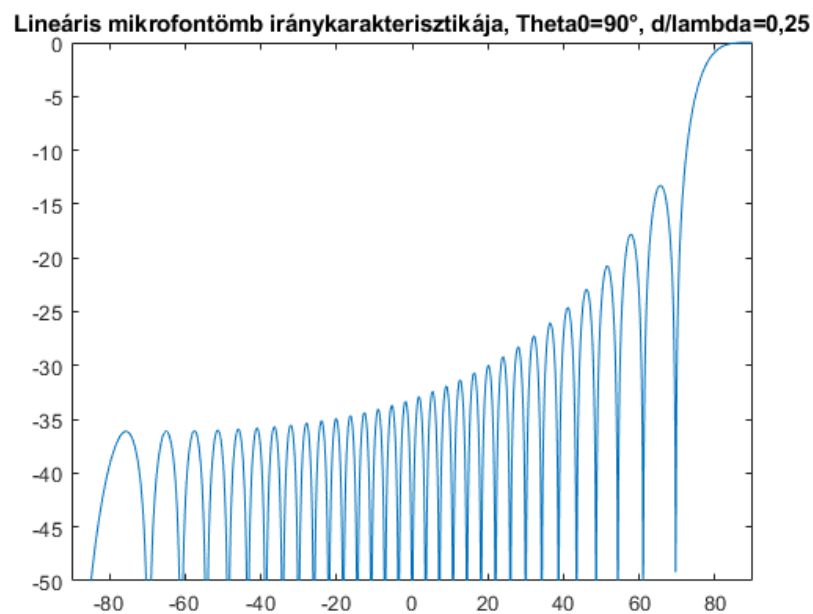


4-11. ábra, lineáris mikrofontömb iránykarakterisztikája,  $\Theta_0 = 30^\circ$



4-12. ábra, lineáris mikrofontömb iránykarakterisztikája,  $\Theta_0 = 90^\circ$

A 4-11. ábra szemlélteti, hogy  $\Theta_0 = 30^\circ$ -os irányú főnyaláb esetén nagy negatív szögek irányában a melléknyalábokat burkoló görbe elkezd emelkedni, azaz egyre kevésbé nyomja el a vett jelet azokból az irányokból a mikrofonok által alkotott tömb.  $\Theta_0 = 90^\circ$  esetén ez azt jelenti, hogy  $-90^\circ$ -os irányból elnyomás nélkül átengedi a jelet a rendszer. Olyan esetben, amikor mindenképpen  $90^\circ$ -os irányból szeretnénk átengedni a vett teljesítményt, ugyanakkor szeretnénk azt ellentétes irányból elnyomni, lecsökkenthetjük a  $\frac{d}{\lambda}$  arányt például 0,5-ről 0,25-re, azaz a mikrofonokat fele akkora távolságra helyezzük el egymástól, mint korábban. Ennek hatása a 4-13.-as ábrán látható, a jobb szemléltetés érdekében 16 helyett 64 mikrofon esetén.



4-13. ábra, lineáris mikrofontömb iránykarakterisztikája,  $\Theta_0 = 90^\circ$ ,  $\frac{d}{\lambda} = 0,25$

Érdekességképpen leolvasható, hogy abban az esetben, ha  $\Theta_0 = 90^\circ$ , vagy  $\Theta_0 = -90^\circ$ -os irányba állítjuk a főnyalábot, akkor nem  $N$ , hanem  $N - 1$  db maximálisan elnyomó irány lesz a  $-90^\circ - 90^\circ$ -os tartományon.

Jelfeldolgozásban a fenti kartézianus koordináta-rendszer mellett gyakran használnak polár koordináta-rendszert is. Sok esetben szemléletesebb, az emberi gondolkodáshoz közelebb álló ábrákat kaphatunk használatával. A két koordináta-rendszer közötti átváltásokat a következő egyenletek definiálják:

$$r = \sqrt{x^2 + y^2} \quad (4.8)$$

$$\varphi = \operatorname{atan} 2(x, y) \quad (4.9)$$

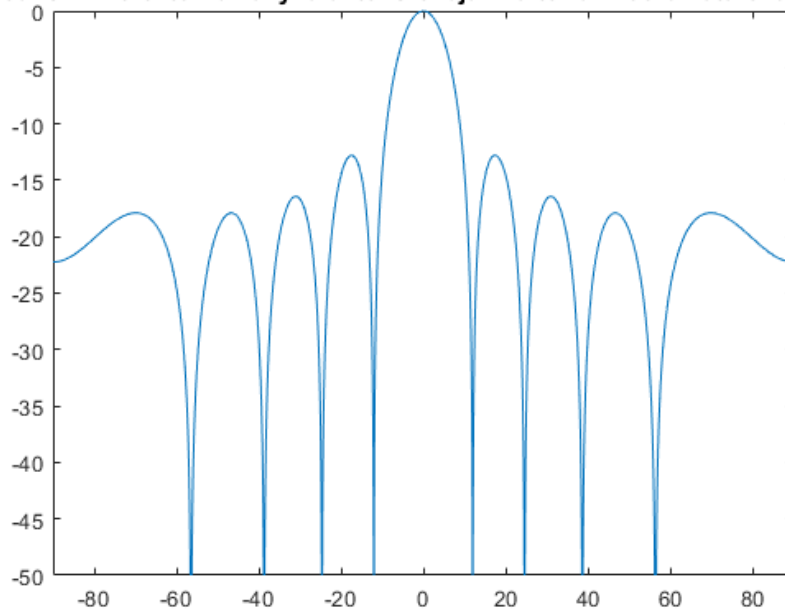
$$\operatorname{atan} 2(x, y) = \begin{cases} \tan^{-1}\left(\frac{y}{x}\right) & , \text{ha } x > 0 \\ \tan^{-1}\left(\frac{y}{x}\right) + \pi & , \text{ha } x < 0, y \geq 0 \\ \tan^{-1}\left(\frac{y}{x}\right) - \pi & , \text{ha } x < 0, y < 0 \\ \frac{\pi}{2} & , \text{ha } x = 0, y > 0 \\ -\frac{\pi}{2} & , \text{ha } x = 0, y < 0 \\ \text{nem definiált, ha } x = 0, y = 0 \end{cases} \quad (4.10)$$

$$x = r \cos \varphi \quad (4.11)$$

$$y = r \sin \varphi \quad (4.12)$$

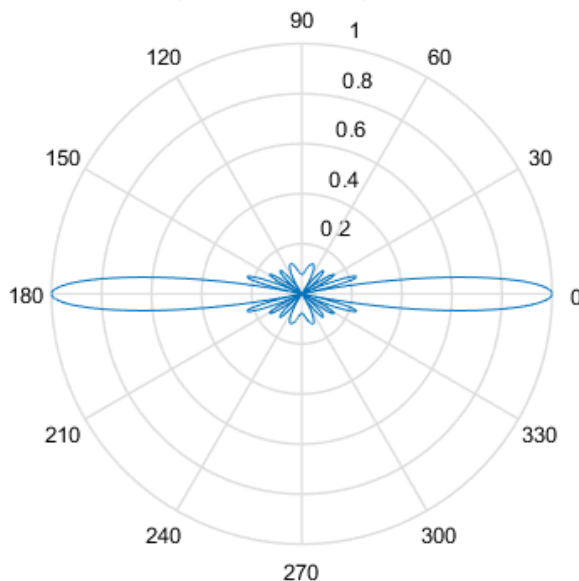
A következő két ábrán szemléltetem a koordinátarendszerek közötti különbséget:

**Lineáris mikrofontömb iránykarakterisztikája - kartezián koordinátarendszerben**



4-14. ábra, karteziánus koordinátarendszer

Lineáris mikrofontömb iránykarakterisztikája - polár koordinátarendszerben



4-15. ábra, polár koordinátarendszer

A fenti megállapítások egy viszonylag egyszerű, vonalban, egymástól egyenlő távolságra elhelyezett mikrofonok alkotta mikrofontömböt feltételeztek. Ugyanakkor már ebben az esetben is láthattuk, hogy milyen soktényezős lehet egy olyan egyszerű feladat, minthogy valamilyen irányból átengedjünk egy bizonyos frekvenciájú jelet, más irányokból pedig elnyomjuk azt. Valós akusztikai esetekben szinte sohasem csupán egyetlen, keskeny frekvenciasávban vizsgálódunk, nem ritka, amikor a legkisebb és legnagyobb frekvencia aránya akár 1:10-hez közelít. Ez azt okozhatja, hogy a kényesen beállított  $\frac{d}{\lambda}$  arány valamilyen irányba eltolódik. Kisebb frekvenciák esetén nem lesz olyan nagy az elnyomás a főnyalábtól távolodva, illetve maga a főnyaláb is szélesebb lesz, nem lesz annyira irányított. Nagyobb frekvenciák esetén először csak akkor jelenik meg egynél több főnyaláb, amikor valamilyen oldalirányba figyelünk, utána pedig már abban az esetben is, ha a főnyaláb közepén helyezkedik el. Továbbá előfordulhat olyan eset is, amikor nem csak egyetlen irányból szeretnénk átengedni a vett jelet, hanem egyszerre több, eltérő irányból is szeretnénk belehallgatni a térbe.

A következő fejezetben ismertetem, hogy irodalomkutatásom során milyen más mikrofonelhelyezéseket ismertem meg, azoknak milyen előnye, illetve milyen hátránya van.

## 4.2. Mikrofontömb elrendezések

Jasna Stamać, Sanja Grubeša, Antonio Petošić [11] munkájuk során MATLAB segítségével szimulálták különböző mikrofonelrendezések esetén az akusztikus kamerák fő- és melléknyalábainak tulajdonságait. Kísérletet végeztek többek között lineárisan, körben, spirálban, négyzet alakban elhelyezett mikrofonok esetén is. Céljuk az volt, hogy megtalálják a legnagyobb erősítésű, legkeskenyebb főnyalábú elrendezést középnyalában. A szimulálás során minden irányban egységes átvitelű MEMS mikrofonokat vettek alapul, illetve munkájuk során 1, 2 és 4 kHz-en végezték el az optimalizálást a legmegfelelőbb elrendezés megtalálásához. Mivel későbbiekben én is ebben a frekvenciatartományban kívántam méréseket végezni, illetve dolgozni, ezért különösen relevánsnak bizonyultak tapasztalataik.

Munkájuk során elsősorban 1 kHz-es frekvencia mellett vizsgálták meg a mikrofonelrendezések tulajdonságait két, különböző esetben. Az elsőben olyan elrendezést alkalmaztak, hogy a legnagyobb távolság két szomszédos mikrofon között kisebb legyen, mint  $\frac{d}{\lambda} = 0,5$ , amit ebben az esetben ( $d < \frac{\lambda}{2} = \frac{c}{f*2} = 0,1715\text{m}$ ) 0,1 m-ben maximalizálták. A második esetben két szomszédos mikrofon közötti távolságnak legalább a hullámhossz felének kellett lennie, ekkor 0,2 m-ben minimalizálták a távolságot.

Ahhoz, hogy megértsük a kísérleteik során kapott eredményeket, először meg kell ismerni néhány antennák tulajdonságaival kapcsolatos fogalmat. A directivity-nek hívott mérőszám azt fejezi ki, hogy mekkora a vett teljesítmény maximuma egy adott pontban, és az átlagosan, minden irányból vett teljesítménynek a hányadosa.

$$D = \frac{P_{max}}{P_{\text{átl}}} \quad (4.13)$$

A „gain”-nek nevezett mennyiség az előzőnek egy olyan kiterjesztése, ami tartalmazza az antenna/forrás veszteségeit is.

$$k = \frac{P_{kibocsátott}}{P_{vett}} \quad (4.14)$$

$$G = k * D \quad (4.15)$$

$$G_{dBi} = 10 * \log_{10} G \quad (4.16)$$

Jasna Stamać, Sanja Grubeša és Antonio Petošić munkájuk során a „gain”-t, azaz az antenna erősítését tartották elsődleges mérőszámnak. Minél nagyobb egy antenna erősítése, annál nagyobb részét vételezi egy adott irányból jövő teljesítménynek, valamint annál jobban képes elnyomni más irányokból jövő teljesítményt. Kísérleteik során bebizonyították, hogy ha a szomszédos mikrofonokat a vizsgálandó frekvencia hullámhosszának felénél messzebbre helyezik el egymástól, akkor keskenyebb főnyalábot kapnak a dedikált irányban, ugyanakkor jóval nagyobb melléknyalábok keletkeznek, ezért a továbbiakban a  $d_{max} = 0,1$  m esetet fogom bemutatni. A legnagyobb antennaerősítést 33 db spirálban elhelyezett mikrofonnal kapták. Ugyanakkor érdekes eredmény, hogy 17 db spirálban elhelyezett mikrofon esetén rosszabb eredményt kaptak, mint 4 db vonalban elhelyezett mikrofon esetében. Figyelembe véve, hogy munkám során mindössze 8 db mikrofon használatára volt lehetőségem, a következő elrendezések erősítését vizsgáltam meg: vonali elrendezés (2, 4 db mikrofon), négyzet csúcspontjaiban való elrendezés (4 db mikrofon), illetve körvonal menti (8 db mikrofon).

vonali, 2 db mikrofon	1,17 dBi
vonali, 4 db mikrofon	4,01 dBi
négyzetes, 4 db mikrofon	1,21 dBi
kör alakú, 8 db mikrofon	7,93 dBi

4-1. táblázat, mikrofonelrendezések erősítése

Legjobb eredményt körben elhelyezett mikrofonok esetén kaphatunk, viszont vonali elrendezések eredményeiből arra következtettem, hogy 8 db mikrofon használata esetén hasonlóan jó erősítést érhetnék el. A vonali elrendezés előnye a körben elhelyezett mikrofonokkal szemben, hogy nagyobb szabadságot biztosít a tömb kialakításakor, egyszerűbb összerakni, illetve akár egyetlen, 8 db mikrofonnal végzett mérés eredményeit is felhasználhatom például egy 6 db mikrofonos bemenettel rendelkező alkalmazáshoz is.

Jasna Stamać, Sanja Grubeša és Antonio Petošić munkája, illetve a mikrofontömb könnyebb elhelyezése miatt a vonali elrendezést választottam a félév során a mérések elvégzésére. A következő fejezetben bemutatom, hogy hogyan végeztem el a méréseket.

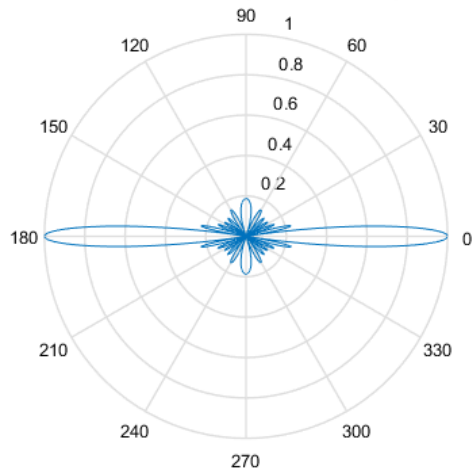


## 5. Mérések

A félév folyamán a méréseket vonali elrendezésű mikrofontömb segítségével végeztem el, melynek okát a 4. fejezetben ismertettem. A mérések elvégzéséhez két különböző terem választottam a következő megfontolásokból. Az első egy körülbelül 30 négyzetméteres, nagyjából négyzet alapú terem volt, benne rengeteg objektummal, asztalokkal, székekkel, szekrényekkel, az ablakokat pedig függöny takarta. A szobában található tárgyak, melyeknek egy része fejmagasságban található, csökkentik a reflexió hatását, mivel megtörik a reflektált hullámot, amelynek így nem tudnak a nagy teljesítményű visszavert hullámai akkora mértékben a mikrofonokig eljutni. A másik terem nagyobb, körülbelül 80-100 négyzetméter alapterületű, téglalap alapú volt, jóval kevesebb benne található tárggyal. Ebben a teremben fejmagasságban nem található semmilyen akadály, így a falakon reflektálódó visszavert hullámok könnyebben juthatnak el a mikrofonokhoz.

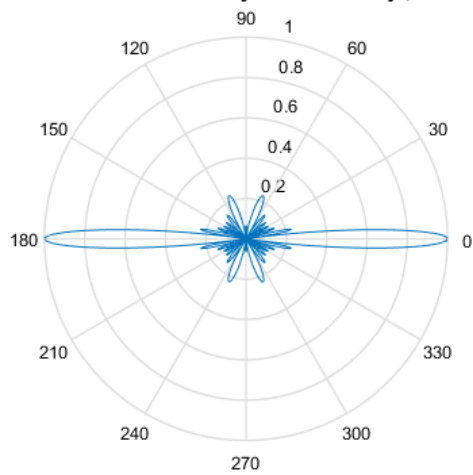
A mikrofonok távolságának meghatározásához előbb meg kellett határoznom, hogy körülbelül milyen frekvenciatartományban szeretném működtetni az alkalmazást, azaz milyen frekvenciatartományban hozzak majd létre iránykarakterisztikus szűrőt. Az emberi beszéd teljesítményének nagy része 300 Hz és 3,3 kHz között található [12], a régebbi telefonhálózatok is ebben a tartományban működtek, ezzel csökkentve a szükséges sávzélességet. A  $\frac{d}{\lambda}$  arány megválasztásakor arra figyeltem, hogy az ne legyen túl nagy, azaz mindenképpen legyen kisebb 1-nél, hiszen másképp jelentős melléknyalábok is megjelennek, illetve ne is legyen túl kicsi, azaz a főnyaláb közepétől számítva 30°-os távolságban már legyen számottevő az elnyomás. Ahhoz, hogy a fenti kényszereket tartani tudjam, csökkentenem kellett a később preferáltan használt frekvenciatartomány átfogását. A vizsgált tartomány minimum frekvenciáját 600 Hz-ben, a maximumát pedig 3 kHz-ben határoztam meg. 3 kHz esetén a következő átviteleket kaptam 8 db-os mikrofontömb esetén a mikrofontávolságokat változtatva:

Lineáris mikrofontömb iránykarakteristikája,  $d = 9 \text{ cm}$



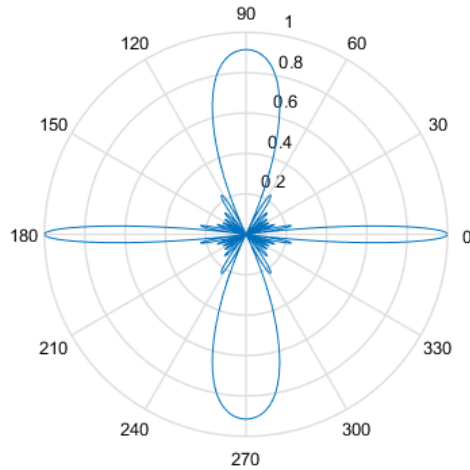
5-1. ábra, mikrofontömb iránykarakteristikája, mikrofontávolság 9 cm

Lineáris mikrofontömb iránykarakteristikája,  $d = 10 \text{ cm}$



5-2. ábra, mikrofontömb iránykarakteristikája, mikrofontávolság 10 cm

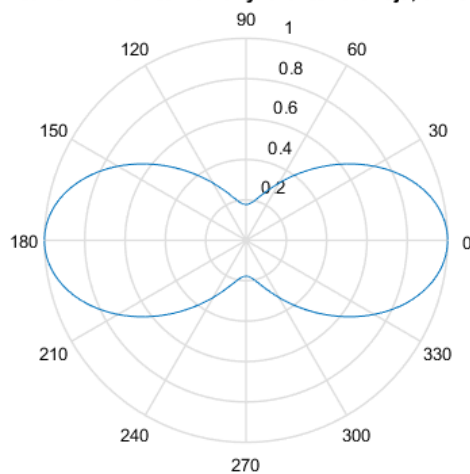
Lineáris mikrofontömb iránykarakterisztikája, d = 11 cm



5-3. ábra, mikrofontömb iránykarakterisztikája, mikrofontávolság 11 cm

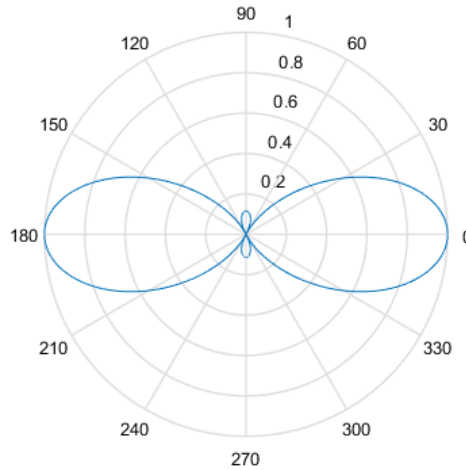
Látható, hogy egészen 10 cm-es mikrofontávolságig nem jelenik meg jelentős melléknyaláb az átviteli karakterisztikán, ugyanakkor 11 cm esetén már egy, a főnyalábnál szélesebb és nagyjából azonos erősítésű melléknyalábokat láthatunk. Ez azt jelenti, hogy körülbelül  $\frac{d}{\lambda} = \frac{d}{c/f} = \frac{10}{34000/3000} = 0,88$ -as aránnyal még megfelelő alkalmazást állíthatunk elő. Ugyanezt a vizsgálatot elvégeztem az alsó, 600 Hz-es határfrekvencián is. Alacsony  $\frac{d}{\lambda}$  arány esetén itt nem a megjelenő melléknyalábok, hanem a túlságosan széles főnyaláb jelenti a fő problémát. Célként tűztem ki, hogy a dedikált iránytól, azaz a főnyaláb közepétől 30°-os távolságban legfeljebb 0,5-szörös erősítés legyen.

Lineáris mikrofontömb iránykarakterisztikája, d = 6 cm



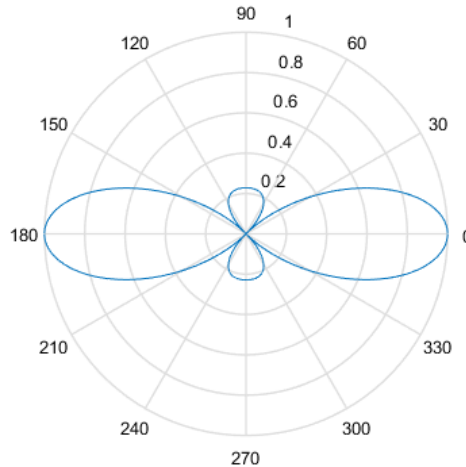
5-4. ábra, mikrofontömb iránykarakterisztikája, mikrofontávolság 6 cm

Lineáris mikrofontömb iránykarakterisztikája,  $d = 8$  cm



5-5. ábra, mikrofontömb iránykarakterisztikája, mikrofontávolság 8 cm

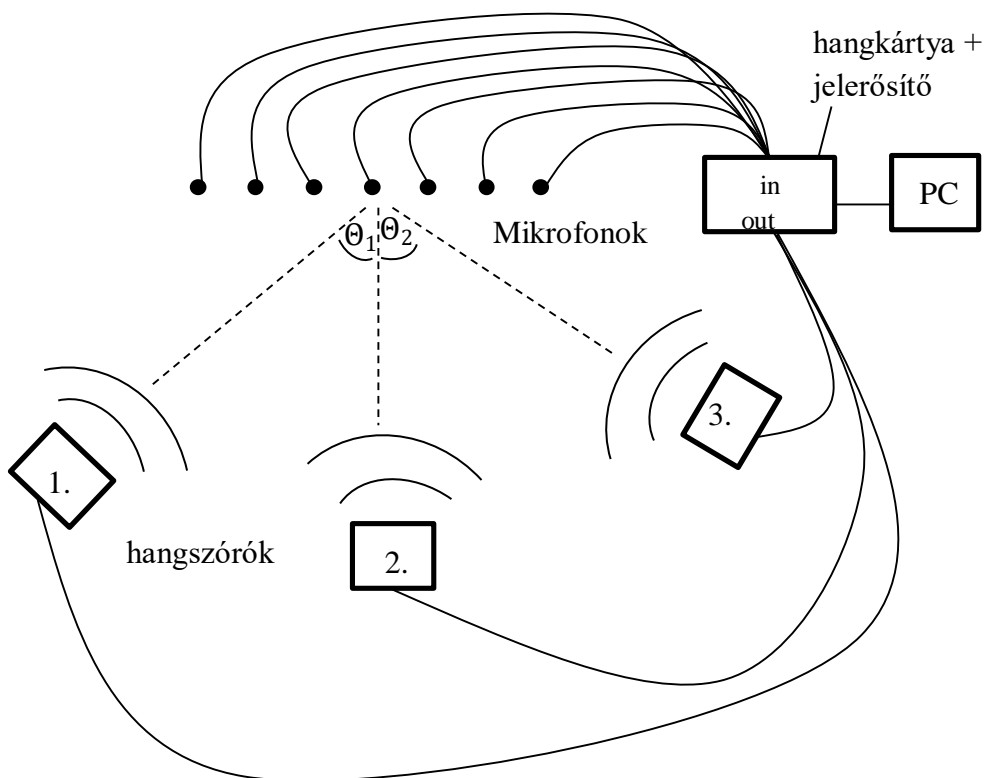
Lineáris mikrofontömb iránykarakterisztikája,  $d = 10$  cm



5-6. ábra, mikrofontömb iránykarakterisztikája, mikrofontávolság 10 cm

Az 5-5. ábra és 5-6. ábra alapján leolvasható, hogy kb. 9-10 cm-es mikrofontávolság esetén érem el az általam előbb támasztott kritériumot. Ennek következtében a méréseim során 10 cm-es mikrofontávolságot használtam. Ennél nagyobb felbontásra, azaz a főnyaláb szélességének néhány fokra csökkentésére sajnos nem volt lehetőségem. Ehhez vagy körülbelül egy nagyságrenddel több mikrofont kellett volna használnom a jelenlegi frekvenciatartomány mellett, vagy a tartományt kellett volna annyira beszűkítenem, hogy az az emberi beszéd teljesítményének olyan kis részét tartalmazta volna, hogy a rekonstruált beszéd érthetlenné vált volna. A következő kérdés a mikrofonok számát illeti. A tanszéken összesen 8 db mikrofon állt rendelkezésre, ugyanakkor a mérések időpontjában

sajnos nem volt elérhető ezek közül mind. Emiatt az első mérésem során 5, a második esetben pedig 7 mikrofon segítségével készítettem el a felvételeket. A felvételek elkészítéséhez 2 db Genius SP-HF1250x hangszórót használtam, melyekkel az előre felvett hangokat játszottam le a mérések során. Az előre felvett anyagok körülbelül 10 másodperc hosszú szöveget tartalmaztak, ehhez kevertem hozzá Audacity segítségével a két csatorna/hangszóró bemenet valamelyikére a zavarásokat. Zavarként használtam a mérések alkalmával 1000 Hz-es szinusz jelet, valamint előállítottam olyan bemenetet is, ahol az előre felvett szöveget két csatornára kevertem, amelyből az egyik csatorna a mérések során mint zavarás volt jelen.



5-7. ábra, mérési elrendezés

Az ábrán látható 3 lehetséges hangszóró hely közül egyszerre kettőben volt elhelyezve sugárzó, összesen 3 lehetséges módon: bal és jobb oldalt, középen és jobb oldalt, illetve középen és bal oldalt. A két oldalsó hely úgy volt beállítva, hogy:  $40^\circ < \theta_1, \theta_2 < 60^\circ$ . A hangsugárzók olyan távol helyeztem el a mikrofonoktól, hogy az már távotérnek számítson, ugyanakkor a megfelelő jel-zaj viszony megtartásának érdekében ne is legyenek túl távol. Peter Delos, Bob Broughton, and Jon Kraft [8] a következőképpen definiálták munkájuk során a távotér fogalmát:

$$\text{Távoltér} > \frac{2 * (N - 1) * d}{\lambda_{min}} \quad (5.1)$$

7 mikrofon esetében körülbelül 6 méter, 5 esetén pedig 3 méter felett beszélhetünk távoterről 10 cm-es mikrofontávolság esetén. A mérések során körülbelül ekkora távolságban helyeztem el a hangsugárzókat a mikrofonoktól.



5-8. ábra, valódi mérési elrendezés

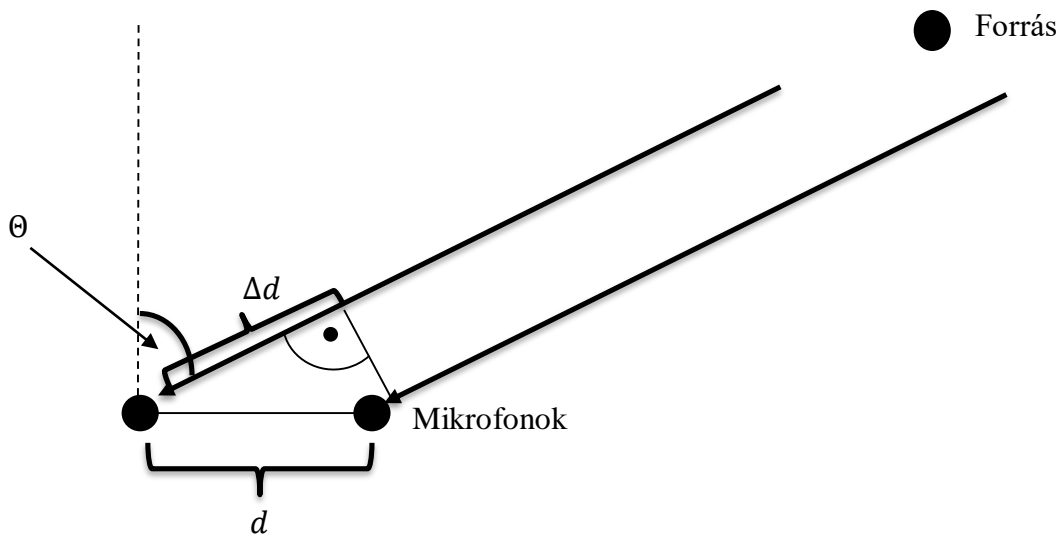
A következő fejezetben bemutatom a megtervezett iránykarakterisztikus szűrőt, melynek segítségével adott frekvencián térbeli, valamint széles frekvenciatartományban térbeli és időbeli szűrést valósítottam meg.

## 6. Nyalábformálás

Az iránykarakterisztikus szűrő megtervezéséhez számos, a területen korábban elvégzett munkát tanulmányoztam. Irodalomkutatásom során céлом az volt, hogy betekintést nyerjek olyan különféle módszerekbe, melyekkel térbeli és időbeli szűrést lehet megvalósítani. Otis Lamont Frost [13], Barry D. Van és Veen és Kevin M. Buckley [14], illetve Antonio Cantoni és L.C. Godara [15] legfőbb céljával egy dedikált forrás jelének minél jobb rekonstrukcióját, valamint más irányokból jövő zajok minél jobb elnyomását tűzték ki célul. Antonio Cantoni és L.C. Godara munkájában jelentős szerepet játszik a beamforming algoritmus, amellyel vételhez és kibocsátáshoz is formálhatunk nyalábot a tér egy kiválasztott irányába, esetünkben azért, hogy csak a számunkra fontos hangforrás teljesítményét engedjük át a szűrőn. A 6.2-es fejezetben bővebben kifejtem, hogy milyen megfontolások, kényszerek alapján lehet különböző frekvenciasávú szűrőket tervezni adott alkalmazásokhoz. Mivel akusztikus eszközök esetén a frekvenciatartomány minimuma és maximuma között körülbelül egy nagyságrendbeli különbség van, ezért szükség lesz szélessávú szűrő használatára munkám során. Ahhoz, hogy az iránykarakterisztikus szűrőt megfelelően megtervezhessem, szükségem van a mikrofontömb elrendezésének ismeretére, illetve a bemenet jelmodelljének létrehozására. A félév során vonali elrendezésű, egymástól egyenlő távolságra elhelyezett mikrofonokat használtam, melynek okait korábban a 4. fejezetben kifejtettem. A következő fejezetben bemutatom a létrehozott jelmodellt, hogy ezek után a megfelelő szűrőmodell létrehozható legyen.

### 6.1. Jelmodell

Attól függően, hogy hány dimenziós térben vizsgálódunk, a hang a forrástól számítva koncentrikus körök, illetve koncentrikus gömbök mentén terjed a mikrofonok felé a valóságban (lásd korábban 4-1. ábra.). Azonban ha távolférr terjedést feltételezünk, ahogy azt jelen esetben megtehetjük (lásd 5-ös fejezet, (5.1)-es egyenlet), akkor a jel terjedése a már ismertetett módon síkhullámmá válik, ahogyan azt a 6-1. ábra ismételtlen szemlélteti.



6-1. ábra, síkhullám terjedése

Ahogy az az ábrán látható, a hang nem egyforma nagyságú utat tesz meg a mikrofonokig, ami segítségünkre lesz a következőkben a szűrő megtervezésekor. A fenti jelterjedés matematikai leírása egy monokromatikus síkhullám és lineáris mikrofontömb használata esetén:

$$x(t) = S * e^{j2\pi ft} * e^{j2\pi fT_0} \quad (6.1)$$

$$S(\theta) = [e^{-j2\pi f\tau_1}, e^{-j2\pi f\tau_2}, \dots, e^{-j2\pi f\tau_L}]^T \quad (6.2)$$

$$\tau_i = \frac{1}{v} * (i - 1) * d * \sin(\theta) \quad (6.3)$$

$i = 1, 2, \dots, L$  – mikrofonok sorszáma

$v$  – hangsebesség

$T_0$  - a hangforrás és a 0. mikrofon pozíciója közötti jelterjedés ideje

$d$  – szomszédos mikrofonok közötti távolság

$\theta$  – forrás iránya a mikrofonokhoz képest

$f$  – monokromatikus síkhullám frekvenciája



Az  $x(t)$  bemenet 3 részre bontható fel a (6.1) egyenlet alapján. Az első az  $e^{j2\pi ft}$ -s tag, mely egy  $f$  frekvenciájú szinuszjel komplex exponenciális segítségével leírt reprezentációja az időben. A jel aktuális értékét  $t$  pillanatban a komplex vektor reális értéke adja meg. A második tag a  $e^{j2\pi f T_0}$ , mely azt a fázisforgatást szemlélteti a komplex vektoron a Fourier-transzformáció eltolási tételének értelmében, melyet a jel a forrástól egy kiválasztott mikrofonig elszenved. A tétel kimondja, hogy egy jelen végzett  $T$  időkésltetés időtartományban  $e^{-j2\pi f T}$  fázisforgatással egyezik meg frekvenciatartományban.

$$\mathcal{F}\{x(t)\} = X(\omega) \quad (6.4)$$

$$\mathcal{F}\{x(t - T)\} = e^{-j\omega T} * X(\omega) \quad (6.5)$$

A harmadik tag az  $S$  vektor, mely a mikrofonok elrendezését tartalmazza. Tetszőlegesen kiválasztunk egy origó pontot, az egyszerűség miatt ezt én az egyik szélső mikrofon pontjába tettem. Az ehhez a ponthoz képesti jelterjedésből származó jelútkülönbségek miatti komplex vektorokra ható fázisforgatást tartalmazza az  $S$  vektor. Ha az origó pontban lévő mikrofon sorszámát 0-nak vesszük, és utána a sorban következő mikrofonok esetében ezt egyesével iteráljuk, akkor feltéve, hogy a szomszédos mikrofonok egyenlő távolságra helyezkednek el egymástól az egyenlet felírható (6.2) és (6.3) szerint. A fentiekből következik, hogy a  $T_0$  idő megegyezik a forrás és a 0. mikrofon közötti jelterjedési idővel. Későbbiekben ez a tag akár el is hanyagolható, hiszen nem befolyásolja a mikrofonok által vett jelek egymáshoz képesti különbségeit.

A 4.-es fejezetben specifikált mikrofontömb, és az ebben az alfejezetben részletezett jelmodell alapján készítettem el az iránykarakterisztikus szűrőt. A jelmodell elkészítésében nagy segítségemre volt Otis Lamont Frost [15] munkája. A következő fejezetben prezenálni fogom az iránykarakterisztikus szűrő modelljét.

## 6.2. Szűrőmodell

Akusztikában a vett jel szűrése alapvetően két csoportra bontható, idő- és frekvenciatartománybéli processzálásra. Az időtartománybéli szűrés egyik nagy előnye, hogy általában nem követel meg blokkos jelfeldolgozást. Ez olyan alkalmazásokban, amelyek valós időben működnek, kis válaszidővel, vagy pedig offline működnek, de megfelelően kis reakcióidővel a bemenetre vonatkozóan, előnyt jelent. Blokkonkénti adatfeldolgozaskor össze kell várni sok esetben akár 10-100 ms nagyságrendbe eső hangmintát, mire megfelelő

mennyiségű adat áll rendelkezésre az elegendő felbontású diszkrét Fourier-transzformációhoz. Vett jel időbeli szűrésére rengeteg módszer létezik. Céltól függően lehet alkalmazni, aluláteresztő, felüláteresztő szűrőket, lyukszűrőket, FIR, IIR szűrőket, stb. Mind ehhez mikrofontömb létrehozására sincs feltétlenül szükség, hiszen elég egyetlen mikrofont is használnunk ezekben az esetekben. Ugyanakkor, ahogyan azt a Bevezetésben ismertettem, előfordulhatnak olyan alkalmazások, ahol nem elég, ha időbeli szűrést alkalmazunk. Vannak esetek, amikor a kiszűrendő jel frekvenciatartománya átfedésben van, vagy azonos a dedikált jel frekvenciaspektrumával, viszont más irányból érkezik. Ekkor kerül előtérbe a mikrofontömb használata, amint azt a 4.1-es fejezetben bemutattam. Ha csupán egyetlen irányból szeretnénk venni a jelet, minden más irányból pedig elnyomni, akkor elég egy vonali elrendezésű mikrofontömb használata is, akár kiegészítve valamilyen időtartománybéli szűréssel a térben szűrt jelen. Ugyanakkor nem minden esetben elegendő, ha csak egy irányból engedjük át a jelteljesítményt, vagy például csak bizonyos frekvenciákat szeretnénk elnyomni adott irányokból. Iránykarakterisztikus szűrő segítségével ezek megoldható problémák, hardveres korlátjainkat figyelembe véve szinte tetszőleges szűrőt állíthatunk elő mind térben, mind időben.

Ha csupán időbeli szűrést alkalmaznánk, akkor az alkalmazásunk nem tudná kezelni az interferencia jelenségét. A jel-zaj viszony mérőszám javítható a segítségével a nem szükséges frekvenciasávokban lévő teljesítmény elnyomásával, viszont általában a zavarforrások sajnos az általunk vizsgált frekvenciatartományban is sugároznak. Különösen igaz ez az akusztikus alkalmazásokban, ahol a vizsgált minimum és maximum frekvencia között gyakran egy nagyságrendbeli különbség is lehet. A szuperpozíció elve miatt a különböző forrásoktól származó, azonos frekvenciájú jelkomponensek összeadódnak, egyedül a terjedés iránya szerint lehet őket megkülönböztetni, ezért fontos, hogy térbeli szűrést is alkalmazzunk.

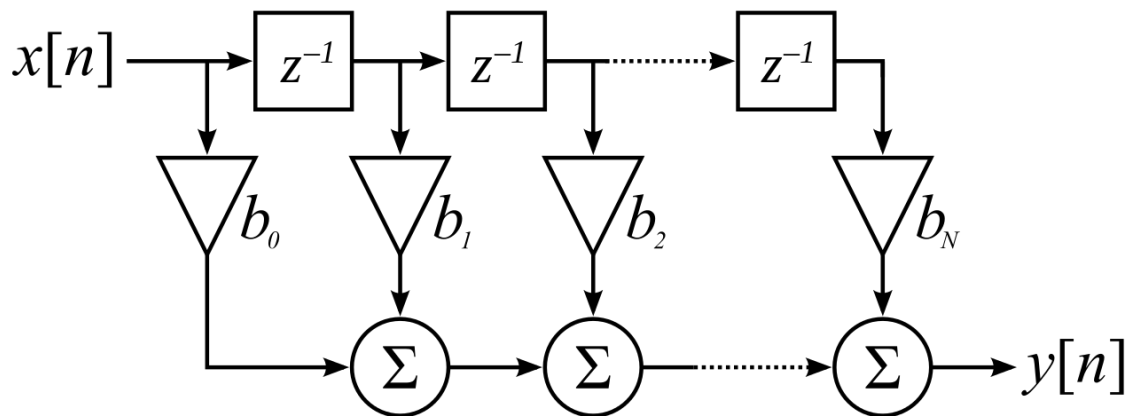
A következő fejezetekben bemutatom az iránykarakterisztikus szűrők modelljét, melyek segítségével végrehajthatjuk a térbeli szűrést is a vett jelen.

### **6.2.1. Keskenysávú szűrő modellje**

Keskenysávú szűrőt olyan alkalmazásokban lehet használni, ahol a vizsgált frekvenciatartomány közepének frekvenciája jóval nagyobb, mint a frekvenciatartomány határainak különbsége. Ezt a középfrekvenciát kommunikációs alkalmazásokban, például rádiózásban, vivőfrekvenciának hívják. Akusztikus alkalmazásokban csak akkor használható, ha

a forrás keskeny frekvenciatartományon belül sugároz, ennél fogva emberi beszéd rekonstrukciójára nem alkalmas, viszont megértése fontos a komplexebb szűrők tervezéséhez.

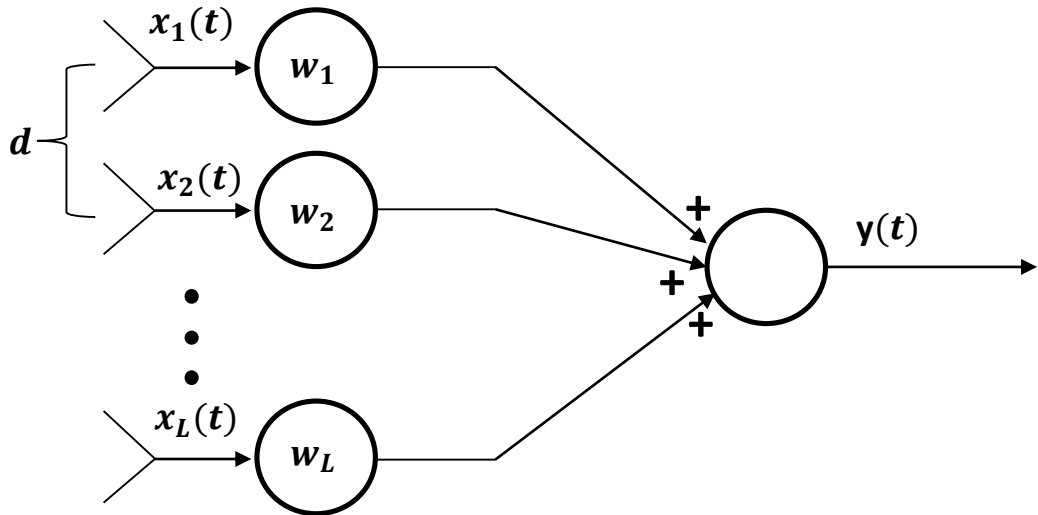
Barry D. Van és Veen and Kevin M. Buckley [14] többek között a keskenysávú szűrő és a FIR szűrő analógiáját ismertetik munkájukban. A FIR szűrő egy véges hosszú gerjesztésre véges hosszú választ ad, azaz meghatározható egy olyan eleme a válasznak, ahonnan kezdve 0 a kimenet értéke. A szűrő blokkvázlata a 6-2.-es ábrán látható.



6-2. ábra, FIR szűrő blokkvázlata [16]

A beérkező mintákat meghatározott számú esetben késleltetjük, miközben minden egyes késleltető blokk kimeneti jelét összeadjuk, ezzel megkapva az  $y[n]$  kimenetet. Amennyiben az  $x[n]$  bemenet  $M$  db mintát tartalmaz, akkor az ábrán látható szűrő válasza  $(M + N)$  hosszúságú lesz. Ez időtartománybéli FIR szűrés esetén annyit tesz, hogy a bemenet minden egyes mintájára  $N$  db késleltető esetén  $(N + 1)$  db kényszert írhatunk elő, egy minta a feldolgozása utáni  $(N * T_s)$ -ik időpillanatig fejt ki hatást a kimenetre (ahol  $T_s$  – mintavételi idő). Az  $x[n]$  bemenet feldolgozása tehát mintánként zajlik, ami nem jelent jelentős plusz válaszidőt a rendszerben, szemben a frekvenciatartománybéli feldolgozással. A késleltető blokkok megcsapolt kimenete ezen kívül áthalad egy szorzón is az összegzés előtt, mely szorzóval alakíthatjuk ki a kívánt szűrőkarakterisztikát. Tervezéskor a szorzók értékeinek megadásával definiáljuk a kényszereket dedikált időpontokban.

A keskenysávú szűrő analógiát mutat a FIR szűrőre azzal az eltéréssel, hogy nem időben, hanem térben szűri meg a bemenő jelet. Ennek érdekében ahelyett, hogy egyetlen bemenetén a jelét késleltetné, több bemenettel rendelkezik, azaz több mikrofonnal történik a felvétel, mely vett jelek egymáshoz képest késnek a különböző megtett jelutak miatt, és ezek jelét súlyozza, majd összegezi, ezzel előállítva a kimenetet. A keskenysávú szűrő blokkdiagramja a 6-3.-as ábrán látható.



6-3. ábra, keskenysávú szűrő blokkvázlata

A 6-3.-as ábrán látható mikrofonok töltik be tehát a késleltető elemek szerepét. Mivel fizikailag nem egy helyen találhatóak, hanem elosztva a térben, ezért a mikrofonok elhelyezésével tetszőleges nagyságú késleltetéseket hozhatunk létre a rendszerben. Ezek után következik, hogy annyi pontra/irányra tudjuk felbontani a teret, ahány mikrofonunk van, hiszen FIR szűrő esetén is  $N$  db késleltetővel ( $N + 1$ ) db pontra tudtuk felbontani az időt, mivel a 0 késleltetés is késleltetés.

A következő egyenletek segítségével bemutatom, hogy milyen módszerrel definiáltam a  $D$  mátrixot, mely a szűrő tetszőleges iránykarakterisztikáját tartalmazza.

$$D = \begin{matrix} e^{-j*0*\Delta\varphi_1} & e^{-j*0*\Delta\varphi_L} \\ e^{-j*1*\Delta\varphi_1} & e^{-j*1*\Delta\varphi_L} \\ \vdots & \vdots \\ e^{-j*(L-1)*\Delta\varphi_1} & e^{-j*(L-1)*\Delta\varphi_L} \end{matrix} \quad \dots \quad (6.6)$$

$$\Delta\varphi_i = \frac{d}{\lambda} * \sin(\text{dir} s_i) * 2\pi \quad (6.7)$$

$$dirs = [\theta_1, \theta_2, \dots, \theta_L] \quad (6.8)$$

A  $D$  mátrix  $L$  db mikrofon esetén egy  $L \times L$ -es mátrix, mivel  $L$  db mikrofon segítségével adjuk meg  $L$  db irányból az átvitelt. Alkalmazásomat úgy alakítottam ki, hogy a mátrix sorai az egyes mikrofonokhoz, oszlopai pedig az egyes irányokhoz tartoznak. Az irányok definiálásához a  $dirs$  vektort hoztam létre. Ebben  $L$  db különböző irány szerepel,  $-90^\circ$  és  $90^\circ$  között, tehát abban a tartományban, amiben a vonali elrendezésű, egymástól egyenlő távolságban lévő mikrofonokkal vizsgálhatom a teret. A keskenysávú szűrő egy dedikált frekvencián működik, melynek hullámhossza  $\lambda$ . A hullámhossz, a mikrofonok ismert távolsága és elrendezése segítségével megállapíthatom, hogy két szomszédos mikrofon között mekkora fázistolást szenved el a jel a jelutak különbsége miatt különböző irányok esetén. Ezeket a fázistolásokat tárolom a  $\Delta\varphi_i$  vektorban. Mivel a mikrofonok egyenlő távolságra vannak egymástól, ezért egyszerűen számolható a fázistolás az origóhoz, azaz a 0. mikrofonhoz képest, ez látható a (6.6)-os egyenletben a  $\Delta\varphi_i$  fázistolások előtti egész szorzókon. Az eddigiekben meghatároztuk, hogy milyen irányokban szeretnénk meghatározni az átvitelt. A következő egyenletekben bemutatom, hogy hogyan kaphatjuk meg a 6-3.-as ábrán látható  $w = [w_1, w_2, \dots, w_L]$  súlyvektort, mellyel a szűrő átvitelét kaphatjuk meg.

$$ref = [a_1, a_2, \dots, a_L] \quad (6.9)$$

$$w * D = ref \quad (6.10)$$

$$w = ref * D^{-1} \quad (6.11)$$

$$y = w * x(t) \quad (6.12)$$

A  $ref$  vektorban az  $L$  db irányra adhatunk meg egy 0 és 1 közötti számot, amellyel meghatározzuk az adott irányból jövő teljesítmény átvitelének mértékét. Szélsőséges esetben akár teljesen el is nyomhatjuk, illetve zavartalanul át is engedhetjük a jelet. Célunk, ahogyan azt a (6.10)-es egyenlet mutatja, hogy a kívánt átvitelt (azaz a súlyvektort) a  $D$  mátrixban meghatározott kényszerekkel megszorozva a kívánt kimenetet, azaz a  $ref$  vektort kapjuk. Ennek az egyenletnek a megoldása látható (6.11)-ben. Az így kapott súlyvektort egy valódi,  $x(t)$  bemeneti vektorral megszorozva kaphatjuk meg a kimenetet,  $y(t)$ -t.

Amennyiben csak a  $D$  mátrixban definiált irányokból érkezik jel, abban az esetben kimenetünk ideális lesz, a  $ref$  vektorban definiáltak szerint fog alakulni. Abban az esetben, ha másmilyen irányokból is érkezik teljesítmény a mikrofontömbhöz, a kimenet tartalmazni fog nem hasznos teljesítményt is. Ezt a jelenséget egy egyszerű esetben mutatom be, amikor 2 mikrofont használunk a rendszerben, illetve 2 különböző irányból érkezik jel a mikrofonokhoz, ahol az egyik irányt rosszul tudjuk, ezért a  $D$  mátrix nem lesz tökéletes.

$$\Delta\varphi_1 = a \quad (6.13)$$

$$\Delta\varphi_2 = b \quad (6.14)$$

$$D = \begin{bmatrix} e^{-j*0} & e^{-j*0} \\ e^{-j*a} & e^{-j*b} \end{bmatrix} \quad (6.15)$$

$$w = [w_1, w_2] \quad (6.16)$$

$$x = \begin{bmatrix} e^{j*k} * e^{-j*0} + e^{j*k} * e^{-j*0} \\ e^{j*k} * e^{-j*a} + e^{j*k} * e^{-j*c} \end{bmatrix} \quad (6.17)$$

Tehát a  $D$  mátrixban  $a$  és  $b$  irányokban definiáltuk az átvitelt. Valamilyen  $ref$  vektor használatával pedig előállítottuk a  $w$  súlyvektort. A (6.17)-es egyenletben látható, hogy  $a$  és  $c$  irányokból érkezik jel a mikrofonokhoz, ahol az  $x$  vektor első sora az első mikrofon által vett jelet, második sora a második mikrofon által vett jelet tartalmazza. Az  $e^{j*k}$  tag a jel terjedését reprezentálja a  $k$ . időpillanatban egy forgó exponenciálissal. A bemenet egyszerűsítve a következőképpen írható fel:

$$x = \begin{bmatrix} 2 * e^{j*k} \\ e^{j*k} * (e^{-j*a} + e^{-j*c}) \end{bmatrix} \quad (6.18)$$

Elvégezve a jelfeldolgozást, a következő kimenetet kapjuk:

$$y = w * x = e^{j*k} * (w_1 * 2 + w_2 * e^{-j*a} + w_2 * e^{-j*c}) \quad (6.19)$$

Ha  $c$  helyett  $b$  irányból érkezett volna a hang, akkor a  $w * D$  szorzatnak az eredményét kaptuk volna, ami maga a  $ref$  vektor. A (6.19)-es egyenletben a kimenetnek egyetlen  $c$

iránytól függő tagja van a zárójelben, ami meghatározza a kimenetben a hibánkat, melynek nagysága a következő:

$$w_2 * e^{-j*c} = w_2 * e^{-j*b} * e^{-j(c-b)} \quad (6.20)$$

Amennyiben nem hibáztunk volna a bemenet irányainak meghatározásában,  $e^{-j(c-b)} = 1$  lenne, viszont minden ettől eltérő eredmény a kívánt kimenetünket a fenti egyenletben látható módon rontja. A legnagyobb hiba akkor lép fel, ha a vizsgált és a valódi irányokhoz tartozó fázistolások közötti különbség  $\pi$ -vel egyezik meg. A fenti eset nem csak akkor fordulhat elő, ha rosszul határozzuk meg a bemenetek irányát, hanem akkor is, ha egyszerre több irányból érkezik bemenet, mint ahány mikrofonunk van, ezért néhány irányban nem tudunk kényszereket előírni.

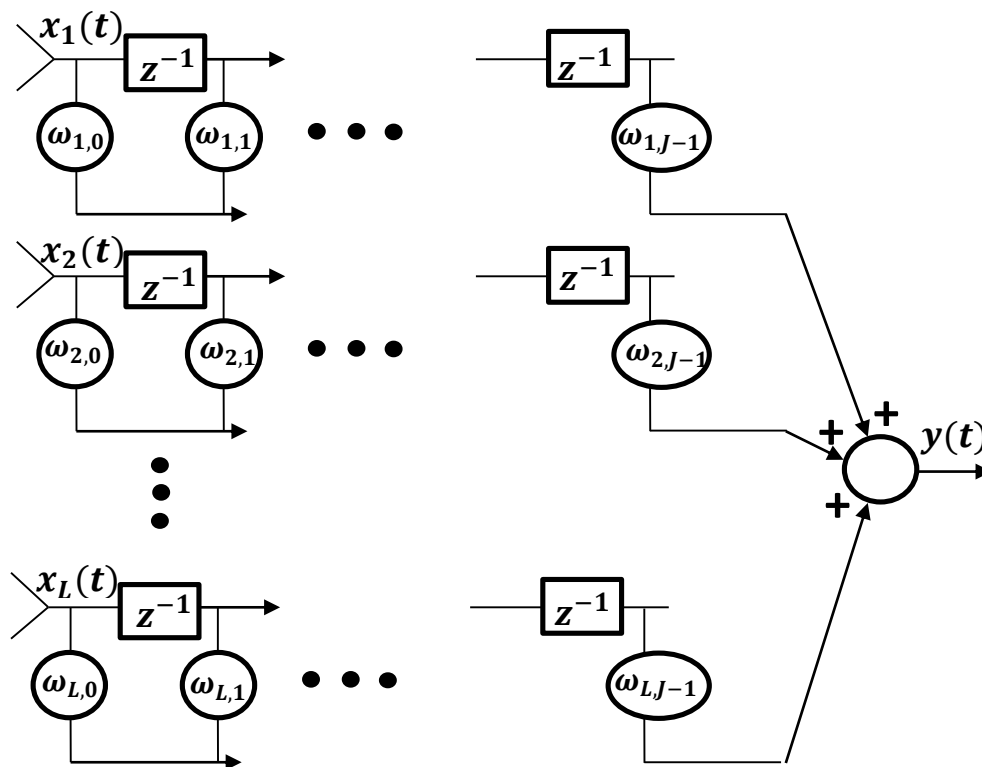
Az egyszerűség kedvéért a  $D$  mátrixban található kényszerek azonos kezdőfázisú esetekre vannak felírva. Az előbbi gondolatmenethez hasonlóan levezethető, hogy a kezdőfázisnak nincsen semmilyen befolyásoló hatása a kimenetre.

A következő alfejezetben bemutatom a szélessávú szűrő működését, amely a keskenysávú szűrőnek egy plusz dimenzióval való kiterjesztése.

### 6.2.2. Szélessávú szűrő modellje

Szélessávú szűrő alkalmazási területei közt megemlíthetjük azokat a rendszereket, ahol a vizsgált frekvenciasáv minimuma és maximuma távol esik egymástól. Mivel az elkészült rendszerben emberi beszéd rekonstrukcióját szeretném minél jobb hatásfokkal végrehajtani, aminek a határfrekvenciái között körülbelül egy nagyságrendbeli különbség van, ezért elengedhetetlen a szélessávú szűrő használata. Míg az előző esetben térben szűrtük meg a jelet  $L$  db mikrofon segítségével, addig most a térbeli szűrésen kívül minden mikrofon jelét egy egyedi időtartománybeli FIR szűrővel is feldolgozzuk.

A szűrő blokkvázlata a 6-4.-es ábrán látható. Mikrofononként  $(J - 1)$  darab késleltetőt helyezünk el, mellyel az időbeli szűrését hajtjuk végre az  $L$  darab bemeneten.



6-4. ábra, szélessávú szűrő blokkvázlata

Az  $L$  db mikrofonnak és  $(J - 1)$  db késleltetőnek köszönhetően  $L$  különböző irányban, irányonként  $J$  db időpontban írhatunk elő kényszereket. Az időtartományban  $J$  db időpont frekvenciatartományban  $J$  db különböző frekvenciakomponenst jelent a spektrumon. A keskenysávú esethez hasonlóan a  $D$  mátrix előállítás a következőképpen történt:



$$D = \begin{matrix}
e^{j*0*\Delta\varphi_{1,\omega_0}} & & e^{j*0*\Delta\varphi_{1,\omega_{J-1}}} & & e^{j*0*\Delta\varphi_{L,\omega_{J-1}}} \\
e^{j*1*\Delta\varphi_{1,\omega_0}} & & e^{j*1*\Delta\varphi_{1,\omega_{J-1}}} & & e^{j*1*\Delta\varphi_{L,\omega_{J-1}}} \\
\vdots & & \vdots & & \vdots \\
e^{j*(L-1)*\Delta\varphi_{1,\omega_0}} & & e^{j*(L-1)*\Delta\varphi_{1,\omega_{J-1}}} & & e^{j*(L-1)*\Delta\varphi_{L,\omega_{J-1}}} \\
e^{j*(0*\Delta\varphi_{1,\omega_0}-\omega_0*T_s)} & & e^{j*(0*\Delta\varphi_{1,\omega_{J-1}}-\omega_0*T_s)} & & e^{j*(0*\Delta\varphi_{L,\omega_{J-1}}-\omega_0*T_s)} \\
e^{j*(1*\Delta\varphi_{1,\omega_0}-\omega_0*T_s)} & & e^{j*(1*\Delta\varphi_{1,\omega_{J-1}}-\omega_0*T_s)} & & e^{j*(1*\Delta\varphi_{L,\omega_{J-1}}-\omega_0*T_s)} \\
\vdots & & \vdots & & \vdots \\
e^{j*((L-1)*\Delta\varphi_{1,\omega_0}-\omega_0*T_s)} & \dots & e^{j*((L-1)*\Delta\varphi_{1,\omega_{J-1}}-\omega_0*T_s)} & \dots & e^{j*((L-1)*\Delta\varphi_{L,\omega_{J-1}}-\omega_0*T_s)} \\
\vdots & & \vdots & & \vdots \\
e^{j*(0*\Delta\varphi_{1,\omega_0}-(J-1)*\omega_0*T_s)} & & e^{j*(0*\Delta\varphi_{1,\omega_{J-1}}-(J-1)*\omega_0*T_s)} & & e^{j*(0*\Delta\varphi_{L,\omega_{J-1}}-(J-1)*\omega_0*T_s)} \\
e^{j*(1*\Delta\varphi_{1,\omega_0}-(J-1)*\omega_0*T_s)} & & e^{j*(1*\Delta\varphi_{1,\omega_{J-1}}-(J-1)*\omega_0*T_s)} & & e^{j*(1*\Delta\varphi_{L,\omega_{J-1}}-(J-1)*\omega_0*T_s)} \\
\vdots & & \vdots & & \vdots \\
e^{j*((L-1)*\Delta\varphi_{1,\omega_0}-(J-1)*\omega_0*T_s)} & & e^{j*((L-1)*\Delta\varphi_{1,\omega_{J-1}}-(J-1)*\omega_0*T_s)} & & e^{j*((L-1)*\Delta\varphi_{L,\omega_{J-1}}-(J-1)*\omega_0*T_s)}
\end{matrix} \quad (6.21)$$

$$\Delta\varphi_{i,j} = \frac{d}{\lambda_j} * \sin(dirs_i) * 2\pi \quad (6.22)$$

$$\lambda = [\lambda_0, \lambda_1, \dots, \lambda_{J-1}] \quad (6.23)$$

$$dirs = [\theta_1, \theta_2, \dots, \theta_L] \quad (6.24)$$

A *dirs* vektorban ebben az esetben is  $L$  db mikrofonhoz van meghatározva  $L$  különböző irány. A fázistolásokat tartalmazó  $\Delta\varphi$  vektor jelen esetben nem egyetlen hullámhosszon adja meg a fázistolás nagyságát két szomszédos mikrofon között, hanem  $L$  irányban, irányonként  $J$  db különböző frekvencián, melyek a  $\lambda$  vektorban vannak megadva. Ezek alapján áll elő a  $D$  mátrix, melynek alakja a (6.21)-es egyenletben látható. Az első  $L$  db sor az  $L$  db mikrofonhoz tartozik, zéró késleltetés esetén. A következő  $L$  db sor hasonlóan, csak már 1 mintányi késleltetéssel, és í. t.  $(J - 1)$  mintányi késleltetésig. A mátrix első  $J$  db oszlopa az első irányhoz tartozó frekvenciákat tartalmazza, második  $J$  db oszlopa a második irányhoz tartozókat, és í. t. az  $L$ . irányhoz tartozó  $J$  db frekvenciához. Megjegyzendő, hogy noha az egyenletek úgy lettek felírva, hogy mindegyik irányban ugyanazon frekvenciákon írjuk elő a kényszereket, ez nem kötelező, tetszőleges FIR szűrőt tervezhetünk minden irányhoz. A következő egyenletekben bemutatom, hogy jelen esetben, hogyan határoztam meg a  $w$  súlyvektort.

$$ref = [a_{1,\omega_0}, a_{1,\omega_1}, \dots, a_{1,\omega_{J-1}}, a_{2,\omega_0}, a_{2,\omega_1}, \dots, a_{2,\omega_{J-1}}, \dots, a_{L,\omega_0}, a_{L,\omega_1}, \dots, a_{L,\omega_{J-1}}] \quad (6.25)$$

$$w * D = ref \quad (6.26)$$

$$w = ref * D^{-1} \quad (6.27)$$

$$y = w * x(t) \quad (6.28)$$

A *ref* vektorban adom meg az egyes irányokhoz és frekvenciákhoz/időpontokhoz tartozó átviteleket olyan sorrendben, amiben a *D* mátrix oszlopai is vannak. A (6.26), (6.27) és (6.28)-as egyenletek teljes analógiát mutatnak az előző alfejezetben definiált (6.10), (6.11) és (6.12)-es egyenletekkel. Ahogyan azt korábban bemutattam keskenysávú szűrő esetében, hasonlóan itt is levezethető, hogy az adott bemenetre kapott kimenet nem függ a *D* mátrixban található elemek kezdőfázisától. Azokban az esetekben, amikor a vizsgált irányokon vagy frekvenciákon kívül is jelen van komponense a bemenetnek, az átvitel a (6.13), (6.14), (6.15), (6.16), (6.17), (6.18), (6.19) és (6.20)-as egyenletekhez hasonlóan itt is megadható.

### 6.2.3. Valós bemenet

Az eddigi fejezetekben a bemenetre a 6.1.-es fejezetben bemutatott jelmodellt alkalmaztam. Ez magában foglalja, hogy a bemenetet egy komplex forgó exponenciálissal modelleztem. Ha megvizsgálunk egy valós jel frekvenciaspektrumát adott felbontással, akkor azt találunk, hogy minden egyes komponenséhez tartozik egy forgó exponenciális valamilyen amplitúdóval és kezdőfázissal. Ehhez a jelmodellhez a mikrofonok által vett skálár bemenet frekvenciaspektrumára van szükségünk, ami viszont blokkos adatfeldolgozású, nagy válaszidejű rendszert eredményezne. Mind keskeny-, mind szélessávú szűrő esetén csupán a pozitív frekvenciakomponenseket vettem eddig figyelembe, aminek következménye, hogy a súlyvektor is komplex volt. Ahhoz, hogy a súlyvektor skalár legyen, a *D* mátrix előállításához nem csak a pozitív, hanem a negatív frekvenciákra is szükség van. A negatív frekvencia fizikailag ugyan értelmetlen, viszont létezése egy jel frekvenciatartománybéli matematikai leírásból következik. Mivel a negatív frekvenciakomponens a pozitív megfelelőjének komplex konjugáltja, ezért a (6.26)-os egyenletben szereplő mátrixszorzás során a *D* mátrix elemeinek imaginárius részei „kiesnek”, így a súlyvektor skalár lesz.

A *D* mátrix előállítása során tehát hasonlóan kell eljárni, mint az a 6.2.1.-es és 6.2.2.-es fejezetekben be lett mutatva, azzal a különbséggel, hogy negatív frekvenciakomponensekre is elő kell írni a kényszereket. A mátrix alakja egy egyszerű esetben, 2 db mikrofon

és mikrofononként 2 pontos FIR szűrő (azaz egy pozitív és hozzá tartozó negatív frekvenciát vizsgálva) esetén a következő:

$$D = \begin{bmatrix} e^{-j*0*\Delta\varphi_{1,\omega_0}} & e^{-j*0*\Delta\varphi_{1,\omega_1}} & e^{-j*0*\Delta\varphi_{2,\omega_0}} & e^{-j*0*\Delta\varphi_{2,\omega_1}} \\ e^{-j*1*\Delta\varphi_{1,\omega_0}} & e^{-j*1*\Delta\varphi_{1,\omega_1}} & e^{-j*1*\Delta\varphi_{2,\omega_0}} & e^{-j*1*\Delta\varphi_{2,\omega_1}} \\ e^{-j*0*\Delta\varphi_{1,\omega_0}-\omega_0*T_s} & e^{-j*0*\Delta\varphi_{1,\omega_1}-\omega_1*T_s} & e^{-j*0*\Delta\varphi_{2,\omega_0}-\omega_0*T_s} & e^{-j*0*\Delta\varphi_{2,\omega_1}-\omega_1*T_s} \\ e^{-j*1*\Delta\varphi_{1,\omega_0}-\omega_0*T_s} & e^{-j*1*\Delta\varphi_{1,\omega_1}-\omega_1*T_s} & e^{-j*1*\Delta\varphi_{2,\omega_0}-\omega_0*T_s} & e^{-j*1*\Delta\varphi_{2,\omega_1}-\omega_1*T_s} \end{bmatrix} \quad (6.29)$$

$$\Delta\varphi_{i,j} = \frac{d}{\lambda_j} * \sin(dirs_i) * 2\pi \quad (6.30)$$

$$\lambda = [\lambda_0, \lambda_1], \text{ ahol } \lambda_0 = -\lambda_1 \quad (6.31)$$

$$dirs = [\Theta_1, \Theta_2] \quad (6.32)$$

Megjegyzendő, hogy a *ref* vektor elemeinek megadása során ügyelni kell arra, hogy az összetartozó negatív és pozitív frekvenciakomponensek azonos értéket kapjanak, máskülönben a súlyvektor nem lesz skalár.

A következő fejezetben bemutatom, hogy az ebben a fejezetben taglalt algoritmus milyen eredményeket adott modellezett, illetve valós körülmények között. Kifejtem továbbá, hogy milyen nehézségek adódtak a *w* súlyvektor kiszámításához szükséges *D* mátrix invertálásakor, és, hogy milyen módszereket, ötleteket alkalmaztam ezek megoldására.

## 7. Eredmények

A rendszer tesztelése során a céloom az volt, hogy meghatározzam, hogy milyen mértékben képes az elkészült algoritmus rekonstruálni a bemenetet. Ehhez szükséges a dedikált jel lehető legkisebb torzítása, illetve rajta kívül minden zavarásnak a minél jobb elnyomása. A későbbi alkalmazások során a rendszer például emberi utasítások, szavak, beszédek jelrekonstrukcióját végezné, ezért tesztjeim során is ez volt a bemenet néhány kivételtől eltekintve.

A következő alfejezetekben be fogom mutatni, hogyan érdemes meghatározni a korábban definiált  $D$  mátrixot (6.2. fejezet), amely a kényszereket tartalmazza. Nem mindegy ugyanis, hogy milyen kényszereket adunk meg, noha elviekben tetszőleges szűrőt tervezhetünk. Ennek matematikai és rendszer felépítéséből adódó okai is vannak. Korábban, az 5. fejezetben már kifejtettem, hogy 5, illetve 7 db mikrofonnal sikerült méréseket végezni két különböző teremben. Ez egy erős korlátot szabott a valódi környezetben való tesztelés során, ugyanakkor még modellezett környezetben is szembesültem bizonyos korlátokkal. Először a keskenysávú szűrő tesztelését, és az ott levont következtetéseket fogom bemutatni, mivel egyszerűsége miatt itt látványosabbak az eredmények, illetve könnyebben levezethetők a problémák okai. Csak ezután fogom bemutatni dolgozatom fő témájának, a szélessávú iránykarakterisztikus szűrőnek az eredményeit.

### 7.1. Keskenysávú szűrő

Keskenysávú iránykarakterisztikus szűrő használatakor egy adott frekvenciát és annak szűk környezetét vagyunk képesek szűrni. Tesztjeim során különböző irányból érkező, különböző erősségű szinusz jeleket használtam modellezett és valós környezetben is. A szűrő feladata, hogy ezen bemenetek közül néhánynak a teljesítményét átengedje a jelalak torzítása nélkül, míg a nemkívántakét elnyomja. A (6.11)-es egyenletben szereplő invertálás műveletnek kitüntetett szerepe van, mely erős függésben van a mátrix tulajdonságaival. Az itt szereplő  $D$  mátrix oszlopai az egyes irányokhoz, sorai az egyes mikrofonokhoz tartoznak. Az mikrofonok elrendezéséből adódó egymáshoz képesti késleltetések, azaz fázisforgatások adóttak, viszont az irányokat úgy kell megválasztani, hogy azok az előbb ismertetett feladatot teljesítsék, illetve a mátrix tulajdonságai is megfelelő jószágúak legyenek. Ennek oka, hogy olyan mátrixoknak, amelyeknek determinánsa 0-hoz közelít és/vagy kondíciószáma nagy, azoknak az invertálása a számítási korlátok miatt pontatlan

lesz, az invertált mátrix elemeinek nagysága elszáll. Ekkor beszélhetünk arról, hogy a mátrix közel szinguláris. Munkám során a kondíciósám általánosan az egyik legfontosabb tulajdonságnak bizonyult. Empirikusan arra a következtetésre jutottam, hogy általában minél nagyobb a mátrix kondíciósáma, annál kisebb, annál inkább 0-hoz közelítő a determinánsa. A következő egyenletekben bemutatom, hogy a kis determináns miért jelent problémát esetemben. Legyen a példában szereplő rendszer egy 2 mikrofonból álló keskenysávú iránykarakterisztikus szűrő.

$$\Delta\varphi_1 = a \quad (7.1)$$

$$\Delta\varphi_2 = b \quad (7.2)$$

$$D = \begin{bmatrix} e^{-j*0} & e^{-j*0} \\ e^{-j*a} & e^{-j*b} \end{bmatrix} \quad (7.3)$$

$$\det(D) = e^{-j*b} - e^{-j*a} \quad (7.4)$$

$$\text{inv}(D) = \frac{1}{e^{-j*b} - e^{-j*a}} \begin{bmatrix} e^{-j*b} & -e^{-j*0} \\ -e^{-j*a} & e^{-j*0} \end{bmatrix} \quad (7.5)$$

$$\text{ref} = [1 \ 0] \quad 7.6$$

$$w_1 = \frac{e^{-j*b}}{e^{-j*b} - e^{-j*a}} \quad (7.7)$$

$$w_2 = \frac{1}{e^{-j*b} - e^{-j*a}} \quad (7.8)$$

A (7.7) és (7.8)-es képletekben látszik, hogy  $\Delta\varphi_1$  és  $\Delta\varphi_2$  fázisforgatások különbsége minél kisebb, annál nagyobbak lesznek a súlyvektor elemei. A különbség maximális értéke  $|\Delta\varphi_1 - \Delta\varphi_2| = \pi$  esetén lesz a súlyvektorok képletében szereplő nevező maximális, azaz ekkor kapunk kis értékeket a súlyvektor elemeire. Ez azért fontos, mert azokban az esetekben, amikor olyan irányból érkezik jel a mikrofonok felé, ahová vagy már nem tudtunk kényszerrel megadni, vagy az a priori tudásunk volt rossz a forrás irányára vonatkozóan, akkor a kimenet értéke a (6.20)-as egyenletben specifikáltak szerint alakul. Itt látható, hogy a hiba nagysága egyenesen arányos a súlyvektor értékével. Tehát ahhoz, hogy ne csak ideális, modellezett környezetben működjön az alkalmazás, ahol tökéletesek az a priori ismereteink, illetve csak néhány ismert irányból érkezik jelteljesítmény, úgy kell

meghatározni a  $D$  mátrixot az irányokra vonatkozóan, hogy annak kondíciószáma kicsi legyen, azaz minél közelebb tartózkodjon 1-hez, determinánsa pedig ne közelítsen a 0-hoz. Természetesen több mikrofon esetén nem ilyen egyszerű a helyzetünk. Egyszerre kell ügyelni a fázisforgatások közti különbségre, miközben a kívánt irányok felé mindenképpen elő kell írunk a kényszereket. A következő alfejezetekben modellezett illetve valóságos környezetben mutatom be a megfelelő iránykarakterisztikus szűrőket, illetve azok jóságát.

### 7.1.1. Modellezett környezet

Modellezett környezetben próbáltam a későbbi, valódihoz hasonló körülményeket teremteni. Ennek érdekében 10 cm-es mikrofontávolságot használtam, illetve a mikrofonok számát 7-nek választottam. Mivel a szűrő csak egy kiválasztott frekvencián működik, ezért a modellezett bemenetnek csak egyetlen frekvenciakomponense van, ahogy a zavarás is csak itt rendelkezik teljesítménnyel.

Az első esetben egy 1000 Hz-es szinusz érkezik a mikrofonokhoz  $-60^\circ$ -os irányból 0,8-as amplitúdóval, míg egy ugyanilyen teljesítményű szinusz zavarásként  $60^\circ$ -ból jön. 1000 Hz esetén a mikrofontávolság és a hullámhossz aránya:

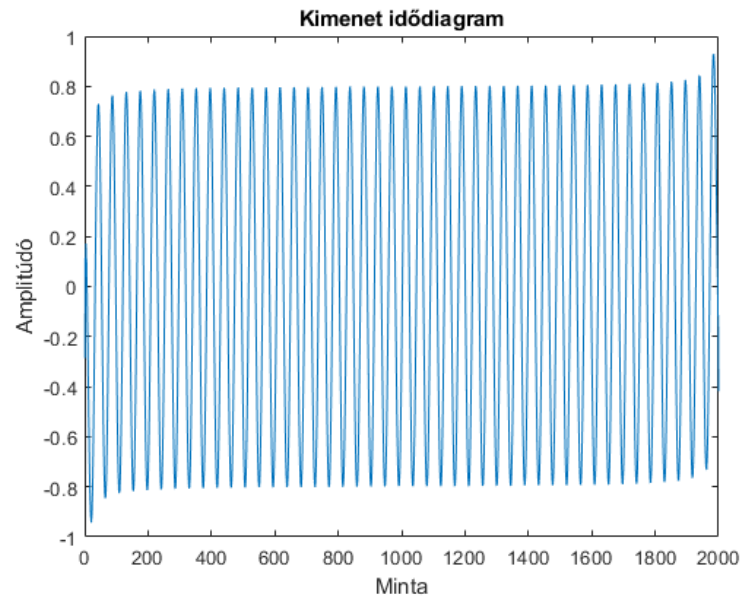
$$\frac{d}{\lambda} = \frac{d}{f/c} = 0,29 \quad (7.9)$$

Ez azt jelenti, hogy  $-90^\circ$  és  $90^\circ$ -os tartományban az elérhető maximális fázisforgatás két szomszédos mikrofon között:

$$\Delta\varphi_{max} - \Delta\varphi_{min} = \frac{d}{\lambda} * 2 * \pi * (\sin(90^\circ) - \sin(-90^\circ)) = 1,18 * \pi \quad (7.10)$$

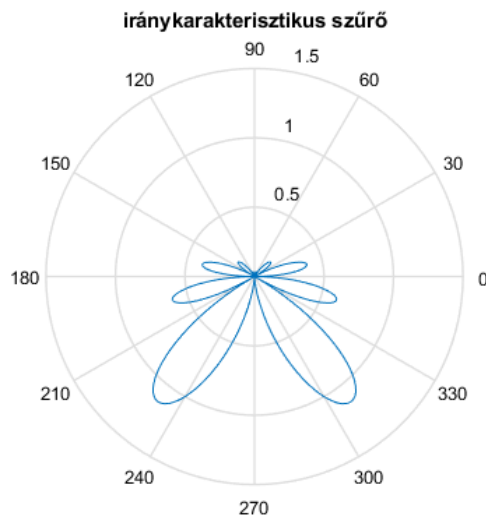
7 db mikrofon esetén 6 db szomszédos mikrofonpár van. Tehát a (7.10)-es egyenletben szereplő maximális különbséget 6 db szomszédos mikrofonpár fázistolásainak összegével érhetjük el. Az irányokat, melyekre a kényszereket előírjuk, tehát úgy érdemes meghatározni, hogy egymástól minél távolabb, egyenlő távolságban legyenek. Ennek két oka van. Egyrészt a nem specifikált irányokban ekkor a legkisebb az átvitel, azaz ekkor a legkisebb a melléknyalábok erősítése. Másrészt pedig ekkor lesz jól kondicionált a mátrix, valamint lesz 0-tól távoli a determinánsa (lásd (7.4)-es egyenlet). Tehát 7 mikrofon esetén ezen irányok esetén kapjuk a legjobban kondicionált mátrixot:  $-90^\circ$ ,  $-60^\circ$ ,  $-30^\circ$ ,  $0^\circ$ ,  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$ . Esetünk ebből a szempontból ideális, hiszen a jel  $-60^\circ$ -os, a zavarás pedig  $60^\circ$ -os

irányból érkezett. Ebben az esetben tehát ezekben az irányokban írtam elő a kényszereket, a *ref* vektort pedig úgy hoztam létre, hogy a dedikált bemenet irányából áteressen, minden máshonnan pedig maximálisan elnyomjon. A kimenet időfüggvénye, az átvitel karakterisztikája, impulzusválasza, illetve a bemenet és kimenet frekvenciaspektruma a következő ábrákon látható.

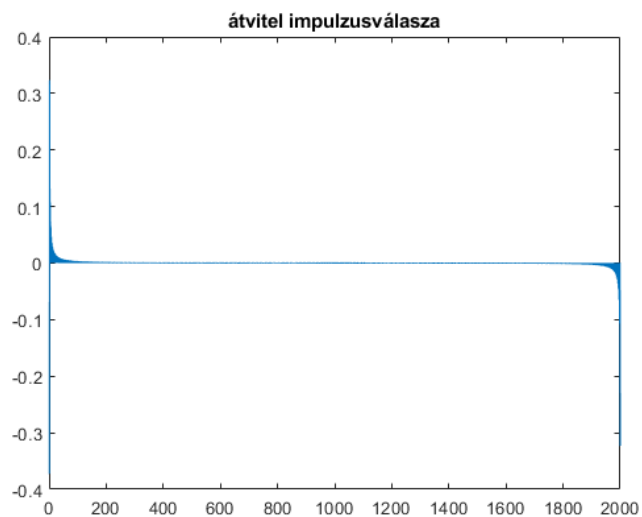


7-1. ábra, kimenet időfüggvénye

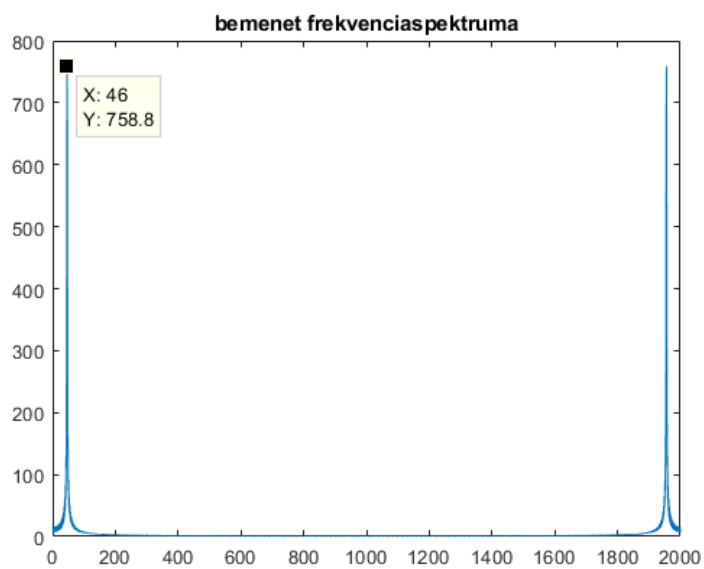
Látható, hogy a kimenet amplitúdója megegyezik a dedikált bemenet amplitúdójával.



7-2. ábra, iránykarakterisztikus szűrő átvitele

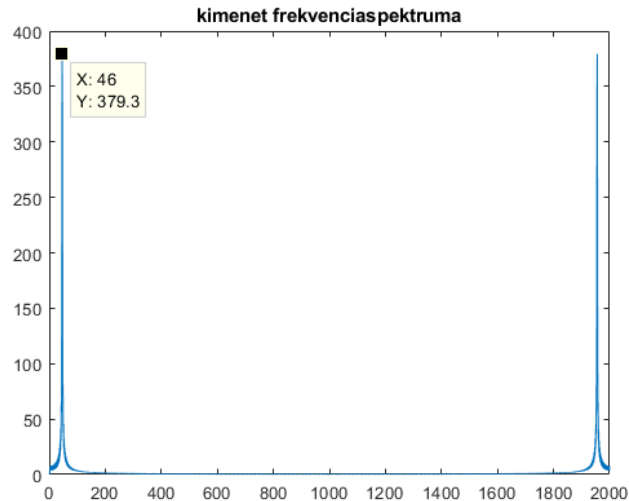


7-3. ábra, átvitel impulzusválasza



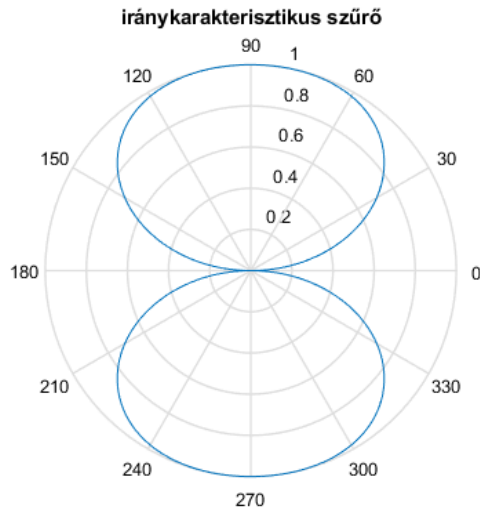
7-4. ábra, bemenet frekvenciaspektruma





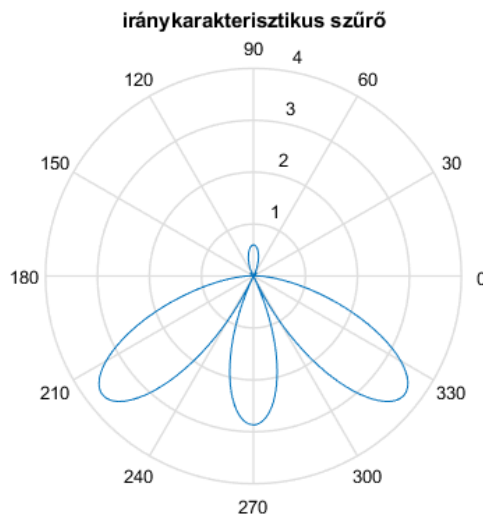
7-5. ábra, kimenet frekvenciaspektruma

A 7-4. ábra és 7-5. ábra mutatja, hogy a kimenet 1000 Hz-es frekvenciaspektrumon lévő amplitúdója megegyezött, mivel a dedikált szinuszforrással megegyező amplitúdójú, de más irányból jövő zavarást sikeresen kiszűrtem. A kimenet időfüggvényében látható szinuszjel amplitúdójának változása az átvitel impulzusválaszának következménye, ezért nem pontosan 0,8 a kimeneten lévő szinusz amplitúdója a teljes tartományon. A 7-2. ábra szemlélteti a szűrő átviteli karakterisztikáját. Láthatjuk, hogy nem teljesen tökéletes eredményt kaptunk, noha a kívánt irányokban a *ref* vektor szerinti átvitelt kapjuk.  $-50^\circ$  környékén láthatunk egy 1-nél nagyobb erősítést, mivel a főnyaláb közepe enyhén eltér a kívántól, valamint viszonylag kisebb melléknyalábok is megtalálhatóak. Ahhoz, hogy semmilyen melléknyalábot ne kapjunk, illetve a főnyalábunk is jó irányba mutasson, valamint maximuma 1 legyen, az szükséges, hogy két szomszédos irányhoz tartozó fázisforgatás különbsége  $\pi$  legyen. A következő ábra 2 mikrofon és  $\frac{d}{\lambda} = 0,5$ -ös arány mellett mutatja a szűrő átvitelét ( $90^\circ$ -os irányból átengedve,  $0^\circ$ -os irányból elnyomva a jelet):



7-6. ábra, szűrő iránykarakterisztikája

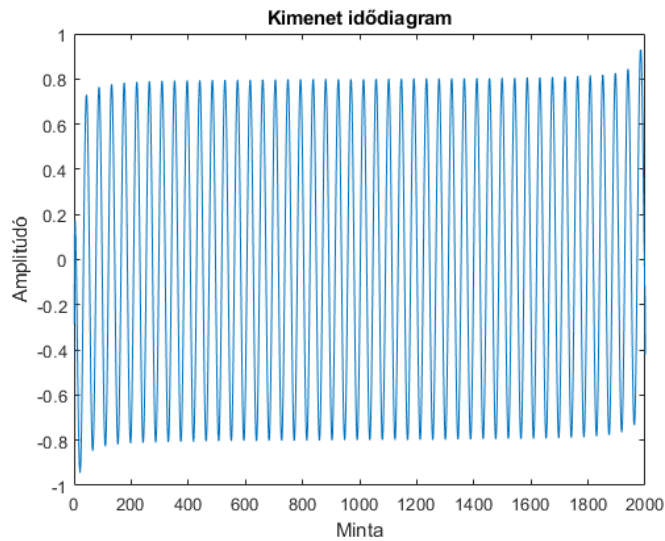
Az átvitel fenti alakja, a két főnyaláb szimmetrikussági okok miatt alakul ki. Abban az esetben, ha nem az alfejezet elején eljártak szerint írjuk elő a kényszereket, magyarán azokat az irányokat, amelyek felől nem érkezik jel, megváltoztatjuk, akkor a következő átvitelt kapjuk például ezen irányok használatával:  $-65^\circ$ ,  $-60^\circ$ ,  $10^\circ$ ,  $15^\circ$ ,  $20^\circ$ ,  $30^\circ$ ,  $60^\circ$ .



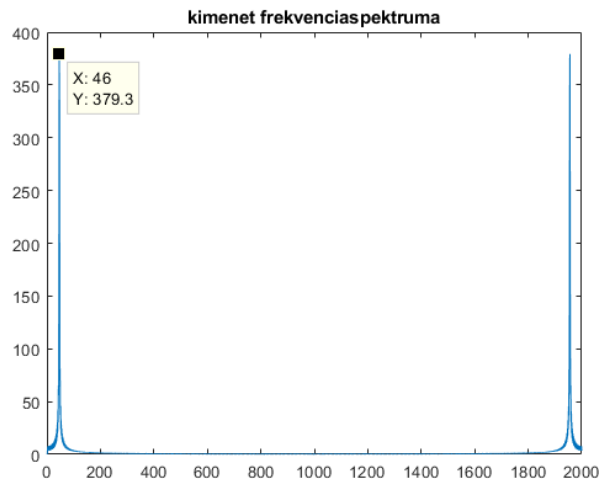
7-7. ábra, iránykarakterisztikus szűrő, nem ideális irányok

Látható, hogy  $-90^\circ$ -os irányban megjelent egy 3-as erősségű melléknyaláb, a főnyaláb rossz irányba mutat, erősítése közepén pedig majdnem 4. Az első, még ideális esetben a  $D$  mátrix kondíciószáma 21 volt, ebben az esetben viszont már 1421. Ugyanakkor jelen modellezett környezetben ez nem jelent problémát, hiszen csak két különböző irányból

érkezik hang, amelyekre az előírt kényszerek pontosan teljesülnek. Ezért a kimenet időfüggvénye és spektruma változatlan.



7-8. ábra, kimenet időfüggvény, nem ideális eset



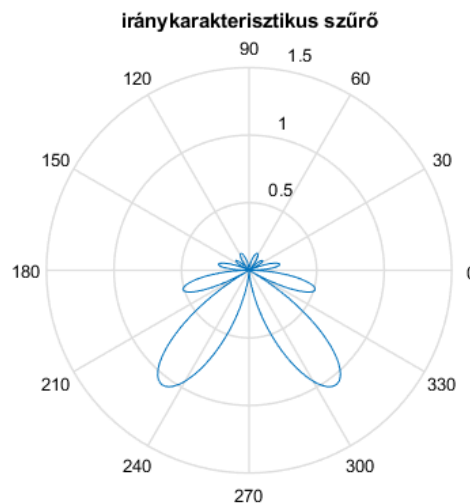
7-9. ábra, kimenet spektruma, nem ideális eset

A következő alfejezetben olyan eseteket mutatok be, amikor valódi felvételek készültek, ahol a dedikált forrás és a zavarás is 1000 Hz-es szinusz jel volt, ugyanakkor természetesen zaj és reflexió is jelen volt a térben.

### 7.1.2. Valós környezet

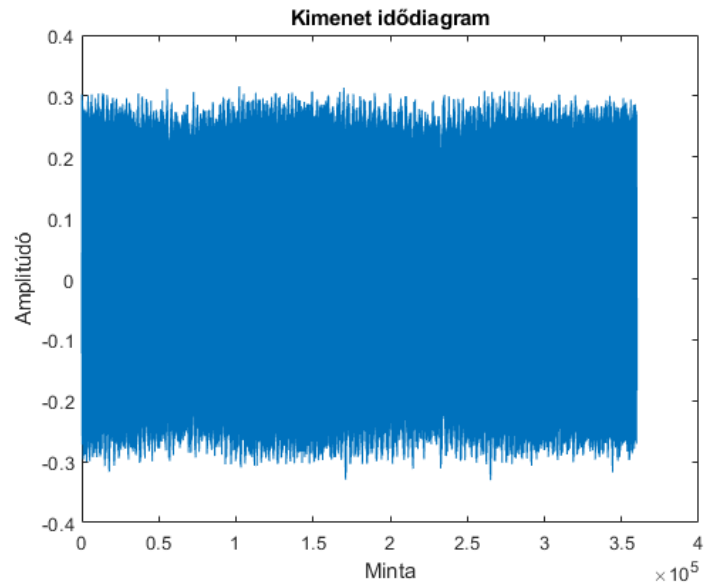
Valós környezetben is készítettem méréseket keskenysávú iránykarakterisztikus szűrő tesztelésére. A modellhez hasonlóan itt is Audacity segítségével állítottam elő 0,3-as amplitúdójú, 1000 Hz-es szinusz jelet két csatornára. Mindkét csatorna jelét egy-egy hangszóró bemenetére vezettem hangkártya segítségével, amíg ugyanezen hangkártyával

a 7 db, vonalban egymástól egyenlő távolságban elhelyezett mikrofonok jeleit vételeztem. A méréseket ebben az esetben csak a nagyobb, reflexióra hajlamosabb teremben végeztem el. A mérések vizsgálata hasonló eredményre vezetett minden esetben, egy tipikus eredmény a következő volt. Szinuszos jel érkezett a mikrofonok felé  $-58^\circ$ -os irányból, ezt tekintettem ekkor dedikált forrásnak, illetve  $48^\circ$ -os irányból, ezt pedig zavarforrásnak. Az eddig ismertettek alapján, ekkora mikrofonszám esetén érdemes úgy megválasztani az irányokat, amelyekre a kényszereket előírjuk, hogy a szomszédosok minél egyenletesebb távolságban legyenek egymástól. Ennek alapján a következő irányokra írtam elő átvitelt:  $-90^\circ$ ,  $-58^\circ$ ,  $30^\circ$ ,  $0^\circ$ ,  $25^\circ$ ,  $48^\circ$ ,  $90^\circ$ . A  $-58^\circ$ -ból érkező jel teljesítményét átengedtem, minden más irányból érkezőt pedig elnyomtam, a kialakult iránykarakterisztikus szűrő a következő ábrán látható.

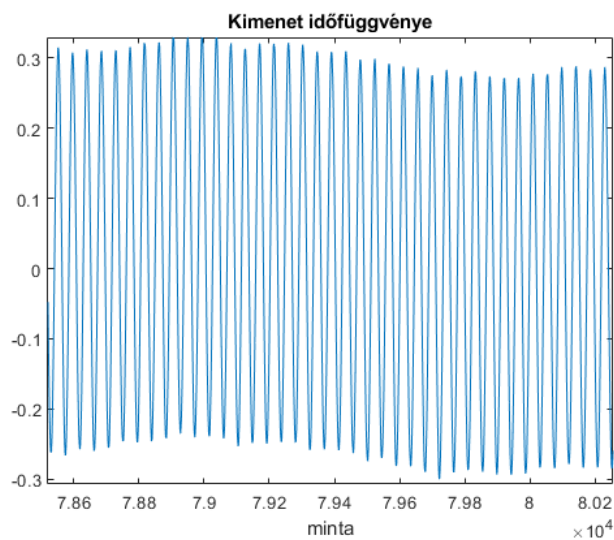


7-10. ábra, iránykarakterisztikus szűrő átvitele, valós eset

Látható, hogy a főnyaláb iránya és nagysága megfelelő, nem találhatóak nagy erősítésű melléknyalábok az átvitelben.



7-11. ábra, kimenet időfüggvény, valós eset



7-12. ábra, kimenet időfüggvény nagyítva, valós eset

A 7-11. ábra és 7-12. ábra mutatja a kimenet időfüggvényét. Itt megfigyelhető, hogy el-  
 lentétben a modellezett esettel, zajos a kimenet, hiszen a bemenet is zajos volt. A kimenet  
 amplitúdója 0,3 alatti egy kicsivel, azaz csak a dedikált bemenet jelenik meg a kimeneten.  
 A következő alfejezetben bemutatom a szélessávú szűrő működését, illetve az ottani fel-  
 merülő nehézségeket a kényszerek felírása közben.

## 7.2. Szélessávú szűrő

Diplomatervem során célom az volt, hogy szélessávú szűrés segítségével rekonstruáljam a dedikált emberi beszédforrást, elnyomjam más források, zavarások teljesítményét, illetve a reflexiók hatását minél jobban csökkentsem. Az 5. fejezetben kifejtettem, hogy a szűrést 600 Hz és 3000 Hz közti frekvenciákon végeztem. Ennek oka, melyet korábban is kifejtettem, hogy az emberi beszéd teljesítményének nagy része ebben a tartományban található, illetve a mikrofonok elrendezése szempontjából is előnyös ez a tartomány, hiszen fizikai korlátok miatt ennél közelebb nemigen tudtam volna őket elhelyezni egymáshoz. A 6.2.2.-es fejezetben kifejtettem a szűrő modelljét, annak felépítését. A modell szerint szélessávú szűrés során minden egyes irányhoz hozzárendelhetünk egy időtartománybeli FIR szűrőt, melyek lehetnek egyediek is, nem szükséges minden mikrofon jelét ugyanazzal a FIR szűrővel szűrni. A 6.2.3.-as fejezetben leírtam, hogy ahhoz, hogy a későbbi alkalmazás a mikrofonok valós bemenetét dolgozhassa föl, szükséges a pozitív frekvenciák mellett a negatívakra is előírni a kényszereket. Ebben az esetben a (6.27)-es egyenletben szereplő súlyvektor elemei valós értéket vesznek fel. Ehhez az is szükséges, hogy a negatív és pozitív frekvencia párokra ugyanakkora átvitelt írjunk elő az egyenletben szereplő *ref* vektorban. Ezt a továbbiakban nem fogom hangsúlyozni, hanem minden esetben nyilvánvalónak fogom venni, hogy az említett pozitív frekvenciák mellett azok negatív párjaira is előírtam az adott példában az átvitelt.

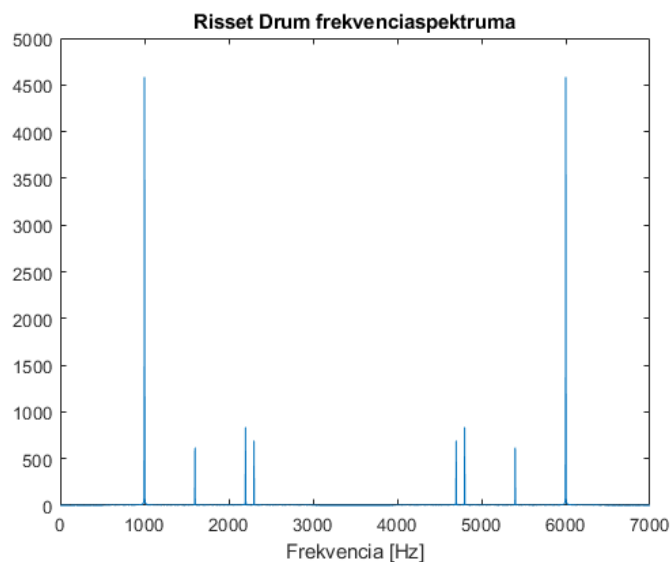
Ahogy keskenysávú szűrő esetén sem volt mindegy, hogy milyen irányokban írom elő a kényszereket az azokat tartalmazó mátrix invertálhatósága miatt, úgy itt is található néhány szabály. Esetemben, 5 vagy 7 mikrofon használatával, illetve  $-90^\circ$  é  $90^\circ$  közötti szögterülettel cél, hogy azokat az irányokat, amik nem tartoznak közvetlen hangforráshoz, úgy állapítsam meg, hogy két szomszédos irányok minél egyenlőbb távolságban legyenek egymástól, ahogyan ezt a 7.1.-es fejezetben kifejtettem. Ezen kívül jelen esetben fontos meghatározni azt is, hogy a használt frekvenciatartományt milyen felbontással érdemes vizsgálni. Munkám során azt tapasztaltam, hogy a mátrix kondíciójának jósága szempontjából 3-400 Hz-ig érdemes csökkenteni a felbontást. Ennél kisebb felbontás esetén a mátrix nem lesz elég jól kondicionált, és a kapott súlyvektor elemeinek értéke végesen növekedni kezd, amely torzítja a bemenetet. Az alkalmazásban a 600 Hz és 3 kHz közötti tartományon kívül eső részeket a Matlab *firls* parancsával, azaz legkisebb négyzetek szerinti FIR szűrővel szűrtem ki a jelből. Mivel így 3 kHz felett nem található jelentős teljesítmény, ezért a felvétel során használt, 44,1 kHz-es mintavételi frekvenciát

sem kötelező továbbra is használni, amit én a későbbiekben 7 kHz-re csökkentettem újramintavételezéssel.

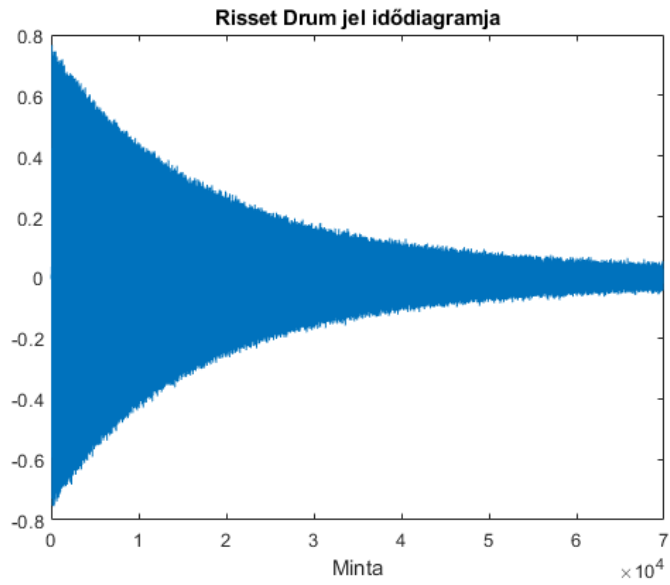
A következő alfejezetekben a modellezett és a valós környezet során szerzett tapasztalataimat mutatom be a szélessávú iránykarakterisztikus szűrővel.

### 7.2.1. Modellezett környezet

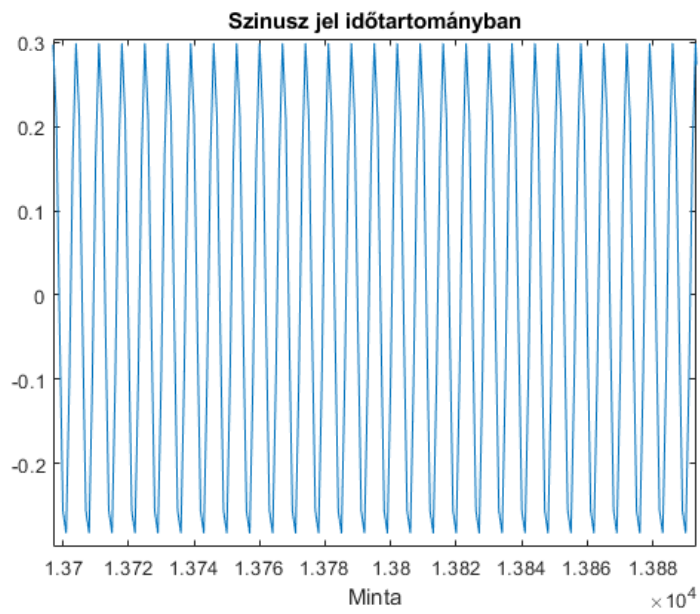
Mivel a méréseket 10 cm-es mikrofontávolsággal, illetve 5 vagy 7 db mikrofon segítségével végeztem, ezért a modellben is ezeket a beállításokat használtam. Három különböző zavarással teszteltem a szűrő működését modellezett környezetben, 1000 Hz-es szinuszzal, Risset Drum jellel [17], illetve emberi beszéddel. A Risset Drum jel egy nagyteljesítményű lecsengő alapharmonikuson kívül tartalmaz néhány, kisebb amplitúdójú, hasonlóan lecsengő enharmonikust, illetve viszonylag széles tartományban kis teljesítményű zajt is. A következő ábrákon a zavarjelek figyelhetőek meg.



7-13. ábra, Risset Drum jel frekvenciaspektruma

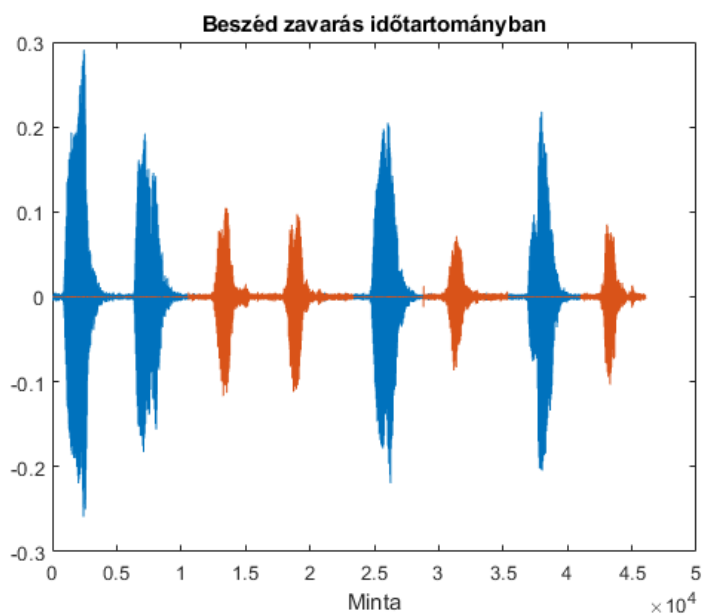


7-14. ábra, Risset Drum jel idődiagramja

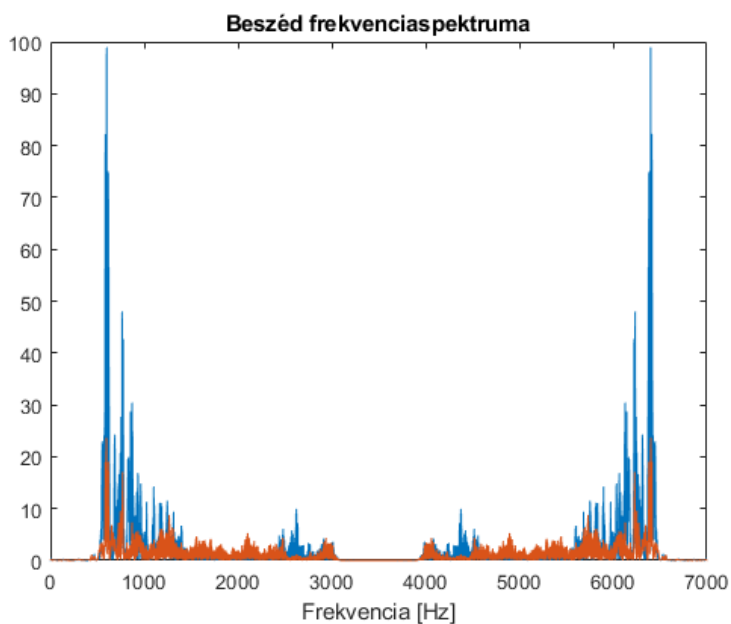


7-15. ábra, 1000 Hz-es színusz idődiagramjának részlete





7-16. ábra, beszéd idődiagramja két különböző csatornán, egyik csatorna dedikált forrás, másik csatorna zavarforrás



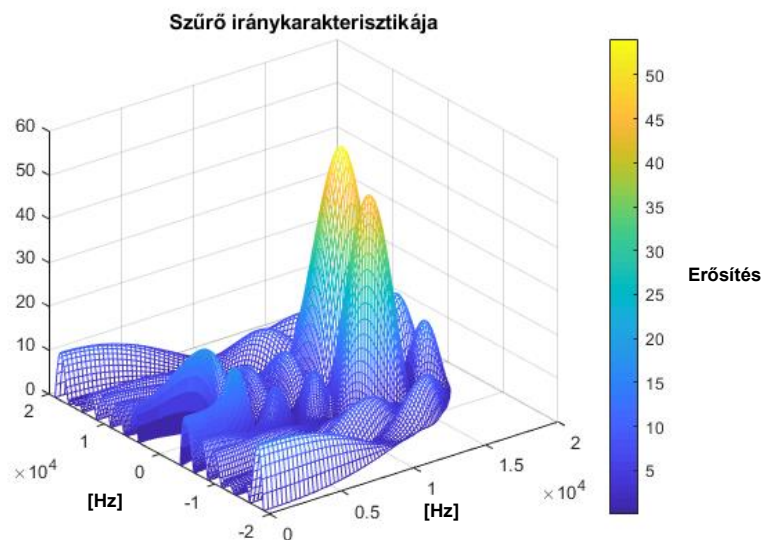
7-17. ábra, beszéd frekvenciaspektruma két különböző csatornán, egyik csatorna dedikált forrás, másik csatorna zavarforrás

A 7-16. ábra és 7-17. ábra egy sztereó mikrofonokkal felvett beszédet mutat (szöveg: „bal, bal, jobb, jobb, bal, jobb, bal, jobb”). A felvételtől komplementer részeket töröltem ki Audacity segítségével, és csendet illesztettem be helyettük. Az iránykarakterisztikus

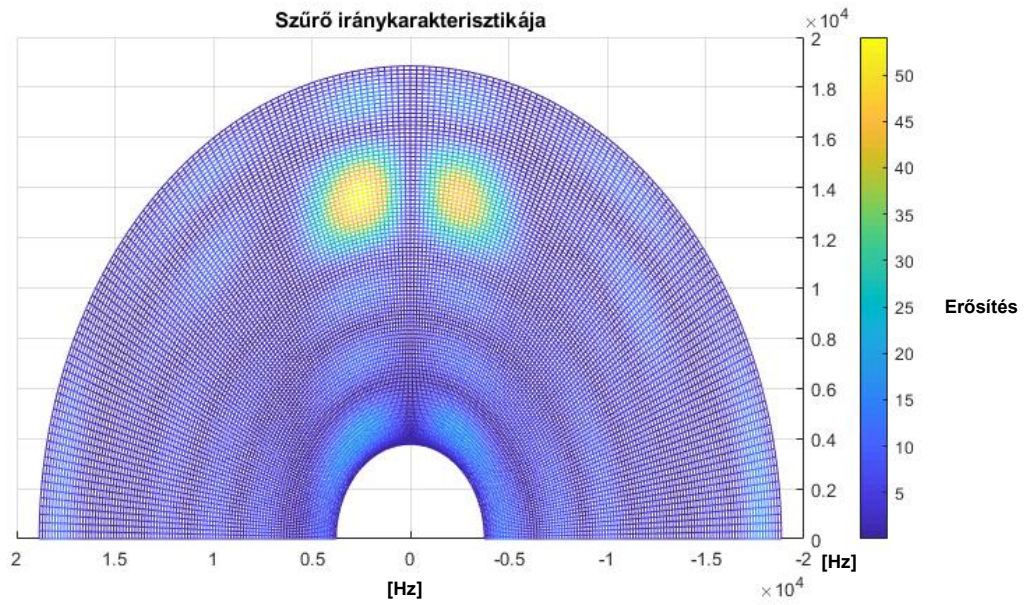
szűrő tesztelésekor az egyik csatornát az egyik hangszóróra, másikat pedig a másokra kevertem, ezek kimenetét vettem fel 5 vagy 7 db mikrofonnal a mérések alkalmával, illetve rekonstruáltam az egyik csatorna jelét, míg a másikat elnyomtam.

Modellezett környezetben való tesztelés során nem találtam számottevő különbséget 5, illetve 7 mikrofonos rendszer között. Mivel két hangforrással teszteltem a rendszert, ezért könnyedén be tudtam állítani a kényszereket az irányokra vonatkozóan mindkét esetben. A továbbiakban minden esetben 7 mikrofonos modellt fogok prezentálni.

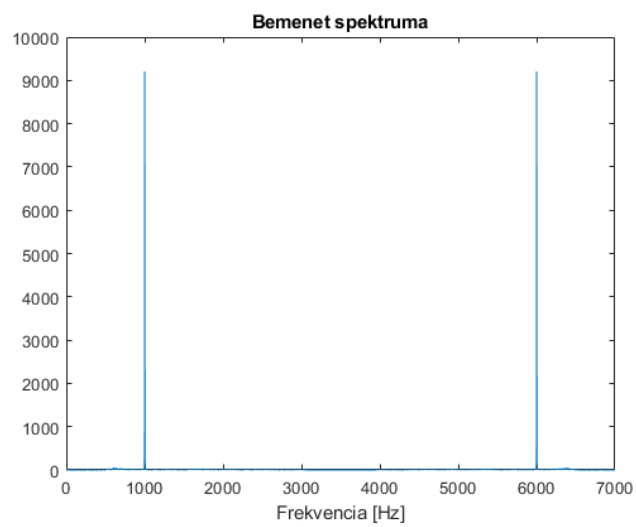
A következő példában 1000 Hz-es, 0,3-as amplitúdójú szinusz jel volt a zavarforrás  $-30^\circ$ -os irányban, a dedikált forrás, azaz a beszéd  $45^\circ$ -os irányból érkezett. Specifikált irányok:  $-90^\circ$ ,  $-60^\circ$ ,  $-30^\circ$ ,  $0^\circ$ ,  $30^\circ$ ,  $45^\circ$ ,  $90^\circ$ , specifikált frekvenciák minden irányban: 600 Hz és 3 kHz között 400 Hz-es felbontással).



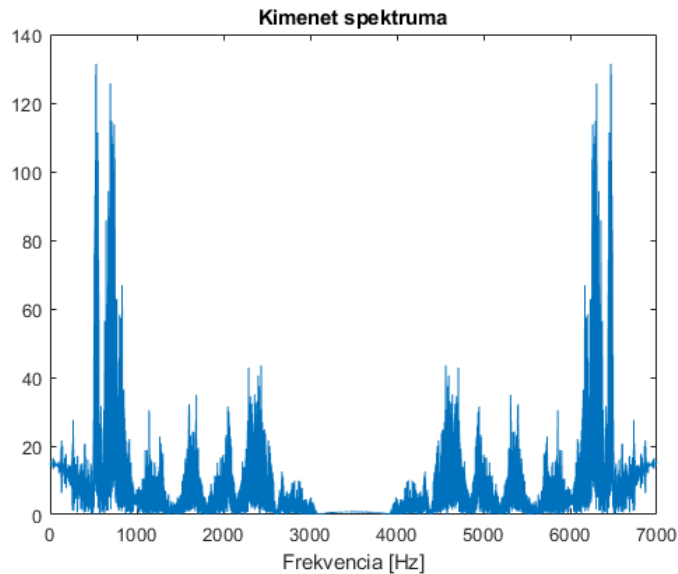
7-18. ábra, iránykarakterisztikus szűrő átvitele



7-19. ábra, iránykarakterisztikus szűrő átvitele, felülnézet



7-20. ábra, Bemenet spektruma, szinusz jel és beszéd

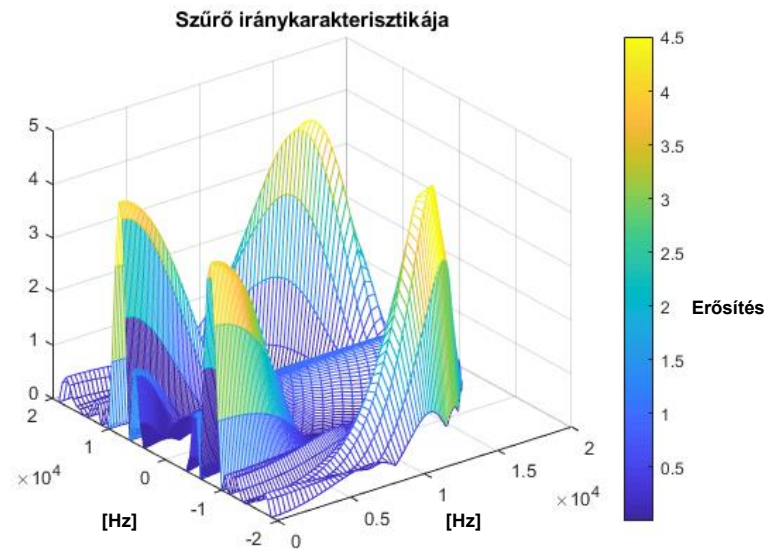


7-21. ábra, kimenet spektruma

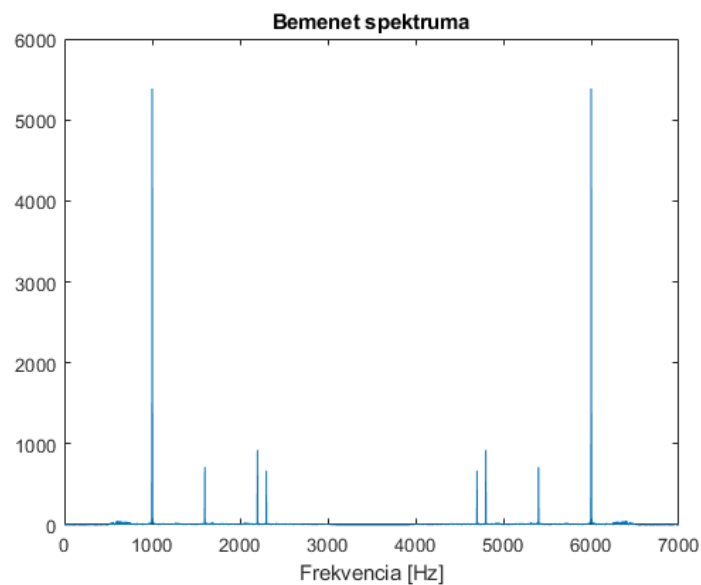
A 7-18. ábra és 7-19. ábra (felülnézeti kép) szemlélteti a szűrő átvitelét. A diagram koncentrikus „félkörökből áll”. A sugár mentén belülről kifelé haladva 25 Hz-es felbontással láthatóak az egyes frekvenciákon az átvitelek. Egy adott frekvencián, azaz egy „félkörön”  $-90^\circ$  (diagram bal fele) és  $90^\circ$  (diagram jobb fele) közötti tartományon fokenkénti felbontással ábrázoltam a tér irányait. A félkörök egyes pontjainak z-koordinátája adja meg az adott helyen az erősítést. Szembeötlő, hogy található benne 40-50-szeres erősítés is. Sajnos ez a jelenség minden olyan esetben előfordult szélessávú szűrő kapcsán, ahol a jelet nem  $0^\circ$ -os irányból engedtem át, hanem valamilyen szélső irányból. A hasonló túllövések nagysága körülbelül arányosan csökken a kényszereket tartalmazó mátrix kondicionáltságának javulásával. Jelen esetben  $10^3$  nagyságrendbe esett a mátrix kondíciószáma, ami a jel rekonstruálása szempontjából, hallhatóság szempontjából még megfelelőnek számít. A bemenet és kimenet spektrumának vizsgálatakor észrevehetjük, hogy az 1 kHz-es zavarás teljes mértékben eltűnt, köszönhetően a 600 és 3000 Hz közötti, 400 Hz-es felbontású mikrofonkénti FIR szűrőnek az elnyomott irányokban. Ugyanakkor olyan frekvenciákon és olyan irányokban, ahová nem írtam elő kényszereket, ott az átvitel szerint erősítések jönnek létre, melyek a spektrumok ábráin is láthatóak. Ilyen mértékű torzítás még nem befolyásolja jelentősen a szöveg érthetőségét, mely meghallgatható a csatolt fájlok között (bemenet: Modell\_Szinusz\_Bemenet.wav, kimenet: Modell\_Szinusz\_Kimenet.wav).

A következő példában a dedikált forrás középen helyezkedik el, azaz  $0^\circ$ -os irányból átengedtem a jelet, minden más irányból pedig elnyomtam azt. A zavarforrás a Risset Drum

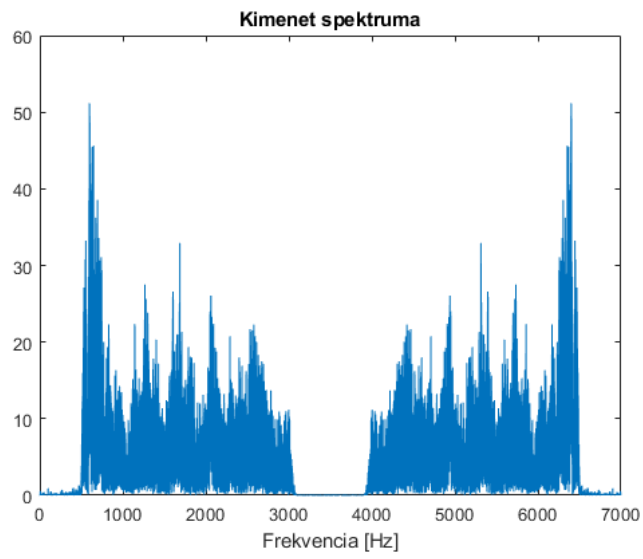
jel volt a teszt során, mely  $55^\circ$ -os irányból érkezett a mikrofonok felé, a következő irányokban írtam elő kényszereket:  $90^\circ$ ,  $55^\circ$ ,  $30^\circ$ ,  $0^\circ$ ,  $-30^\circ$ ,  $-60^\circ$ ,  $-90^\circ$ . Olyan irányokban, ahol elnyomom a teljesítményt, a következő frekvenciákra határoztam meg a kényszereket: 600, 1000, 1600, 1800, 2200, 2300, 3000 Hz. A Risset Drum jel harmonikus frekvenciái 1000, 1600, 2200 és 2300 Hz-en találhatóak meg, a többi frekvenciát úgy választottam meg, hogy nagyjából egyenletesen fedjék le a köztes tartományokat. A kényszereket tartalmazó mátrix kondíciószáma jelen esetben is  $10^3$ -os nagyságrendbe esett.



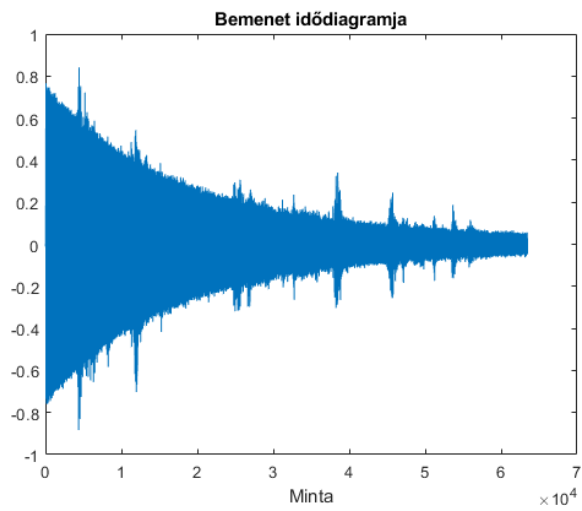
7-22. ábra, iránykarakterisztikus szűrő átvitele



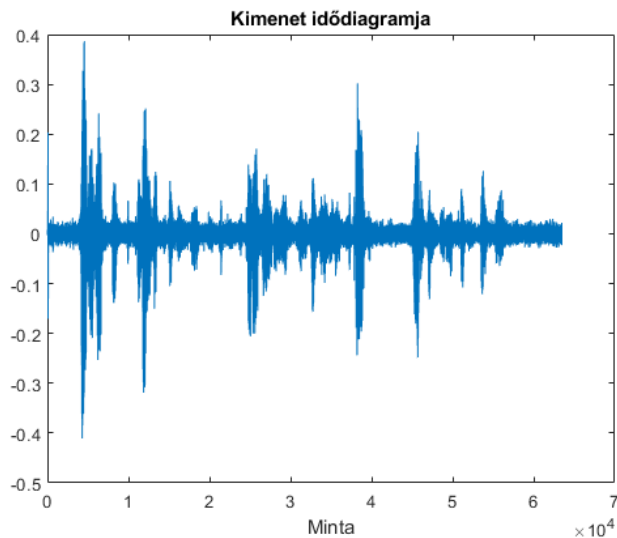
7-23. ábra, bemenet frekvenciaspektruma



7-24. ábra, kimenet frekvenciaspektruma



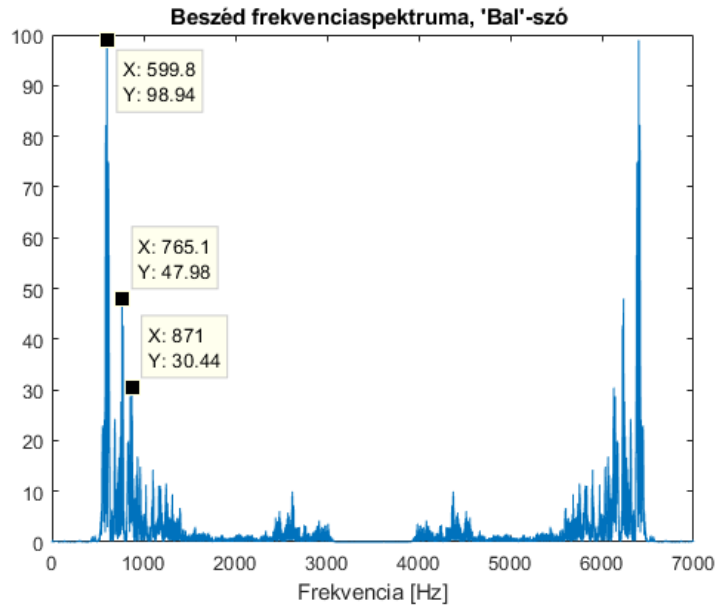
7-25. ábra, bemenet idődiagramja



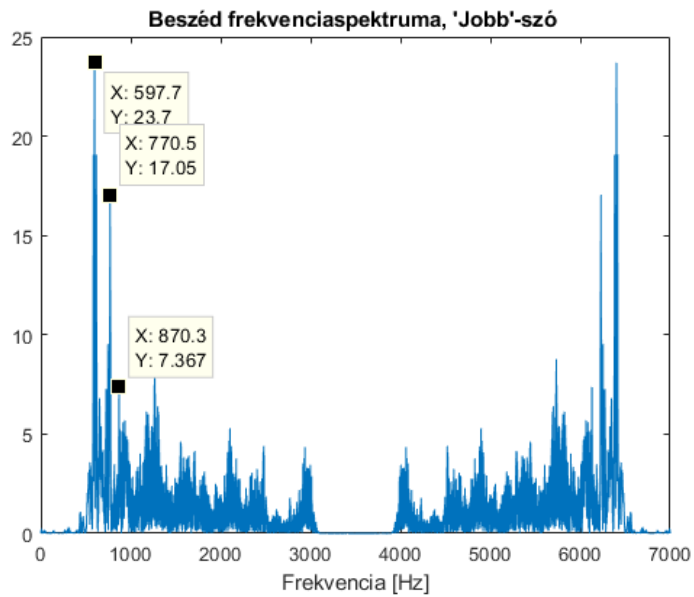
7-26. ábra, kimenet idődiagramja

A szűrő átviteli diagramján láthatjuk, hogy jelen esetben már csak legfeljebb 4-szeres erősítésű pont található az átvitelben. A frekvenciaspektrumok megtekintésekor észrevehetjük, hogy a Risset Drum jel harmonikus frekvenciáinak teljesítménye eltűnt a kimenetből. A hangfájlokat meghallgatva hallható, ahogy a bemenetben a lecsengő harmonikusok után hallható szélessávú zaj megmaradt a kimenetben is, ezt nem sikerült kiszűrni a jelből. (A bemenet és a kimenet meghallgatható az alábbi csatolt fájlokban: bemenet: Modell\_Risset\_Bemenet.wav, kimenet: Modell\_Risset\_Kimenet.wav).

Modellezett környezetben teszteltem azt is, hogy milyen mértékben lehetséges két beszédforrás közül az egyiket torzítatlanul venni, míg a másikat elnyomni. Jelen esetben a dedikált forrás  $0^\circ$ -os irányból érkezett a mikrofonok felé, míg a zavarforrás  $75^\circ$ -os szögben helyezkedett el. A két szöveg jobb megismeréséhez előbb nézzük meg azok frekvenciaspektrumait.



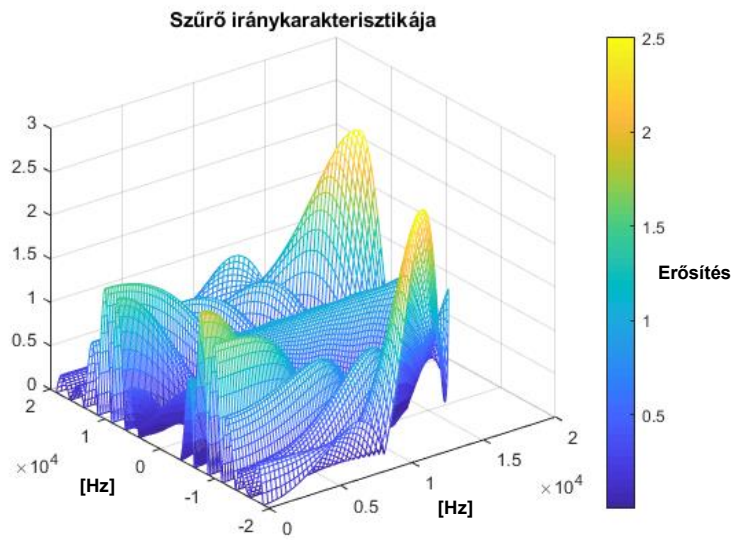
7-27. ábra, „bal” szó frekvenciaspektruma, zavarforrás



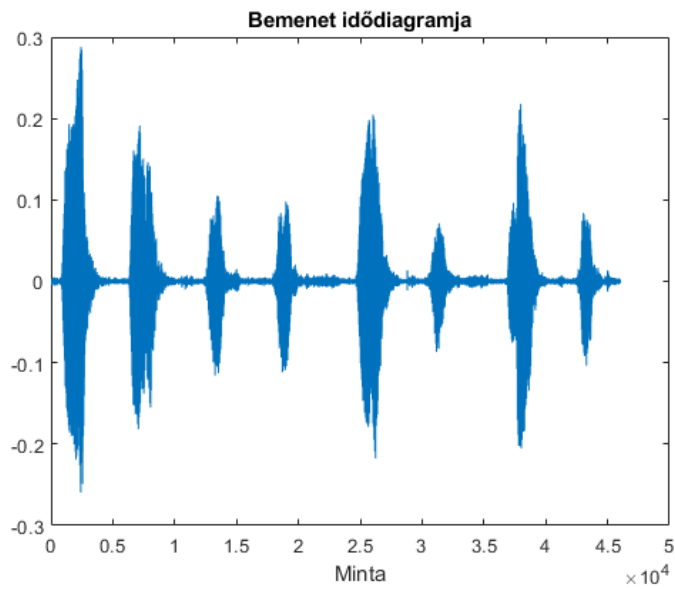
7-28. ábra, „jobb” szó frekvenciaspektruma, dedikált forrás

A felső ábrán látható spektrum a „bal” szó többszöri ismétlése alapján generálódott, az alsó hasonlóan, csak a „jobb” szó ismételtetésével. Mivel mindkét frekvenciaspektrum az én beszédemet tartalmazza, ezért a jelentős teljesítményű frekvenciakomponensek fizikai okból kifolyólag azonosak. Jelen helyzetben egy olyan szűrőt kellett terveznem, mely minden frekvencián elnyom  $75^\circ$ -os irányból,  $0^\circ$ -os irányból viszont torzítatlanul veszi a jelet. Ehhez a 600 és 3000 Hz közti tartományt egyenletesen osztottam fel, 400 Hz-es felbontással.

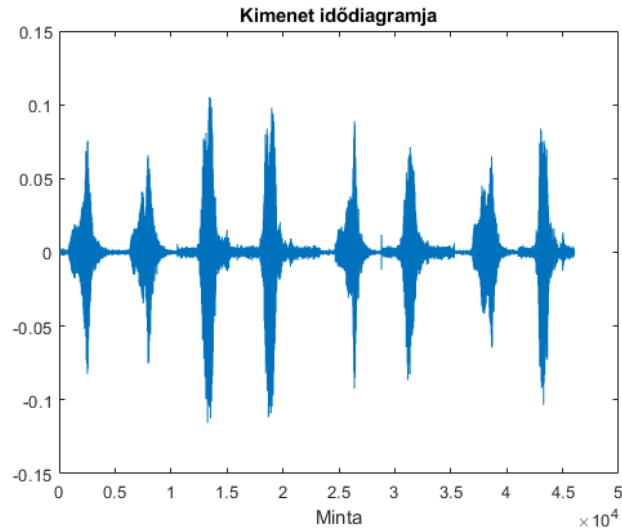




7-29. ábra, iránykarakterisztikus szűrő átvitele

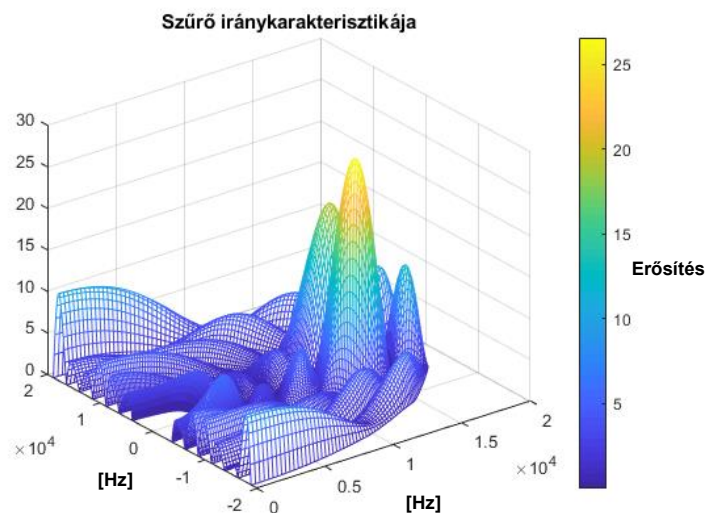


7-30. ábra, bemenet idődiagramja

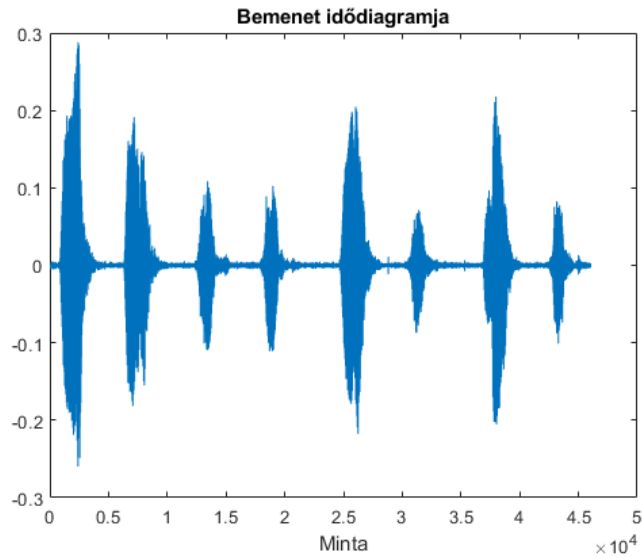


7-31. ábra, kimenet idődiagramja

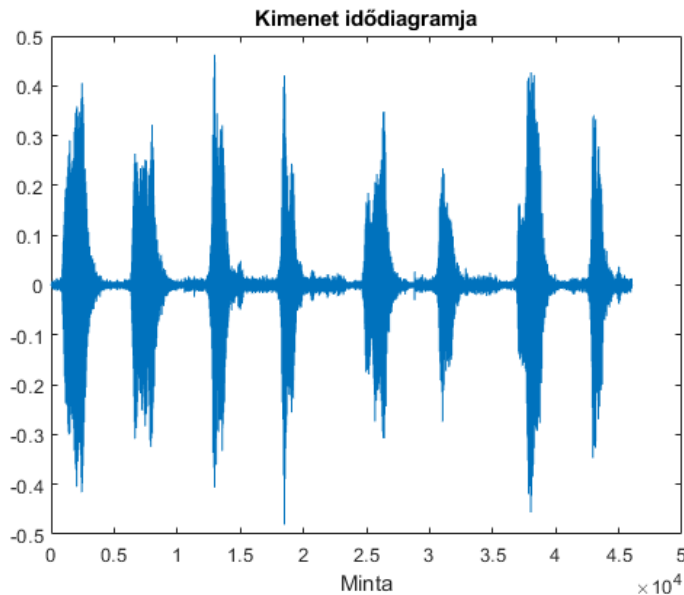
Látható, hogy jelen esetben, amikor az elnyomott irányokban a frekvenciák egyenletesen vannak elosztva, akkor nincsen nagy túllövése az átvitelnek. Az idődiagramokon a következő szöveg van felvéve: „bal, bal, jobb, jobb, bal, jobb, bal, jobb”. A „bal” szó minden esetben  $75^\circ$ -os irányból, a „jobb” szó pedig középről érkezik. Látható, hogy az elnyomandó jel amplitúdóját körülbelül negyedére sikerült lecsökkenteni anélkül, hogy a dedikált jel jelentősen torzult volna (a hangfájlok meghallgathatók a csatolt fájlok között: bemenet: Modell\_Beszed\_Bemenet.wav, kimenet: Modell\_Beszed\_Kimenet.wav). Abban az esetben, ha a dedikált jel nem középről, hanem mondjuk  $-30^\circ$ -os irányból érkezik, akkor a következő eredményt kapjuk (hangfájlok meghallgathatók: bemenet: Modell\_Beszed2\_Bemenet.wav, kimenet: Modell\_Beszed2\_Kimenet.wav).



7-32. ábra, iránykarakterisztikus szűrő átvitele



7-33. ábra, bemenet idődiagramja



7-34. ábra, kimenet idődiagramja

Látható, hogy ekkor megint csak megjelennek 10-es nagyságrendbe eső erősítési pontok az átvitelben, melyek miatt a kimenet torzított lesz. Az elmondott szavak még mindig jól érthetők, viszont az idődiagramokon látható, hogy a „bal” szavak teljesítménye nem csökkent, hanem inkább növekedett némileg, továbbá a „jobb” szavak teljesítménye is nagymértékben nőtt.

Modellezett környezetben elmondható, hogy olyan esetekben, ahol néhány harmonikus frekvencia adja a zavarforrás teljesítményének nagy részét, az iránykarakterisztikus szűrő megfelelően nyomja el a zavarforrást, és rekonstruálja az emberi beszédet, azok irányától

függetlenül. Ugyanakkor, ha a zavarforrás is emberi beszéd, akkor mindenképpen szükséges, hogy a dedikált forrás a mikrofonokhoz képest középen helyezkedjen el, mivel ekkor a zavarjel amplitúdója negyedére csökkenthető, mely már hallható különbséget jelent.

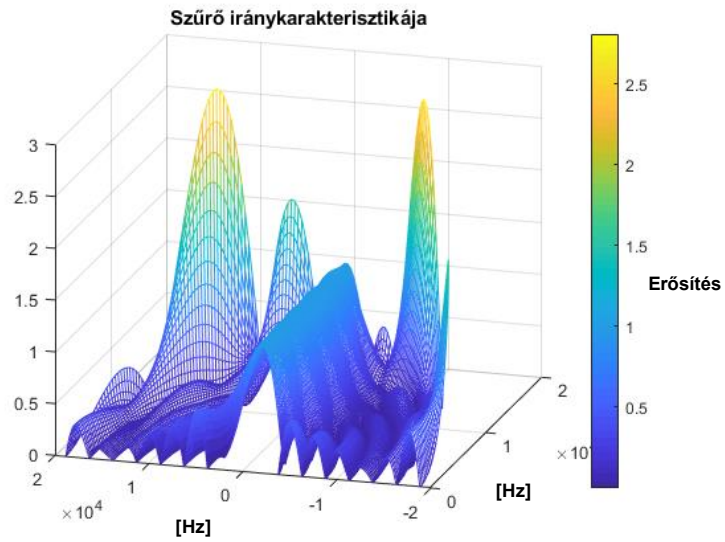
A következő alfejezetben bemutatom azokat az eredményeket és következtetéseket, amelyeket valós környezetben való tesztelés közben szereztem.

### **7.2.2. Valós környezet**

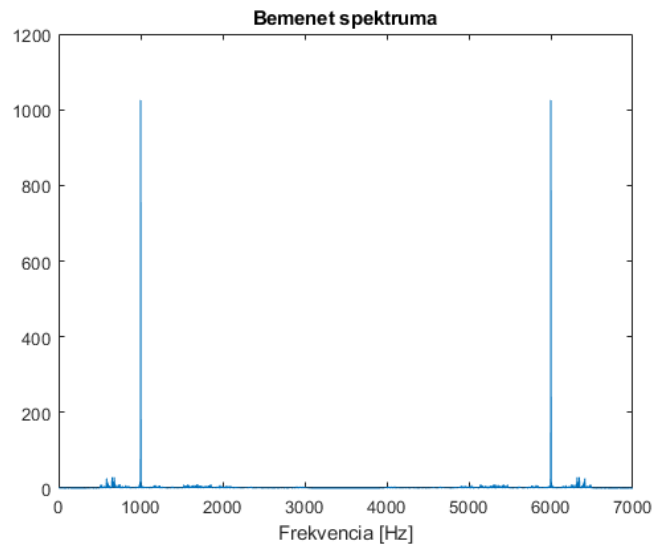
Valós környezetben sikerült méréseket végezni egy kisebb, kevésbé visszhangos teremben 5 db vonalban elhelyezett mikrofon segítségével, illetve egy nagyobb, visszhangosabb teremben 7 db mikrofonnal (lásd 5. fejezet). Diplomatervem során sajnos csak kétféle különböző zavarforrás hatását tudtam kipróbálni méréseim alkalmával, 1000 Hz frekvenciájú szinuszhullámot, illetve azt, amikor egy második beszédforrás volt jelen a dedikált mellett. Különbség a modellezett esethez képest, hogy itt nem Matlab segítségével állítom elő az egyes mikrofonok egymáshoz képest késleltetett jeleit, hanem valós felvételeket használtam. Modellezett környezetben nem volt semmiféle egyéb háttérzaj a zavarforráson kívül, valamint nem voltak visszavert hanghullámok sem. A következő példákban látható lesz, hogy valós esetben a visszavert, reflektált hullámokat nem tudtam teljesen kiszűrni a jelből. Ennek egyik oka, amit a fejezet korábbi részében kifejtettem, hogy ha túl sok kényszert írok elő, azaz túl sok irányban és/vagy túl sok frekvencián kívánom megadni az átvitelt, akkor az ezeket tartalmazó mátrix rosszul kondicionált lesz, valamint 0-hoz közeli determinánssal fog rendelkezni. Ennek nem kívánt hatása, hogy a súlyvektor skalár elemei nagy értéket fognak felvenni, ami olyan irányokban és frekvenciákon, melyekre nem írtam elő átvitelt, kiemelés fog eredményezni. Ez a mátrix kondicionáltságának függvényében akár szingularitási problémát is okozhat invertáláskor. Emiatt az általam használt mikrofontömbök esetén körülbelül 6-8 db különböző frekvencián tudom még úgy előírni az átvitelt, hogy az ne okozzon akkora kiemelés, mely a dedikált jelet túlzottan eltorzítaná. A következő példák során minden esetben ezek a frekvenciák a 600 és 3000 Hz közötti tartomány 400 Hz-es felbontásának elemei voltak. Ez a felbontás megfelelt az 1000 Hz-es szinuszos, és a beszéd zavarforrások esetén is.

Az első példában  $57^\circ$ -os irányból érkezett egy 1000 Hz-es, 0,3-as amplitúdójú szinuszhullám zavarásként, a dedikált forrás, azaz a beszédforrás középen helyezkedett el. A

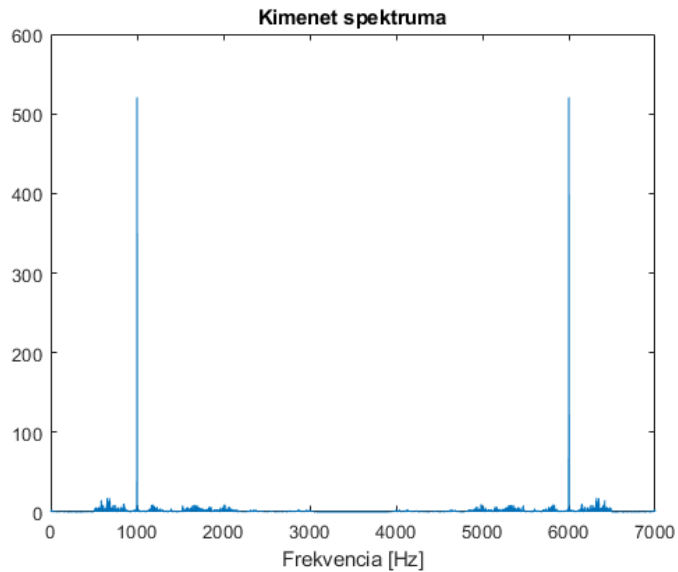
felvétel a kisebb teremben készült, 5 db mikrofonnal, az irányok, amelyekre a kényszereket előírtam:  $90^\circ$ ,  $57^\circ$ ,  $0^\circ$ ,  $-45^\circ$ ,  $-90^\circ$ .  $0^\circ$ -os irányból minden frekvencián torzítatlanul átengedtem, a többi irányból pedig elnyomtam a jelet.



7-35. ábra, iránykarakterisztikus szűrő átvitele



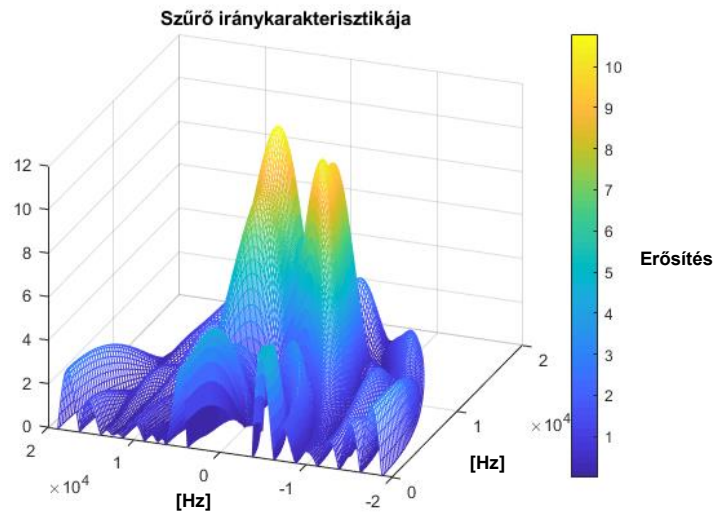
7-36. ábra, bemenet frekvenciaspektruma



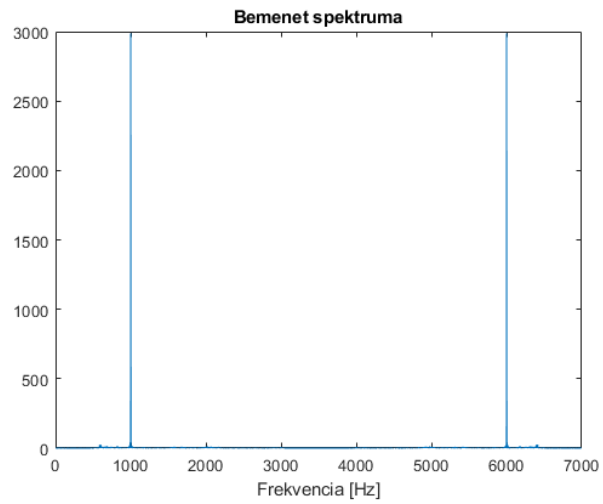
7-37. ábra, kimenet frekvenciaspektruma

Az átvitel már a korábbi alfejezetben megismertekhez hasonlóan alakul, mivel szimmetrikus, ezért nem tartalmaz nagy kiemeléseket semmilyen ponton. A bemenet és a kimenet spektrumát megvizsgálva szembeötlő, hogy a szinuszos zavarásjel amplitúdója mindössze felére csökkent, ellentétben a modellezett esetekkel, ahol gyakorlatilag teljesen eltűnt. Ennek oka a teremben jelen lévő reflexió, mely során a forrástól a hanghullámok nem csak direkt úton terjednek, és jutnak el a mikrofontömbig (3.1.1. fejezet). Mivel a visszavert hullámok egyszerre számos irányból érkeznek a mikrofonok felé, ezért szűrésük egy általánosan minden irányban és minden frekvencián elnyomó iránykarakterisztikus szűrővel lenne lehetséges, viszont túl sok irány és frekvencia átvitelének specifikálása a kényszereket tartalmazó mátrix szingularitását okozná. A hangfájlok meghallgathatók a csatolt dokumentumok között (bemenet: Valos\_Szinusz\_Bemenet.wav, kimenet: Valos\_Szinusz\_Kimenet.wav).

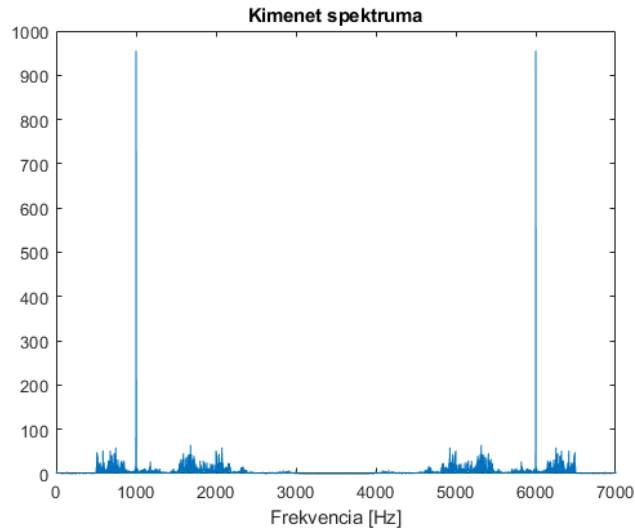
A következő esetben a dedikált forrás iránya megváltozott, 50°-os irányból érkezett a beszéd. Kényszereket a következő irányokban írtam elő: 90°, 57°, 0°, -50°, -90°, -50°-os irányból átengedve a jelet, más irányokból pedig elnyomva.



7-38. ábra, iránykarakterisztikus szűrő átvitele



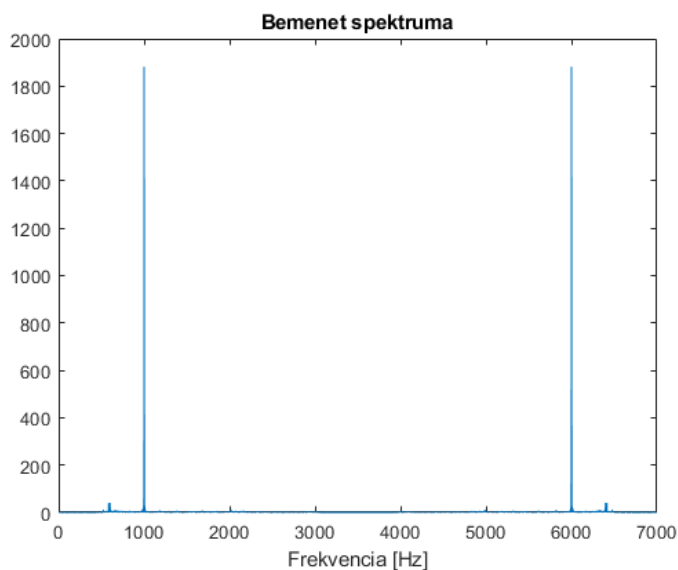
7-39. ábra, bemenet frekvenciaspektruma



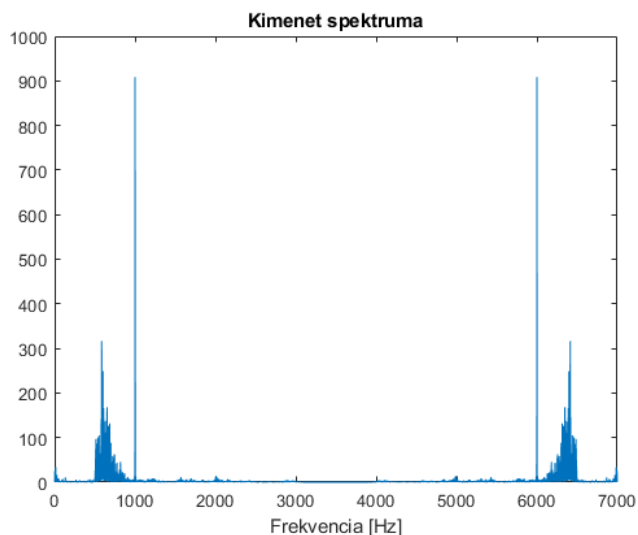
7-40. ábra, kimenet frekvenciaspektruma

A 7-39. ábra és 7-40. ábra szemléltette frekvenciaspektrumokról leolvasható, hogy jelen esetben a szinuszos zavarás amplitúdóját kevesebb, mint harmadára sikerült csökkenteni, ugyanakkor most is viszonylag jelentős teljesítmény jutott el a mikrofonokig visszavert hullámként, melyet nem sikerült szűrni. (A hangfájlok meghallgathatók a csatolt fájlok között: bemenet: Valos\_Szinusz2\_Bemenet.wav, kimenet: Valos\_Szinusz2\_Kimenet.wav). A nagyobb, visszhangosabb teremben végzett mérés során a szinuszhullám  $48^\circ$ -os irányból érkezett, a beszéd  $0^\circ$ -os irányból. A következő ábrákon a bemenet és a kimenet frekvenciaspektrumát megtekintve látható, hogy a szinusz amplitúdóját a visszhangosabb teremben is körülbelül 50%-kal sikerült csökkenteni, a visszavert hullámok nagyobb teljesítményének ellenére. (A hangfájlok meghallgathatók a csatolt fájlok között: bemenet: Valos\_Szinusz3\_Bemenet.wav, kimenet: Valos\_Szinusz3\_Kimenet.wav).



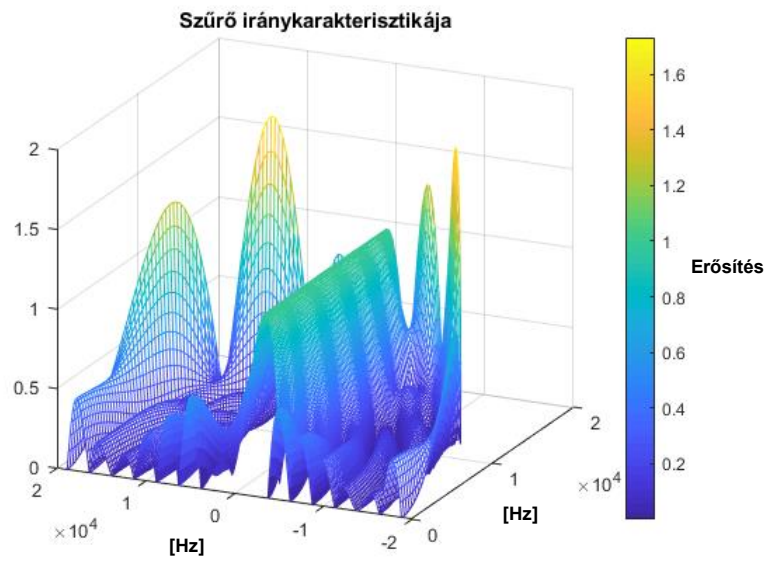


7-41. ábra, bemenet frekvenciaspektruma

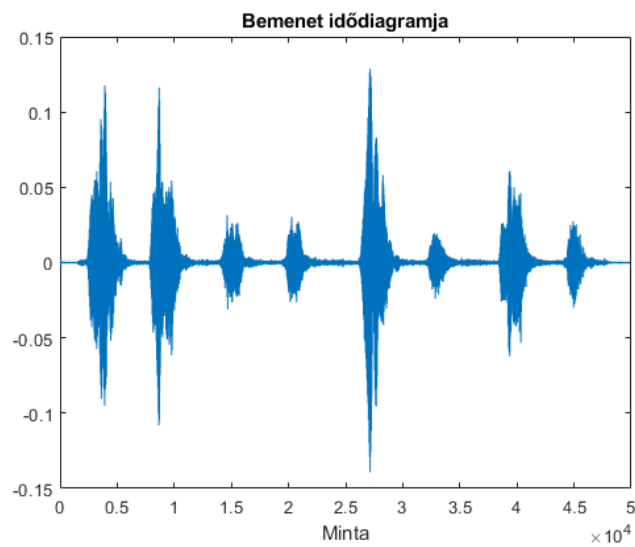


7-42. ábra, kimenet frekvenciaspektruma

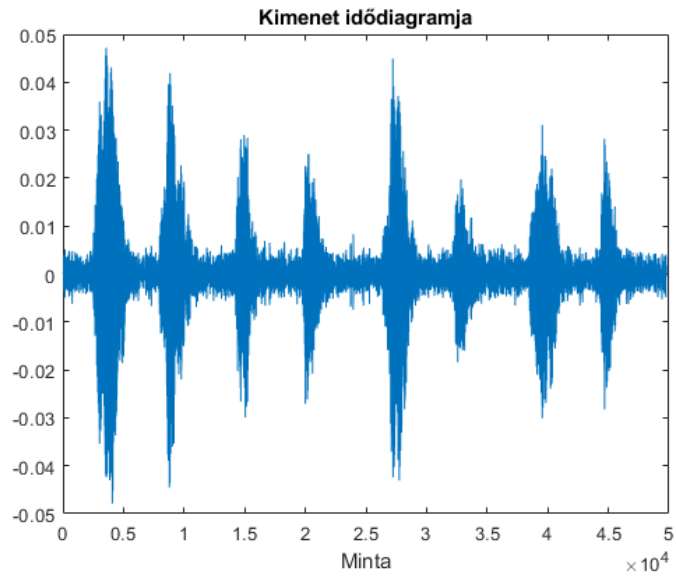
Méréseket végeztem olyan esetben is, amikor a zavarforrás is beszéd volt, hasonlóan a 7.2.1. fejezetben ismertetett modellhez. Egy, a kisebb teremből származó esetben a zavaróként jelen lévő második beszédforrás jellegéből adódóan a teljes frekvenciatartományban megtalálható volt, valamint reflektált hullámok is nehezítették a jelrekonstrukciót. Ennek ellenére a zavarforrás amplitúdóját sikerült kevesebb, mint felére csökkenteni, anélkül, hogy a dedikált forrás jelét jelentősen torzítottam volna (szöveg: „bal, bal, jobb, jobb, bal, jobb, bal, jobb”, dedikált forrás: „jobb” szó, középről érkezett, zavar: „bal” szó, -50°-os irányból érkezett). A kényszereket a 90°, 45°, 0°, -50°, -90°-os irányokban írtam elő.



7-43. ábra, iránykarakterisztikus szűrő átvitele



7-44. ábra, bemenet idődiagramja



7-45. ábra, kimenet idődiagramja

A hangfájlok meghallgathatók a csatolt fájlok között: bemenet: Valos\_Beszed\_Bemenet.wav, kimenet: Valos\_Beszed\_Kimenet.wav.

## 8. Összefoglalás

Diplomatervem során célom az volt, hogy egy olyan alkalmazást készítssek, amellyel térben és időben szűrhetem a mikrofontömb által vett jelet. A lehetséges felhasználások között megemlíthetők hangvezérelt rendszerek, mint például okos otthonok, illetve olyan konferenciák, előadások, interjúk felvételeinek feldolgozása, melyek zajos környezetben készültek. Az iránykarakterisztikus szűrő segítségével végrehajtott jelfeldolgozás során a jelet terhelő zaj elnyomása, illetve a dedikált bemenet torzításmentessége volt a célom.

A félév során tanulmányoztam a mikrofontömbök tulajdonságait, felépítésük szempontjából, melynek során az egymástól egyenlő távolságban, vonalban elhelyezett elrendezést választottam méréseim elvégzéséhez.

Elkészítettem a szűrő modelljét mind keskenysávú, mind szélessávú esetben a bemeneti jelmodellem alapján. Ezek működését Matlab segítségével létrehozott modellben, többféle zavarjellel is teszteltem. Felvételeket készítettem hasonló zavarforrások használatával több különböző teremben, melynek segítségével a valóságban is megvizsgáltam az iránykarakterisztikus szűrő működését. Elmondható, hogy modellezett környezetben, természetes zaj és visszhangok nélkül megfelelő mértékben sikerült elnyomnom a zavar teljesítményét. Valós esetben azonban láthatóak voltak az alkalmazás korlátai, ugyanakkor sikerült a dedikált hangforrást, a beszédet érthetőbbé tennem, a zaj teljesítményét érzékelhetően elnyomnom ekkor is.

## **9. Köszönetnyilvánítás**

Diplomatervem végén szeretném megragadni az alkalmat, hogy köszönetet mondjak mindazoknak, akik segítettek munkámban a félév során. Kiemelném konzulensem, dr. Orosz György segítségét és ösztönzését a szakdolgozatom precíz elkészítéséhez.

## 10. Irodalomjegyzék

- [1] D. Newman, „Voice Control Coming to a Conference Room Near You,” *Future of work*, May 2018.
- [2] B. S. S. GmbH, „www.behringer.com,” Behringer, július 2000. [Online]. Available: [https://downloads.music-group.com/software/behringer/ECM8000/ECM8000\\_C\\_Specs.pdf](https://downloads.music-group.com/software/behringer/ECM8000/ECM8000_C_Specs.pdf). [Hozzáférés dátuma: 28. május 2020.].
- [3] UA-101, „www.roland.com,” Roland Corporation, 10. december 2005.. [Online]. Available: [https://www.roland.com/global/support/by\\_product/ua-101/owners\\_manuals/](https://www.roland.com/global/support/by_product/ua-101/owners_manuals/). [Hozzáférés dátuma: 28. május 2020.].
- [4] Audacity, „www.audacityteam.org,” Audacity, [Online]. Available: <https://www.audacityteam.org/help/documentation/>.
- [5] G. K. Systems, „www.manualslib.com,” [Online]. Available: <https://www.manualslib.com/products/Genius-Kye-Systems-Sp-Hf1250x-3486197.html>. [Hozzáférés dátuma: 28 Szeptember 2020].
- [6] G. Nándor, „Akusztikus Lokalizáció Mikrofontömb segítségével,” *Budapesti Műszaki Egyetem, Méréstechnikai és Információs tanszék*, Január 2020.
- [7] D. J. K. Kenneth Jenkins, „Digital Signal Processing: Adaptive Signal Processing,” January 2006.
- [8] B. B. a. J. K. Peter Delos, „Phased Array Antenna Patterns—Part 1: Linear Array Beam Characteristics and Array Factor,” *Analog Dialogue*, p. vol. 54., május 2020.
- [9] B. B. a. J. K. Peter Delos, „Phased Array Antenna Patterns—Part 2: Grating Lobes and Beam Squint,” *Analog Dialogue*, p. vol. 54., június 2020.
- [10] P. Frenyó, „Beltéri akusztikus lokalizáció és forrásazonosítás,” *Budapesti Műszaki és Gazdaságtudományi Egyetem, Méréstechnikai és Információs Tanszék, Msc Diplomatervezés*, Budapest, 2014.
- [11] S. G. A. P. Jasna Stamać, „Designing the Acoustic Camera using MATLAB with respect to different types of microphone arrays,” *2nd International Colloquium on Smart Grid Metrology (SMAGRIMET)*, 9-12. Április 2019..

- [12] J. Rodman, „The Effect of Bandwidth on Speech Intelligibility,” Polycom, Szeptember 2006. [Online]. Available: [http://docs.polycom.com/global/documents/whitepapers/effect\\_of\\_bandwidth\\_on\\_speech\\_intelligibility\\_2.pdf](http://docs.polycom.com/global/documents/whitepapers/effect_of_bandwidth_on_speech_intelligibility_2.pdf). [Hozzáférés dátuma: 26 November 2020].
- [13] O. L. Frost, „An Algorithm For Linearly Constrained Adaptive Array Processing,” *Proceedings of the IEEE*, VOL. 60, NO. 8, 1972, augusztus.
- [14] K. M. B. Barry D. Van és Veen, „Beamforming: A versatile approach for spatial filtering,” *IEEE ASSP Magazine*, 1988, április.
- [15] L. G. Antonio Cantoni, „Fast Algorithms for Time Domain Broadband Adaptive Array Processing,” *IEEE TRANSACTIONS ON AEROSPACE AND ELECTRIC SYSTEMS VOL. AES-18, NO.5*, szeptember 1982.
- [16] „<https://en.wikipedia.org/>,” Wikipedia, [Online]. Available: [https://en.wikipedia.org/wiki/Finite\\_impulse\\_response](https://en.wikipedia.org/wiki/Finite_impulse_response). [Hozzáférés dátuma: 28. május 2020.].
- [17] „<https://manual.audacityteam.org/>,” Audacity, 19 június 2020. [Online]. Available: [https://manual.audacityteam.org/man/risset\\_drum.html](https://manual.audacityteam.org/man/risset_drum.html). [Hozzáférés dátuma: 10 november 2020].
- [18] D. A. Greensted, „Delay Sum Beamforming,” *The Lab Book Pages*, október 2012.