



UNIVERSIDADE ESTADUAL DE CAMPINAS  
Faculdade de Engenharia Elétrica e de Computação

Lucas Hernandez Pureza

**Desenvolvimento de um sensor acústico com  
padrão diretivo ajustável para monitoramento  
remoto.**

Campinas

22 de agosto de 2023



UNIVERSIDADE ESTADUAL DE CAMPINAS  
Faculdade de Engenharia Elétrica e de Computação

Lucas Hernandez Pureza

## **Desenvolvimento de um sensor acústico com padrão diretivo ajustável para monitoramento remoto.**

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica, na Área de Eletrônica, Microeletrônica e Optoeletrônica.

Orientador: Prof. Dr. Bruno Sanches Masiero

Este exemplar corresponde à versão final da tese defendida pelo aluno Lucas Hernandez Pureza, e orientada pelo Prof. Dr. Bruno Sanches Masiero

---

Campinas  
22 de agosto de 2023

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Área de Engenharia e Arquitetura  
Rose Meire da Silva - CRB 8/5974

P973d Pureza, Lucas Hernandez, 1995-  
Desenvolvimento de um sensor acústico com padrão diretivo ajustável para monitoramento remoto / Lucas Hernandez Pureza. – Campinas, SP : [s.n.], 2023.

Orientador: Bruno Sanches Masiero.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Eletroacústica. 2. Harmônicos esféricos. 3. Microfone. 4. Códigos de processamento digital de sinal. 5. Processamento de sinal acústico. I. Masiero, Bruno Sanches, 1981-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações Complementares

**Título em outro idioma:** Development of an acoustic sensor with adjustable directive pattern for remote monitoring

**Palavras-chave em inglês:**

Electro-acoustics

Spherical harmonics

Microphone arrays

Signal processing - Digital technique - Data processing

Acoustic imaging

**Área de concentração:** Eletrônica, Microeletrônica e Optoeletrônica

**Titulação:** Mestre em Engenharia Elétrica

**Banca examinadora:**

Bruno Sanches Masiero [Orientador]

Fabiano Fruett

William D'Andrea da Fonseca

**Data de defesa:** 22-06-2023

**Programa de Pós-Graduação:** Engenharia Elétrica

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0003-0101-0369>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3286948067830774>

## COMISSÃO JULGADORA - DISSERTAÇÃO DE MESTRADO

**Candidato:** Lucas Hernandez Pureza **RA:** 262736

**Data da Defesa:** 22 de junho de 2023

**Título da tese:** "Desenvolvimento de um sensor acústico com padrão diretivo ajustável para monitoramento remoto".

Prof. Dr. Bruno Sanches Masiero (Presidente)

Prof. Dr. William D'Andrea da Fonseca

Prof. Dr. Fabiano Fruett

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

*Dedico essa dissertação ao Dr. Geyverson Teixeira de Paula, que me mostrou que Engenharia Elétrica não é uma chatice completa.*

# Agradecimentos

Agradeço a todos que compartilharam um pouco de si para me impulsionar na minha jornada.

Agradeço a minha família, o ninho criado por Mercedes Hernandez Moraes (minha mãe) e por meu pai, Walter Pureza Júnior, que com amor e coragem, me ensinaram a seguir sempre por um caminho de crescimento e benevolência.

Agradeço aos meus amigos, uma lista injusta de se fazer pois não quero rotular preferidos, mas agradeço imensamente cada pessoa que dividiu um sorriso comigo, abraçou minhas causas visões e se divertiu comigo nesse amadurecimento diário que é a vida.

Agradeço aos pesquisadores da Unicamp participantes do grupo de pesquisa para o desenvolvimento tecnológico da empresa Taesa em conjunto com o instituto Eldorado, que esteve presente comigo por 2 anos discutindo ideias e tecnologias, e especialmente ao Dr. Fabiano Fruett, Gabriel Gialucca, Fernanda Caldas e, especialmente-mente ao Dr. Bruno Masiero, meu orientador que me instiga e dá lenha aos pensamentos que desenvolvo.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Agradeço a minha namorada, Elisa Mara do Nascimento, pois compartilhamos de perto a loucura um do outro no término de um mestrado.

Enfim, dedico essa tese à imaginação, mentes abertas são o combustível para o progresso.

# Resumo

Em busca de um monitoramento de um local em análise, tecnologias *Ambisonics* parece um caminho promissor para a caracterizar ambientes sonoros, retirar informações de áudio espacial e avaliar parâmetros auditivos de grande interesse em um ambiente de estudo. Pensando por esse lado, essa dissertação acompanha a descrição técnica de um arranjo tetraédrico de microfones *wireless*, com transmissão via *WiFi*, capaz de gravar em tempo-real e de forma contínua sinais de áudio (rotulados como sinais de Formato B), sinais capazes de reconstruir e monitorar o ambiente sonoro.

Primeiramente é descrito a teoria que envolve todo o processamento acústico, introduzindo o conceitos de ondas planas e harmônicos esféricos para a descrição de ambientes sonoros em um espaço esférico, e depois é explicado como essas considerações teóricas podem ser combinadas para reconstruir um ambiente sonoro a partir de microfones espalhados de forma equiangular em uma esfera. O protótipo então é idealizado, construído e explicado em detalhes no Capítulo 4, com o processamento dos sinais feito na linguagem *Python* em um microcontrolador *Raspberry Pi 3*, e os sinais transmitidos via *WiFi* para um servidor que fará o processamento para a obtenção dos sinais de Formato B. No Capítulo 5 são feitos experimentos para avaliar a qualidade e diretividade dos microfones empregados, além de um experimento para avaliar a filtragem espacial utilizando de um cenário composto por duas fontes tonais. Na parte reservada para as conclusões, é apontado a boa resposta da filtragem espacial, ao qual teve-se uma diferença de 18,8 dB da componente de frequência de interesse entre os sinais filtrados espacialmente. A partir do controle e ajuste do padrão polar, pôde-se monitorar os sinais simplesmente ponderando os canais de Formato B em uma faixa de frequência restrita.

**Palavras-chaves:** Formato B; Eletroacústica; Harmônicos esféricos; Filtragem áudio espacial; Áudio espacial.

# Abstract

In search of a monitoring system of an environment in analysis, Ambisonics technology is a promissory way to characterize sound fields, retrieve spatial audio information, and evaluate auditory parameters of interest in a given space under study. Thinking in that direction, this dissertation technically describes the construction of a wireless tetrahedral array of microphones with Wi-Fi transmission capable of recording (in real-time and continuously) audio signals in B-Format. These signals can reconstruct and manage the sound field around the sensor.

First, the necessary theory of acoustics is presented, introducing the concepts of plane waves and spherical harmonics; later, it explains how these theoretical considerations can be combined to reconstruct a sound field from microphones scattered equiangularly on a sphere. The prototype is idealized, built, and explained in detail in Chapter 4, with the processing carried out in Python by a microcontroller Raspberry Pi 3, and signal transmission via Wi-Fi, connecting sensor and server (operational system in charge of the processing to obtain the B-Format signals). In Chapter 5, experiments are made to evaluate the microphones' quality and directivity. Further, one experiment evaluates the spatial filter using a scenario composed of two tonal sound sources. The last chapter points to the result of the spatial filtering experiment, which had a distinction of 18.8 dB of the magnitudes of the frequency components of interest between the spatial filtered signals. With the control and adjustment of the polar pattern, one can manage signals by weighting the B-Format channels in a limited frequency range.

**Keywords:** B-Format; Eletroacoustic; Spherical harmonics; Spatial audio filtering; Spatial audio.

*“He has realized at last, that imaginary guitar notes and imaginary vocals  
exist only in the imagination of the imaginer!”  
(Frank Zappa)*

# Lista de ilustrações

Figura 2.1 – Processo de geração da vibração e ruído do transformador. Fonte: (YING <i>et al.</i> , 2018).	18
Figura 3.1 – Representação da manifestação do som como fenômeno natural. Fonte: (Wikipedia contributors, 2022).	22
Figura 3.2 – Variação da intensidade sonora em relação com a distância de propagação. Fonte: (EARGLE, 2005).	22
Figura 3.3 – Relação do sistema de coordenadas esféricas com o sistema de coordenadas cartesianas. Fonte: (RAFAELY, 2018).	23
Figura 3.4 – Relação entre a pressão sonora e velocidade de partícula (A) e a relação entre velocidade da partícula, deslocamento dessa partícula de ar e o gradiente da pressão (B). Fonte: (EARGLE, 2005).	24
Figura 3.5 – Derivação do padrão cardioide pela combinação do componente de pressão (omnidirecional) e componente de gradiente (bidirecional). Fonte: (EARGLE, 2005).	24
Figura 3.6 – Diretividade de um microfone hiper cardioide para diversas frequências. Fonte: (DPA Microphones, 2020).	25
Figura 3.7 – Gráfico tridimensional em forma de balão dos harmônicos esféricos para $n = 0$ até $n = 4$ . Fonte: (RAFAELY, 2018).	28
Figura 3.8 – Posicionamento no plano cartesiano de cada microfone presente no arranjo tetraedro — círculos maiores representam que o microfone está acima do plano bidimensional representado na figura, círculos menores indica que sua posição está inferior ao plano.	33
Figura 3.9 – Padrão polar tridimensional do sinal W para diferentes frequências. Fonte: (SCHULZE-FORSTER, 2018).	33
Figura 3.10 – Padrão polar tridimensional do sinal X para diferentes frequências. Fonte: (SCHULZE-FORSTER, 2018).	34
Figura 3.11 – Resposta em frequência dos filtros $V_{n,\alpha}$ para harmônicos esféricos de ordem até 4, considerando um campo sonoro sobre uma esfera rígida. $r$ é a distancia das cápsulas para o centro efetivo do arranjo. Fonte: (POLETTI, 2005)	35
Figura 3.12 – Resposta em frequência dos filtros para a saída omni W e as de figura de oito X, Y, Z para um arranjo de raio=1,47cm. Fonte: (GERZON, 1975).	35
Figura 3.13 – Esquemática do processamento dos sinais nos filtros de correção. Adaptada de (FARINA, 2006).	37

Figura 3.14–Resposta em frequência e fase dos filtros equalizadores determinados a partir de medidas experimentais para o caso de um arranjo com microfones cardioides ( $r = 1,5$ cm). Adaptada de (FARINA, 2006). . .	38
Figura 3.15–Respostas em frequências puras e equalizadas de um arranjo de teste. Adaptada de (HELLER; BENJAMIN, 2012). . . . .	40
Figura 4.1 – Resposta em frequência normalizada em 1 kHz do IMP23ABSUTR. Autor:(STMicroelectronics NV, 2020) . . . . .	43
Figura 4.2 – Resposta em frequência normalizada em 1 kHz do IMP23ABSUTR em faixa de ultrassom. Autor:(STMicroelectronics NV, 2020) . . . . .	43
Figura 4.3 – Base/suporte feito utilizando software baseado em estereolitografia. . .	44
Figura 4.4 – Peças para uma estrutura esférica feitas utilizando software baseado em estereolitografia.. . . . .	44
Figura 4.5 – Estrutura completa do arranjo tetraédrico projetado. . . . .	45
Figura 4.6 – Diagrama de conexões entre o array de microfones e a caixa de conexões.	45
Figura 4.7 – Fotografia da caixa de conexões e fonte de alimentação regulada para microfones. . . . .	46
Figura 4.8 – Diagrama de funcionamento do sistema de captação acústica. . . . .	46
Figura 4.9 – Os 4 sinais (amplitude $\times$ tempo) referentes a gravação dos 4 canais do arranjo. . . . .	48
Figura 4.10–Diagrama explicativo de como a biblioteca Socket opera. . . . .	48
Figura 4.11–Fonte e microfones posicionados para o ensaio em sala anecoica. Fonte: Autor . . . . .	50
Figura 4.12–Suporte auxiliar para a caracterização dos microfones. Fonte: Autor . .	50
Figura 4.13–Respostas em frequência de cada microfone sobre as mesmas condições (curva azul é o primeiro microfone da placa, de cima para baixo, curva laranja para o segundo microfone, amarelo para o terceiro e roxo para o último). Fonte: Autor . . . . .	51
Figura 4.14–Espectrograma da varredura de frequência gravada por um dos microfones com distorções não-lineares. Fonte: Autor . . . . .	52
Figura 4.15–Espectrograma da varredura de frequência gravada por um dos microfones com corte no tempo feito. Fonte: Autor . . . . .	53
Figura 4.16–Respostas em frequência de cada microfone depois do corte no tempo feito (curva azul é o primeiro microfone da placa, de cima para baixo, curva laranja para o segundo microfone, amarelo para o terceiro e roxo para o último). Fonte: Autor . . . . .	53
Figura 4.17–Resposta em frequência dos 4 equalizadores para os canais de Formato A. Fonte: Autor . . . . .	54
Figura 5.1 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 1 kHz.	55

Figura 5.2 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 2 kHz.	56
Figura 5.3 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 6,35 kHz.	56
Figura 5.4 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 10 kHz.	56
Figura 5.5 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 20 kHz.	57
Figura 5.6 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 40 kHz.	57
Figura 5.7 – Resposta em Frequência da fonte de teste A e B. . . . .	58
Figura 5.8 – Resposta em Frequência do sinal filtrado em $(\phi = 0^\circ, \theta = 0^\circ)$ e do sinal filtrado em $(\phi = 270^\circ, \theta = 0^\circ)$ utilizando padrão polar $\alpha = 0, 5$ . . . . .	59
Figura 5.9 – Resposta em Frequência do sinal filtrado em $(\phi = 0^\circ, \theta = 0^\circ)$ e do sinal filtrado em $(\phi = 270^\circ, \theta = 0^\circ)$ utilizando padrão polar $\alpha = 0$ . . . . .	61

# Lista de tabelas

Tabela 3.1 – Harmônicos Esféricos $Y_n^m(\theta, \phi)$ para $n = 1, \dots, 3$ . . . . .	27
Tabela 4.1 – Orçamento do sistema de captação acústica. . . . .	47
Tabela 5.1 – Valor da magnitude das componentes de frequência dos sinais filtrados espacialmente com padrão polar $\alpha = 0, 5$ . . . . .	59
Tabela 5.2 – Valor da magnitude das componentes de frequência dos sinais filtrados espacialmente com padrão polar $\alpha = 0$ . . . . .	60

# Lista de Abreviaturas e Siglas

**RMS** Root Mean Square ou Raiz do valor quadrático médio.

**LFU** Left-Front-Up ou Esquerda-Frente-Cima.

**RFD** right-front-down ou Direita-Frente-Baixo.

**LBD** left-back down ou Esquerda-Atrás-Baixo.

**RBU** right-back-up ou Direita-Atrás-Cima.

**FIR** Finite Impulse Response ou Resposta ao Impulso Finito.

**DOA** Direction of Arrival ou Direção de Chegada.

**STFT** Short Time Fourier Transform ou Transformada de Fourier de Tempo Curto.

# Sumário

<b>1</b>	<b>Prefácio: Ode ao Som</b>	<b>16</b>
<b>2</b>	<b>Introdução</b>	<b>17</b>
2.1	O som em análise	17
2.2	Objetivos	19
2.3	Estrutura do documento	19
<b>3</b>	<b>Revisão Teórica</b>	<b>21</b>
3.1	Representação do som como onda plana	21
3.2	Diretividade dos microfones	23
3.3	Harmônicos esféricos	25
3.4	Do Formato A para o Formato B	29
3.5	Filtros Corretivos	34
3.6	Aplicações do Formato B.	39
3.6.1	Filtragem espacial.	40
3.6.2	Direção de chegada de uma onda sonora.	41
<b>4</b>	<b>Metodologia</b>	<b>42</b>
4.1	Hardware	42
4.2	Software	47
4.2.1	Filtros equalizadores do arranjo	49
<b>5</b>	<b>Experimentos</b>	<b>55</b>
5.1	Diretividade do microfone IMP23ABSUTR	55
5.2	Filtragem espacial	58
<b>6</b>	<b>Conclusão</b>	<b>62</b>
6.1	Perspectivas futuras	63
	<b>Referências</b>	<b>64</b>
	<b>Anexos</b>	<b>67</b>
	<b>ANEXO A Código do Cliente</b>	
	(Gravação multicanal e transmissão WiFi)	68
	<b>ANEXO B Código do Servidor</b>	
	(Processando o Sinal e Salvando Arquivos)	70

# 1 Prefácio: Ode ao Som

O mundo fala, responde e avisa. No movimento de cada partícula terrestre, choques acontecem e reverberam em resposta — as vezes grita, as vezes suspira. Os sons têm essa habilidade única de respeitar muito a natureza. Cada barulho é resultado de parâmetros caóticos e medidos (na medida do possível) que interagem de uma forma que nunca mais irão se interagir. Esse som, em uma gravação, é uma lembrança, fotografias de segundos que, a partir das nossas tecnologias, esforçam em reproduzir de maneira fiel como foi o início, meio e fim de um som no universo, no espaço.

E não é estranho o ser humano usar o som para desembrulhar o mundo. O próprio teste de eletrocardiograma descreve o ritmo pulsante do nosso coração e, se a frequência amostrada no exame se divergir de dados de um coração saudável, isso indica que algo descompassa seu coração — melhor seria uma manutenção.

E existe vários outros equipamentos que usam som: sonares, detectores de presença, doppler vascular. E todos avaliam dados acústicos, uma leitura mais detalhada de algum problema que esteja acontecendo, pois se a simples mudança de rota de um elétron causa uma nova vibração no ambiente (mesmo que quase imperceptível), as mudanças perceptíveis nos ruídos que algo normalmente emite podem indicar uma mudança de estado — alertando uma falha na estrutura.

## 2 Introdução

Esse capítulo é dedicado a apresentar o tipo de sensor que será desenvolvido nessa pesquisa, apontando diversos cenários que ele pode atuar e quais são os objetivos que precisam ser alcançados para culminar em um sensor *ambisonic wireless* com diretividade manipulável propício para monitorar ambientes sonoros diversos.

### 2.1 O som em análise

O mundo é a divisão de vários ambientes diversos, locais distintos, cada qual com o seu nicho e seu valor. Transitamos entre zonas rurais, florestais, urbanas, e colhemos o melhor de cada lugar em prol do desenvolvimento humano.

A análise de ambientes sonoros está evoluindo a passos largos, dado que sinais acústicos carregam informações exclusivas de seres vivos, lugares e equipamentos. Sensor acústico *ambisonic* é o tipo de tecnologia capaz de caracterizar um ambiente sonoro a partir do processamento e combinação de sinais de áudio captados em microfones distribuídos em um certo padrão em uma esfera, de tal forma que cada microfone gravaria os sinais sonoros de incidência angular distinta, tentando cobrir todo o espaço esférico tridimensional.

Combinar tecnologia *ambisonic* com monitoramento ecológico possibilita o recolhimento de informações sobre bioma, mudança climática, entre outras particularidades da paisagem sonora dentro de um biossistema. Foi feito na Polônia um estudo baseado em gravações *ambisonic* de grande porte do ambiente sonoro da Floresta Bialowieza, uma das maiores partes remanescentes da imensa floresta primal que uma vez cobria grande parte da Europa. As gravações e medidas foram feitas em locais característicos, ao qual o nível do fundo acústico de tais locais foram registrados em diversas condições climáticas. A maioria das gravações foram feitas em ambientes remotos, reservas naturais distantes da influência humana (MAŁECKI *et al.*, 2016).

Outro exemplo de ambiente sonoro que necessita monitoramento seria subestações elétricas. Um grande participante na paisagem sonora de uma subestação são os transformadores, peças chave de um sistema elétrico de potência, sendo os responsáveis pela elevação e diminuição da tensão elétrica em uma linha de transmissão ou em redes de distribuição. Em geral, a vibração do transformador é resultado de dois fenômenos: (i) forças eletromagnéticas geradas das interações das correntes nos enrolamentos e (ii) as forças no núcleo de ferro causadas pela magnetostricção das lâminas de aço (silício)

(QIANG *et al.*, 2021). Essas duas fontes de vibração (núcleo e enrolamentos) produzem o ruído do transformador, o que é detalhado no esquema da Figura 2.1.

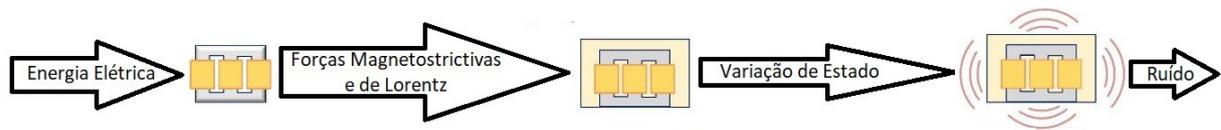


Figura 2.1 – Processo de geração da vibração e ruído do transformador. Fonte: (YING *et al.*, 2018).

A vibração no núcleo terá a frequência fundamental igual ao dobro da frequência do sistema de potência (60 Hz no Brasil), e sua amplitude é proporcional à tensão RMS ao quadrado de tal rede (HONG; HUANG, 2014). Já a vibração no enrolamento tem amplitude proporcional à corrente da carga ao quadrado, com sua frequência sendo o dobro da frequência da corrente da carga ligada ao transformador. Logo, se existe flutuação na tensão da rede ou na sua frequência, a vibração no núcleo flutuará também, e o mesmo é pensado para o enrolamento — corrente na carga flutua, os enrolamentos vibrariam proporcionalmente com essa flutuação. Mas se algo está danificado na estrutura do transformador, essa resposta linear entre os parâmetros da rede e a vibração se altera e se torna não-linear, como é mostrado em (HONG; HUANG, 2014).

Visando o monitoramento urbano, outro estudo focado em *ambisonic* foi feito nos Estados Unidos. Em prol no desenvolvimento de uma cidade inteligente, foi desenvolvido um sistema para monitorar, analisar e mitigar a poluição sonora urbana. A partir de medidas acústicas calibradas, a combinação de sensores de baixo custo conseguem reconhecer fontes sonoras individuais em tempo real. Métodos de ciência cidadã são usados para ajudar os residentes urbanos se conectarem com agências municipais e entre si, compreender mais sobre as pegadas de ruído (*noise footprint*) que produzem e facilitar denúncias e auto regulação (SALAMON *et al.*, 2018). Os dados coletados podem ser usados para traçar estratégias aplicáveis em legislação envolvendo controle de ruídos para reduzir a poluição sonora.

A combinação entre sensores *ambisonic* e ambientes sonoros que se beneficiam com um monitoramento contínuo parece um par poderoso. E entender o processamento e estrutura desses sensores é de grande importância para analisar informações recorrentes e dados que entregam características próprias do ambiente sonoro e dos participantes dele.

Pela necessidade de tecnologias acessíveis e um processamento rápido e inteligente de dados acústicos, foi desenvolvido um sensor acústico *ambisonic wireless* de baixo custo capaz de realizar filtragem sonora espacial para separar e monitorar sons de distintas fontes sonoras em um ambiente, recolhendo dados continuamente para beneficiar sistemas de monitoramento que visam estudar, fiscalizar ou agir no tipo de ambiente em

que atuam.

## 2.2 Objetivos

O objetivo desta pesquisa é mostrar didaticamente como construir um arranjo tetraédrico capaz de entregar dados acústicos via *WiFi* em tempo real. O propósito é monitorar diversas fontes sonoras que participam de um ambiente sonoro a partir da filtragem espacial, gerando sinais de áudio distintos, cada um referente ao ângulo entre sensor e fonte sonora em observação. Para isso, é necessário um estudo sobre: ambientes sonoros e decomposição de um campo de ondas sonoras; tecnologias e teoria por dentro de sensores *ambisonics* de primeira ordem; *hardware* e *software* necessário para atender as especificações necessárias para ter um sensor *wireless* que recolha dados de maneira autônoma.

Os objetivos específicos necessários para atingir o resultado esperado são listados a seguir.

1. Ter embasamento teórico em acústica: entender os modelos de ambientes sonoros, harmônicos esféricos, microfones direcionais, conversão de Formato A para Formato B (formato de áudio com 4 canais, 1 canal W omnidirecional e 3 que indicam a intensidade vetorial X,Y,Z do som no arranjo) e filtragem espacial de um campo sonoro.
2. Caracterizar os microfones utilizados no arranjo, com o objetivo de aproximar a resposta em frequência de cada cápsula para que tenham resposta coincidente, assim criando filtros digitais para sua equalização.
3. Desenvolvimento de *hardware* capaz de operar gravações multicanais.
4. Desenvolvimento do algoritmo de processamento digital a ser implementado no *hardware* que entregue dados via *WiFi*.
5. Realizar experimentos para avaliar a filtragem espacial do arranjo, com discussão entre os pontos fortes e fracos.

## 2.3 Estrutura do documento

Primeiramente, será mostrado a teoria sobre o desenvolvimento do arranjo, começando pelas considerações teóricas envolvendo acústica, propagação de ondas sonoras em uma esfera e gravações *ambisonics* (formato de áudio com caráter tridimensional).

Depois serão discutidos como cada componente foi escolhido para a construção do protótipo, evidenciando as suas especificações e como elas fortalecem o arranjo. Com o *hardware* projetado, detalhar sobre o *software* e como o algoritmo trabalha, em função da teoria, para entregar os dados acústicos necessários.

Um terceiro capítulo apresenta as caracterizações dos microfones do arranjo, buscando avaliar as respostas em frequência e resposta diretiva do modelo escolhido para o arranjo. Um cenário de filtragem espacial com duas fontes tonais é feito para avaliar a qualidade dos sinais resultantes.

O último capítulo é destinado para discussão final, mostrando como um cenário prático iria se beneficiar desse arranjo, em que cenários teriam ambiguidades nos dados que o arranjo recolheria e o que deve ser feito para atingir melhores resultados.

## 3 Revisão Teórica

Antes de entrarmos na parte prática e toda construção do arranjo de microfones projetado caracterizar o ambiente sonoro a sua volta, é necessário estar bem detalhado algumas considerações teóricas da acústica, e como podemos manipular esses detalhes ferramentais para, digitalmente, processar as gravações afim de tomar controle do espaço do som e retirar informações direcionais, pensando nas fontes que queremos monitorar em volta do arranjo.

### 3.1 Representação do som como onda plana

Existe inúmeros fatores que influenciam na forma como uma fonte sonora emite sons — de como a sua estrutura vibra e reverbera ou até a mesmo a temperatura do seu corpo. Então, às vezes, é bastante complexo caracterizar teoricamente a propagação sonora influenciada pela atividade que algo ou alguém produz, cabendo aos teóricos ajudarem os engenheiros com modelos que aproximam um cenário acústico por fórmulas matemáticas que possibilitam uma manipulação mais simples. Uma solução matemática para o fenômeno de propagação sonora é considerar que a fonte emite uma onda plana de uma única frequência, e utilizaremos dessa consideração daqui pra frente.

Imagine uma fonte sonora monocromática, ou seja, emitindo somente uma frequência. A perturbação periódica dessa fonte gerá ondas mecânicas em direção a um ouvinte, ilustrado na Figura 3.1. A pressão sonora é a amplitude do seno mostrado na Figura 3.1, que representa a variação da pressão atmosférica exercida por uma coluna de ar em uma área do espaço em função do tempo. Nas cristas da função senoidal, a pressão tem seu maior valor no ouvinte, e nos vales, tem-se a menor pressão.

O modelo de ondas planas considera então que essa emissão será feita em um plano perpendicular à direção de emissão sonora, e que os valores de pressão no plano perpendicular são iguais para todos os pontos desse plano. Além disso, assume-se que a intensidade terá somente influência com a distância entre o ouvinte e a fonte. Isso é melhor exemplificado na Figura 3.2, mostrando a relação entre distância e intensidade sonora. Curiosidade: um som com potência de 1 W transmitido a 1 m terá uma pressão sonora de 5,96 Pa (EARGLE, 2005).

Depois de atíçar a imaginação para enxergarmos como o som atua no ambiente, iremos em frente com uma solução matemática que satisfaça a equação de onda homogênea em coordenadas esféricas. Considere uma onda plana com amplitude unitária

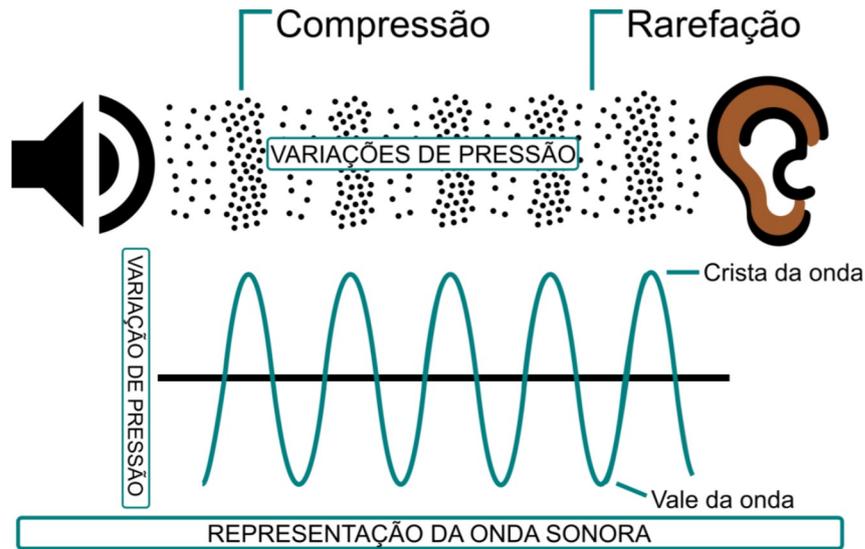


Figura 3.1 – Representação da manifestação do som como fenômeno natural. Fonte: (Wikipedia contributors, 2022).

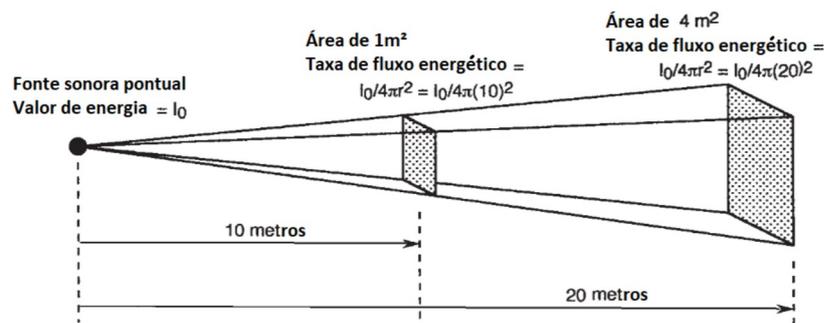


Figura 3.2 – Variação da intensidade sonora em relação com a distância de propagação. Fonte: (EARGLE, 2005).

e monocromática, com ângulo de incidência  $(\theta_k, \phi_k)$ , sendo  $\theta_k$  o ângulo azimutal de incidência e  $\phi_k$  (a Figura 3.3 mostra estes ângulos em um plano cartesiano), a inclinação de incidência; e o vetor de onda  $\vec{k} = (k, \theta_k, \phi_k)$  em coordenadas esféricas, sendo  $k = \omega/c$  ( $\omega$  é a frequência angular em rad/s e  $c$  a velocidade do som, aproximadamente 340 m/s).

A expressão que calcula a pressão sonora em um ponto  $r = (r, \theta, \phi)$  causada por essa onda é:

$$p(\vec{k}, \vec{r}, \theta, \phi) = e^{-i\vec{k} \cdot \vec{r}} \quad (3.1)$$

Vale salientar também que na “viagem” que uma onda sonora faz pelo espaço, as partículas de ar não acompanham essa onda na viagem, mas se movimentaram periodicamente com uma certa velocidade e direção no espaço, e pode-se relacionar a velocidade da partícula de ar  $u(t)$  com o deslocamento da partícula com a seguinte função:

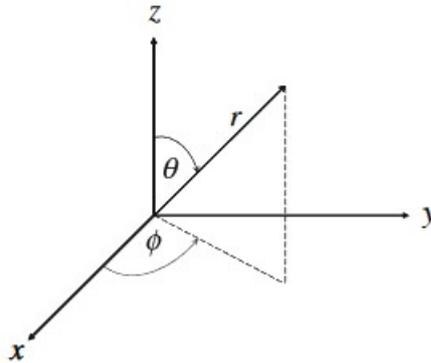


Figura 3.3 – Relação do sistema de coordenadas esféricas com o sistema de coordenadas cartesianas. Fonte: (RAFAELY, 2018).

$$u(t) = i\omega x(t), \quad (3.2)$$

sendo  $x(t)$  o valor máximo do deslocamento da partícula em função do tempo e  $i$  causando um deslocamento de 90 graus na fase.

Os microfones direcionais (aqueles que são sensíveis à direção de propagação do som) atuam sobre a diferença da pressão de ar entre dois pontos de um transdutor,  $dp/dl$ , chamado de gradiente da pressão. A Figura 3.4 mostra bem a relação entre pressão sonora, velocidade da partícula de ar e o gradiente de pressão.

## 3.2 Diretividade dos microfones

A maioria dos microfones direcionais em uso profissional são membros da família de cardioides de primeira ordem. O termo primeira ordem se refere à equação de sua resposta polar e a inclusão de um termo cossenoidal para a primeira ordem (EARGLE, 2005).

A partir das combinações entre pressão e gradiente, discutidos no final da seção passada, foi desenvolvido, no passado, um tipo especial de microfone: o microfone cardioide ou unidirecional. A Figura 3.5 mostra bem como se chega no coração da resposta polar de microfone, com a combinação da pressão omnidirecional com o elemento gradiente (presente no termo cosseno da equação), intensificando o som de frente e minimizando o som que vem atrás dele, como é modelado pela seguinte equação:

$$\rho = 0,5(1 + \cos\theta) \quad (3.3)$$

O microfone então passa a ter uma característica diretiva atribuída a ele. Mas

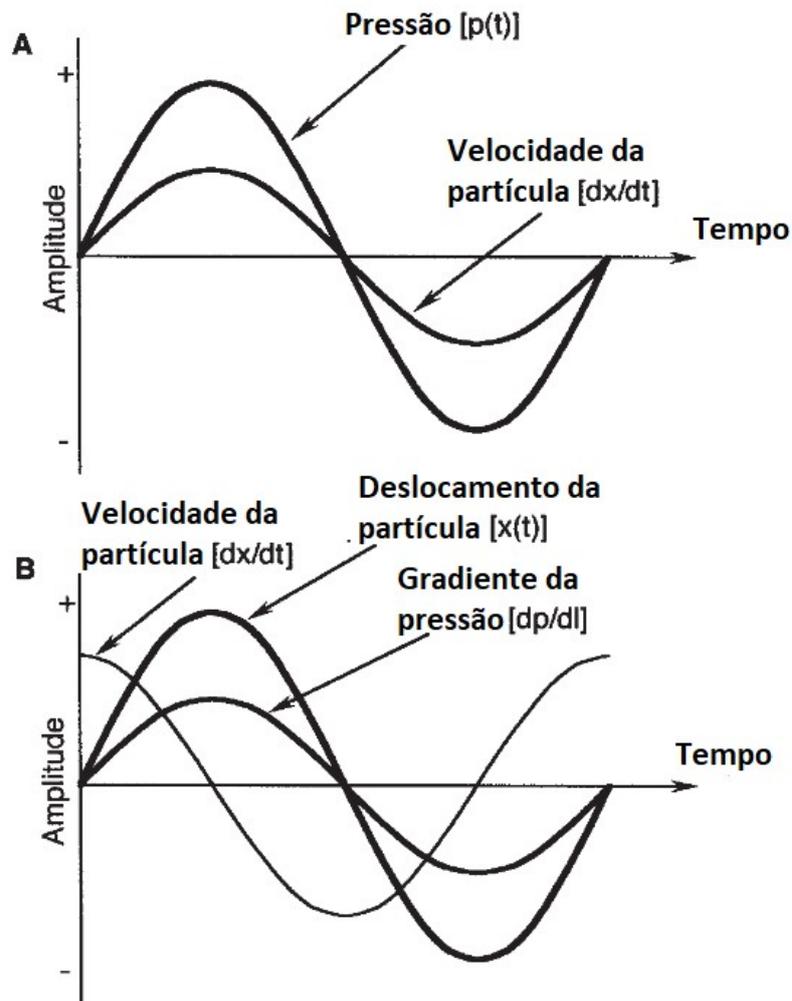


Figura 3.4 – Relação entre a pressão sonora e velocidade de partícula (A) e a relação entre velocidade da partícula, deslocamento dessa partícula de ar e o gradiente da pressão (B). Fonte: (EARGLE, 2005).

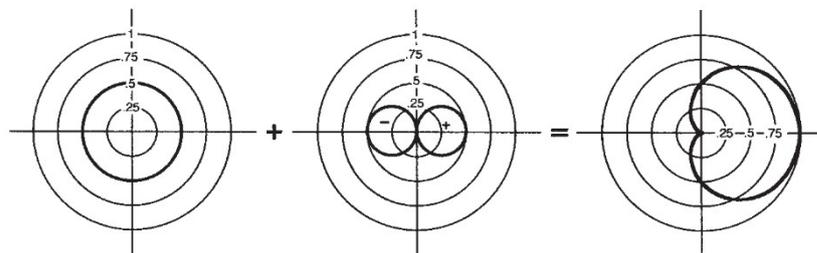


Figura 3.5 – Derivação do padrão cardioide pela combinação do componente de pressão (omnidirecional) e componente de gradiente (bidirecional). Fonte: (EARGLE, 2005).

existe um porém à cerca da diretividade de um microfone direcional: ela pode variar conforme a frequência que ela capta. Quando o ambiente sonoro apresenta baixas frequências

(na casa entre 20 a 200 Hz), a resposta polar do microfone tende a ser mais perto do omnidirecional — resultado do grande comprimento de onda nessas frequências, logo o componente de gradiente terá pouco relevância pois a onda não teve uma grande diferença de fase e amplitude entre os dois pontos do transdutor. Enquanto que para frequências maiores, apesar da resposta começar a se tornar mais direcional, sua resposta começa a ficar bem mais complexa, criando-se lóbulos laterais que prejudicam a resposta direcional. Podemos observar esse fenômeno na Figura 3.6, que mostra como a resposta polar de um microfone se modifica conforme a frequência.

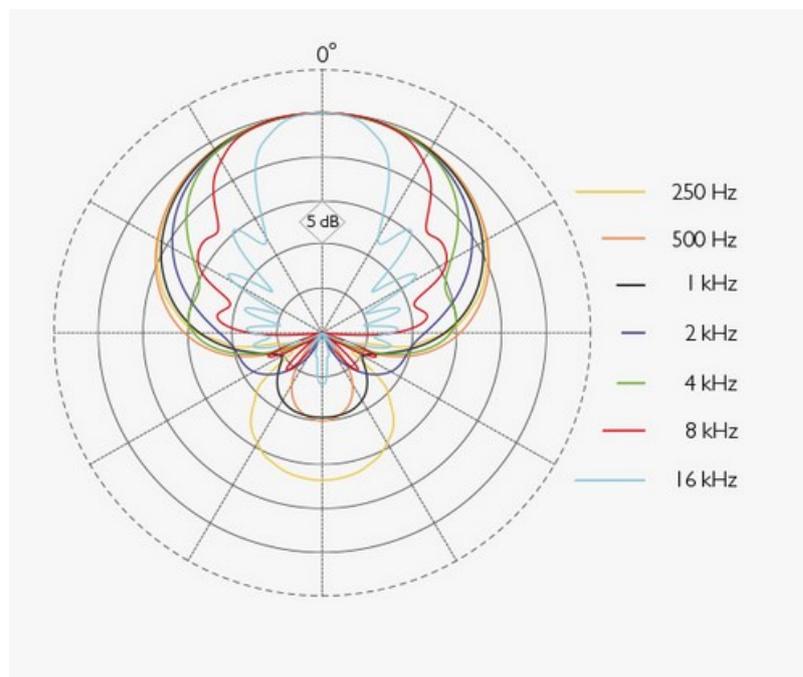


Figura 3.6 – Diretividade de um microfone hiper cardioide para diversas frequências. Fonte: (DPA Microphones, 2020).

Esse obstáculo à cerca da frequência é o maior desafio a ser contornado quando conversamos sobre direcionalidade para altas frequências, e mais adiante será comentado como que, na utilização do tetraedro, essa complicação atrapalha a resposta direcional do arranjo.

### 3.3 Harmônicos esféricos

Agora imagine uma função descrita em um espaço esférico unitário. A forma como tal função varia angularmente ( $\theta, \phi$ ) em uma superfície esférica unitária pode ser caracterizada com funções básicas de ponderação dependentes de um azimute e de uma inclinação. Essas funções básicas são chamadas de harmônicos esféricos, e são definidas como:

$$Y_n^m(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta) e^{i\phi}, \quad (3.4)$$

em que o símbolo  $(...)!$  representa a função fatorial,  $P_n^m(...)$  são as funções de Legendre associadas,  $m \in \mathbb{Z}$  é um inteiro indicando o grau da função e  $n \in \mathbb{N}$  é um número natural que representa a ordem da função.

Os valores de  $m$  e  $n$  podem ser entendidos como fatores de influência que os termos associados a  $\theta$  e  $\phi$  exercem sobre os harmônicos esféricos — quanto maior a ordem  $n$ , mais bem descrito é a sua função na esfera.

O grau  $n$  determina a maior potência dos termos cossenoidais e senoidais que controlam os harmônicos esféricos sobre  $\theta$ , enquanto  $m$  determina a dependência sobre  $\phi$  pelo termo exponencial (RAFAELY, 2018). Vale apontar que graças ao termo exponencial complexo, o valor absoluto de  $Y_n^m(\theta, \phi)$  será constante ao longo de  $\phi$ , sendo assim, mais prático representar harmônicos esféricos como partes imaginárias e reais em vez de sua representação em magnitude e fase.

A Tabela 3.1 mostra os valores dos harmônicos esféricos de ordem 0 até ordem 3. A simplicidade da ordem 0 é notável, pois se os harmônicos esféricos são expressões que representam as variações de uma função em uma esfera, um valor constante para toda esfera indica que independente da direção da fonte no espaço, o valor energético sonoro será o mesmo. No mundo da engenharia de som, os microfones omnidirecionais fazem exatamente esse trabalho, gravando o som de todo o ambiente sem se preocupar de onde o som está vindo.

Se ainda está meio nebulosa a interpretação de harmônicos esféricos, a sua representação gráfica consegue trazer mais clareza quanto a sua resposta em um espaço tridimensional. A Figura 3.7 mostra como as funções da Tabela 3.1 são representadas em coordenadas cartesianas quando aumentamos a ordem e o grau, sendo que os pontos com a cor azul esverdeado são valores positivos para os harmônicos e o roxo são valores negativos (o lado esquerdo, que é a parte imaginária de  $Y_n^m$  são para  $-n \leq m \leq -1$ , o meio representa  $m = 0$  e o lado direito, parte real, são para  $1 \leq m \leq n$  - e de cima para baixo,  $n$  vai de 0 até 3).

Dois propriedades importantes que devem ser ressaltadas antes de seguirmos em frente são sobre seu complexo conjugado e ortogonalidade.

A integral do produto de duas funções de harmônicos esféricos em toda a superfície da esfera unitária terá o seguinte valor:

$$\int_0^{2\pi} \int_0^\pi [Y_n^m(\theta, \Phi)] * Y_{n'}^{m'}(\theta, \Phi) \sin(\theta) d\theta d\Phi = \delta_{nn'} \delta_{mm'}, \quad (3.5)$$

$n = 0$	$Y_0^0(\theta, \phi) = \sqrt{\frac{1}{4\pi}}$
$n = 1$	$Y_1^{-1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\phi}$
	$Y_1^0(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos \theta$
	$Y_1^1(\theta, \phi) = -\sqrt{\frac{3}{4\pi}} \sin \theta e^{i\phi}$
$n = 2$	$Y_2^{-2}(\theta, \phi) = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{-2i\phi}$
	$Y_2^{-1}(\theta, \phi) = \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{-i\phi}$
	$Y_2^0(\theta, \phi) = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1)$
	$Y_2^1(\theta, \phi) = -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi}$
	$Y_2^2(\theta, \phi) = -\sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{2i\phi}$
$n = 3$	$Y_3^{-3}(\theta, \phi) = \sqrt{\frac{35}{64\pi}} \sin^3 \theta e^{-3i\phi}$
	$Y_3^{-2}(\theta, \phi) = \sqrt{\frac{105}{32\pi}} \cos \theta \sin^2 \theta e^{-2i\phi}$
	$Y_3^{-1}(\theta, \phi) = \sqrt{\frac{21}{64\pi}} (5 \cos^2 \theta - 1) \sin \theta e^{-i\phi}$
	$Y_3^0(\theta, \phi) = \sqrt{\frac{7}{16\pi}} (5 \cos^3 \theta - 3 \cos \theta)$
	$Y_3^1(\theta, \phi) = -\sqrt{\frac{21}{64\pi}} (5 \cos^2 \theta - 1) \sin \theta e^{i\phi}$
	$Y_3^2(\theta, \phi) = \sqrt{\frac{105}{32\pi}} \cos \theta \sin^2 \theta e^{2i\phi}$
	$Y_3^3(\theta, \phi) = -\sqrt{\frac{35}{64\pi}} \sin^3 \theta e^{3i\phi}$

Tabela 3.1 – Harmônicos Esféricos  $Y_n^m(\theta, \phi)$  para  $n = 1, \dots, 3$ .

sendo  $\delta_{nm'}$  a função Delta de Kronecker, que tem valor unitário para  $n = n'$  de mesma ordem ou mesmo grau  $m = m'$  e zero para qualquer outra combinação.

A partir do sistema ortogonal que os harmônicos esféricos de valores complexos formam sobre a superfície de uma esfera unitária, qualquer combinação de  $Y_n^m$  forma um sistema ortogonal novamente (JOHANN-MARKUS, 2009), incluindo o conjunto de harmônicos esféricos reais.

Os harmônicos esféricos são funções complexas graças a sua componente exponencial, enquanto as funções associadas a Legendre são todas reais (RAFAELY, 2018). O complexo conjugado dos harmônicos esféricos tomam a seguinte forma:

$$[Y_n^m(\theta, \Phi)]^* = -(1)^m Y_n^{-m}(\theta, \Phi) \quad (3.6)$$

Logo, se podemos reajustar o sistema com novas combinações de  $Y_n^m$  graças à propriedade de ortogonalidade, um novo sistema só com valores reais pode ser obtido se

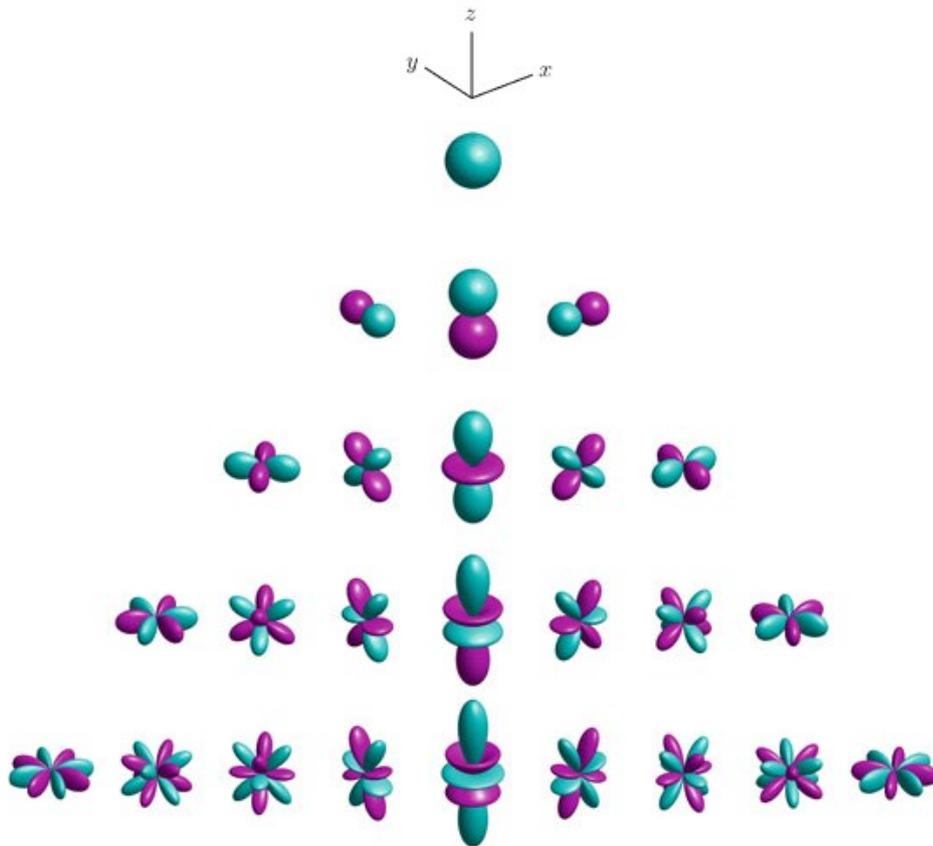


Figura 3.7 – Gráfico tridimensional em forma de balão dos harmônicos esféricos para  $n = 0$  até  $n = 4$ . Fonte: (RAFAELY, 2018).

definirmos uma superfície tal qual:

$$S_n^m(\theta, \Phi) = \begin{cases} \frac{-1^m}{\sqrt{2}}(Y_n^m + Y_n^{m*}) & m > 0 \\ Y_n^m & m = 0 \\ \frac{-1}{j\sqrt{2}}(Y_n^m - Y_n^{m*}) & m < 0 \end{cases} \quad (3.7)$$

Dessa forma, a equação responsável pela parte azimutal do espaço esférico terá as seguintes condições de contorno:

$$\text{trg}_m(\Phi) = \begin{cases} \sqrt{2} \cos(m\Phi) & m > 0 \\ 1 & m = 0 \\ -\sqrt{2} \sin(m\Phi) & m < 0 \end{cases} \quad (3.8)$$

Quando estamos operando com gravações, sendo valores amostrados no domínio do tempo, a maioria das vezes é usado harmônicos esféricos com valores reais, pois o processamento pode ser feito com sucesso inteiramente no domínio do tempo (JOHANNMARKUS, 2009).

### 3.4 Do Formato A para o Formato B

A Equação 3.1 pode evoluir um pouco mais. A pressão sonora vindo de uma fonte de onda plana pode ser reescrita como sendo a soma de harmônicos esféricos e funções esféricas de Bessel, como é descrito a seguir:

$$p(x, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m(k) J_n(kr) Y_n^m(\theta, \Phi) \quad (3.9)$$

em que  $J_n(kr)$  é a função esférica de Bessel de primeira ordem,  $k$  representa o número de onda,  $Y_n^m(\theta, \Phi)$  são os harmônicos esféricos e  $A_n^m(k)$  é a amplitude complexa associada a cada harmônico esférico, também conhecido como os coeficientes dos harmônicos esféricos (JOHANN-MARKUS, 2009).

A funcionalidade das funções esféricas de Bessel é como solução de equações diferenciais (lembrando que a equação de onda é uma equação diferencial), mas para além disso, sua importância nesse projeto como ferramenta matemática é de melhor descrever o campo sonoro, representando como o som varia em relação ao raio da superfície esférica, algo que não é levado em conta pelos harmônicos esféricos. Combinações lineares de suas funções de primeira e segunda ordem podem descrever uma equação de onda. Para mais detalhes, leia a Seção 2.2 de (RAFAELY, 2018).

Para isolar os coeficientes dos harmônicos esféricos e calcular-los, se faz o seguinte passo: ao multiplicar os dois lados da Equação 3.9 por  $Y_n^{m'*}(\theta, \Phi)$  e integrarmos a função sobre a esfera, teremos (POLETTI, 2005):

$$A_n^m(k) = \frac{1}{J_n(kr)} \int_0^{2\pi} \int_0^{\pi} p(r, \theta, \Phi, k) Y_n^{m'*}(\theta, \Phi) d\theta d\Phi \quad (3.10)$$

Os zeros de  $J_n(kr)$  produzem um problema de equalização (POLETTI, 2005), pois se temos um denominador tendendo a zero, será produzido respostas infinitas, e uma forma de contornar esse problema é posicionando microfones cardioides apontando radialmente para fora da esfera.

Uma outra forma de evitar o problema de equalização listado acima é posicionar os microfones sobre a superfície de uma esfera rígida. Neste caso, é necessário considerar novas condições de contorno, e para isso utilizaremos a técnica chamada de decomposição em esfera rígida (*Solid Sphere Decomposition*).

Basicamente, o campo resultante será o campo sonoro sem uma esfera (mas descrito em um espaço esférico), representado pela Equação 3.10, somado com um campo resultante do espalhamento que o som irradia para fora da esfera sólida, condição de

contorno que assume total rigidez na esfera (impedância acústica tendendo ao infinito) e que a velocidade radial total é igual a zero na superfície esférica.

Depois de algumas considerações e manipulações matemáticas (FERNANDEZ-GRANDE, 2016), esse contorno irá acarretar na seguinte fórmula simplificada para os coeficientes dos harmônicos esféricos:

$$A_n^m(k) = -i(kr)^2 H_n'(kr) \int_0^{2\pi} \int_0^\pi p(r, \theta, \Phi, k) Y_n^{m*}(\theta, \Phi) d\theta d\Phi \quad (3.11)$$

em que  $H_n'(kr)$  é a função de Hankel de ordem  $n$ .

Lembra do teorema de Nyquist? Da mesma forma que um sinal precisa ser limitado em banda de frequência para que a reconstrução de sua forma de onda seja perfeita, um pensamento análogo pode ser feito para a amostragem de um campo sonoro. A função do campo sonoro será representada por um número finito de funções base que dependerão do número de microfones distribuídos na esfera. Como a representação do som na esfera é feita a partir de harmônicos esféricos, então a reconstrução desse ambiente sonoro é limitado pela ordem desses harmônicos. O desenvolvimento de um sistema de amostragem espacial usando microfones envolve uma negociação do número de microfones (conhecidos nesse contexto como pontos de amostragem), ou seja, menos microfones resulta em um sistema menos complexo. Mas com mais microfones, podemos ter uma maior acurácia na reconstrução do ambiente sonoro amostrado.

Para se ter pelo menos uma característica tridimensional na caracterização do campo sonoro, é preciso que a função seja descrita pelo menos até harmônicos de primeira ordem, pois os harmônicos esféricos de primeira ordem têm um formato de figura de oito, tal qual como vetores de ponderação para um espaço vetorial tridimensional. A fórmula a seguir relaciona o número  $Q$  de microfones necessários para alcançar a representação de harmônicos esféricos de ordem  $N$  (ABHAYAPALA, 2002):

$$Q \geq (N + 1)^2 \quad (3.12)$$

Logo, se a ordem dos coeficientes harmônicos esféricos é limitado para  $N = 1$ , o número de microfones necessários será no mínimo 4, sendo possível melhorar a resolução se acrescentarmos mais microfones e distribuímos eles pela esfera.

O ambiente sonoro será então amostrado usando um arranjo composto de 4 microfones encapsulados e arrançados em um formato tetraédrico. As cápsulas podem ser interpretadas como sendo os vértices, arestas e faces do tetraedro (SCHULZE-FORSTER, 2018). Deixaremos exposto aqui os vetores da posição das cápsulas, pois os ângulos em que estão distribuídas é de extrema importância para o funcionamento do equipamento:

$$r_1 = (r; 2, 19; 0), \quad r_2 = \left(r; 0, 96; \frac{\pi}{2}\right) \quad r_3 = (r; 2, 19; \pi) \quad e \quad r_4 = \left(r; 0, 96; \frac{3\pi}{2}\right). \quad (3.13)$$

O áudio em Formato A é definido como a gravação de áudio feito por cada um dos microfones sem processamento algum, indiferentes em relação de onde está chegando as ondas sonoras, sendo somente a representação no tempo da pressão sonora registrada na cápsula. Logo, um vetor  $s_A$  em cada componente é o valor de cada sinal de Formato A distribuído no arranjo pode ser descrito como:

$$s_A = [s_{LFU}, s_{RFD}, s_{LBD}, s_{RBU}] \quad (3.14)$$

A posição de cada microfones é dado, respectivamente (esquerda pra direita), pela Equação 3.13, e serão chamados daqui pra frente como sendo LFU (*left-front-up*), RFD (*right-front-down*), LBD (*left-back down*) e RBU (*right-back-up*), sendo a terminologia usada pela maioria dos fabricantes de tal tipo de sensor.

Um áudio de Formato B é um tipo de áudio derivado do Formato A, que representa um campo sonoro a partir de 4 canais de áudio: W, X, Y e Z. Então, o primeiro passo é convertermos os sinais de Formato A dos 4 microfones para o Formato B, propício para ser usado em aplicações envolvendo ambientes sonoros. Este formato de áudio traz a representação espacial do som no arranjo, cobrindo tanto o plano horizontal, quanto sons que vem acima e abaixo da origem dessa esfera.

Como limitamos o número de pontos amostrais até 4, a integral da Equação 3.11 pode ser reescrita como sendo uma soma que vai de  $l = 1$  a  $l = 4$ , como é mostrado na equação a seguir:

$$A_n^m(k) = -j(kr)^2 H_n'(kr) \sum_{l=1}^{L=4} g_l p(r, \theta, \Phi, k) Y_n^m(\theta_l, \Phi_l)^*, \quad (3.15)$$

em que  $g_l$  é o fator de quadratura introduzido na conversão de uma integral para uma soma (JOHANN-MARKUS, 2009).

A matriz com os termos dos harmônicos esféricos pode ser combinada com uma matriz de autovetores  $\Phi$ . Essa matriz contém os autovetores  $[Y_1^*, Y_2^*, Y_3^*, Y_4^*]$ , tal que  $Y_l(\theta_l, \Phi_l) = [Y_0^0, Y_1^{-1}, Y_1^0, Y_1^1]^T$  são os respectivos harmônicos esféricos para uma série de ordem  $N = 1$ . Se tomarmos a versão de  $Y_m^N$  com valores reais (graças à propriedade de ortogonalidade discutida na seção anterior) e os ângulos do microfones como sendo os apresentados na Equação 3.13, essa matriz calculada substituirá a matriz dos coeficientes harmônicos na Equação 3.15, podendo ser reescrita como:

$$\Phi = \frac{1}{4\pi} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \quad (3.16)$$

Se antes tínhamos uma equação matricial com esse formato:

$$\begin{bmatrix} A_0^0 \\ A_1^{-1} \\ A_1^0 \\ A_1^1 \end{bmatrix} = \begin{bmatrix} V_{0,\alpha} & 0 & 0 & 0 \\ 0 & V_{1,\alpha} & 0 & 0 \\ 0 & 0 & V_{1,\alpha} & 0 \\ 0 & 0 & 0 & V_{1,\alpha} \end{bmatrix} \begin{bmatrix} Y_{1,0}^0(\theta_l, \Phi_l) & Y_{2,0}^0(\theta_l, \Phi_l) & Y_{3,0}^0(\theta_l, \Phi_l) & Y_{4,0}^0(\theta_l, \Phi_l) \\ Y_{1,1}^{-1}(\theta_l, \Phi_l) & Y_{2,1}^{-1}(\theta_l, \Phi_l) & Y_{3,1}^{-1}(\theta_l, \Phi_l) & Y_{4,1}^{-1}(\theta_l, \Phi_l) \\ Y_{1,1}^0(\theta_l, \Phi_l) & Y_{2,1}^0(\theta_l, \Phi_l) & Y_{3,1}^0(\theta_l, \Phi_l) & Y_{4,1}^0(\theta_l, \Phi_l) \\ Y_{1,1}^1(\theta_l, \Phi_l) & Y_{2,1}^1(\theta_l, \Phi_l) & Y_{3,1}^1(\theta_l, \Phi_l) & Y_{4,1}^1(\theta_l, \Phi_l) \end{bmatrix} \begin{bmatrix} s_{LFU} \\ s_{RFD} \\ s_{LBD} \\ s_{RBU} \end{bmatrix} \quad (3.17)$$

fazendo essa mudança de base da matriz  $Y_n^m(\theta_l, \Phi_l)$  por  $\Phi$  na Equação 3.15, então os coeficientes dos harmônicos esféricos podem ser calculados pela seguinte equação matricial mais simples:

$$\begin{bmatrix} A_0^0 \\ A_1^{-1} \\ A_1^0 \\ A_1^1 \end{bmatrix} = \frac{1}{4\pi} \begin{bmatrix} V_{0,\alpha} & 0 & 0 & 0 \\ 0 & V_{1,\alpha} & 0 & 0 \\ 0 & 0 & V_{1,\alpha} & 0 \\ 0 & 0 & 0 & V_{1,\alpha} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} s_{LFU} \\ s_{RFD} \\ s_{LBD} \\ s_{RBU} \end{bmatrix} \quad (3.18)$$

sendo  $V_{0,\alpha}$  e  $V_{1,\alpha}$  funções derivadas das funções de Bessel e Hankel, em que nesse contexto são chamadas de filtro corretivos de não-coincidência, e serão discutidos na Seção 3.5.

A motivação para o uso do formato tetraedro é a sua capacidade de cobrir boa parte da superfície esférica, baixo custo e processamento mais rápido, atendendo frequências médias e baixas. Outro fator relevante é a simplificação que acontece ao usar 4 microfones, o que leva os sinais em Formato B (W, X, Y e Z) a serem iguais aos coeficientes harmônicos esféricos de primeira ordem  $[A_0^0, A_1^{-1}, A_1^0, A_1^1]$ , diminuindo bastante a complexidade do processamento necessário para encontrarmos os coeficientes harmônicos (essa simplificação vem a custo de menos precisão na caracterização do ambiente sonoro).

O arranjo com geometria tetraédrica da pesquisa, com sua casca esférica e cápsulas distribuídas em sua superfície, tem os microfones arranjados de tal forma que cada um está posicionado a 45 graus relativo aos eixos cartesianos, como é mostrado na Figura 3.8.

Solucionando a Equação 3.18, os quatro sinais em Formato A são combinados, como é mostrado na Equação 3.19, gerando os quatro sinais de Formato B, sendo um com

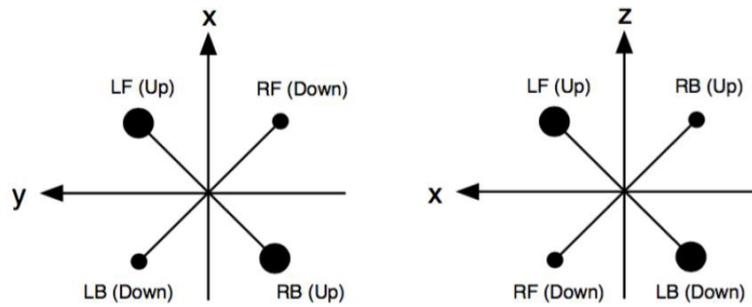


Figura 3.8 – Posicionamento no plano cartesiano de cada microfone presente no arranjo tetraedro — círculos maiores representam que o microfone está acima do plano bidimensional representado na figura, círculos menores indica que sua posição está inferior ao plano.

característica omnidirecional (sinal W) e três sinais equivalentes a um sensor proporcional à velocidade de partícula nos eixos cartesianos (X, Y e Z) (CHRISTOF; MIHAILO, 2009).

$$\begin{cases} W = s_{LFU} + s_{RFD} + s_{LBD} + s_{RBU} \\ X = s_{LFU} + s_{RFD} - s_{LBD} - s_{RBU} \\ Y = s_{LFU} - s_{RFD} + s_{LBD} - s_{RBU} \\ Z = s_{LFU} - s_{RFD} - s_{LBD} + s_{RBU} \end{cases} \quad (3.19)$$

Em altas frequências, a resposta diretiva dos sinais de Formato B tomam formas distorcidas. A Figura 3.9 mostra como a resposta do canal Omnidirecional W varia com a frequência que a fonte emite, e a Figura 3.10 mostra o mesmo fenômeno para o canal de figura de oito X.

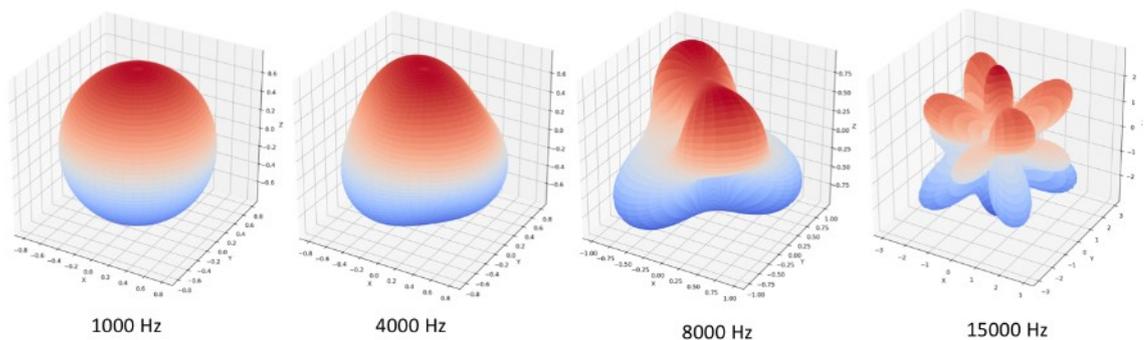


Figura 3.9 – Padrão polar tridimensional do sinal W para diferentes frequências. Fonte: (SCHULZE-FORSTER, 2018),

Será discutido, na próxima seção, formas de melhorar o desempenho do Formato B apresentando técnicas para equalizar os sinais e corrigir a não-coincidência da resposta diretiva do arranjo.

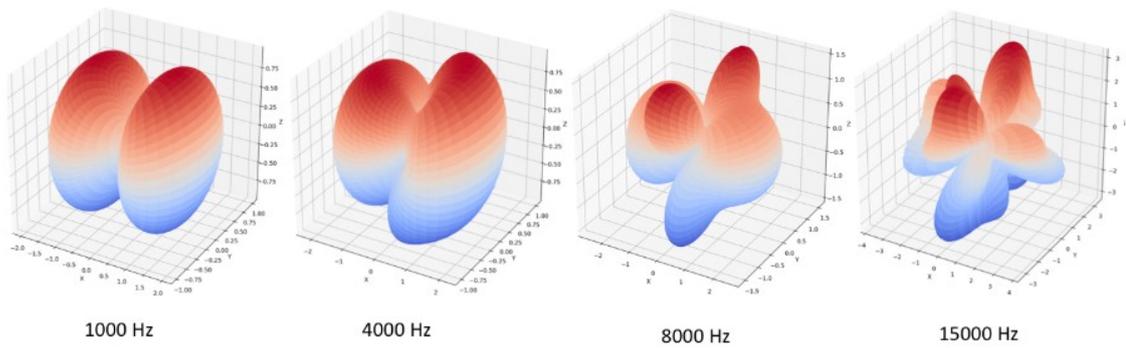


Figura 3.10 – Padrão polar tridimensional do sinal X para diferentes frequências. Fonte: (SCHULZE-FORSTER, 2018),

### 3.5 Filtros Corretivos

Olhe de novo a Equação 3.18 e foque a sua atenção na matriz que possui os termos  $V_{0,\alpha}$  e  $V_{1,\alpha}$ . Essa função é o resultado de combinações das funções de Bessel e de Hankel para caracterizar um campo sonoro livre com espalhamento ocasionado pela esfera rígida de raio  $r$  do arranjo:

$$V_{n,\alpha} = -i(kr)^2 H'_n(kr) \quad (3.20)$$

Sabendo que  $n$  é referente a ordem do harmônicos esféricos, e que para um arranjo tetraédrico teremos  $n$  indo de 0 até 1, podemos achar os valores de  $V_{0,\alpha}$  (filtro para a componente omnidirecional W) e  $V_{1,\alpha}$  (filtro para as componentes de figura de oito X, Y e Z), que terão a resposta em frequência descrita pela Figura 3.11:

Michael A. Gerzon, um grande pioneiro de áudio e das tecnologias de campos sonoros, avisa: “*acima de uma frequência limitante  $F = c/\pi r = 10,8/r$  kHz ( $r$  em cm), o diagrama polar se torna severamente contaminado por harmônicos esféricos de maior ordem, e descobriu-se que é melhor equalizar a saída omnidirecional (harmônicos esféricos de ordem nula) e as de figura de oito (harmônicos esféricos de primeira ordem) por uma resposta aproximadamente plana para campos sonoros homogêneos e aleatórios. Cálculos mostram que antes de qualquer equalização, as saídas omni e figura de oito de um arranjo tetraédrico regular de cardioides com  $r = 1,47$ cm terão uma resposta em frequência em campos sonoros aleatórios igual a mostrada pela Figura 3.13 (frequência limitante em 7,35 kHz). Isso deve convencer os céticos da falácia de assumir que ‘aproximadamente coincidente’ é bom o suficiente! No objetivo de garantir perfeita coincidência antes da frequência limitadora, as respostas de fase devem também serem compensadas pelas mudanças causadas pelos fatores da Equação 3.20*” (GERZON, 1975).

Então é necessário equalizar essas saídas com filtros que corrijam a não-coincidência

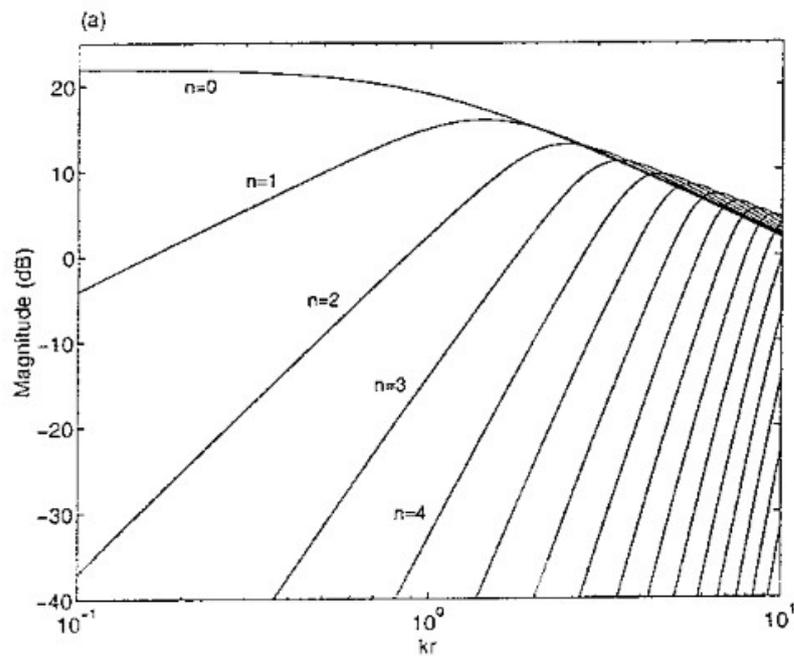


Figura 3.11 – Resposta em frequência dos filtros  $V_{n,\alpha}$  para harmônicos esféricos de ordem até 4, considerando um campo sonoro sobre uma esfera rígida.  $r$  é a distancia das cápsulas para o centro efetivo do arranjo. Fonte:(POLETTI, 2005)

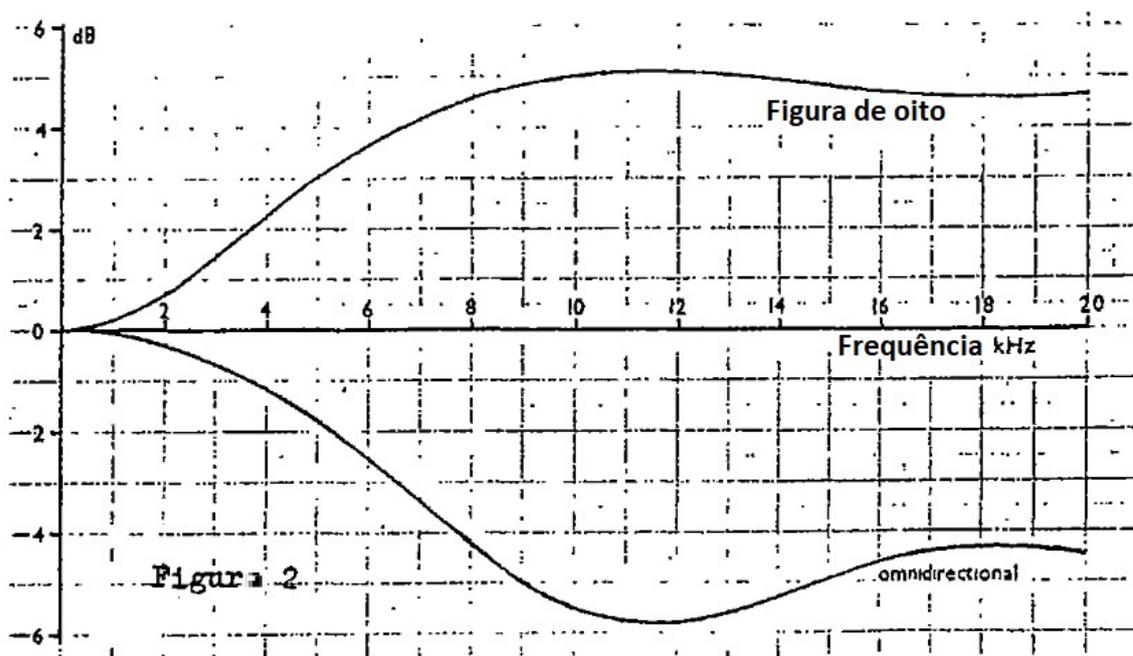


Figura 3.12 – Resposta em frequência dos filtros para a saída omni W e as de figura de oito X,Y,Z para um arranjo de raio=1,47cm. Fonte: (GERZON, 1975).

dos sinais, sendo que a solução teria então o seguinte formato:

$$\begin{cases} W_{\text{Final}} = W(t) * h_W(t) \\ X_{\text{Final}} = X(t) * h_X(t) \\ Y_{\text{Final}} = Y(t) * h_Y(t) \\ Z_{\text{Final}} = Z(t) * h_Z(t) \end{cases} \quad (3.21)$$

, sendo  $h_W(t), h_X(t), h_Y(t)$  e  $h_Z(t)$  funções de transferência para equalizar os sinais de Formato B.

Existe várias maneiras para se achar esses filtros de correção. Gerzon (1975) propõem a obtenção a partir da hipótese que um microfone esférico com cápsulas uniformemente distribuídas, com todas as saídas correspondendo a harmônicos esféricos de mesma ordem, precisam da mesma correção na resposta a frequência. Em (CHRISTOF; MIHAÏLO, 2009), os autores respondem à pesquisa de Gerzon, apontando que a determinação de tais filtros a partir somente da teoria é problemático, pois há várias nuances que fogem das considerações ideais assumidas (coincidência das cápsulas, precisão do posicionamento, efeitos de difração, etc.), assim os filtros determinados divergiriam consideravelmente dos filtros idealizados. Os dois propõem, então, filtros adaptativos a partir da média dos mínimos quadrados na vantagem que a perfeita simetria possa ser desconsiderada no cálculo dos filtros de correção para W, X, Y e Z, assumindo a dependência complexa com a frequência das respostas direcionais de W e X ( $D_W(\theta, \Phi, w)$  e  $D_x(\theta, \Phi, w)$ ), que podem ser medidas experimentalmente.

Mas Farina (2006) propõem um caminho mais rápido para corrigir o sinal de Formato B e resolver a Equação 3.21, aproveitando das facilidades que o desenvolvimento computacional nos proporciona, além de tirar a solução do plano teórico e levando para as adversidades da realidade, corrigindo desvios existentes para que os microfones tenham respostas em frequência e fase mais próximas entre si, pelo menos na posição em que ocorreu as medições.

O método proposto por ele começa com medidas em uma sala anecoica dos sinais dos microfones do arranjo. O primeiro passo é fazer medidas separadas de cada cápsula do arranjo e de um microfone para medida de referência, todos na mesma posição (ou lado a lado), e assim calcular a função de transferência entre as cápsulas (LFU, RFD, LBD ou RBU) e o microfone de referência, servindo como molde para os filtros equalizadores, garantindo que a resposta em frequência e fase das cápsulas sejam idênticas entre si. Sendo que cada microfone precisa ter a saída igual ao microfone de referência, os filtros para cada cápsula terão resposta em frequência igual à divisão da resposta em frequência do microfone de referência com a resposta em frequência da cápsula que será equalizada, respeitando a propriedade de convolução do domínio da frequência:

$$\begin{cases} h_{LFU} = \frac{\text{MicRef}(\omega)}{S_{LFU}(\omega)} \\ h_{RFD} = \frac{\text{MicRef}(\omega)}{S_{RFD}(\omega)} \\ h_{LBD} = \frac{\text{MicRef}(\omega)}{S_{LBD}(\omega)} \\ h_{RBU} = \frac{\text{MicRef}(\omega)}{S_{RBU}(\omega)} \end{cases} \quad (3.22)$$

O segundo passo é fazer a conversão para Formato B, processando os sinais das cápsula através desses filtros equalizadores, combinando em uma matriz os sinais resultantes filtrados e aplicando os sinais na Equação 3.19, o que resulta nos sinais de Formato B já corrigidos. Em seguida são feitas três gravações, posicionando o microfone de referência em cada eixo do plano cartesiano, para que cada canal do Formato B (W, X, Y, Z) possa ser equiparado com o sinal do microfone de referência:

$$\begin{cases} h_W = \frac{\text{MicRef}_W(\omega)}{W(\omega)} \\ h_X = \frac{\text{MicRef}_X(\omega)}{X(\omega)} \\ h_Y = \frac{\text{MicRef}_Y(\omega)}{Y(\omega)} \\ h_Z = \frac{\text{MicRef}_Z(\omega)}{Z(\omega)} \end{cases} \quad (3.23)$$

O cálculo da função de transferência entre os 4 canais de Formato B é feito para as entradas X, Y e Z com os sinais das cápsulas já pré-filtradas (Equação 3.22), e com a gravação do microfone de referência servindo como saída, é calculado toda a estrutura de pós-filtragem, uma final correção do sinal. O pós-filtro para o canal W é diminuído por 3 dB (coeficientes multiplicados por 0,707) para compensar o ganho resultante da soma dos sinais de Formato A para o canal omnidirecional. Ao final, teremos 8 filtros FIR, como é mostrado na Figura 3.13.

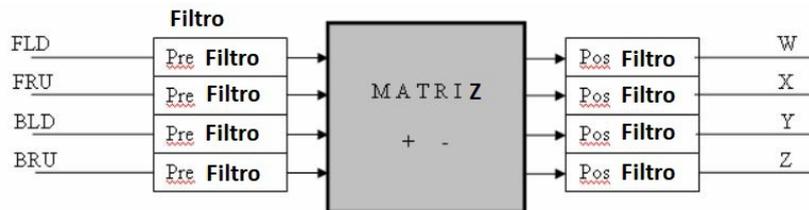


Figura 3.13 – Esquemática do processamento dos sinais nos filtros de correção. Adaptada de (FARINA, 2006).

A Figura 3.14 mostra como os filtros obtidos a partir de medições reais de Farina se diferem bastante dos filtros teóricos propostos por Gerzon da Figura 3.12. Qualquer consideração minimalista enquanto à natureza do arranjo precisa ser colocada na mesa.

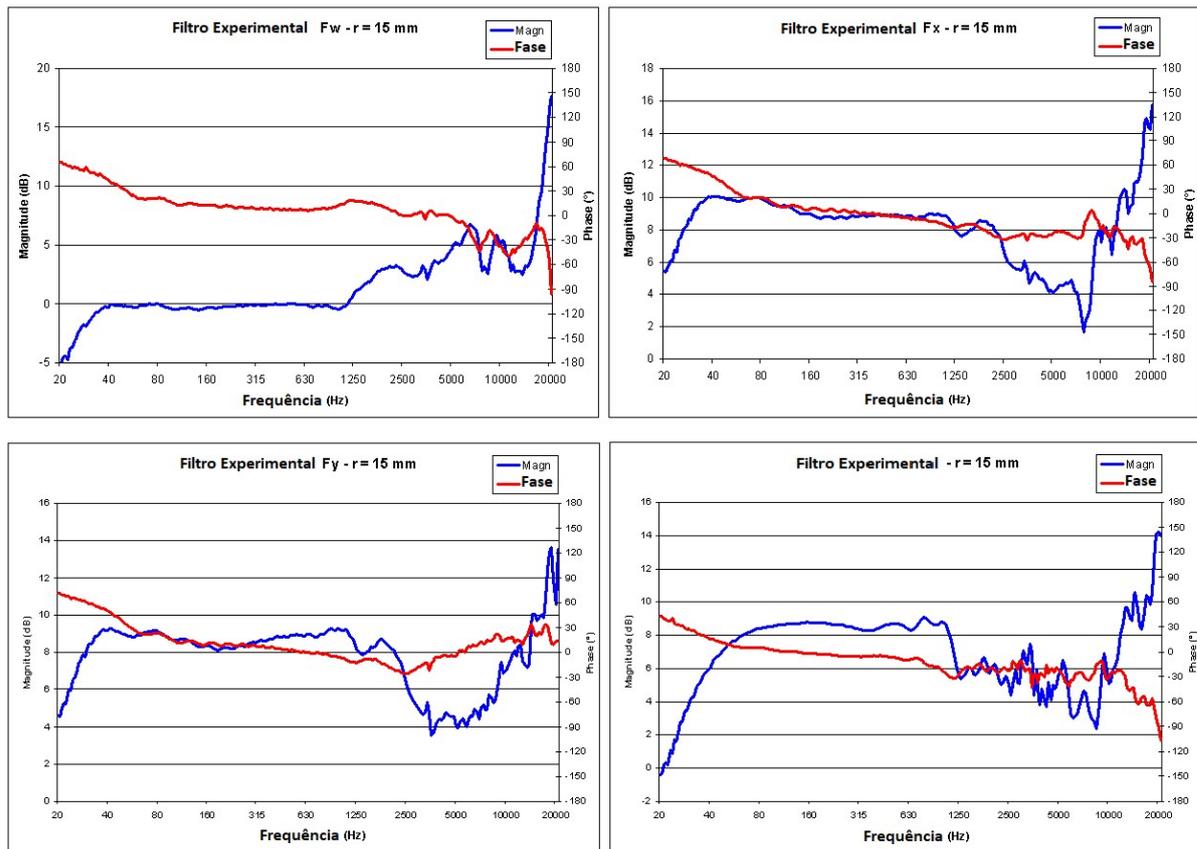


Figura 3.14 – Resposta em frequência e fase dos filtros equalizadores determinados a partir de medidas experimentais para o caso de um arranjo com microfones cardioides ( $r = 1, 5$  cm). Adaptada de (FARINA, 2006).

Mas uma coisa que não é considerada nas técnicas citadas até agora é o fator difuso que um ambiente sonoro pode ter. Se o arranjo tetraédrico é um sensor acústico que busca ter uma resposta em frequência estável para qualquer direção de chegada, e se o cenário em análise é um ambiente onde várias fontes estão emitindo sons de grande intensidade, e esses sons são refletidos em outros equipamentos, ou no chão, é crítico calibrar o arranjo para que ele tenha uma resposta em frequência plana para qualquer incidência sonora, independente do ângulo.

Historicamente, reverberação acústica em salas tem sido modelada como ruído isotrópico (independentes da direção) esférico. Tal campo sonoro pode se construído por combinações de ruídos não-correlacionados em todas as direções e de mesma intensidade. Em acústica de salas, esse campo ruidoso modelado é referido como um campo sonoro “difuso”(ELKO, 2004).

A resposta em campo difuso de um sistema acústico é o valor eficaz da resposta do arranjo em todas as direções da esfera, relativo a um microfone omnidirecional (HELLER; BENJAMIN, 2012). Para um único microfone com simetria rotacional em seu eixo de máxima sensibilidade, o **ISO/IEC 60248-4** (IEC-60268-4, 2010) define a resposta em

campo difuso como sendo:

$$M_{\text{diff}}^2(f) = \frac{1}{2} \int_0^\pi M(\theta)^2 d\theta \quad (3.24)$$

No caso do arranjo tetraédrico, as respostas do arranjo não são rotacionais simetricamente em volta de um eixo, então é empregado uma forma generalizada das expressões de fator diretivo dado em (SALVADOR-BARRERA *et al.*, 2007) e (KINSLER *et al.*, 1982), culminando na Equação 3.25:

$$M_{\text{diff}}^2(f) = \frac{1}{4\pi} \int_{4\pi} |M(\Omega, f)|^2 d\Omega \quad (3.25)$$

, sendo  $M_{\text{diff}}^2(f)$  a resposta em campo difuso do arranjo,  $M(\Omega, f)$  é o ganho complexo de uma saída particular do arranjo (um dos componentes de Formato B; W, X, Y e Z) para uma onda de frequência  $f$ , com direção de chegada  $\Omega$ .

Em (HELLER; BENJAMIN, 2012), o ajuste da resposta do arranjo é feito a partir da resposta de campo difuso, e não só da resposta de campo livre de cada sinal de formato B, aproximando a Equação 3.25 pela medição das respostas em um campo difuso em várias direções distintas e calculando a soma ponderada dessas respostas — isso asseguraria que, independente da direção, os sinais terão resposta plana em frequência.

Pode-se simular um campo difuso se distribuirmos vários alto falantes em uma sala com bastante reverberação, para que o som seja homogêneo sobre qualquer direção de incidência.

A Figura 3.15 mostra um ajuste real de um arranjo tetraédrico feito por (HELLER; BENJAMIN, 2012). Primeiro é medido a resposta em campo difuso de cada cápsula, em seguida, com essa resposta, calcula-se a resposta para cada sinal de Formato B e em seguida equalizadores de fase mínima (HELLER; BENJAMIN, 2012) são aplicados para “empurrar” a resposta dos sinais de Formato B para uma resposta plana em campo difuso.

O resultado é uma resposta equalizada de campo difuso plana até uma certa frequência limitante, sendo que a limitação em frequências mais altas seja resultado da resposta em campo difuso particulares de cada microfone do arranjo.

A próxima seção trará algumas aplicações que o arranjo tetraédrico pode atuar, apoiando-se nas vantagens de ter uma característica diretiva no sinal de áudio gravado.

## 3.6 Aplicações do Formato B.

O principal objetivo do sistema a ser desenvolvido é entregar os canais do Formato B utilizando um arranjo tetraédrico *wireless*, mas existe uma gama de formas

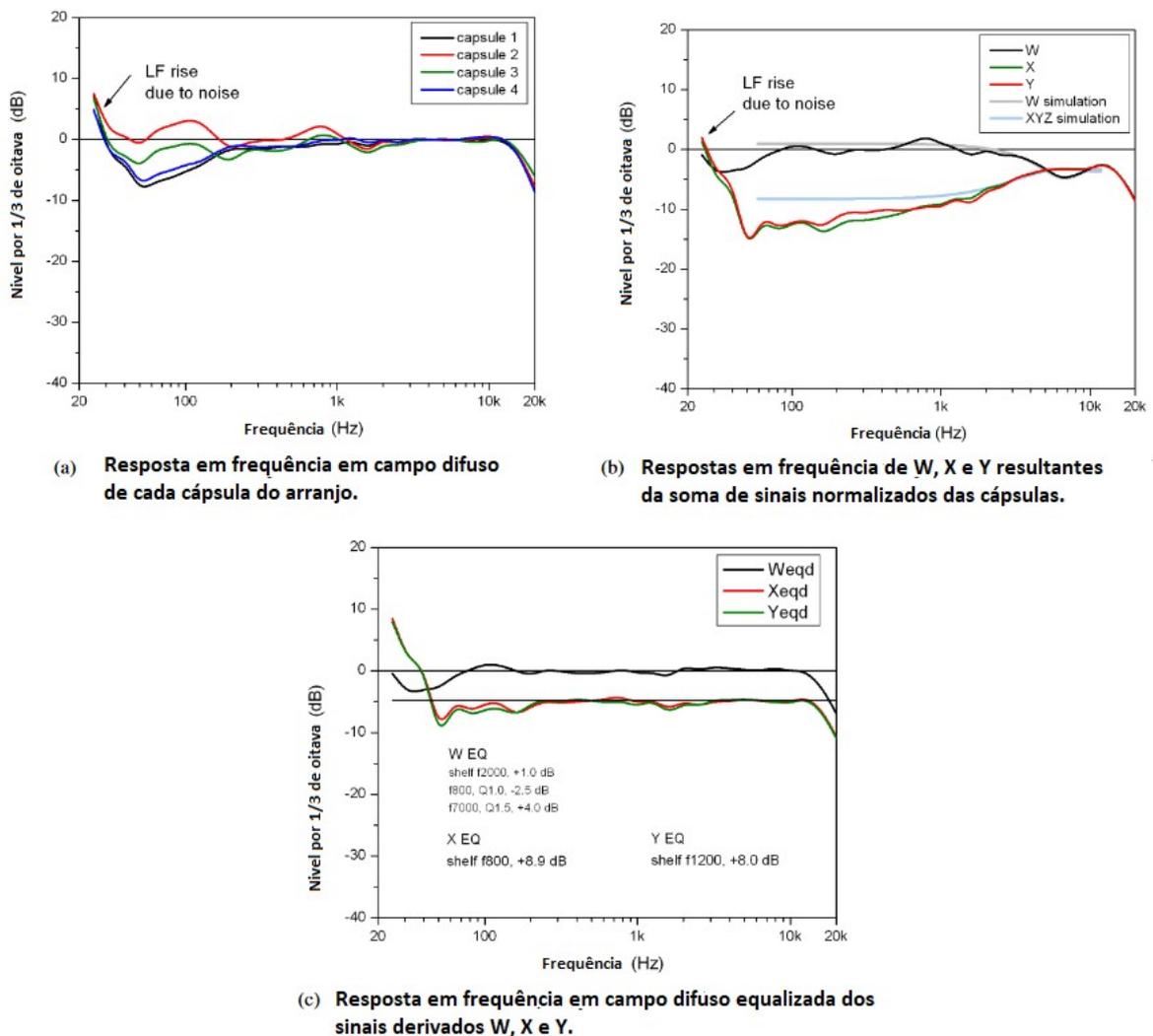


Figura 3.15 – Respostas em frequências puras e equalizadas de um arranjo de teste. Adaptada de (HELLER; BENJAMIN, 2012).

de utilizar a informação espacial que o Formato B entrega, e essa seção irá abordar resumidamente sobre duas aplicações.

### 3.6.1 Filtragem espacial.

Com a obtenção dos sinais de Formato B, pode-se aplicar técnicas (que podem ser até triviais) de manipulação do campo sonoro gravado. Com a separação dos sinais vindos de várias fontes sonoras, podemos avaliar o comportamento acústico direcionado, utilizando do padrão diretivo manipulável para dar ênfase em uma direção.

A filtragem espacial do ambiente sonoro é a técnica de microfone virtual (SAN- NUM, 2018), que consiste na projeção no ambiente sonoro de um canal referente à gravação de um microfone apontando para a direção  $(\theta, \phi)$ . A ideia consiste na manipulação do sinal com padrão polar omnidirecional W com os de figura de oito X, Y e Z, como mostra

a equação a seguir:

$$M(\theta, \phi) = \alpha\sqrt{2}W + (1 - \alpha)(\cos \phi \cos \theta X + \sin \phi \cos \theta Y + \sin \theta Z) \quad (3.26)$$

, sendo  $\alpha$  o padrão polar ajustável.

Uma boa forma de avaliar a gravação do ambiente sonoro é imaginar que cada fonte em si disponibiliza uma assinatura em Formato B para o tetraedro, com seu próprio  $W$  e as ponderações da direção do som para cada coordenada cartesiana  $X$ ,  $Y$  e  $Z$ . Se mais fontes são colocadas no ambiente, é como se fosse somado ao microfone *ambisonic* mais assinaturas no canais de Formato B, cada qual com a sua direção de chegada.

Podemos “direcionar” o Formato B a partir da ponderação dos canais  $X$ ,  $Y$  e  $Z$  para uma direção conhecida, como é mostrado na Equação 3.26. Dessa forma, é realçado o som vindo dessa direção de interesse e atenuando os sons de outras direções.

### 3.6.2 Direção de chegada de uma onda sonora.

Uma aplicação bastante promissora do arranjo tetraédrico é sua combinação com aprendizado de máquina para que seja determinado qual é a direção que está vindo o som captado pelo arranjo. A direção de chegada (DOA, do inglês *Direction of Arrival*) é determinada pelos vetores de intensidade sonora, com a difusão do campo sonoro sendo estimado pela relação entre a intensidade sonora e a densidade geral de energia (KUECH *et al.*, 2008). Ahonen (AHONEN *et al.*, 2007) propôs uma técnica que utiliza dessas considerações, com estimações dos vetores de intensidade feitas a partir dos sinais de um arranjo que possui harmônicos esféricos de primeira ordem como o seu padrão direcional.

Para encontrarmos os vetores de intensidade de cada eixo do Formato B, primeiramente, é preciso fazer as transformadas de Fourier de tempo curto (STFT, do inglês *Short Time Fourier Transform*) para cada canal de Formato B, em seguida, multiplicar os canais  $X$ ,  $Y$  e  $Z$  com o conjugado da STFT do canal  $W$ , retirar a parte real do resultado da multiplicação e dividir por 2 o produto final. Com isso, teremos os seguintes sinais no domínio da Frequência-Tempo:

$$\begin{cases} I_X(f, t) = (1/2)\text{Re}\{W^*(f, t)X(f, t)\} \\ I_Y(f, t) = (1/2)\text{Re}\{W^*(f, t)Y(f, t)\} \\ I_Z(f, t) = (1/2)\text{Re}\{(W^*(f, t)Z(f, t)\} \end{cases} \quad (3.27)$$

Utilizando esses parâmetros calculados, pode-se estimar a direção de chegada de um som, permitindo apontar onde está a fonte sonora pelos vetores de intensidade.

## 4 Metodologia

Agora que foi deixado mais claro a natureza do equipamento, esta seção será a ponte entre teoria e prática. Desde a escolha dos componentes até os experimentos a serem feitos, a funcionalidade do arranjo para que ele atenda aos requisitos necessários para caracterização do ambiente sonoro é resultado do cuidado com a teoria aplicada, e cada ponto terá uma explicação fundamentada nessa teoria.

### 4.1 Hardware

O primeiro passo para a construção do protótipo é determinar quais componentes farão parte do seu corpo, e o melhor caminho é começar pelo transdutor que captará o sinal sonoro, ou seja, o modelo que iremos utilizar nos 4 microfones que irá compor o arranjo.

O microfone precisa ser escolhido pensando em como e onde ele vai atuar. A seguir é listado os requisitos que contam na escolha do componente:

- Resposta em frequência: o microfone precisa ter uma resposta de frequência que atenda aos valores comuns de sonoridade. Um microfone que atenda uma banda de 20Hz até 40kHz saciaria essa demanda.
- Razão sinal-ruído: sendo a razão sinal-ruído a diferença em decibéis entre a faixa dinâmica e o ruído de fundo (conhecido também como *noise floor*), busca-se microfones que tenham uma boa faixa dinâmica e pouco ruído, ou seja, a faixa de amplitude em que o microfone consegue captar sons sem distorção e ignorando o ruído autogerado. Uma maior faixa dinâmica implica uma maior diferenciação entre sons de menor intensidade e os de maior, permitindo monitorar sinais sonoros de baixa e alta intensidade. Equipamentos com razão sinal-ruído maior que 60 dB são considerados aptos para entregar sinais de qualidade.

Considerando essas demandas, foi escolhido o microfone IMP23ABSUTR (ST-Microelectronics NV, 2020) para o projeto, já que ele atende ao que é exigido pelo protótipo. A Figura 4.1 mostra a sua resposta em frequência. É possível notar que o microfone escolhido possui uma resposta quase unitária até 10 kHz, o que já um ponto forte para o Formato B, sendo essas as frequências que ele já atua bem. Mas outro ponto forte é que ele pode atuar em frequências de ultrassom, como é mostrado na Figura 4.2. Mesmo

que o ideal seria uma resposta constante, podemos ajustar a resposta em frequência para testar cenários em ultrassom futuramente.

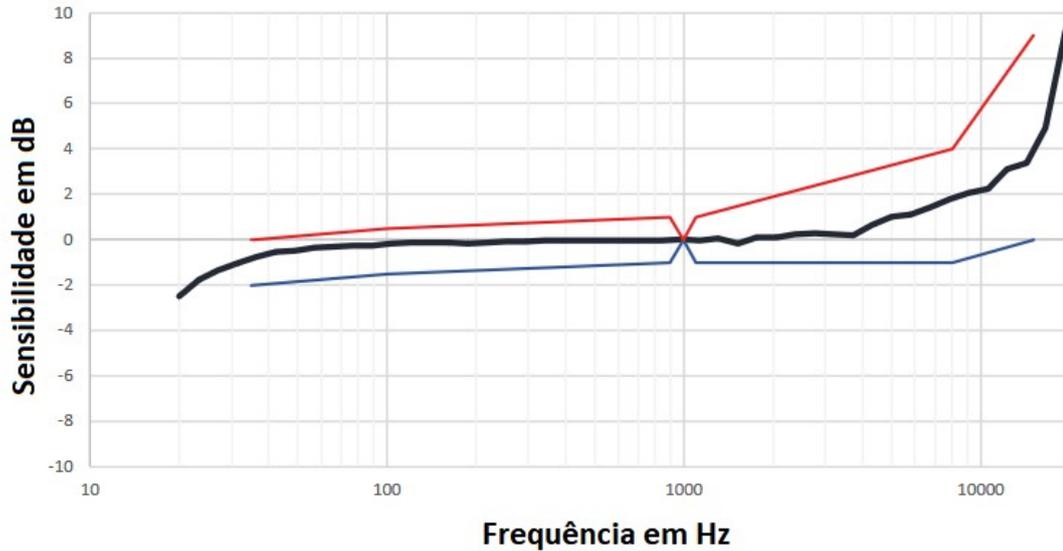


Figura 4.1 – Resposta em frequência normalizada em 1 kHz do IMP23ABSUTR. Autor:(STMicroelectronics NV, 2020)

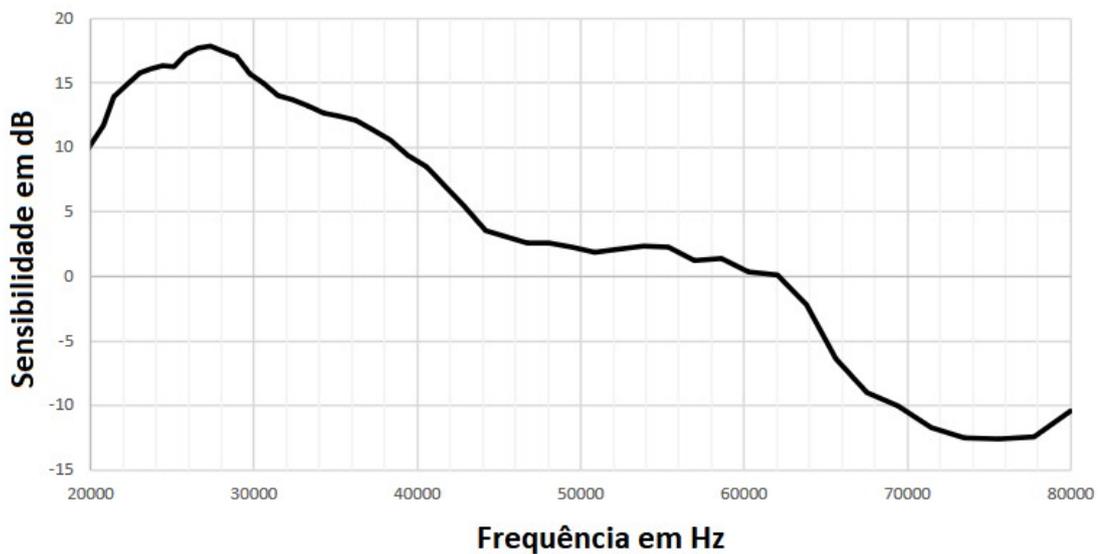


Figura 4.2 – Resposta em frequência normalizada em 1 kHz do IMP23ABSUTR em faixa de ultrassom. Autor:(STMicroelectronics NV, 2020)

O IMP23ABSU tem um ponto de sobrecarga acústica de 130 dB SPL, além de 64 dB de razão sinal-ruído.

Os microfones são fixados em uma estrutura montada a partir da confecção de uma esfera polimérica produzida com impressora 3D, contendo orifícios dispostos no vértice de um tetraedro (inscrito na esfera). Os arquivos **.stl** (formato de arquivo comumente

usado em impressão 3D e outros programas que facilitam o design de projetos e desenhos técnicos) são disponibilizados aqui<sup>1</sup>. Na Figura 4.3 pode-se ver o modelo tridimensional da base a ser confeccionada e na Figura 4.4, as duas metades de uma esfera de diâmetro de 36 mm com orifícios distribuídos levando em consideração as posições estabelecidas pela Equação 3.13.



Figura 4.3 – Base/suporte feito utilizando software baseado em estereolitografia.

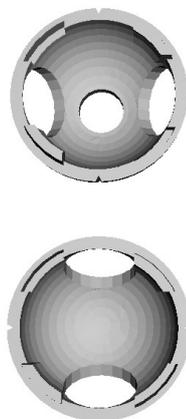


Figura 4.4 – Peças para uma estrutura esférica feitas utilizando software baseado em estereolitografia..

A estrutura do protótipo é mostrada na Figura 4.5.

Foi então projetada uma caixa de conexões para que as saídas dos circuitos internos dos microfones sejam conectados a *bornes* do tipo fêmea P10 (cabos comumente utilizados em saídas de instrumentos elétricos, como a guitarra), e para que tenham um circuito de alimentação regulada que possa energizar todos eles. Os microfones são alimentados a partir de uma bateria recarregável de 3,7 V. A Figura 4.6 mostra a topologia

<sup>1</sup> <[https://drive.google.com/drive/folders/1O8RJ\\_HVllWBXNXQ6uB0VUwYW8qIRo5zD](https://drive.google.com/drive/folders/1O8RJ_HVllWBXNXQ6uB0VUwYW8qIRo5zD)>



Figura 4.5 – Estrutura completa do arranjo tetraédrico projetado.

do circuito projetado para conexão e alimentação dos microfones e a Figura 4.7 mostra a caixa de alimentação projetada.

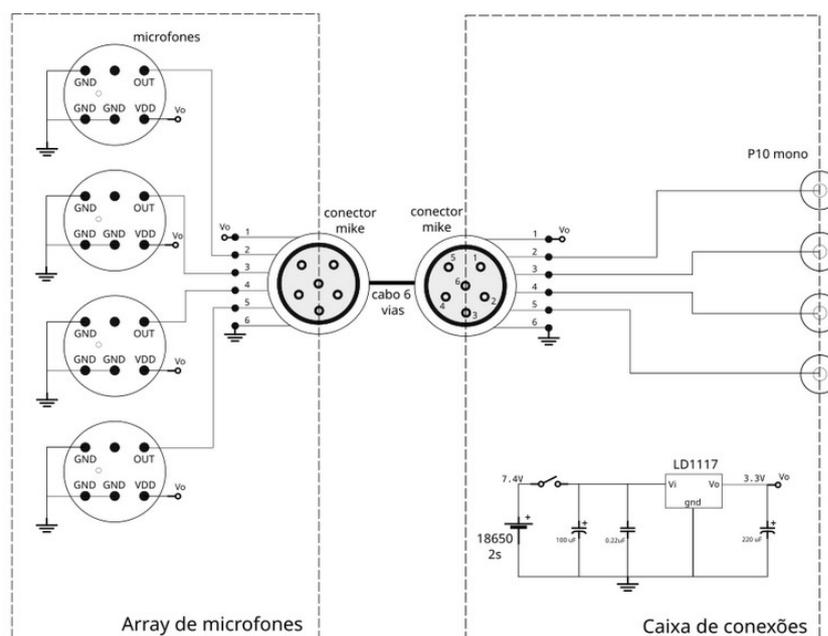


Figura 4.6 – Diagrama de conexões entre o array de microfones e a caixa de conexões.

Os sinais são processados em um microcontrolador Raspberry Pi 3, escolhido por ter uma velocidade de *clock* de 1,2 GHz (mais do que o suficiente para processamento de sinais de áudio em ultrassom), capacidade de comunicação *WiFi* e 1 Gb de memória RAM. Mas o ponto mais favorável para desenvolvimento do projeto é a sua ligação com a linguagem Python, uma linguagem fácil de ser manipulada e a que é compõem o código do *software* do sistema projetado.

Por último, utilizando uma das portas USB 2 do RaspBerry Pi 3, é conec-



Figura 4.7 – Fotografia da caixa de conexões e fonte de alimentação regulada para microfones.

tado uma interface de áudio Scarlett 8i18, trabalhando como pré-amplificador e conversor analógico-digital. A vantagem de usar a Scarlett é sua capacidade de gravar 8 canais simultaneamente, além de facilitar a comunicação de dados entre arranjo e microcontrolador com sua transmissão via USB. Outro fator interessante é sua taxa de amostragem, que pode ser colocada até 192 kHz, taxa esta propícia para estudos em ultrassom.

A Figura 4.8 detalha o diagrama do projeto, apontando a função de cada componente e o fluxo de operação do sistema.



Figura 4.8 – Diagrama de funcionamento do sistema de captação acústica.

A Tabela 4.1 detalha os itens e os preços deles — o orçamento para o sistema de captação de áudio em Formato B de baixo custo e *wireless*. Se retirarmos a interface de áudio (equipamento utilizado para receber os sinais de áudio em 4 canais) e se for disponibilizado uma impressão 3D gratuita (como é disponibilizado em algumas instituições de ensino superior), o preço do protótipo cai para 395,39 R\$.

Tabela 4.1 – Orçamento do sistema de captação acústica.

Item	Preço unitário	Unidade	Preço total
Analog MEMS Microphone	12,97 R\$	4	51,88 R\$
Impressão 3D das peças (hora de impressão)	150,00 R\$	2	300,00 R\$
Scarlett 8i18	2.399,00 R\$	1	2.399,00 R\$
Cabo 15 cm com duas extremidades P10 mono	20,00 R\$	4	80,00 R\$
Conector jack P10 mono fêmea de painel	17,99 R\$	4	71,96 R\$
Par de conector Mike 6 vias macho/femea par	12,00 R\$	1	12,00 R\$
Metro de cabo manga blindado 6 vias x 26 awg	11,00 R\$	3	33,00 R\$
Regulador de tensão 3V3	25,00 R\$	1	25,00 R\$
Bateria Li-ion 6800msh 3,7V recarregável	25,00 R\$	2	50,00 R\$
Suporte para baterias	17,50 R\$	1	17,50 R\$
Caixa plástica para montagem eletrônica	50,00 R\$	1	50,00 R\$
Chave gangorra liga desliga	4,05 R\$	1	4,05 R\$
<b>TOTAL</b>			<b>3.094,39 R\$</b>

## 4.2 Software

Agora que todo o sistema já foi planejado, é hora de programar o Raspberry Pi 3 para a gravação multicanal simultânea, que é feita em Python utilizando a biblioteca Pyaudio. Introduzindo de forma rápida, Pyaudio é a ligação entre aplicações em Python com o PortAudio (biblioteca para entrada/saída de dados de áudio em C ou C++) (PHAM, 2022), permitindo uma fácil programação e processamento de sons em um ambiente de desenvolvimento de Python.

O código escrito no Raspberry Pi 3 é disponibilizado no Anexo A, e as gravações de áudio são feitas para 4 canais com codificação de 16 bits para cada amostra de áudio coletada de cada canal. Logo, são  $4 \times 16$  ou 8 bytes por amostra de áudio, em uma taxa de amostragem de 192000 Hz. Os parâmetros mais importantes para controlar a gravação de áudio em Python são os seguintes:

- **CHUNK**: Tamanho do *buffer* de amostras de áudio por comando de gravação.
- **FORMAT**: Tipo de conversão digital utilizada, no código original é usado 16 bits de números inteiros para a amostragem.
- **CHANNELS**: Número de canais para a gravação simultânea.
- **RATE**: Taxa de amostragem da gravação.

O código grava os 4 canais simultaneamente, como é mostrado na Figura 4.9 mas não terá gente disponível para ir no local, gravar manualmente e retirar os 4 sinais em Formato A para processar. Então é utilizado a biblioteca Socket para administrar a

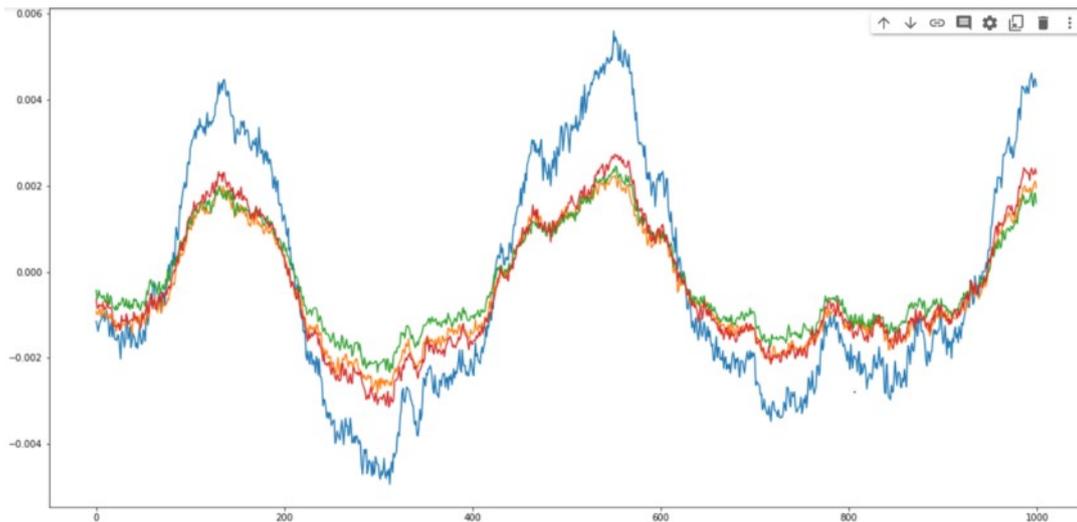


Figura 4.9 – Os 4 sinais (amplitude  $\times$  tempo) referentes a gravação dos 4 canais do arranjo.

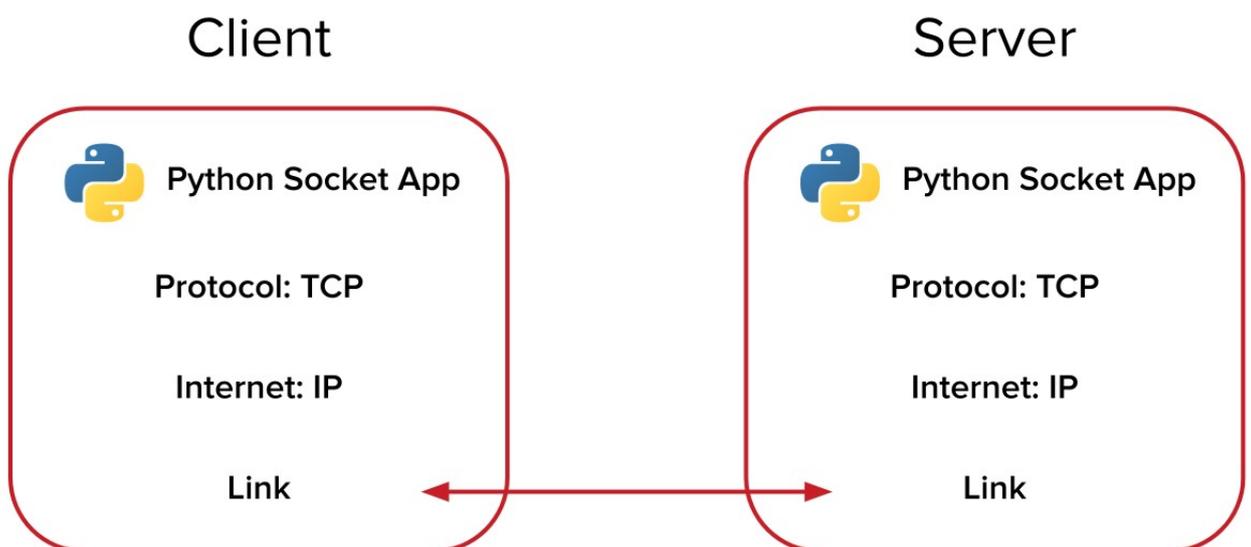


Figura 4.10 – Diagrama explicativo de como a biblioteca Socket opera.

transmissão via WiFi. A Figura 4.10 é um diagrama de como funciona o elo entre servidor e cliente. Se ambos os terminais estiverem na mesma conexão *WiFi* e é conhecido o IP (*Internet Protocol* ou, em português, Protocolo de Rede) de cada terminal, a conexão é possível para a transmissão de dados, e para o nosso sistema, o cliente será o terminal de gravação (Raspberry Pi 3) e o servidor será o terminal de processamento (computador com sistema operacional Windows).

O maior desafio da transmissão é conseguir manipular o tamanho do *buffer* (pacote de amostras que são gravadas por ciclo). Foi escolhido 2048 bytes por *buffer* como taxa de transmissão a partir de vários testes com valores entre 256 a 4096, além de

consultas em fóruns de programação (a comunidade de programadores indica fazer blocos com tamanho de potências de 2), o valor que conseguiu manter um fluxo contínuo de transmissão foi o de 2048 bytes.

As mensagens de rotina "CHUNK COMPLETE", "CHUNK INCOMPLETE", "READY" e "FILE SENT" são mensagens que, codificadas, tem o mesmo tamanho em bytes e servem para guiar a transmissão de dados, caso algum pacote se perca na transmissão ou quando o sistema está pronto para mais aquisições. Esse cuidado possibilita que não haja áudios incompletos no final da transmissão.

Feita as considerações a cerca da amostragem e transmissão, iremos agora direcionar a atenção para o código do servidor, que é mostrado na íntegra no Anexo B.

O código primeiro aguarda a conexão com a Raspberry Pi, e após a confirmação, começa a receber 2048 bytes por *buffer* de áudio. Completada a transmissão que no total representam 10 segundos de gravação, é encerrado a rotina de transferência de dados e começa o processamento do áudio: separando os 4 canais dos microfones, convoluindo os sinais de Formato A com os filtros calculados a partir dos passos da Seção 3.5 e fazendo as operações com os vetores resultantes seguindo o sistema linear da Equação 3.21 para obter os sinais de Formato B, que então são concatenados entre si e salvos no formato FLAC (acrônimo de *Free Lossless Audio Codec*, que significa Codec de Áudio Livre Sem Perdas). Após salvo, o programa repete o *loop*.

### 4.2.1 Filtros equalizadores do arranjo

Com o auxílio do Instituto de Pesquisa Eldorado, foi disponibilizado uma sala semi anecoica para que fosse feita a caracterização dos microfones que serão utilizados no arranjo. A importância desse ensaio é que, a partir da gravação dos microfones em um ambiente controlado e utilizando da mesma entrada (fonte sonora posicionada na mesma distância de cada microfone, na mesma inclinação e azimute) para cada microfone, podemos caracterizar cada microfone e aplicar correções para garantir a coincidência de resposta ao sinal em todos os microfones, como é sugerido por Farina (FARINA, 2006) e descrito na Seção 3.5.

A Figura 4.11 mostra como foram posicionados a fonte e os microfones e, a partir de um nível a laser, há a garantia que o centro do microfone esteja apontando diretamente para o centro da fonte.

Cada cápsula é empilhada como é mostrado na Figura 4.12, e com uma manivela colocada no suporte que mantém os microfones fixos, podemos ajustar a altura, assim, para cada medida referente a cada microfone, mantenha-se a posição de referência com a fonte, no mesmo plano.



Figura 4.11 – Fonte e microfones posicionados para o ensaio em sala anecoica. Fonte: Autor



Figura 4.12 – Suporte auxiliar para a caracterização dos microfones. Fonte: Autor

A fonte sonora reproduz então uma varredura linear de frequência, de 0 a 96 kHz, cobrindo a grande banda de interesse, para que seja gerada a resposta em frequência de cada cápsula.

As gravações foram feitas utilizando o ITA-Toolbox (BERZBORN *et al.*, 2017) do MatLAB, uma biblioteca poderosa para estudos em acústica. Com ela, pode-se amostrar dados de vários canais, apresentar tais dados no domínio do tempo, da frequência, ou em espectrograma, além de outras ferramentas que entregam robustez e informação na coleta de dados.

A Figura 4.13 mostra a resposta em frequência e pode-se observar que, por ser um Tweeter, a fonte consegue entregar um sinal de qualidade somente para frequências acima de 2 kHz e abaixo de 40 kHz. Não foram utilizados outros falantes para completar as outras faixas pois o foco da análise seria em uma faixa entre média e ultrassom, além de que a não-coincidência do arranjo de Formato B é um fenômeno que ocorre perto de 3 kHz a 5 kHz para um arranjo com raio de 3 cm.

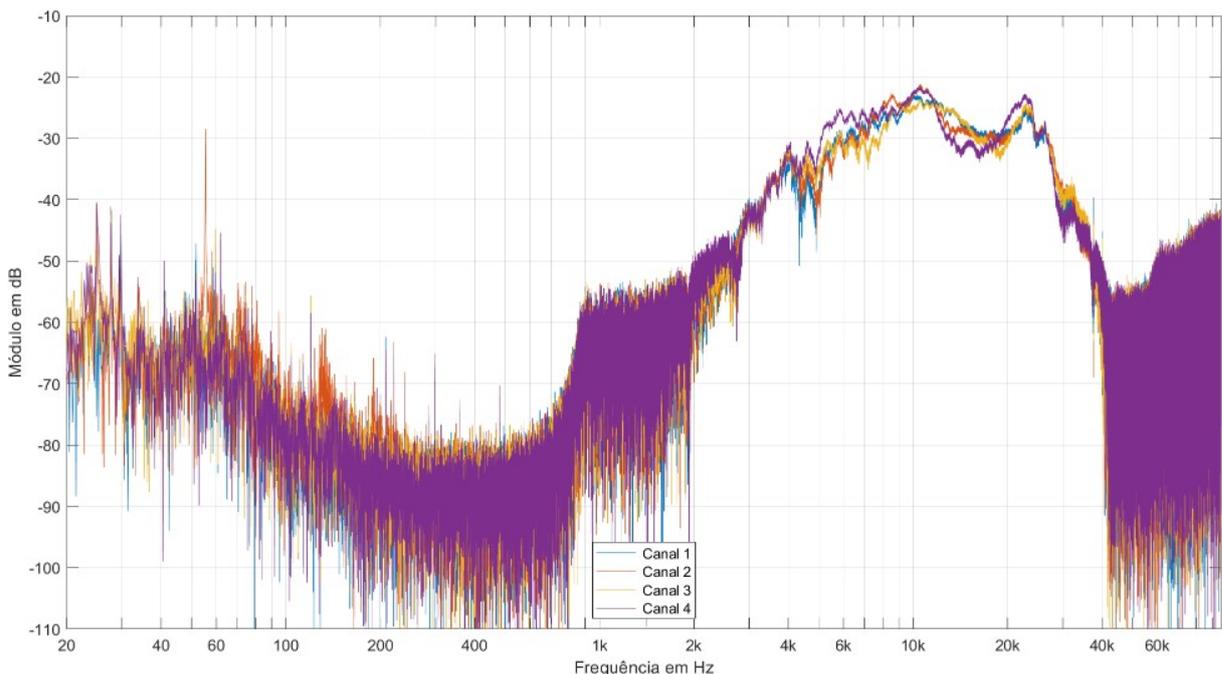


Figura 4.13 – Respostas em frequência de cada microfone sobre as mesmas condições (curva azul é o primeiro microfone da placa, de cima para baixo, curva laranja para o segundo microfone, amarelo para o terceiro e roxo para o ultimo). Fonte: Autor

Vale apontar que, para uma melhor caracterização do sinal captado, um janelamento no tempo para retirar a influência de componentes não-lineares do sinal é extremamente recomendado — não queremos sujar a resposta medida. A Figura 4.14 mostra o espectrograma do impulso referente a varredura de frequência. Os riscos entre os segundos 9 e 10 são as distorções não-lineares da varredura linear causadas pela dis-

torção do alto-falante, sendo que o sinal de interesse está somente na faixa colorida na esquerda da imagem. Logo o corte será entre 0 e 0,1 segundos (corte feito simplesmente selecionando somente as primeiras amostras até ter-se 0,1 segundos de gravação), focando somente a faixa de maior energia do gráfico.

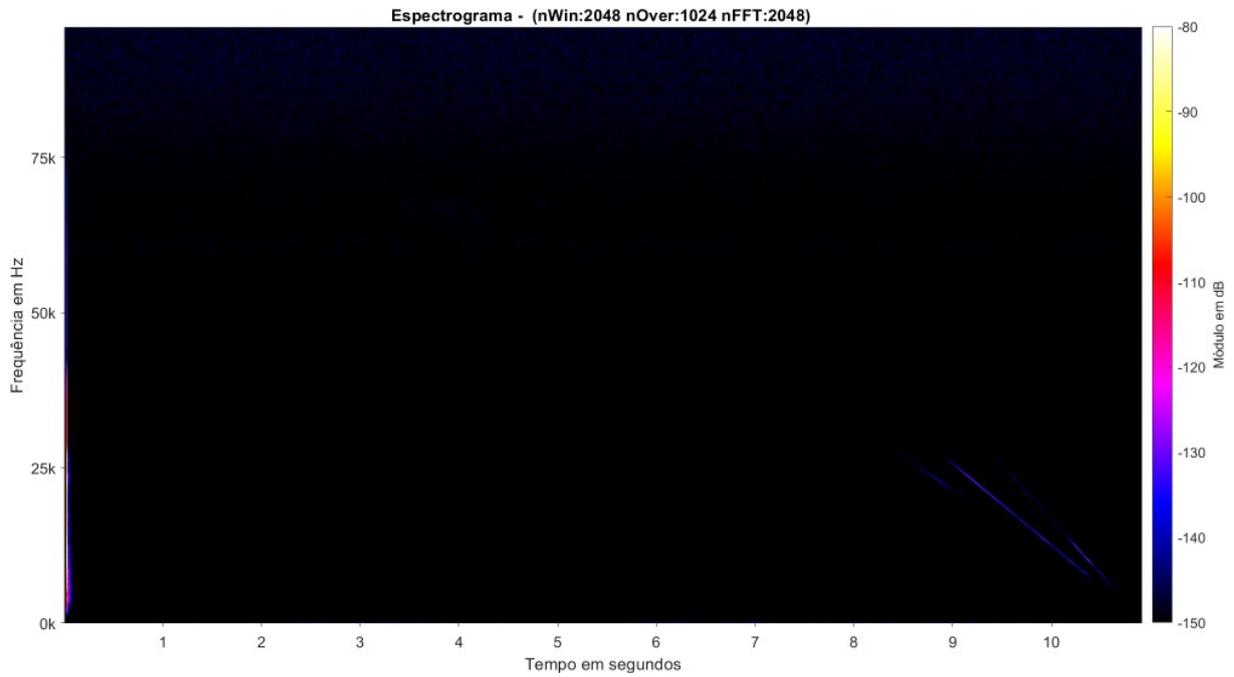


Figura 4.14 – Espectrograma da varredura de frequência gravada por um dos microfones com distorções não-lineares. Fonte: Autor

Com o corte feito, a Figura 4.15 mostra somente o sinal de interesse, sem outros componentes entrando no meio da resposta que buscamos analisar.

O gráfico em resposta em frequência da Figura 4.16 nos mostra que, mesmo em uma pequena escala, os microfones em si respondem quase de forma similar, mas ainda não são coincidentes, tanto em amplitude como em fase. Pela Equação 3.22, iremos utilizar as gravações dos 4 sinais como entradas dos equalizadores e teremos uma mesma saída para todos, que é a média aritmética das 4 respostas medidas.

Os 4 equalizadores para os sinais de Formato A são mostrados na Figura 4.17. Foi escolhido uma resposta unitária para a faixa de frequência de 0 a 2 kHz. Entre 2 kHz e 40 kHz, a faixa de frequência de melhor resolução no equipamento, foi utilizado a resposta criada a partir da média dos sinais dos microfones, como anteriormente descrito. Para frequências maiores que 40 kHz é dado um ganho unitário.

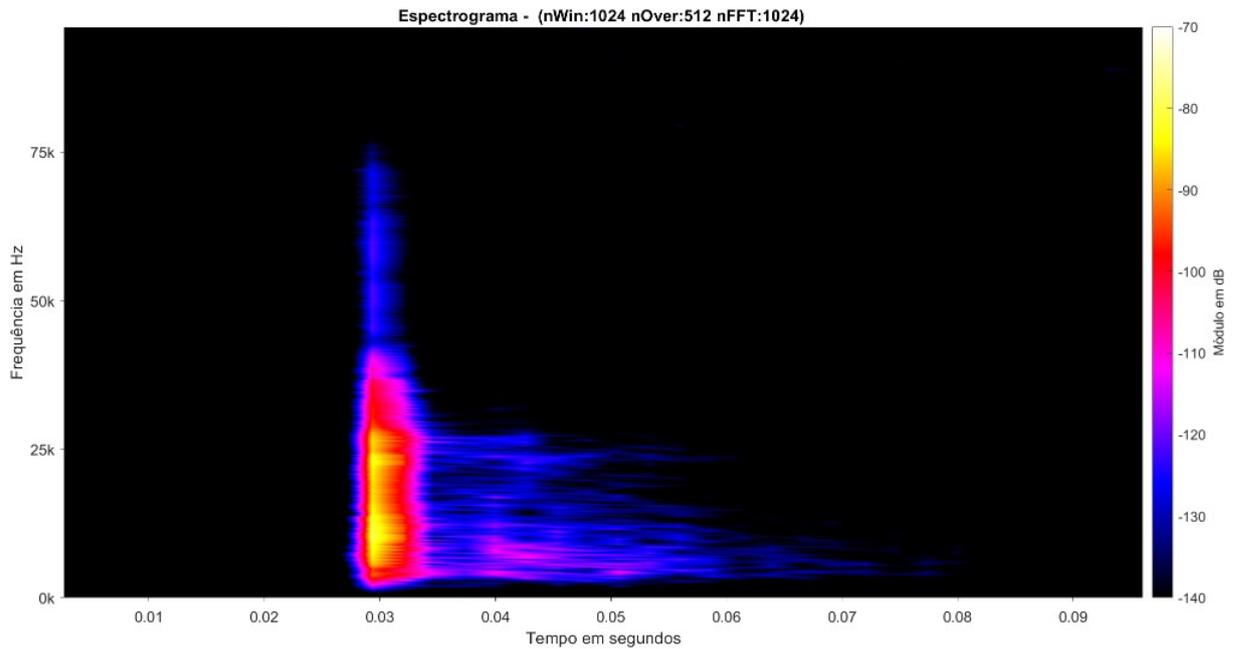


Figura 4.15 – Espectrograma da varredura de frequência gravada por um dos microfones com corte no tempo feito. Fonte: Autor

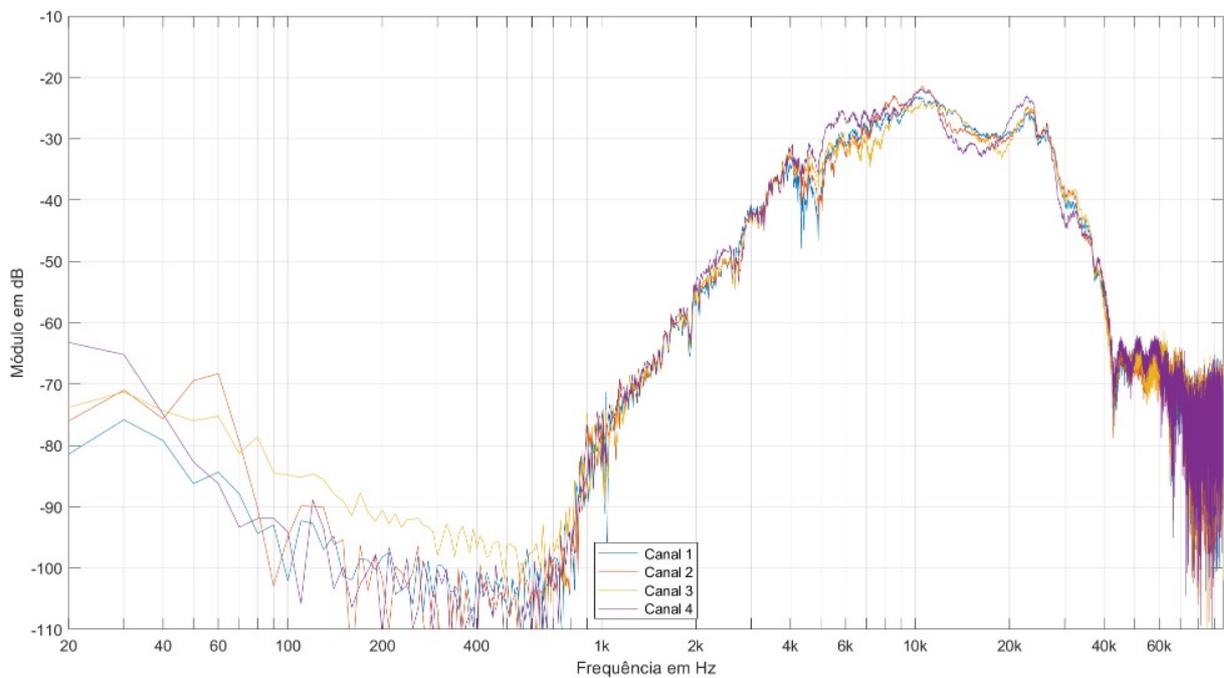


Figura 4.16 – Respostas em frequência de cada microfone depois do corte no tempo feito (curva azul é o primeiro microfone da placa, de cima para baixo, curva laranja para o segundo microfone, amarelo para o terceiro e roxo para o último). Fonte: Autor

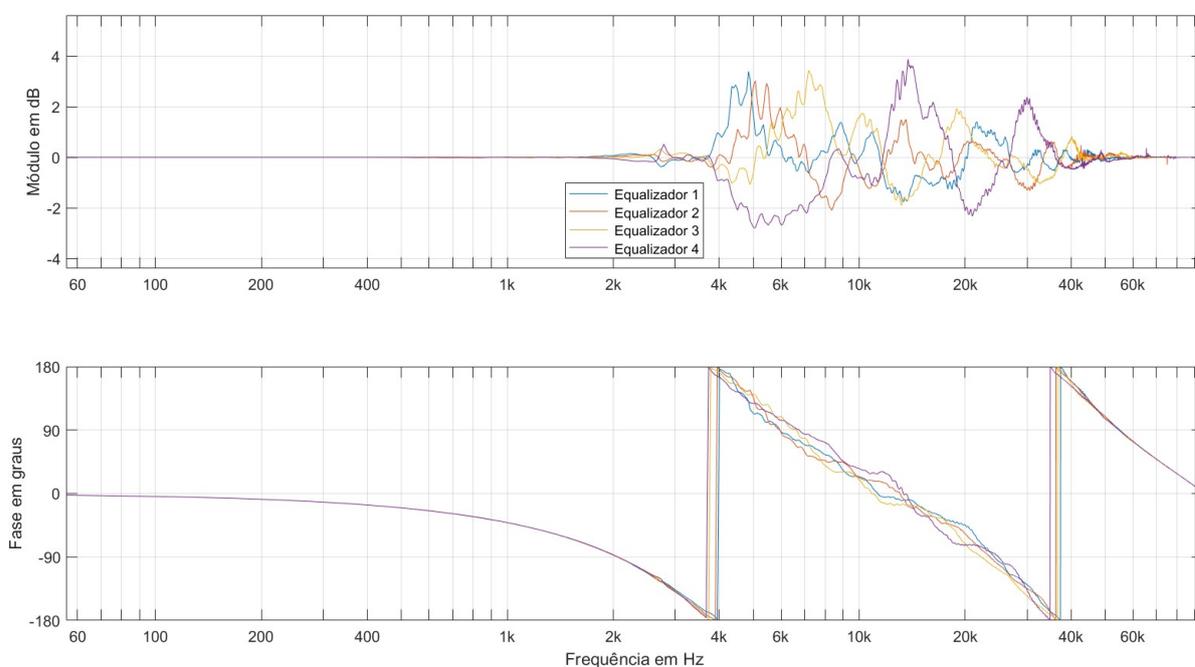


Figura 4.17 – Resposta em frequência dos 4 equalizadores para os canais de Formato A.  
 Fonte: Autor

## 5 Experimentos

Essa seção é destinada aos experimentos e testes feitos para avaliar o arranjo e as capacidades que ele detém. A seguir será demonstrado dois experimentos: um para avaliar a diretividade de cada cápsula e o outra para avaliar a filtragem espacial.

### 5.1 Diretividade do microfone IMP23ABSUTR

Para medir a diretividade do microfone, foi administrado na sala anecoica do Instituto Eldorado um experimento e seus passos são simples: primeiro é colocado um microfone a uma distancia fixa de uma fonte sonora que emitirá uma varredura de frequência que vai até 96 kHz — queremos observar os níveis de pressão sonora no microfone para diversas bandas de frequência. Logo, após gravado a função de transferência entre fonte e microfone nessa posição, o microfone é girado em passos de  $15^\circ$  e feito outra varredura pela fonte e mais uma gravação, até que tenha-se girado o microfone  $180^\circ$ , totalizando então 12 gravações, espelha-se essas medidas nos  $180^\circ$  faltantes para caracterizar a diretividade do microfone em um círculo completo de  $360^\circ$

Assim temos a resposta do microfone para cada frequência em cada ângulo no plano  $xy$ , sendo que os gráficos foram normalizados em função do nível de pressão em  $0^\circ$  (frente do microfone). Analisamos as funções de transferência em bandas de oitava para termos uma visão mais aproximada de como age cada vizinhança de frequência, e não uma frequência só.

A seguir são mostrados, nos gráficos das Figuras 5.1 até 5.6, as respostas de diretividade do microfone IMP23ABSUTR para várias bandas de frequência.

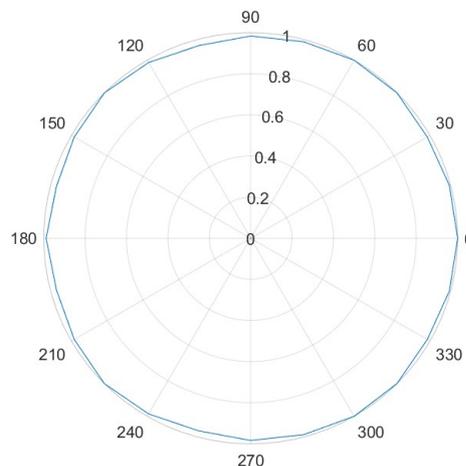


Figura 5.1 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 1 kHz.

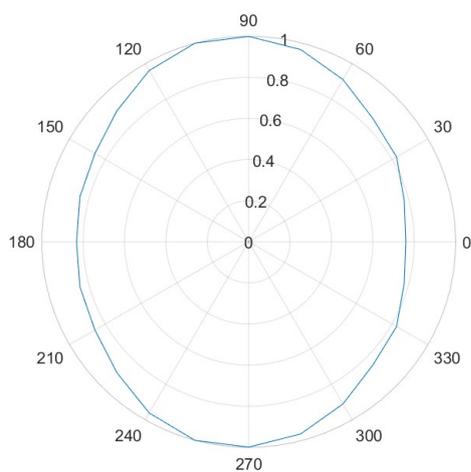


Figura 5.2 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 2 kHz.

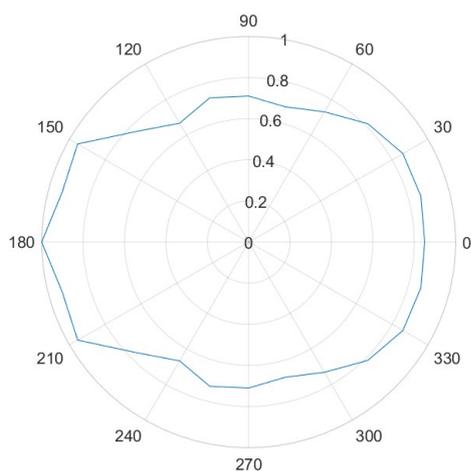


Figura 5.3 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 6,35 kHz.

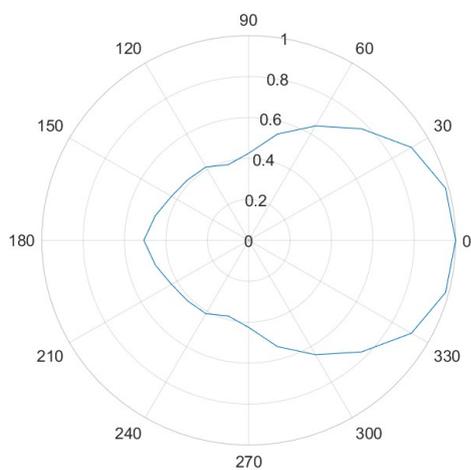


Figura 5.4 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 10 kHz.

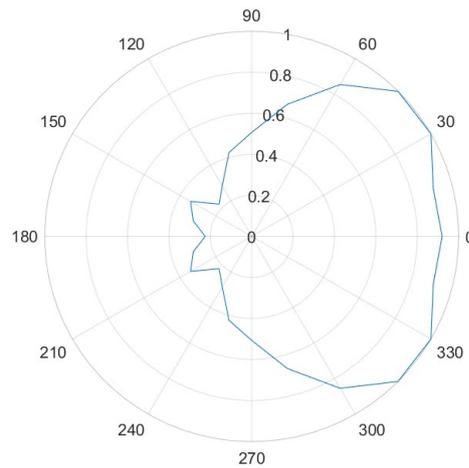


Figura 5.5 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 20 kHz.

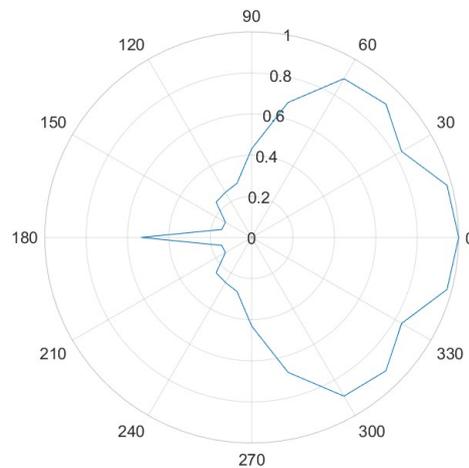


Figura 5.6 – Diretividade do microfone MEMS IMP23ABSUTR para sinais de 40 kHz.

Com esses gráficos, temos uma visão resumida de como a diretividade do microfone trabalha em baixas, médias e altas frequências.

Por exemplo, na Figura 5.1, na vizinhança de 1 kHz, o microfone se comporta de uma maneira bastante omnidirecional, graças ao grande comprimento de onda presente nesses sinais de baixa frequência. As medidas foram feitas com uma fonte Tweeter, que trabalha bem em altas frequências, mas sinais em baixa frequência têm energia bastante reduzida nesse tipo de amplificador de som, logo sinais de menor frequência não entregam informações confiáveis.

Já na Figura 5.2 e Figura 5.4, referente a 2 kHz e 10 kHz, pelo gráfico se observa que é a partir das médias frequências que o microfone começa a ter um comportamento mais direcional, evidenciando a perda de pressão captada na medida que o ângulo entre fonte e microfone aumenta.

As faixas de alta frequência responderam surpreendentemente bem, como é visto nas Figura 5.5 e Figura 5.6, 20 kHz e 40 kHz, respectivamente. À medida que a

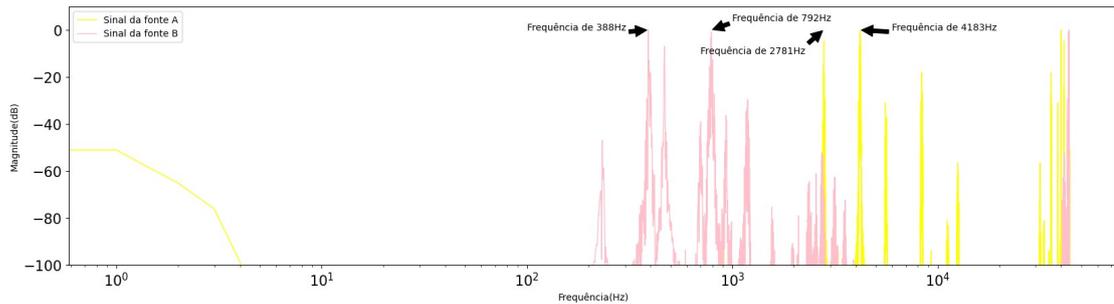


Figura 5.7 – Resposta em Frequência da fonte de teste A e B.

frequência aumenta, o padrão diretivo do microfone continua direcional, mas seria necessário experimentos com passos menores para caracterizar melhor a resposta diretiva do sinal em tais frequências — sendo o teste feito em passos de  $15^\circ$ , os lóbulos laterais que formam em altas frequências podem ter sido não evidenciados pela aproximação dos resultados.

## 5.2 Filtragem espacial

A seguir é mostrado um teste para avaliar a filtragem espacial. Um teste foi feito utilizando de duas fontes sonoras espaçadas em 1 metro do arranjo tetraédrico. As fontes vão emitir tonalidades distintas uma da outra para avaliar a filtragem espacial dos sinais.

A fonte A é colocada de frente para o eixo principal do arranjo ( $\phi = 0^\circ, \theta = 0^\circ$ ), ou seja, no eixo positivo de X (componente de maior frequência de 4183 Hz), e a fonte B colocada a direita do arranjo em ( $\phi = 270^\circ, \theta = 0^\circ$ ), ou seja, no eixo negativo de Y (componente de maior frequência de 388 Hz), e em seguida é gravada a cena em uma câmara semi-anecoica.

As respostas em frequência (normalizadas em si mesmas para referenciar as componentes frequenciais de maior magnitude) dos dois sinais são mostradas na Figura 5.7.

Os sinais dos microfones são convertidos para o Formato B e, pela Equação 3.26, aplica-se um filtro espacial das direções conhecidas das fontes para obter dois sinais diferentes tal que o sinal filtrado para focar os sons de frente pro arranjo é igual ao canal  $0,5X + 0,5W$  (combinação de figura de oito com omnidirecional para obter um padrão polar de cardioide) do Formato B e os sons que vem da direita do arranjo é igual ao  $-0,5Y + 0,5W$ .

A Figura 5.8 mostra a resposta em frequência do sinal filtrado “mirando” para a posição conhecida da fonte A (sinal verde) e o sinal filtrado na direção da fonte B (sinal

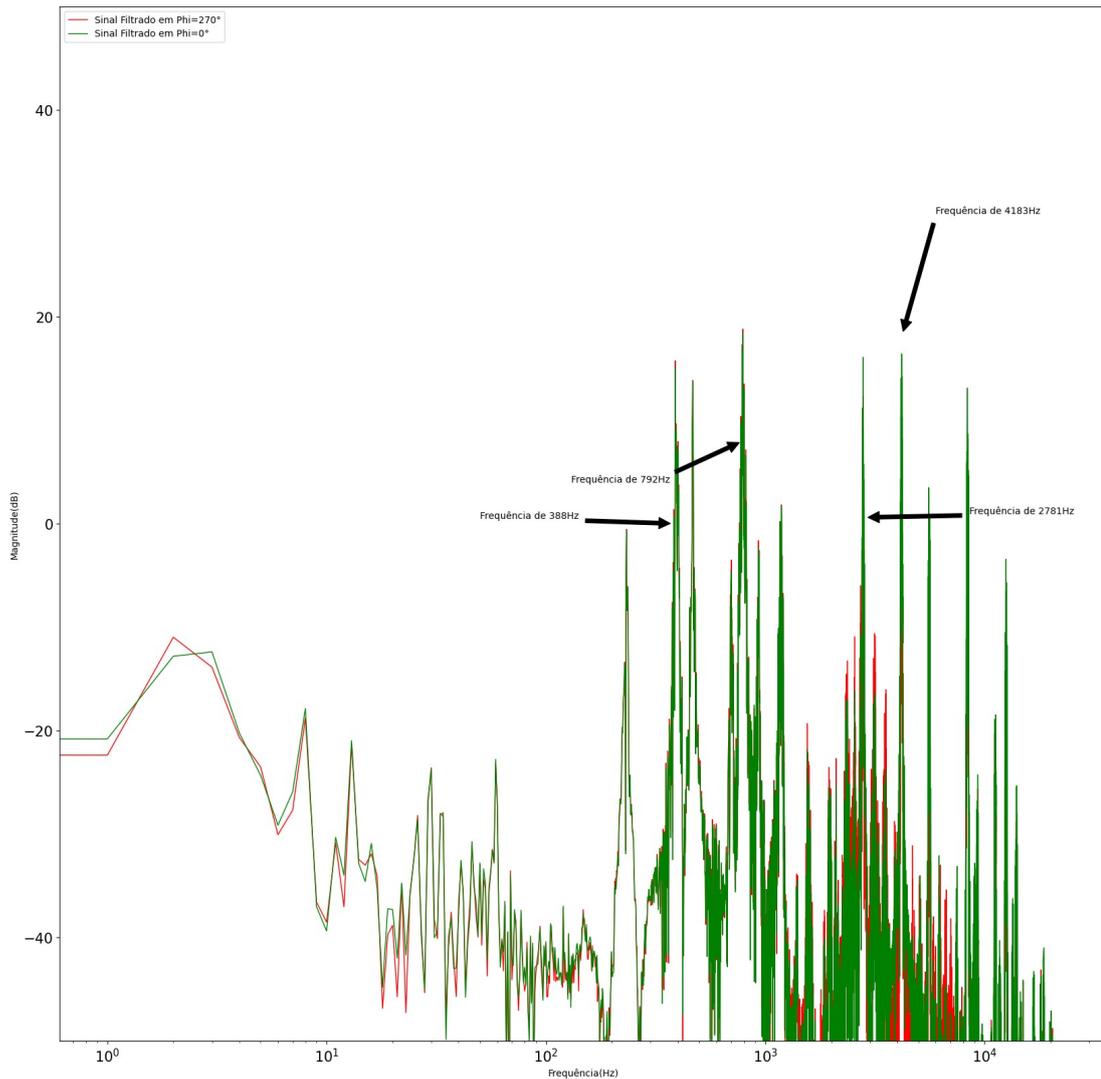


Figura 5.8 – Resposta em Frequência do sinal filtrado em  $(\phi = 0^\circ, \theta = 0^\circ)$  e do sinal filtrado em  $(\phi = 270^\circ, \theta = 0^\circ)$  utilizando padrão polar  $\alpha = 0,5$ .

vermelho), sobrepostos, quando se utiliza um padrão polar  $\alpha = 0,5$  (diretividade de um cardioide, que valoriza sinais vindo de frente e cancela sinais que vem da direção oposta) na Equação 3.26.

Os valores em decibéis das componentes de frequência são mostrados na Tabela 5.1 :

Tabela 5.1 – Valor da magnitude das componentes de frequência dos sinais filtrados espacialmente com padrão polar  $\alpha = 0,5$ .

Frequência	Sinal Filtrado 0°	Sinal Filtrado 270°	Omnidirecional W
388 Hz	3,0 dB	3,8 dB	6,6 dB
792 Hz	18,3 dB	18,8 dB	21,7 dB
2781 Hz	11,2 dB	7,3 dB	10,2 dB
4183 Hz	16,4 dB	8,2 dB	12,9 dB

Se espera que as componentes de baixa frequência (388 e 792 Hz) tenham valores maiores no sinal filtrado em  $270^\circ$ . Em 388 Hz, a diferença de 0,83 dB entre os dois sinais filtrados mostra que o arranjo dá somente uma leve preferência a direção. Isso se deve ao caráter dos microfones que compõem o arranjo e de como essa baixa frequência ressoa no arranjo — basicamente essa frequência tem magnitude relativamente igual em toda a esfera.

Mas quando focamos a atenção para as frequências acima de 2 kHz, começa a existir uma melhor distinção entre os sinais, sendo que para a frequência principal da Fonte A, 4183 Hz, há uma diferença de 8,14 dB entre os sinais filtrados.

O padrão polar é um parâmetro manipulável na Equação 3.26, logo pode-se controlar ele de forma a aproveitar melhor as características do espaço a ser caracterizado. Já que é uma situação de câmara semi-anecoica, não teremos som vindo de outras direções a não ser das fontes que foram colocadas na cena, então se fazemos  $\alpha = 0$ , estamos restringindo a resposta diretiva na parte positiva do eixo do equipamento, dando maior preferência para os sons diretos nesse ângulo de incidência e diminuindo os valores de pressão dos ângulos laterais, mas aumentando bastante o ganho dos sons que vierem no sentido oposto do filtrado, basicamente um padrão polar de figura de oito (não há a combinação com o sinal omnidirecional). A Figura 5.9 mostra a resposta para esse tipo de configuração e a Tabela 5.2 mostra os valores de magnitude das frequências de interesse.

Tabela 5.2 – Valor da magnitude das componentes de frequência dos sinais filtrados espacialmente com padrão polar  $\alpha = 0$ .

Frequência	Sinal Filtrado $0^\circ$	Sinal Filtrado $270^\circ$	Omnidirecional W
388 Hz	-30,4 dB	-11,5 dB	6,6 dB
792 Hz	-16,8 dB	9,1 dB	21,7 dB
2781 Hz	11,9 dB	-4,4 dB	10,2 dB
4183 Hz	18,1 dB	1,4 dB	12,9 dB

Nessa configuração, pode-se ver que os sinais filtrados se “separaram” de forma mais clara, de tal forma que em 388 Hz, a diferença de magnitude entre os dois sinais filtrados é de 18,87 dB, e em 4183 Hz, a diferença de magnitude entre os dois sinais filtrados é de 16,73 dB.

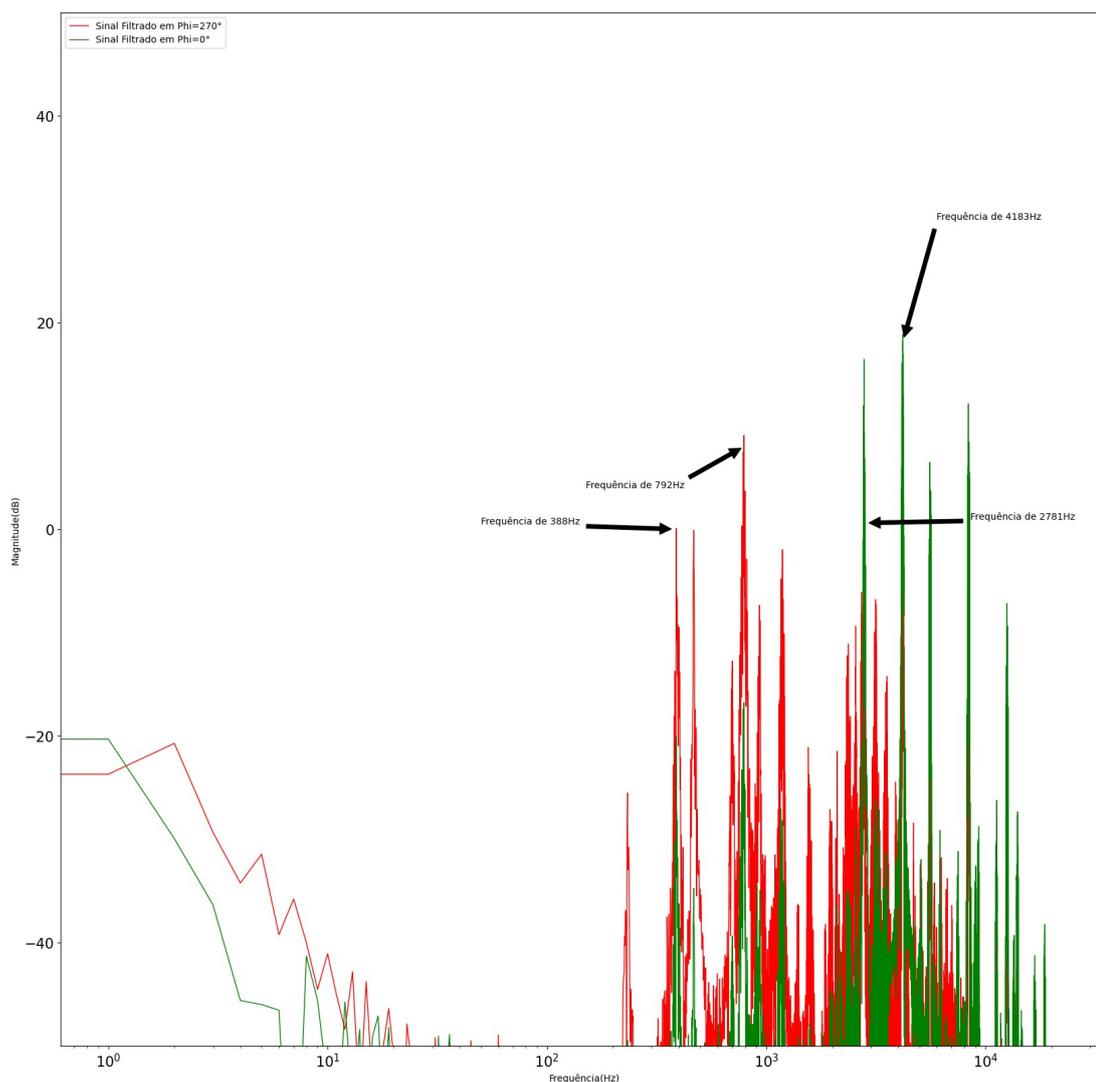


Figura 5.9 – Resposta em Frequência do sinal filtrado em  $(\phi = 0^\circ, \theta = 0^\circ)$  e do sinal filtrado em  $(\phi = 270^\circ, \theta = 0^\circ)$  utilizando padrão polar  $\alpha = 0$ .

## 6 Conclusão

A utilização de tecnologia *Ambisonic* é bastante tentadora — a possibilidade de poder entregar características espaciais para gravações sonoras, recurso antes restrito a características de tempo e frequência, abre as portas para aplicações bastante interessantes. Mas experimentos indicam que caracterizar o ambiente sonoro utilizando um arranjo tetraédrico pode ser um grande desafio e detalhes específicos devem sempre estar presentes na fabricação do protótipo.

Primeira conclusão retirada seria, obviamente, a sua complicação em relação à frequência. Comprimentos de onda grandes, referentes às baixas frequências (frequências menores que 2 kHz), são, por natureza, ondas com comprimento de onda extenso, o que leva os equipamentos não tão ter respostas tão diretivas. Médias frequências (frequências entre 2 kHz e 5 kHz) são as melhores — com o aumento da frequência, os sensores começam a ter uma resposta mais diretiva. Mas se uma frequência atinge valores maiores que um limiar entre 5 kHz e 6 kHz, para um tetraedro de 36 mm de diâmetro, a diretividade do sinal de Formato B se deteriora, devido ao jeito que ondas de alta frequências se espalham pela esfera, gerando lóbulos em torno do arranjo. Esta distorção resulta da necessidade de harmônicos esféricos de maior ordem para descrever ondas de alta frequência. Uma maneira de melhorar isso é com uma melhor caracterização do espaço sonoro a partir de um arranjo de maior ordem. Apesar de um arranjo tetraédrico ser um equipamento com geometria mais simples, processamento de sinais simplificado e menos canais para tomar conta, a sua caracterização do ambiente sonoro peca em relação à veracidade do que está realmente acontecendo em volta dele. Se apoiando em apenas nos primeiros harmônicos esféricos para sua descrição espacial, o ambiente sonoro é “fotografado” com uma péssima “câmera”. Se aumentamos o número de cápsulas no arranjo, conseguimos utilizar harmônicos de maior ordem e grau para conseguir descrever com maior confiabilidade o ambiente sonoro a que propomos monitorar.

A filtragem espacial realizada na seção 5.2 mostra que algumas manipulações triviais podem ser realizadas com os canais de Formato B. A Equação 3.26 foi testada e comprovou-se que a partir do controle e ajuste do padrão polar, pode-se monitorar os sinais e separa-los simplesmente ponderando os canais de Formato B. Foi possível com a gravação espacial, diferenciar em aproximadamente 18 dB as frequências tonais (em uma faixa de frequência específica) que as fontes estavam emitindo no espaço.

## 6.1 Perspectivas futuras

Infelizmente, devido à inesperada ruptura do protótipo, sendo inviável construir um novo em tempo hábil, não foi possível realizar a medição da resposta diretiva de cada canal de Formato B para observar o comportamento da diretividade do arranjo com a frequência do som captado. Indica-se, portanto, esta caracterização como atividade futura, já que a partir desta informação sobre a resposta diretiva do arranjo tetraédrico, pode-se traçar uma estratégia eficiente em relação aos filtros corretivos a ser implementados para equalizar a resposta de cada canal Formato B dentro da faixa de frequência de interesse.

## Referências

- ABHAYAPALA, D. B. W. T. D. Theory and design of high order sound field microphones using spherical microphone array. *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002. Disponível em: <<https://doi.org/10.1109/ICASSP.2002.5745011>>. Citado na página 30.
- AHONEN, J.; PULKKI, V.; LOKKI, T. Teleconference application and B-format microphone array for directional audio coding. *Journal of the Audio Engineering Society*, 2007. Citado na página 41.
- BERZBORN, M.; BOMHARDT, R.; KLEIN, J.; RICHTER, J.-G.; VORLANDER, M. The ita-toolbox: An open source matlab toolbox for acoustic measurements and signal processing. *Institute of Technical Acoustics, Kopernikusstr*, 2017. Disponível em: <[www.ita-toolbox.org](http://www.ita-toolbox.org)>. Citado na página 51.
- CHRISTOF, F.; MIHAILO, K. Design and limitations of non-coincidence correction filters for soundfield microphones. *126th Audio Engineering Society Convention*, 2009. Citado 2 vezes nas páginas 33 e 36.
- DPA Microphones. *HOW TO READ MICROPHONE SPECIFICATIONS*. 2020. [Online;accessed 22-Março-2023]. Disponível em: <<https://www.dpamicrophones.com/mic-university/how-to-read-microphone-specifications>>. Citado 2 vezes nas páginas 9 e 25.
- EARGLE, J. *The Microphone Handbook*. [S.l.: s.n.], 2005. Citado 5 vezes nas páginas 9, 21, 22, 23 e 24.
- ELKO, G. *Differential Microphone Arrays in Audio Signal Processing for Next-Generation Multimedia Communication Systems*. [S.l.]: Springer International Publishing, 2004. 12–65 p. Citado na página 38.
- FARINA, A. A-format to b-format conversion. *Retrieved from <http://pcfarina.eng.unipr.it/Public/B-format/A2B-conversion/A2B.htm>*, 2006. Citado 6 vezes nas páginas 9, 10, 36, 37, 38 e 49.
- FERNANDEZ-GRANDE, E. Sound field reconstruction using a spherical microphone array). *The Journal of the Acoustical Society of America*, v. 139, n. 3, p. 1168–1178, 03 2016. ISSN 0001-4966. Disponível em: <<https://doi.org/10.1121/1.4943545>>. Citado na página 30.
- GERZON, M. A. The design of precisely coincident microphone arrays for stereo and surround sound. *50th Audio Engineering Society Convention*, 1975. Citado 4 vezes nas páginas 9, 34, 35 e 36.
- HELLER, A. J.; BENJAMIN, E. M. Calibration of soundfield microphones using the diffuse-field response. *Audio Engineering Society 133rd Convention, San Francisco, USA.*, 2012. Citado 4 vezes nas páginas 10, 38, 39 e 40.

- HONG, K.; HUANG, H. Power transformer fault diagnosis based on vibration correlation analysis. *Proceedings of the ASME 2014 International Mechanical Engineering Congress and Exposition IMECE2014*, 2014. Citado na página 18.
- IEC-60268-4, I. E. C. *Sound system equipment*. [S.l.: s.n.], 2010. Citado na página 38.
- JOHANN-MARKUS, B. THE B-FORMAT MICROPHONE REVISED. *AMBISONICS SYMPOSIUM*, 2009. Disponível em: <<https://iaem.at/ambisonics/symposium2009/proceedings/ambisym09-batke-bambimicrovised.pdf>>. Citado 4 vezes nas páginas 27, 28, 29 e 31.
- KINSLER, L. E.; FREY, A. R.; COPPENS, A. B.; SANDERS, J. V. *Fundamentals of Acoustics*. [S.l.]: John Wiley and Sons, Inc., 1982. 12–65 p. (Springer Topics in Signal Processing). ISBN 9783319995618. Citado na página 39.
- KUECH, F.; KALLINGER, M.; SCHULTZ-AMLING, R.; GALDO, G. del; AHONEN, J.; PULKKI, V. Directional audio coding using planar microphone arrays. p. 37–40, 2008. Citado na página 41.
- MAŁECKI, P.; OZGA, A.; PIECHOWICZ, J. Soundscape analysis based on ambisonic recordings executed in a primeval forest. *22nd International Congress on Acoustics, ICA, Buenos Aires – 5 to 9 September, 2016*, 2016. Citado na página 17.
- PHAM, H. *Documentation PyAudio 0.2.12*. 2022. <<https://pypi.org/project/PyAudio/>>. [Online; acessada em 30/05/2022.]. Citado na página 47.
- POLETTI, M. A. Three-Dimensional Surround Sound Systems Based on Spherical Harmonics. *J. Audio Eng. Soc.*, v. 53, p. 1004–1025, 2005. Disponível em: <<http://www.aes.org/e-lib/browse.cfm?elib=13396>>. Citado 3 vezes nas páginas 9, 29 e 35.
- QIANG, H.; JINGKAI, N.; SONGYANG, Z.; WEIMIN, X.; SHENGCHANG, J.; XIN, C. Study of transformer core vibration and noise generation mechanism induced by magnetostriction of grain-oriented silicon steel sheet. *Hindawi Shock and Vibration Volume 2021*, 2021. Citado na página 18.
- RAFAELY, B. *Fundamentals of Spherical Array Processing*. Springer International Publishing, 2018. (Springer Topics in Signal Processing). ISBN 9783319995618. Disponível em: <<https://books.google.com.br/books?id=WtNwDwAAQBAJ>>. Citado 6 vezes nas páginas 9, 23, 26, 27, 28 e 29.
- SALAMON, J.; BELLO, J.; SILVA, C.; NOV, O.; DUBOIS, R.; ARORA, A.; MYDLARZ, C.; DORAISWAMY, H. Sonyc: A system for the monitoring, analysis and mitigation of urban noise pollution. *Communications of the ACM*, 05 2018. Citado na página 18.
- SALVADOR-BARRERA, F.; RASMUSSEN, K.; JACOBSEN, F. On experimental determination of the random-incidence response of microphones. *The Journal of the Acoustical Society of America*, 2007. Citado na página 39.
- SANNUM, V. *Filtering with spatial parameters in B-format audio streams*. 17 p. Dissertação (Mestrado) — Department of Electrical and Information Technology Lund University, 2018. Citado na página 40.

SCHULZE-FORSTER, K. *The B-Format – Recording, Auralization, and Absorption Measurements*. 31-33 p. Dissertação (Mestrado) — Division of Applied Acoustics, 6 2018. Citado 4 vezes nas páginas 9, 30, 33 e 34.

STMicroelectronics NV. *High-performance MEMS microphone with extended frequency response up to 80 kHz for ultrasound applications*. [S.l.], 2020. Citado 3 vezes nas páginas 10, 42 e 43.

Wikipedia contributors. *Som - Wikipedia*. 2022. [Online;accessed 22-Março-2023]. Disponível em: <<https://pt.wikipedia.org/wiki/Som>>. Citado 2 vezes nas páginas 9 e 22.

YING, L.; WANG, D.; WANG, J.; WANG, G.; WU, X.; LIU, J. Power transformer spatial acoustic radiation characteristics analysis under multiple operating conditions. *Energies* ([www.mdpi.com/journal/energies](http://www.mdpi.com/journal/energies)), 2018. Citado 2 vezes nas páginas 9 e 18.

## Anexos

# ANEXO A – Código do Cliente

## (Gravação multicanal e transmissão WiFi)

```

#Importando Bibliotecas necessarias
import pyaudio
import socket,time

#Configuracao para gravacao e transmissao de dados
FORMAT2 = 'utf-8'
CHUNK = 256
FORMAT = pyaudio.paInt16
CHANNELS = 4
RATE = 44000
audio = [ ]
frame1 = [ ]
frame = [ ]
CODE = 'utf-8'

#Mensagens de rotina para comunicacao cliente-servidor
CHUNK_COMPLETE='CHUNK_COMPLETE'
CHUNK_INCOMPLETE='CHUNK_NOMLETE'
READY='READYFORYOULOV'
FILE_SENT='FILE_SENT'

#Configuracao da comunicacao WiFi
socketCliente = socket.socket(socket.AF_INET,socket.SOCK_STREAM)
host_ip= '192.168.25.25'
port = 5050
backlog = 1
socketEndereco = (host_ip,port)
socketCliente.connect(socketEndereco)
print('Conectado com:',socketEndereco)

#Rotina de Gravacao e Transmissao
try:
    while True:
        print("* recording")
        pyaudio_instance = pyaudio.PyAudio()
        stream = pyaudio_instance.open(format=FORMAT,\
            channels = CHANNELS,rate = RATE,input = True,\
            input_device_index = 2,frames_per_buffer = CHUNK)

```

```
for i in range(0,1000):
    data = stream.read(CHUNK)
    frame.append(data)
print("* done recording")
stream.stop_stream()
stream.close()
pyaudio_instance.terminate()
print("* sending")
for i in range(0,1000):
    datasend = frame[i]
    socketCliente.sendall(datasend)
    NEXT = True
    cond = socketCliente.recv(14).decode(FORMAT2)
    while NEXT:
        if cond==CHUNK_INCOMPLETE:
            socketCliente.sendall(datasend)
            cond = socketCliente.recv(14).decode(FORMAT2)
        if cond==CHUNK_COMPLETE:
            NEXT=False
        if cond==READY:
            NEXT=False
        if cond!=CHUNK_INCOMPLETE and cond!=CHUNK_COMPLETE:
            cond = socketCliente.recv(14).decode(FORMAT2)
print("* done sending")
frame = []
print("* Iniciando proximo...")
NEXT = True

#Rotina para aprovacao do servidor para um novo ciclo de gravacao
while NEXT:
    perm = socketCliente.recv(14).decode(FORMAT2)
    if perm==READY:
        data = 0
        time.sleep(5)
        NEXT=False
    else:
        perm = socketCliente.recv(14).decode(FORMAT2)

except KeyboardInterrupt:
print("Fechando")
```

## ANEXO B – Código do Servidor (Processando o Sinal e Salvando Arquivos)

```

#Importando Bibliotecas necessarias
import socket
import numpy as np
from scipy.io import loadmat
from scipy import signal
from datetime import datetime
import time
import soundfile as sf

#Configuracao para transmissao de dados
Servidor = socket.socket(socket.AF_INET, socket.SOCK_STREAM)
Servidor.bind((socket.gethostbyname(socket.gethostname()), 5050))
Servidor.listen()
print('Aguardando conexao')
conn, ender = Servidor.accept()
print('Conectado')
frame = [ ]
RATE = 44000
FORMAT = 'utf-8'

#Mensagens de rotina para comunicacao cliente-servidor
CHUNK_INCOMPLETE= 'CHUNK_NOMLETE'
CHUNK_COMPLETE= 'CHUNK_COMPLETE'
READY= 'READYFORYOULOV'
factor = ( 22050/(192000/2))

#Carregando Filtros calculados no MATLAB
Equalizer1 = loadmat('Equalizer1.mat', mat_dtype=True)
Equalizer2 = loadmat('Equalizer2.mat', mat_dtype=True)
Equalizer3 = loadmat('Equalizer3.mat', mat_dtype=True)
Equalizer4 = loadmat('Equalizer4.mat', mat_dtype=True)
Eq1 = Equalizer1['Equalizer1'].flatten().astype(np.int16)
Eq2 = Equalizer2['Equalizer2'].flatten().astype(np.int16)
Eq3 = Equalizer3['Equalizer3'].flatten().astype(np.int16)
Eq4 = Equalizer4['Equalizer4'].flatten().astype(np.int16)
Eq1 = signal.resample(Eq1, round(len(Eq1) / factor))/500
Eq2 = signal.resample(Eq1, round(len(Eq2) / factor))/500
Eq3 = signal.resample(Eq1, round(len(Eq3) / factor))/500

```

```
Eq4 = signal.resample(Eq1, round(len(Eq4) / factor))/500

try:
    while True:

        #Recebendo arquivos de audio do cliente
        print('Recebendo arquivo')
        for i in range(0,1000):
            data = conn.recv(2048)
            CONTROL1 = True
            while CONTROL1:
                if len(data)<2048:
                    conn.sendall(CHUNK_INCOMPLETE.encode(FORMAT))
                else:
                    conn.sendall(CHUNK_COMPLETE.encode(FORMAT))
                    frame.append(data)
                    CONTROL1=False
                if not data:
                    conn.sendall(READY.encode(FORMAT))
                    time.sleep(10)
            print('Arquivo recebido')
            print('Processando arquivo')\\

        #Adquirindo data em tempo real para o arquivo a ser salvo
        current_datetime = str(datetime.now())
        date = str([current_datetime[11:13],
                    current_datetime[14:16],current_datetime[17:19]])
        newdata = b''.join(frame)

        #Conversao de bytes para um vetor e separacao do arquivo em 4 canais
            distintos

            data_array = np.frombuffer(newdata, dtype='int16')
            dataLF = data_array[1::4]
            dataRF = data_array[0::4]
            dataLB = data_array[2::4]
            dataRB = data_array[3::4]

        #Equalizacao dos sinais de Formato A
            fdataLF = signal.oaconvolve(dataLF, Eq1, 'same')
            fdataRF = signal.oaconvolve(dataRF, Eq2, 'same')
            fdataRB = signal.oaconvolve(dataRB, Eq3, 'same')
            fdataLB = signal.oaconvolve(dataLB, Eq4, 'same')

        #Conversao para Formato B
            W= dataLF + dataRF + dataLB + dataRB
```

```
X= dataLF + dataRF - dataLB - dataRB
Y= dataLF - dataRF + dataLB - dataRB
Z= dataLF - dataRF - dataLB + dataRB
FinalData = np.hstack((W,X,Y,Z))
print('Arquivo processado')

#Salvando Arquivo no HD interno
print('Salvando Arquivo')
sf.write(date+"_flac", FinalData, 44000)
print('Arquivo Salvo')
frame = []
conn.sendall(READY.encode(FORMAT))
time.sleep(10)

except KeyboardInterrupt:
    print('interrompido!')
```