

ESCOLA DE ENGENHARIA DE SÃO CARLOS  
UNIVERSIDADE DE SÃO PAULO

Vitor Christiano Nagashima

**ANÁLISE DE ÁUDIO CAPTADO POR SMARTPHONES PARA  
MEDIR A FORÇA DE UM GOLPE DE TAE-KWON-DO**

PROJETO DE FORMATURA

Curso de Engenharia Elétrica com  
ênfase em Sistemas de Energia e Automação

ORIENTADOR: Eduardo Simões

São Carlos  
2020

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,  
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS  
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da  
EESC/USP com os dados inseridos pelo(a) autor(a).

N147a Nagashima, Vitor Christiano  
Análise de áudio captado por smartphones para  
medir a força de um golpe de tae-kwon-do / Vitor  
Christiano Nagashima; orientador Eduardo do Valle  
Simões. São Carlos, 2020.

Monografia (Graduação em Engenharia Elétrica com  
ênfase em Sistemas de Energia e Automação) -- Escola de  
Engenharia de São Carlos da Universidade de São Paulo,  
2020.

1. áudio digital. 2. captação de áudio. 3.  
esporte de combate. 4. aparador. 5. Ableton Live. I.  
Título.

# FOLHA DE APROVAÇÃO

Nome: Vitor Christiano Nagashima

Título: “Análise de áudio captado por smartphones para medir a força de um golpe de TAE-KWON-DO”

Trabalho de Conclusão de Curso defendido e aprovado  
em 15 / 12 / 2020,

com NOTA 10 (dez , ), pela Comissão Julgadora:

Prof. Dr. Eduardo do Valle Simões - Orientador SSC/ICMC/USP

Prof. Titular Ivan Nunes da Silva - SEL/EESC/USP

Mestre Edvaldo Vasconcelos Soares Maciel - Doutorando  
IQSC/USP

Coordenador da CoC-Engenharia Elétrica - EESC/USP:  
Prof. Associado Rogério Andrade Flauzino

## **Agradecimento**

Dedico este trabalho ao meu anjo da guarda, Felipe Bastardo, à Manu, Balde, Sheyla e Clárinha, por me ensinarem lições que eu jamais poderia aprender em uma sala de aula.

Agradeço à minha família por propiciar todas as condições para que eu pudesse me formar, à República Irmãos Metralha por todas as vitórias e derrotas vividas juntos, ao meu orientador professor Eduardo Simões pela humanidade e tranquilidade nas conversas e ao professor André Salata por me ajudar a enxergar e sedimentar os ensinamentos da engenharia elétrica na música.

Agradeço ao meu amigo e irmão de espírito, Miguel “Micareta” Cabral, por unir nossas maiores paixões neste projeto acadêmico.

Mas acima de tudo agradeço à música, que sendo aquilo que mais amo, me aproximou de Deus e também de mim mesmo.

# Sumário

<b>Sumário</b>	<b>3</b>
<b>Índice de figuras</b>	<b>5</b>
<b>Índice de equações</b>	<b>7</b>
<b>Índice de tabelas</b>	<b>7</b>
<b>Índice de gráficos</b>	<b>7</b>
<b>Gráfico 4.1 - Variação do tempo de reação do atleta em 12 medições</b>	<b>40</b>
<b>Resumo</b>	<b>7</b>
<b>Abstract</b>	<b>8</b>
<b>1. Introdução e Justificativa.</b>	<b>9</b>
1.1 Objetivo	11
<b>2. Revisão Bibliográfica</b>	<b>10</b>
2.1 Sensores Sonoros:	12
2.1.1 Microfone Dinâmico:	12
2.1.2 Microfone Condensador:	13
Microfone Condensador Cardióide Audio-Technica AT2020:	13
Microfone Condensador com tecnologia MEMS:	15
2.2 Interface de áudio USB Focusrite iTrack Solo:	17
2.3 Software de análise e manipulação de áudio utilizado:	17
<b>3 Fundamentação teórica:</b>	<b>21</b>
3.1 Intensidade Sonora	21
3.1.1 Variação da intensidade sonora com a distância	21
3.1.2 A escala decibel	21
3.1 Sinal de áudio no domínio do tempo - envelope ADSR:	22
3.1.1 Envelope ADSR esperado de um golpe em um aparador:	24
3.2 Sinal de áudio no domínio da frequência - timbre:	25
3.2.1 Timbre esperado de um golpe em um aparador:	26
3.3 Sinal analógico:	26
3.4 Sinal digital:	27

3.5 Processamento digital de sinais:	27
3.5.1 Teorema da amostragem:	28
3.5.2 Profundidade de bit de áudio (audio bit depth):	30
3.5.3 Ponto flutuante:	31
<b>4 Experimentação e resultados:</b>	<b>33</b>
4.1 Montagem do experimento	33
4.2 Análise de resultados	35
4.2.1 Diferença do pico de volume entre golpes fortes e fracos captados por um mesmo dispositivo	35
4.2.2 Comparação do envelope ADSR gravado por cada dispositivo	36
4.2.3 Comparação do espectro de frequência dos áudios captados	38
4.2.4 Velocidade de reação do atleta	39
<b>5. Conclusão</b>	<b>41</b>
5.1 Considerações sobre o Curso de Graduação	41
<b>Referências bibliográficas</b>	<b>42</b>

# Índice de figuras

Figura 2.1 - Diagrama do padrão polar do AT2020	13
Figura 2.2 - Resposta em frequência do AT2020	14
Figura 2.3 - Microfone condensador Audio-Technica AT2020	14
Figura 2.4 - Estrutura básica de um microfone MEMS	15
Figura 2.5 - Matriz de microfone MEMS Infineon E2002 encontrado no iPhone 4	16
Figura 2.6 - Vista frontal da interface de áudio USB Focusrite iTrack Solo	17
Figura 2.7 - Arranjo musical criado na "session view" do Ableton Live 10	18
Figura 2.8 - Glue Compressor, plugin nativo com modelagem analógica desenvolvido pela Ableton em colaboração com a Cytomic	18
Figura 2.9 - Seleção de um trecho de 39 ms em azul claro	19
Figura 2.10 - Spectrum, plugin nativo para análise de espectro de frequência de uma amostra de som	19
Figura 3.1 - Diagrama esquemático de um envelope ADSR	22
Figura 3.2 - Envelope ADSR de amplitude de uma amostra de som	24
Figura 3.3 - Sintetizador digital que originou a forma de onda da Figura 3.2	24
Figura 3.4 - Onda sonora de um bumbo de bateria	25
Figura 3.5 - Espectro de frequências do bumbo de bateria da Figura 3.2	26
Figura 3.6 - Representação de um sinal digital (curva vermelha) no domínio do tempo	27
Figura 3.7 - Sinal amostrado, em azul a cópia do sinal original e em verde as imagens replicadas que precisam ser filtradas	28
Figura 3.8 - Nota lá (A1) de um piano com um trecho selecionado de 190 ms	29
Figura 3.9 - Zoom no trecho de 190 ms com seleção em trecho de 3 ms	29
Figura 3.10 - Zoom no trecho de 3 ms	30
Figura 3.11 - Espectro de frequências de um áudio gravado com 16 bits de profundidade	31
Figura 4.1 - Pedestal com microfone e smartphones fixados	33
Figura 4.2 - Espaço onde foram feitas as medições, com marcações no chão indicadas na imagem pelas linhas em verde	34
Figura 4.3 - Atleta golpeando o aparador sobre a marcação de 2 metros	34
Figura 4.4 - Sinal gravado pelo Apple iPhone 8 (preto) comparado com a referência, amostra com duração de 350 ms	36
Figura 4.5 - Sinal gravado pelo Samsung J8 (preto) comparado com a referência, amostra com duração de 350 ms	36

Figura 4.6 - Sinal gravado pelo XIaomi Mi 8 (preto) comparado com a referência, amostra com duração de 350 ms	37
Figura 4.7 - Detalhamento do envelope ADSR da referência em uma amostra de 53 ms	37
Figura 4.8 - Comparativo entre espectro de frequênciadas amostras captadas	38

## **Índice de equações**

(I) Intensidade sonora em uma superfície	21
(II) Intensidade sonora de fonte sonora isotrópica puntiforme	21
(III) Nível sonoro na escala decibel	22
(IV) Demonstração da dobra de intensidade sonora a cada 3 dB	22

## **Índice de tabelas**

Tabela 4.1 - Variação do pico de volume entre golpes fracos e fortes por dispositivo	35
Tabela 4.2 - Média do tempo de reação do atleta em cada dispositivo	39

## **Índice de gráficos**

Gráfico 4.1 - Variação do tempo de reação do atleta em 12 medições	40
--	----

## **Resumo**

O objetivo deste trabalho é avaliar a possibilidade de utilizar smartphones na prática e evolução de um atleta de Tae-Kwon-Do mas aplicável a qualquer esporte de combate que possua golpes a serem proferidos em sacos de pancada e aparadores. Os smartphones foram utilizados como sensores para captar ondas sonoras através de seus gravadores nativos sobre as quais tentou-se inferir a força dos golpes e a velocidade de reação do atleta após escutar uma campainha que ditava quando o golpe deveria ser executado. Na revisão bibliográfica e fundamentação teórica foram levantados aspectos do áudio analógico e digital que serviram de base para as análises dos resultados. Não foi possível medir a força dos golpes devido ao processamento interno de áudio que acontece nos smartphones mas mesmo assim levantaram-se possibilidades para utilizar smartphones no desenvolvimento da prática do esporte.

Palavras-Chave: Áudio digital, Aparador, Captação de áudio, Esporte de combate.

## **Abstract**

The objective of this work is to evaluate the possibility of using smartphones in the practice and evolution of an athlete of Tae-Kwon-Do but applicable to any combat sport that has blows to be delivered in punching bags and pads. The smartphones were used as sensors to capture sound waves through their native recorders, trying to infer the force of the blows and the athlete's speed of reaction after hearing a bell that dictated when the blow should be executed. In the bibliographic review and theoretical foundation, aspects of analog and digital audio were raised, and served as a basis for the analysis of the results. It was not possible to measure the strength of the blows due to the internal audio processing that takes place on smartphones, but even so possibilities were raised for using smartphones in the development of the sport.

Keywords: Digital audio, Punching pad, Audio capture, Combat sport.

# **1. Introdução e Justificativa.**

A utilização da tecnologia na prática e desenvolvimento esportivo já é uma realidade. Ela pode ser aplicada no auxílio da arbitragem, como no futebol com o VAR (“Video Assistant Referee” ou “Árbitro Assistente de Vídeo), tecnologia utilizada durante a Copa do Mundo da FIFA de 2018 para auxiliar na tomada de decisões complexas pelo corpo de arbitragem [16]. Também pode ser aplicada durante os treinos e desenvolvimento dos atletas.

O Tae-Kwon-Do é uma arte marcial e um dos esportes de combate mais praticados no mundo. Apesar de seus fundamentos milenares, utiliza da tecnologia principalmente nas competições onde o uso de coletes e capacetes com sensores eletrônicos é obrigatório para a marcação de pontos. Porém, assim como em muitos esportes, o equipamento eletrônico utilizado é caro e muito pouco acessível para praticantes dentro do mercado nacional, principalmente de nível amador. Os kits para competições vendidos no mercado são compostos normalmente por coletes, caneleiras, capacetes e luvas com sensores piezoeletricos capazes de indicar os golpes recebidos pelo atleta bem como a força aplicada do golpe no equipamento. É possível também adquirir kits mais simples com menos sensores apenas para treinos, como é o caso do par de coletes LED vendidos pela empresa brasileira SulSport no valor de R\$ 3.400,00 [5].

Mesmo com tecnologia e produção nacional, tais equipamentos continuam sendo inviáveis para a grande maioria das pessoas que praticam o esporte em um país como o Brasil onde o valor do salário mínimo em fevereiro de 2020 era de R\$ 1.045,00, valor bem abaixo do kit eletrônico de treino mais barato do mercado. O Tae-Kwon-Do é um esporte de grande expressividade em número de atletas, mercado e abrangência internacional, o que justifica o desenvolvimento de tecnologias nacionais que contribuem com o desenvolvimento dos praticantes do esporte por valores acessíveis.

Sendo um esporte de combate, o Tae-Kwon-Do também utiliza em suas rotinas de treino instrumentos como sacos de pancada e aparadores, e nesse ponto encontra uma intersecção com outros esportes de combate como boxe e tantas outras artes marciais. Cada uma delas possui seus desafios únicos sobre equipamentos e ambientes que realmente possam estimular os atletas. O desenvolvimento de equipamentos e tecnologias mais acessíveis para a prática dos esportes de combate poderia alavancar o Brasil que é o 29º no ranking de número absoluto de medalhas conquistadas em olimpíadas (sem contabilizar as medalhas herdadas devido a casos de doping na Rússia nos Jogos Olímpicos de 2008) mesmo sendo o 6º país mais populoso do planeta.

Se o preço dos equipamentos é uma barreira para a introdução da tecnologia e da prática de alto nível esportivo, existe um tipo de computador muito potente que é dotado de inúmeros sensores

de alta qualidade e está presente no dia a dia da população brasileira: os smartphones. Ao todo são 234 milhões de smartphones no país em junho de 2020 [10], mais de um aparelho por habitante e uma grande oportunidade para utilizar a tecnologia como aliada do desenvolvimento esportivo.

Durante ensaios gravando um atleta de Tae-Kwon-Do proferindo golpes em um saco de pancadas, notou-se que a captação do áudio do chute por um microfone condensador profissional possuía características muito similares ao áudio captado pelo microfone de um smartphone [1]. Tal fato trouxe a possibilidade de descartar o uso de um sistema de hardware dedicado para medir a força do impacto de um golpe em um saco de pancadas ou aparador, e utilizar apenas o microfone embarcado de um smartphone comum e disponível no mercado nacional. Isso demonstra a importância do desenvolvimento de tecnologias que tornem a prática esportiva de alto nível mais acessível para a população, bem como evidenciar que essa tecnologia poderia ser inteiramente desenvolvida em um equipamento muito mais acessível para a população como um smartphone comum.

## **1.1 Objetivo**

O objetivo deste trabalho é utilizar smartphones comuns para medir a força de um golpe aplicado em um aparador durante a prática de um esporte de combate, no caso o Tae-Kwon-Do. Foi testada a hipótese de que as características principais do áudio captado pelos smartphones são as mesmas entre vários modelos de celulares e poderiam servir para inferir a força de um golpe em um aparador, permitindo o desenvolvimento posterior de softwares que utilizassem o hardware embarcado desses smartphones na avaliação da prática do esporte, tornando o uso da tecnologia mais acessível para a realidade da população brasileira.

Foi avaliado o comportamento das ondas sonoras de chutes frontais aplicados em um aparador, captadas por três modelos diferentes de smartphones (Apple iPhone 8, Samsung J8 e Xiaomi Mi8) e por um microfone condensador profissional de referência. A análise foi feita tanto da amplitude da onda (domínio do tempo) como do espectro de frequência de cada amostra, utilizando um software de manipulação de áudio digital de alto nível (Ableton Live 10). Ao comparar as diversas amostras dentro do software, esperava-se que fosse possível inferir a força do impacto de cada golpe.

## **2. Revisão Bibliográfica**

### **2.1 Sensores Sonoros:**

Neste projeto foi considerado o uso de um sensor sonoro específico conhecido como microfone. O microfone é um dispositivo transdutor eletroacústico, que converte ondas longitudinais mecânicas de pressão sonora em um sinal elétrico, dentro do espectro audível que compreende frequências de 20Hz a 20kHz [9]. A intensidade sonora captada pelo microfone é diretamente proporcional à corrente elétrica alternada gerada. Na maioria dos microfones a onda mecânica vibra um diafragma em torno de um ponto de referência, a forma como este sinal captado é convertido em sinal elétrico define o tipo do microfone [15].

Existe uma variedade de transdutores utilizados como microfones para conversão eletroacústica, como o dinâmico (ou eletrodinâmico, ou eletromagnético ou de bobina móvel), piezoresistivo, piezoelétrico, óptico, spintrônico, e condensador (ou capacitivo) [15]. Abaixo estão listados os dois tipos de microfone mais usuais no cotidiano das pessoas: dinâmico e condensador.

#### **2.1.1 Microfone Dinâmico:**

O microfone dinâmico, também conhecido como microfone de bobina móvel, possui uma bobina conectada fisicamente a um diafragma de plástico que vibra de acordo com as ondas sonoras captadas. Essa pequena bobina vibra em torno de um ímã permanente posicionado dentro de seu eixo longitudinal, induzindo uma corrente no condutor [15].

Como possui um componente físico que vibra para gerar o sinal elétrico correspondente à onda sonora captada, os microfones dinâmicos estão sujeitos aos efeitos da inércia da vibração da bobina. Tal característica faz com que esses microfones tenham sua faixa de frequência de operação reduzida, tendo dificuldade em captar as ondas sonoras mais agudas que tem como característica a baixa energia e transientes muito rápidos.

No entanto é também um tipo de microfone de baixo custo, que possui um sinal elevado de saída, suporta bastante pressão sonora e é mais resistente às intempéries. Esse tipo de microfone é muito utilizado na microfonação de amplificadores e para usos gerais de voz, sendo o mais conhecido mesmo entre pessoas alheias às tecnologias da engenharia de áudio. Os microfones dinâmicos são importantes para contextualizar esse dispositivo pois são os mais conhecidos, mas neste projeto serão descartados em virtude dessa resposta prejudicada que possuem com altas frequências e com transientes rápidos como é um caso de um golpe proferido em um aparador. A seguir será apresentado o tipo de microfone escolhido para os testes práticos.

### 2.1.2 Microfone Condensador:

O microfone condensador possui um par de placas paralelas sendo uma delas fixa e outra móvel acoplada ao diafragma que capta a vibração das ondas sonoras. Essas placas são polarizadas com cargas opostas e de mesmo módulo, caracterizando um capacitor que tem sua capacidade variada de acordo com a distância entre as placas. Esse capacitor variável conecta-se em paralelo com um resistor e a tensão nos terminais desse resistor caracteriza o sinal de saída captado [15]. Os microfones condensadores possuem uma grande faixa de operação em frequência sendo geralmente utilizados em aplicações profissionais de áudio hi-fi, porém requerem alimentação externa em corrente CC para polarização do capacitor.

#### a) Microfone Condensador Cardióide Audio-Technica AT2020:

Como referência de captação de áudio foi utilizado um microfone do tipo condensador com padrão polar cardióide da marca japonesa Audio-Technica, modelo AT2020 [12]. O padrão polar cardióide significa, na prática, que as ondas sonoras devem ser captadas na direção frontal pois o sinal sofre atenuação quando captado de forma lateral ou fica praticamente inaudível quando captado pela parte traseira do condensador como se observa no diagrama do padrão polar, Figura 2.1.

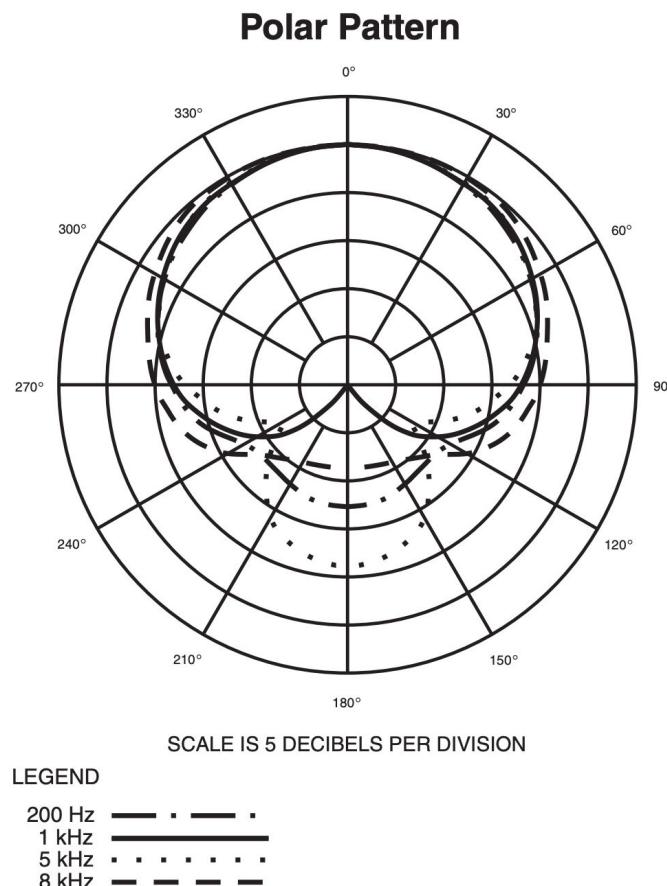
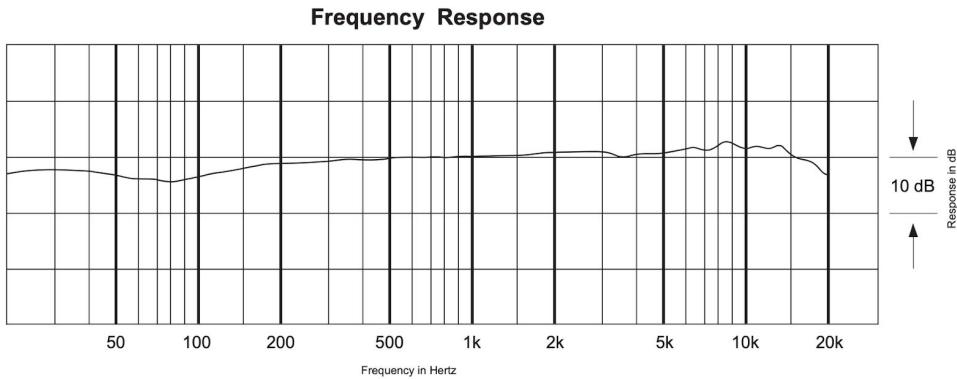


Figura 2.1 - Diagrama do padrão polar do AT2020 [12].



*Figura 2.2 - Resposta em frequência do AT2020 [12].*

Como visto na Figura 2.2, o AT2020 oferece uma boa resposta de frequência, praticamente sem alterar o timbre da onda sonora captada [12], além de possuir uma intensidade sonora máxima de entrada de 144dB, o que em termos práticos seria algo como uma casa de máquinas em operação. Isso, de certa forma, é uma garantia de que o AT2020 não será prejudicado nem atingirá a saturação do sinal captado de um golpe em um aparador. Para funcionar o AT2020 necessita de uma alimentação CC polarizando as placas paralelas do capacitor [12], essa alimentação é feita através da interface de áudio que será apresentada na próxima sessão.

O microfone AT2020 pode ser visto na Figura 2.3, e possui as seguintes características [12]:

- Padrão polar: cardióide
- Resposta de frequência: 20 a 20.000 Hz
- Sensibilidade: -37 dB
- Impedância: 100 ohms
- Intensidade sonora máxima de entrada: 144 dB
- Conector de saída: XLR



*Figura 2.3 - Microfone condensador Audio-Technica AT2020.*

### b) Microfone Condensador com tecnologia MEMS:

Os MEMS (Micro Electro Mechanical Systems) são dispositivos eletromecânicos microscópicos que consideram não apenas as características elétricas dos semicondutores mas também suas características e aplicações mecânicas. Foram desenvolvidos na década de 60 pelo engenheiro americano Harvey C. Nathanson e desde o início do século XXI tem ganhado muita força com a expansão do mercado de smartphones e em aplicações na indústria automobilística, além de aplicações em diversas outras indústrias como na aeroespacial, aeronáutica e na medicina [2] [11].

Os microfones MEMS possuem um sistema de funcionamento que os caracteriza como um microfone capacitivo porém com dimensões microscópicas. Em geral possuem um diafragma da ordem de poucos milímetros que movimenta uma placa móvel conectada a um eletrodo, logo abaixo um *gap* de ar seguido de uma placa fixa também conectada a um eletrodo. A placa fixa possui um micro perfuramento que impede que o *gap* de ar funcione como um amortecedor, interferindo na dinâmica da onda sonora captada.

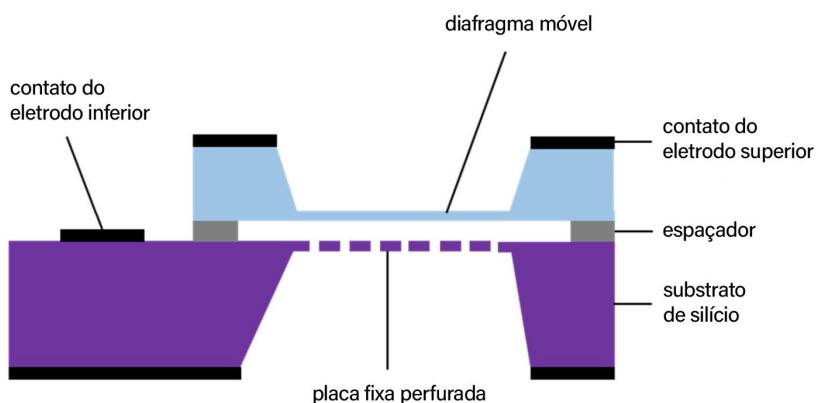


Figura 2.4 - Estrutura básica de um microfone MEMS [2].

As empresas montadoras de smartphones não especificam as características dos microfones em seus datasheets e limitam-se, normalmente, apenas em dizer que seus aparelhos possuem microfone integrado. De acordo com Zawawi et al [2], uma das primeiras empresas a aderir à tecnologia dos microfones MEMS foi a Apple, Inc desde o desenvolvimento do iPhone 4. A utilização de dispositivos MEMS foi um dos fatores de sucesso do aparelho [3] que possuía diversos sensores muito bem integrados ao software e desencadeou a utilização desse tipo de componente nos modelos subsequentes da companhia. O iPhone 4 contava com três microfones MEMS sendo dois deles utilizados para captação do áudio externo e um posicionado na traseira do equipamento apenas para captação e posterior eliminação de ruído [3].

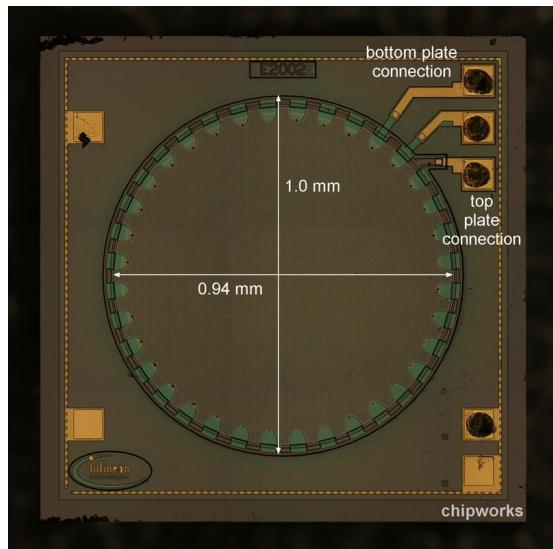


Figura 2.5 - Matriz de microfone MEMS Infineon E2002 encontrado no iPhone 4 [3].

De acordo com [3], as empresas envolvidas no desenvolvimento dos microfones MEMS do iPhone 4 foram a Knowles Technology e Infineon Technologies sendo que a fabricação desses componentes da Knowles é feita pela Sony Semiconductor Kyushu Corp e a Infineon fornece a matriz de seus microfones MEMS para três fabricantes asiáticos de microfones: AAC Acoustic Technologies Holdings Inc., BSE Co. Ltd. e Hosiden Corp. Esses três fornecedores junto com a Knowles e a Analog Devices constituem os 5 maiores fornecedores de microfones MEMS do mercado.

Dado o sucesso do iPhone 4 e do tamanho dos fornecedores envolvidos nesse projeto lançado em 2011, foi inferido neste projeto que todos os smartphones utilizados nos testes possuem microfones MEMS embarcados em seu hardware o que é um indício teórico que reforça a possibilidade de utilizar os smartphones comerciais em aplicações específicas de captação de áudio como o som de um golpe proferido em um aparador ou saco de pancadas, já que esse tipo de microfone tem apresentado resultados notáveis não apenas em custo mas na qualidade do áudio captado [2]. Não foi possível determinar se os smartphones utilizados no projeto possuem realmente microfones MEMS pois as montadoras não divulgam esse tipo de informação, ela só pode ser obtida a partir de auditorias feitas por equipes especializadas [3] [17] como é o caso da empresa francesa System Plus Consulting que divulgou um relatório parcial em 2018 sobre os microfones MEMS das fabricantes Goertek, Knowles e AAC Technologies presentes no modelo Apple iPhone X [17]. Para ter acesso ao relatório completo e detalhado é necessário pagar o valor de 4490,00 € (euros).

## 2.2 Interface de áudio USB Focusrite iTrack Solo:

Como citado anteriormente, para que o microfone AT2020 opere normalmente ele precisa de uma alimentação CC que polarize as placas do seu capacitor. Essa alimentação é feita através do próprio cabo de som que é um cabo平衡ado com conectores XLR conectado a uma interface de áudio da marca Focusrite, modelo iTrack Solo. Essa alimentação CC é de +48V e possui um nome específico, muito utilizado na indústria de áudio profissional chamado *phantom power* [13] [15].



Figura 2.6 - Vista frontal da interface de áudio USB Focusrite iTrack Solo [14].

A interface de áudio funciona como um conversor A/D que capta o sinal analógico emitido pelo AT2020 e o converte para um sinal digital com frequência de amostragem máxima de 96kHz e densidade de bit máxima de 24-bit. De acordo com o Teorema de Nyquist, essa taxa de amostragem é suficiente para captar todas as frequências do espectro audível (de 20 a 20.000 hertz) sem que ocorra aliasing [7] [15]. A amostragem do sinal analógico será melhor explorada na fundamentação teórica deste projeto.

A Focusrite iTrack Solo conecta-se a um computador via porta USB, onde é possível monitorar, gravar, manipular e analisar o áudio digital em software específico.

## 2.3 Software de análise e manipulação de áudio utilizado:

Existem diversos softwares que permitem trabalhar com áudio digital em alto nível. No mercado eles são comumente chamados de “DAW’s” (Digital Audio Workstation) sendo alguns dos nomes mais conhecidos o Logic Pro [18], Cubase [19], Pro Tools [20] e o software utilizado neste projeto, o Ableton Live 10 [4].

Ableton é uma empresa baseada em Berlim, fundada em 1999 e que lançou em 2001 a primeira versão de seu software, o Live [21]. Ao invés de desenvolver inúmeros produtos, a empresa concentra-se em melhorar constantemente o próprio Live e suas próprias soluções de integração de hardware (Push e Link). No aspecto criativo, a Ableton se diferenciou no mercado das DAW's ao criar o conceito de "session view" que aproxima o fluxo de trabalho da música computacional com a maneira como uma banda constrói um arranjo musical [4].



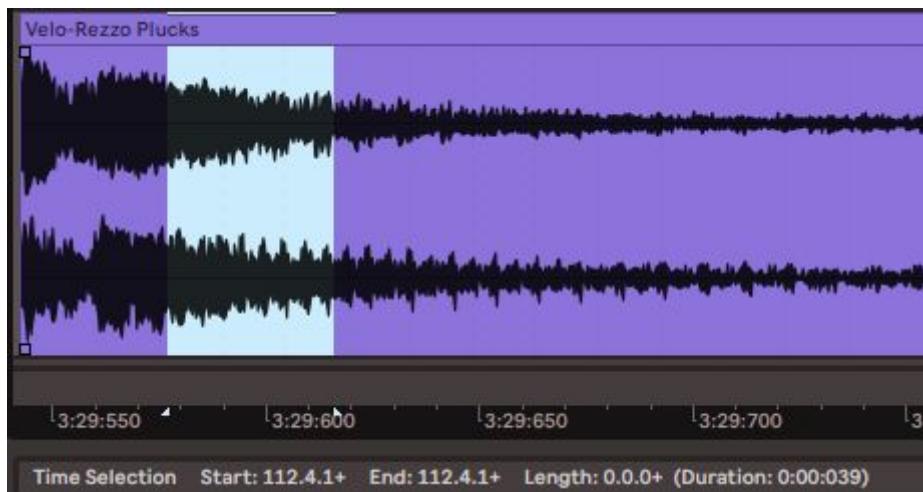
Figura 2.7 - Arranjo musical criado na "session view" do Ableton Live 10

A versão completa do software, o Live Suite, é vendido no site da Ableton por US\$ 749, aproximadamente R\$ 4.000,00 no câmbio de novembro de 2020. É um software robusto que mistura os conceitos da música computacional e eletrônica com emulações de equipamentos analógicos como o plugin nativo "Glue Compressor" que emula o compressor do canal master das mesas SL4000 série E da montadora Solid State Logic, modelo de mesas de som profissional lançado em 1979 e que possui valor inestimável na atualidade [22] [4].

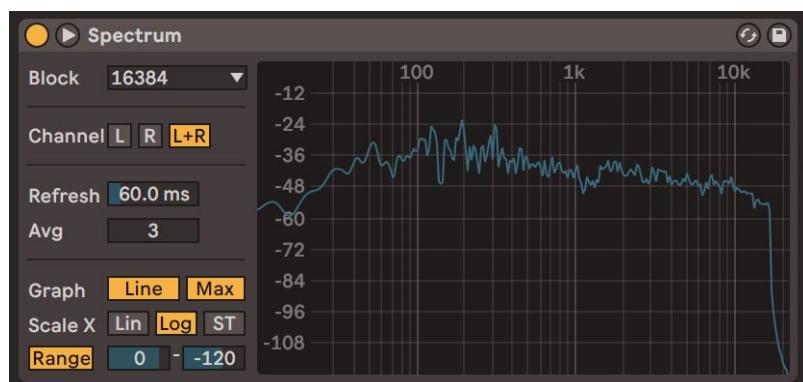


Figura 2.8 - Glue Compressor, plugin nativo com modelagem analógica desenvolvido pela Ableton em colaboração com a Cytomic [4]

Além de desenvolver soluções que focam na criatividade para que qualquer música possa ser criada de forma fluida dentro do Live [21], o software também oferece ferramentas que permitem trabalhar com o áudio em alto nível. O Live 10 é um software que opera seu processamento interno com uma densidade de bits de 32-bit, superior à da interface de áudio e fundamental para um processamento que a própria Ableton chama de “Operação Neutra”, ou seja, não altera as características do sinal processado [4]. Além disso dentro do software é possível manipular e monitorar o áudio de inúmeras formas tanto no domínio do tempo quanto no domínio da frequência. Os principais motivos da escolha deste software para análise das amostras de som são a possibilidade de medir a duração de qualquer amostra com uma resolução máxima de 1 ms, a possibilidade de analisar o espectro de frequência do som e a garantia de que o software não faça manipulações indesejadas no áudio como por exemplo, a diminuição da profundidade de bit e consequente distorção do que foi captado.



*Figura 2.9 - Seleção de um trecho de 39 ms em azul claro*



*Figura 2.10 - Spectrum, plugin nativo para análise do espectro de frequência de uma amostra de som.*

Assim como as outras DAW's, este é um excelente software para compreender de forma didática o comportamento de um sinal digital processado. Este é um adendo ao projeto, pois as DAW's são uma poderosa ferramenta potencial para o ensino da engenharia elétrica nas áreas de sinais e sistemas, dado à sua interface visual intuitiva que conecta os elementos do dia a dia (sons e música) com conceitos avançados do processamento digital de sinais.

### **3 Fundamentação teórica:**

Para avaliar a possibilidade de utilizar um smartphone como sensor na captação de um golpe em um aparador ou saco de pancadas, é necessário entender alguns aspectos que caracterizam tanto os golpes quanto os processos de conversão analógica/digital que ocorrem nas câmeras e microfones utilizados no experimento.

#### **3.1 Intensidade Sonora**

De acordo com [9], a intensidade (I) de uma onda sonora em uma superfície é a média de energia transferida por unidade de área e pode ser descrita matematicamente como:

$$I = \frac{P}{A} \quad (\text{I})$$

Onde P é a potência da fonte sonora que incide sobre uma determinada área A. No caso deste projeto, a potência sonora emitida pela fonte é diretamente proporcional à força do golpe proferido no aparador e inversamente proporcional ao tempo de interação entre atleta e aparador. De acordo com os dados de [1], esse tempo de interação para um atleta experiente varia em poucos centésimos ou milésimos de segundo entre múltiplos golpes. Portanto, dada as condições deste projeto a potência da fonte sonora será considerada como uma variação direta da força do golpe aplicado no aparador.

##### **3.1.1 Variação da intensidade sonora com a distância**

A forma como a intensidade sonora varia com a distância à partir de uma fonte sonora real em um ambiente real, pode ser bastante complexa. Se a fonte sonora for um falante de um monitor de áudio ela é projetada para que o som seja unidirecional, se o ambiente onde a fonte está situada possuir muitas reflexões, a distribuição de frequências emitidas pelo falante.

No entanto, considerando-se uma fonte sonora isotrópica e puntiforme, é possível aproximar a superfície pela qual transfere-se a potência sonora como uma esfera de raio  $r$ , como segue:

$$I = \frac{P}{4\pi r^2} \quad (\text{II})$$

Logo, a intensidade sonora varia inversamente proporcional ao quadrado da distância da fonte sonora.

##### **3.1.2 A escala decibel**

A escala decibel é uma escala logarítmica que compara a intensidade de uma fonte sonora com o limiar da audição humana. O ouvido humano é capaz de identificar uma variação altíssima de intensidades sonoras que vão desde sons ambientes até níveis dolorosos de emissão sonora como estar ao lado do motor de um jato ligado [9] [15].

O nível sonoro ( $\beta$ ) é calculado como sendo:

$$\beta = (10 \text{ dB}) \log \frac{I}{I_0} \quad (\text{III})$$

Onde  $I_0$  é o padrão de referência da menor intensidade sonora percebida pelo ouvido humano ( $I_0 = 10^{-12} \text{ W/m}^2$ ). Se  $I=I_0$ , então a equação acima resulta em  $\beta = 10 \log 1 = 0 \text{ dB}$ .

Através de manipulação matemática também é possível notar que se comparados dois sinais  $a$  e  $b$  sendo que  $a$  possui o dobro da potência sonora de  $b$ , a relação do nível sonoro entre os dois sinais é dada por:

$$\beta = (10 \text{ dB}) \log \frac{I_a}{I_b}$$

$$\beta = (10 \text{ dB}) \log \frac{\frac{P_a}{P_b}}{\frac{T_b}{T_a}}$$

$$\beta = (10 \text{ dB}) \log = 3,01 \text{ dB} \quad (\text{IV})$$

Portanto, a cada 3 dB de ganho o sinal dobra sua potência e a cada 3 dB de atenuação o sinal diminui sua potência pela metade.

### 3.1 Sinal de áudio no domínio do tempo - envelope ADSR:

Existem dois aspectos fundamentais que podem caracterizar qualquer onda sonora, considerando que as mesmas sempre são sinais contínuos e com duração bem definida. O primeiro aspecto é a análise do som no domínio do tempo, popularmente chamado de envelope ADSR na acústica e na música eletrônica.

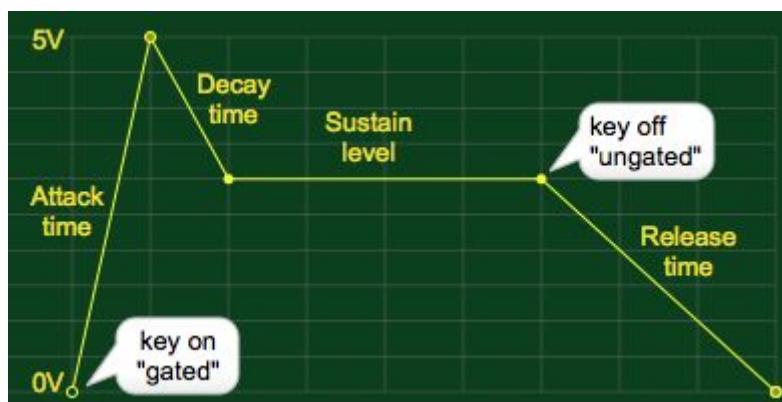


Figura 3.1 - Diagrama esquemático de um envelope ADSR [15].

ADSR é um acrônimo de Attack, Decay, Sustain, Release, ou em português: Ataque, Decaimento, Sustentação e Relaxamento e descreve um envelope de amplitude de um sinal de áudio [15]. O conceito de envelope surgiu durante o desenvolvimento dos primeiros

sintetizadores analógicos na década de 60 quando os engenheiros da época buscavam uma forma de humanizar o som sintetizado e aproximá-lo da experiência de tocar e escutar um instrumento real [23]. Foi proposto por Vladimir Ussachevsky e desenvolvido pela primeira vez pelo engenheiro Robert Moog [15] [23], inicialmente como um envelope que modulava a tensão de um sinal e dessa forma modulava diretamente a amplitude desse mesmo sinal.

Da síntese para a análise, é possível aplicar o conceito do envelope ADSR para entender a dinâmica de uma amostra de som, como é descrito a seguir:

1. Ataque (attack): é o tempo que uma onda sonora leva para sair do silêncio (0 dB) até sua amplitude máxima. Em um instrumento como uma flauta doce o ataque é praticamente instantâneo pois assim que o ar passa pela flauta ela já começa a emitir som. Instrumentos com ataque muito longo são bastante utilizados na construção da harmonia em vários estilos contemporâneos de música eletrônica além de trilhas sonoras cinematográficas que tentam causar um pico de tensão emocional no espectador;
2. Decaimento (decay): após o pico de amplitude a onda sonora perde uma parcela de sua energia até entrar em seu regime de sustentação. Um excelente exemplo de variação do decaimento é tocar um piano sem nenhum pedal pressionado e com o pedal da direita pressionado. O pedal da direita quando pressionado permite que as cordas do piano vibrem livremente o que diminui essa queda na energia da vibração e portanto aumenta o tempo de decaimento e também de sustentação e relaxamento, como descrito a seguir;
3. Sustentação (sustain): utilizando-se a analogia do piano, o período de sustentação seria exatamente o período em que a tecla de uma nota permanece pressionada. Os sintetizadores popularmente chamados de "teclados" são muito utilizados no mercado musical desde a segunda metade do século XX [15]. Como seu som origina-se de um circuito eletrônico oscilador é possível sintetizar timbres característicos com ataque praticamente zero, decaimento zero, sustentação máxima (a nota não perde amplitude enquanto permanecer pressionada) e relaxamento praticamente zero. O relaxamento de uma onda sonora é descrito a seguir;
4. Relaxamento (release): é o tempo de duração da onda sonora após a liberação da nota (fim da sustentação). Em instrumentos musicais convencionais o relaxamento se inicia após soltar a tecla do piano, tocar a corda do violão ou parar de assoprar uma flauta doce. Na acústica o relaxamento das ondas sonoras é muito importante pois é nele que costuma estar contida a cauda de reverberação resultante da interação do sistema que envolve a fonte sonora com a acústica do ambiente e o dispositivo de captação do som (que pode ser um microfone ou o próprio ouvido humano)

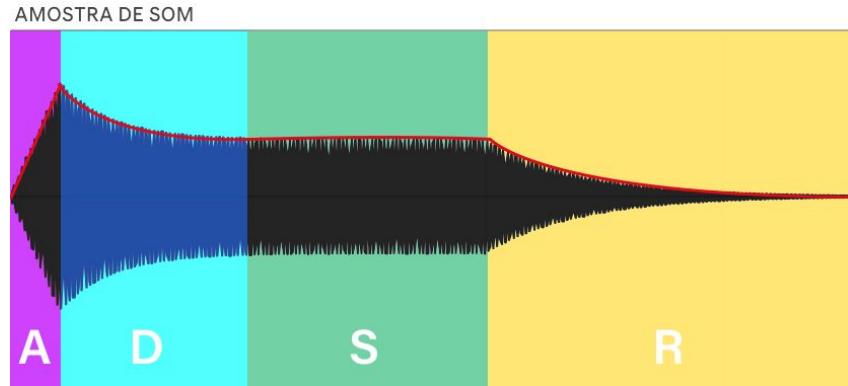


Figura 3.2 - Envelope ADSR de amplitude de uma amostra de som.

A amostra de som da Figura 3.2 foi sintetizada digitalmente dentro do próprio Ableton Live, é uma onda puramente senoidal com frequência de 496 Hz (corresponde à nota "si" da terceira oitava do teclado). Possui um ataque de 25 ms, decaimento de 217 ms, nível de sustentação de -6 dB e período de relaxamento de 400 ms.

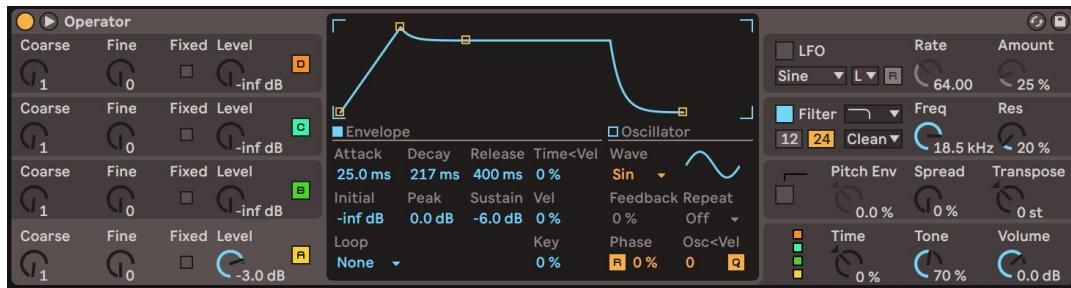


Figura 3.3 - Sintetizador digital que originou a forma de onda da Figura 3.2

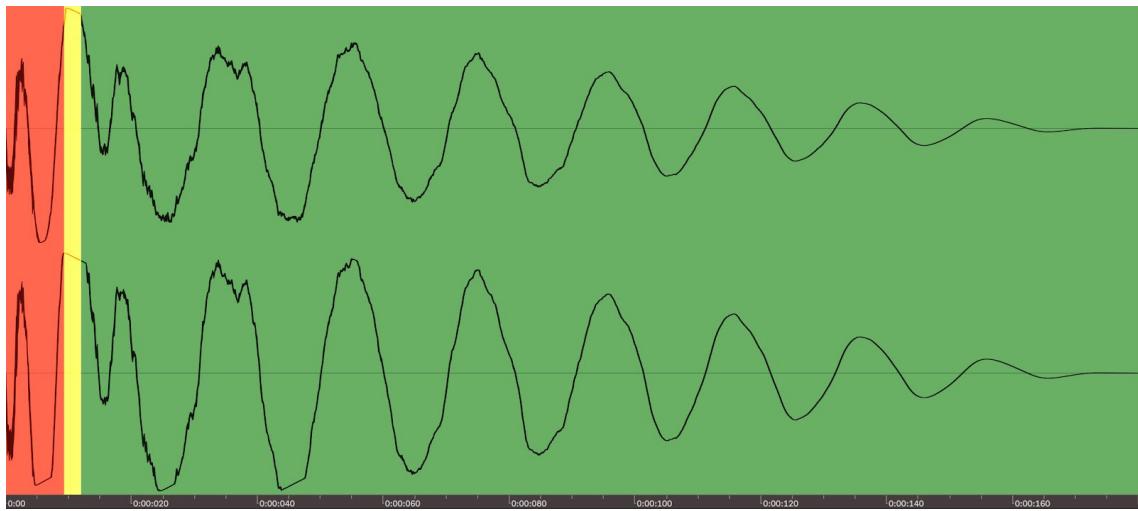
### 3.1.1 Envelope ADSR esperado de um golpe em um aparador:

Antes da experimentação existem algumas características que podem ser levadas em consideração sobre o envelope ADSR esperado de um golpe em um aparador. No universo da produção musical é muito comum que o bumbo da bateria (sintetizado ou gravado) seja chamado de kick ("chute" em inglês) [24]. Isso porque é um som que costuma ter algumas características principais:

1. Ataque muito rápido, da ordem de décimos, centésimos ou até milésimos de segundo;
2. Período de decaimento curto, da ordem de milésimos de segundo;
3. Período de sustentação praticamente zerado;
4. Período de relaxamento curto, da ordem de décimos ou até centésimos de segundo.

No caso de um bumbo de bateria, espera-se que a onda seja potencializada para suas finalidades musicais. No caso de um aparador ou saco de pancadas, o objeto foi construído para dissipar o

máximo de energia possível então é esperado que as ondas dos golpes possuam um período mais curto de relaxamento.



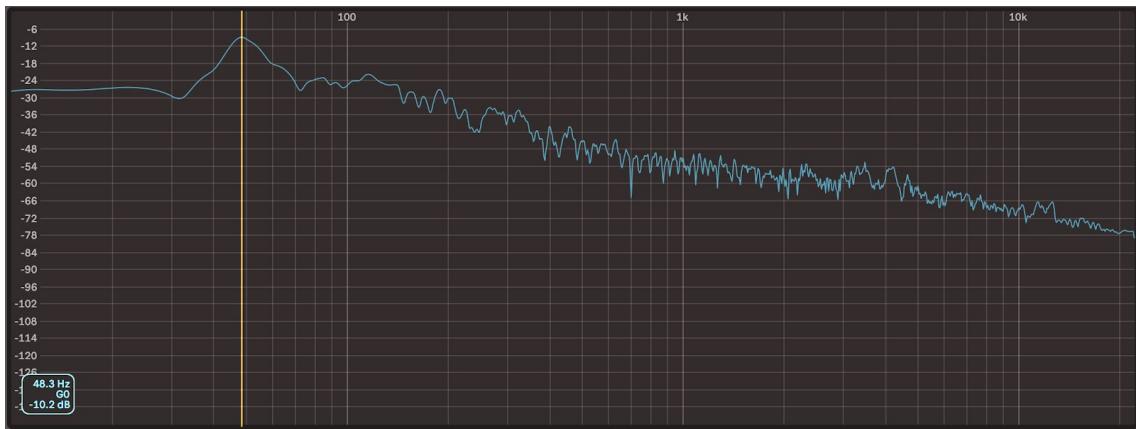
*Figura 3.4 - Onda sonora de um bumbo de bateria.*

Na Figura 3.4 acima observa-se a onda sonora de um bumbo de bateria com ataque de 9 ms (faixa vermelha), decaimento de 4 ms (faixa amarela), sustentação praticamente zerada e relaxamento de 167 ms (faixa verde). Nesse tipo de som chega a ser difícil enxergar e medir o envelope pois o período de decaimento e de sustentação costumam tender a zero visto que é um som extremamente percussivo e carregado de frequências mais graves (que no gráfico aparecem como ondas maiores e com picos e vales mais espaçados, que dificultam a visualização do envelope ADSR)

Em esportes de combate, normalmente deseja-se que o golpe seja o mais rápido possível (que resulta em um ataque rápido) e com o menor período possível de contato do atleta com a superfície golpeada (que resulta em um período de sustentação muito curto).

### **3.2 Sinal de áudio no domínio da frequência - timbre:**

É possível obter duas ondas sonoras com o mesmo envelope ADSR e que podem inclusive estar tocando a mesma nota musical e ainda assim o resultado para os ouvidos ser totalmente diferente. Essa discrepância que ocorre, por exemplo, entre instrumentos musicais é resultado da diferença entre os sinais no domínio da frequência. As notas são determinadas pela frequência fundamental de um sinal mas o conjunto de harmônicos que acompanham essa fundamental é o que faz o ser humano identificar se uma nota específica vem de um piano, de um violão ou de uma flauta. A essas características únicas de cada instrumento ou fonte sonora é dado o nome de timbre [15].



*Figura 3.5 - Espectro de frequências do bumbo de bateria da Figura 3.2.*

Na Figura 3.5 observa-se o espectro de frequência do mesmo bumbo de bateria da Figura 3.4. Destacado pela linha amarela está o harmônico principal deste bumbo, que o próprio software identifica como sendo um pico de -10,2 dB (áudio digital) na frequência de 48,3 Hz que corresponde à nota sol da oitava zero. Esse espectro é resultante dos picos em cada frequência, mas na curta duração de 181 ms dessa amostra de som esse espectro sofre variações.

### **3.2.1 Timbre esperado de um golpe em um aparador:**

Mantendo-se o comparativo do bumbo de bateria com a onda sonora de um golpe em um aparador, é possível fazer uma avaliação prévia da combinação do envelope ADSR com o timbre do golpe ou do bumbo:

1. O ataque inicial possui transientes rápidos e acentuados compostos principalmente por ondas de frequências mais altas que rapidamente dissipam sua energia;
2. A fundamental é uma frequência subgrave que possui mais energia e, portanto, oscila por mais tempo que as frequências agudas;
3. O instante inicial da onda sonora pode definir o momento em que o golpe atingiu o aparador;
4. O primeiro pico da fundamental pode ajudar a definir a força aplicada no aparador pelo golpe.

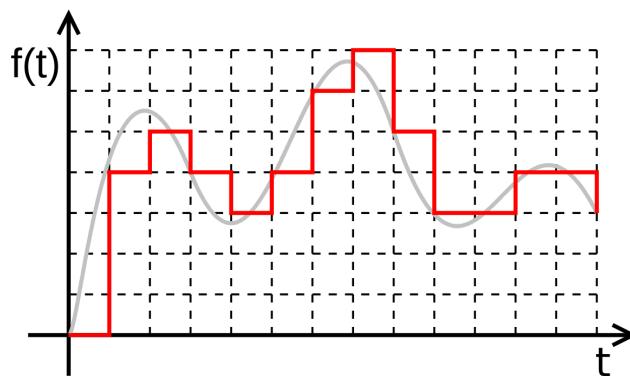
## **3.3 Sinal analógico:**

Um sinal é formalmente definido como uma função de uma ou mais variáveis, a qual veicula informações sobre a natureza de um fenômeno físico [6]. No caso de sinais analógicos, trata-se de um tipo de sinal contínuo que pode assumir infinitos valores de função dado um determinado

intervalo de tempo. Dois exemplos de sinal contínuo são a variação da tensão em volts durante um intervalo de tempo de um componente em um circuito elétrico e a variação da temperatura ao longo de 24 horas na cidade de São Carlos. Além destes exemplos, são também sinais analógicos as ondas sonoras propagadas em um ambiente por uma fonte sonora.

### 3.4 Sinal digital:

Sinal digital é a representação de um sinal analógico em um domínio discreto, onde dado um intervalo de tempo a função pode assumir um número finito de valores.



*Figura 3.6 – Representação de um sinal digital (curva vermelha) no domínio do tempo.*

A digitalização de um sinal possui a desvantagem da perda de informação contida em um sinal analógico durante a sua discretização mas esse aspecto será tratado nas próximas sessões. As vantagens da abordagem digital em relação ao processamento analógico de sinais possuem dois pontos principais [6]:

1. Flexibilidade, pois o mesmo hardware digital pode processar sinais de diferentes formas apenas com adaptações via software, visto que o processamento digital de sinais acontece através de operações matemáticas digitais e armazenamento em memória;
2. Repetitividade, pois como o sinal assume valores finitos bem determinados de função, operações digitais podem realizar o mesmo processamento inúmeras vezes e de maneira exata. Os sistemas analógicos podem sofrer variações de parâmetros devido aos efeitos físicos que ocorrem nos componentes do sistema como por exemplo, variações no timbre de amplificadores de áudio valvulados, muito utilizados em guitarras.

### 3.5 Processamento digital de sinais:

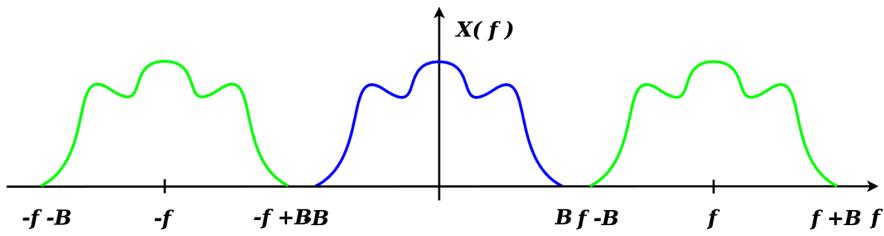
O Processamento Digital de Sinais (DSP) é um dos aspectos principais deste projeto, pois visa captar um sinal analógico (onda sonora propagada pelo golpe) e convertê-lo em um sinal

discreto que possa ser processado e monitorado digitalmente [6] [7]. A conversão analógica-digital (A/D) de sinais de áudio está totalmente presente no dia a dia das pessoas quando fazem uma simples chamada de telefone para outra pessoa através de um celular [6]. A perda de informação contida em um sinal no processo de conversão A/D é contornável desde que obedecidas as premissas do Teorema da Amostragem.

### 3.5.1 Teorema da amostragem:

O teorema da amostragem é uma comprovação matemática de que é possível reconstruir perfeitamente um sinal analógico a partir de uma sequência de amostras uniformemente espaçadas (sinal discreto), desde que o sinal analógico esteja contido dentro de uma banda de frequências finita e que a taxa de amostragem seja superior ao dobro da maior frequência contida no sinal amostrado [7].

A frequência de amostragem ou taxa de amostragem é a taxa que determina o número de amostras registradas de um sinal analógico em um determinado período de tempo. O espectro do sinal amostrado é uma reprodução do espectro do sinal original, com inúmeras cópias do mesmo deslocadas de  $\omega_M$  (sendo  $\omega_M$  a frequência máxima do sinal amostrado), o que implica na necessidade de aplicar um filtro passa-baixa (LPF ou "Low Pass Filter" como é normalmente chamado na produção musical) para obtenção do sinal original à partir do sinal discreto. Esse processo de filtragem para eliminar as imagens do sinal é também chamado de filtro anti-aliasing.



*Figura 3.7 – Sinal amostrado, em azul a cópia do sinal original e em verde as imagens replicadas que precisam ser filtradas.*

Considerando uma onda sonora e o espectro audível humano que vai de 20 Hz a 20 kHz, a taxa de amostragem mínima para reprodução de um sinal de áudio seria de 40 kHz, porém se essa taxa de amostragem fosse exatamente 40 kHz, as regiões de frequências próximas de  $|\omega|$  sofreriam distorções em função da utilização de um filtro passa-baixa real e não ideal [7].

Por esse motivo é necessário haver uma banda de transição que aumente a distância das imagens e crie uma zona de segurança onde as distorções do filtro não interfiram nas frequências do sinal original.

Na indústria, o sample rate mais utilizado é de 44100 Hz. Esse valor específico tem a ver não apenas com o teorema da amostragem ou com a banda de transição, mas com motivações históricas e tecnológicas que envolvem a capacidade de armazenamento dos CD's, os sistemas de transmissão analógica NTSC e PAL e a gravação de mídias via adaptadores PCM.

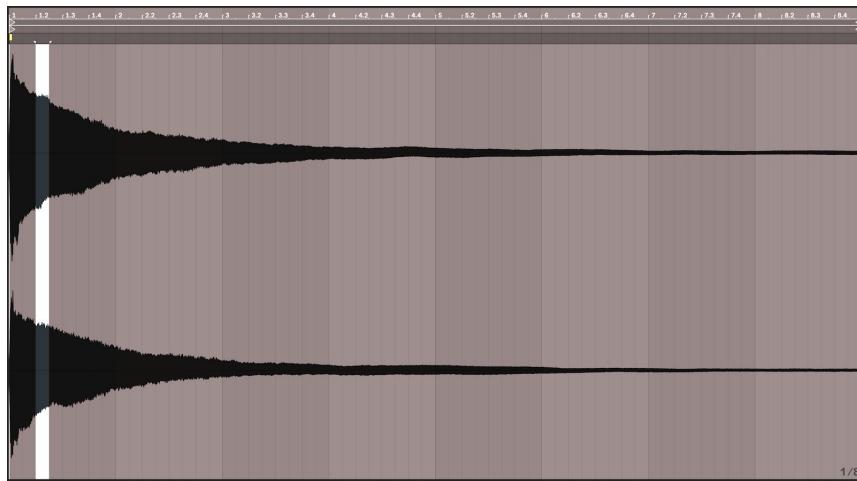


Figura 3.8 – Nota lá (A1) de um piano com um trecho selecionado de 190 ms.

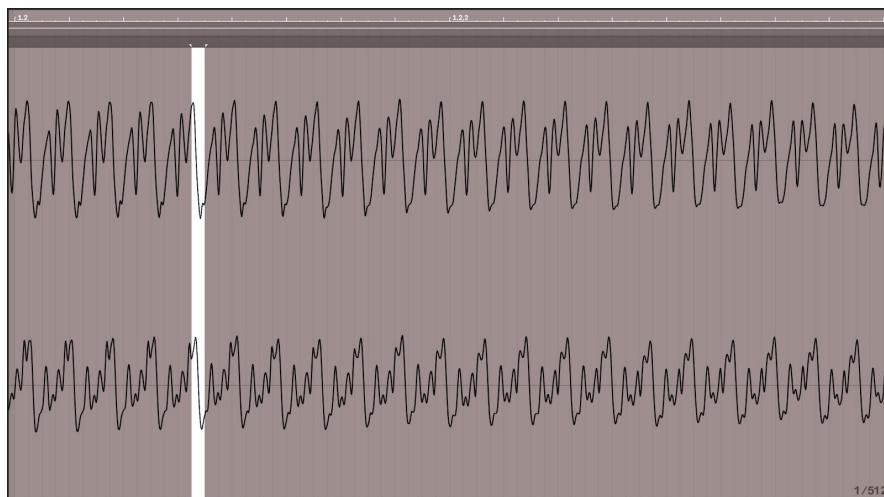
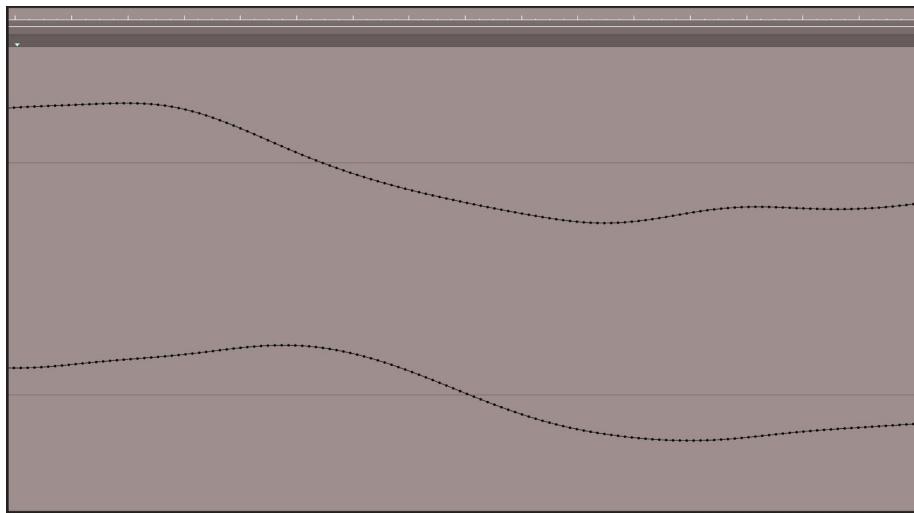


Figura 3.9 – Zoom no trecho de 190 ms com seleção em trecho de 3 ms.



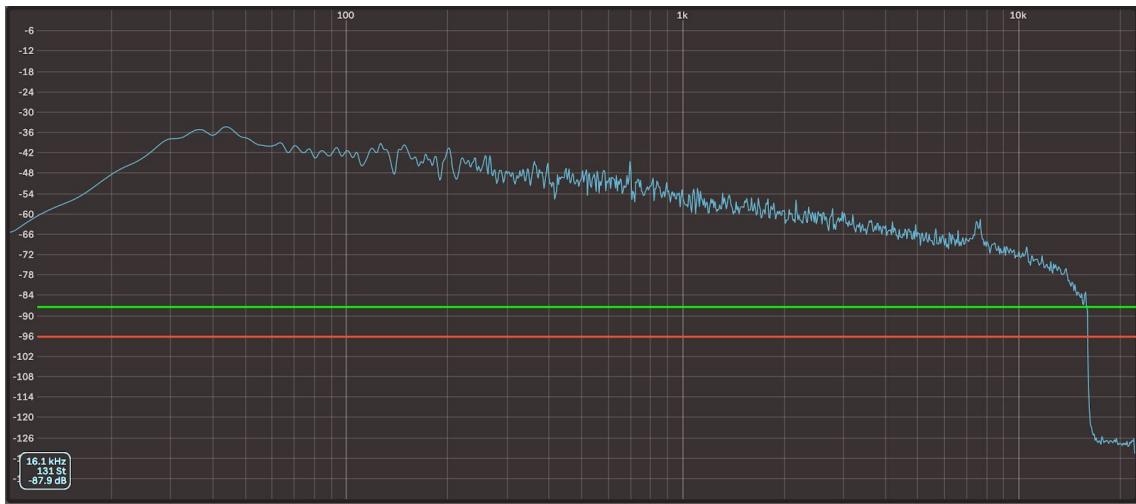
*Figura 3.10 - Zoom no trecho de 3 ms.*

Nas três figuras acima é possível observar o envelope de uma nota lá tocada em um piano, aumentando-se o zoom atinge-se um ponto onde o próprio software indica os pontos de sampling do sinal. Para efeito visual o Ableton Live 10 une cada ponto da amostra com um segmento de reta, mas o sinal trabalhado dentro do software é um sinal discreto e na prática esses pontos estão ligados por degraus discretizados e por uma interpolação de sinal discreto quando reproduzido [4].

### **3.5.2 Profundidade de bit de áudio (audio bit depth):**

No áudio digital, bit depth é o número de bits de informação em cada amostra de áudio, o que está diretamente relacionado com a resolução do sinal discretizado. Quando um valor analógico assume um valor discreto na vizinhança de seu valor real, o sinal sofre uma distorção também chamada de erro de quantização e que resulta na geração de um ruído de quantização no sinal [8]. Na engenharia de áudio trabalha-se com um valor de 6 dB por bit. Mídias comuns como arquivos MP3 e CDs possuem 16 bits de profundidade e uma relação sinal/ruído de 96 dB. Isso significa que sons com amplitude abaixo de -96 dB sumirão do sinal no processo de conversão A/D.

É muito comum que no processo de produção musical se trabalhe com 24 bits de profundidade em uma relação sinal/ruído de 144 dB para que o áudio possa ainda sofrer perdas de resolução no processo de masterização e para que não se percam elementos que ficaram muito baixos na etapa de mixagem mas façam parte do contexto musical.



*Figura 3.11 - Espectro de frequências de um áudio gravado com 16 bits de profundidade.*

A Figura 3.11 apresenta o espectro de frequência do som da mata durante a noite em São Sebastião/SP, gravado em um iPhone 8 com profundidade de 16 bits. A linha laranja indica o nível de -96 dB, valor mínimo esperado para uma amostra com essa profundidade de bit, a linha verde indica o nível de -87,9 dB que é o nível de volume mais baixo identificado como parte do sinal captado.

O eixo ordenado está com seu alcance de -144 dB (limite para um sinal com 24-bit depth) até 0dB. Por volta de -126 dB é possível observar o ruído de quantização que acompanhou a gravação. Apesar desse ruído ser praticamente inaudível, ele ocupa espaço de memória e também ocupa espaço na mixagem da música, interferindo no processo de finalização de uma música, por exemplo. Existem formas de atenuar esse ruído de quantização como a utilização de um dither, mas este já foge das considerações que envolvem este projeto.

### 3.5.3 Ponto flutuante:

Diferente da intensidade sonora medida em um meio físico, as ferramentas que operam com áudio digital possuem um limite máximo de amplitude do sinal estabelecido em 0 dB. Isso acontece pois o nível de intensidade sonora no ambiente depende do equipamento de reprodução do som, por isso essas ferramentas garantem um nível máximo de saída de sua interface digital. O problema é que a limitação digital de um sinal, diferente de um sistema eletrônico analógico que possui semicondutores e demais componentes com curvas contínuas de saturação/limitação, é extremamente agressiva pois ceifa o sinal de forma brusca criando um pico com formato de onda quadrada [25].

Na produção musical o fenômeno da limitação digital é extremamente indesejável, visto que cria harmônicos desagradáveis para o ouvido humano e literalmente tira a definição das ondas sonoras (este fenômeno é popularmente chamado de "clipagem digital"). Para evitar esse

problema durante o processo criativo, algumas DAWs (incluindo o Ableton Live) operam com ponto flutuante no canal master e também fazem uma conversão para 32-bit depth quando se opera dentro do projeto. O Ableton Live permite que o sinal ultrapasse para até +6 dB, alocando 1 bit para esse overflow e alterando a resolução da faixa de 0 a -192 dB para a faixa de +6 a -186 dB em seu roteamento interno. É uma boa prática evitar a operação nessa amplitude de sinal para que não ocorram outras distorções nos processamentos dos canais individuais nem do próprio canal master.

## 4 Experimentação e resultados:

### 4.1 Montagem do experimento

Para a montagem do experimento foram utilizados além do microfone condensador Audio Technica AT2020 (que funcionará como referência de captação do áudio), três modelos de smartphones comuns no mercado atual: 1 Apple iPhone 8, 1 Samsung J8 e 1 Xiaomi Mi 8. A escolha desses modelos foi motivada principalmente pelo fato de serem aparelhos de montadoras diferentes (Apple, Samsung e Xiaomi), com sistemas operacionais diferentes (iOS e Android), desenvolvidos em países diferentes (Estados Unidos, Coréia do Sul e China) e que tem uma boa penetração no mercado, sendo comuns entre os usuários de smartphones [10].

O AT2020 foi conectado diretamente à interface de áudio Focusrite iTrack Solo via cabo XLR. Esta por sua vez estava conectada no computador através de uma entrada USB. O áudio captado pelo microfone foi gravado diretamente no Ableton Live.

O experimento consiste em gravar golpes proferidos por um atleta experiente em um aparador convencional da marca Sulsport utilizando-se o AT2020 e também os três modelos de smartphone simultaneamente e repetir esse experimento em três distâncias diferentes (1, 2 e 4 metros). O microfone foi fixado em um pedestal próprio com suporte do tipo "shock mount" para evitar que o mesmo captasse as vibrações do solo. Os smartphones foram fixados no próprio pedestal do microfone, de forma a ficarem o mais próximos possível e equidistantes da fonte sonora. Dada a velocidade do som de aproximadamente 344 m/s [15], a diferença de alguns centímetros entre os pontos de captação não exerce influência significativa sobre o experimento. O fato de um dos aparelhos estar mais próximo da parede poderia aumentar a intensidade da reverberação incidindo na captação do áudio mas como a fonte sonora se trata de um golpe que gera uma onda sonora com transientes muito fortes e duração muito curta, o efeito da reverberação do espaço também foi desconsiderado. O local escolhido foi a céu aberto com interferências do som ambiente a fim de simular um ambiente normal de treino.



Figura 4.1 - Pedestal com microfone e smartphones fixados.

Havia também uma fonte sonora externa ao sistema (falante) emitindo um som senoidal de 1,05 kHz (correspondente à nota "dó" da 5<sup>a</sup> oitava). Este som servia apenas como sinal para que o atleta golpeasse o aparador nos momentos certos e para que pudesse ser medida a velocidade de reação do atleta. Essa velocidade de reação foi medida posteriormente e está conectada diretamente com aspectos da gamificação do esporte discutidos em [1].

Foram feitas 2 (duas) medições para cada dispositivo, variando-se a distância entre o aparador e os smartphones e a força do chute (forte ou fraco). Na próxima seção são apresentados os resultados obtidos.



*Figura 4.2 - Espaço onde foram feitas as medições, com marcações no chão indicadas na imagem pelas linhas em verde.*



*Figura 4.3 - Atleta golpeando o aparador sobre a marcação de 2 metros.*

## 4.2 Análise de resultados

### 4.2.1 Diferença do pico de volume entre golpes fortes e fracos captados por um mesmo dispositivo

O primeiro critério a ser avaliado foi a diferença que um golpe fraco e um golpe forte possuem em termos do seu pico de volume. A média dos resultados encontra-se na tabela a seguir:

*Tabela 4.1 - Variação do pico de volume entre golpes fracos e fortes por dispositivo.*

	Distância: 1 metro			Distância: 2 metros			Distância: 4 metros		
	FRACO	FORTE	$\Delta$ (dB)	FRACO	FORTE	$\Delta$ (dB)	FRACO	FORTE	$\Delta$ (dB)
	PICO (dB)	PICO (dB)		PICO (dB)	PICO (dB)		PICO (dB)	PICO (dB)	
AT2020	-5,8	-1,9	3,9	-10,0	-2,4	7,7	-14,1	-6,3	7,8
Apple iPhone 8	-3,5	-3,2	0,3	-4,2	-3,8	0,4	-4,7	-3,3	1,5
Samsung J8	0,4	0,3	-0,1	-0,6	0,3	0,9	0,1	0,4	0,3
Xiaomi Mi 8	0,1	0,0	-0,1	0,2	0,0	-0,2	-0,7	-0,2	0,5

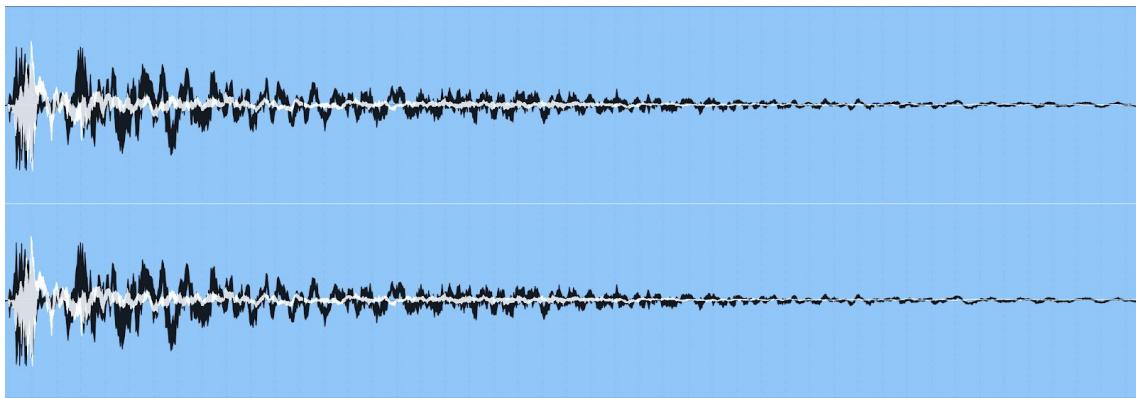
Nessa primeira análise já é possível identificar que a variação do pico de volume nos smartphones é muito baixa em relação à referência do AT2020. Como as gravações foram todas feitas de forma simultânea, foi o mesmo golpe que determinou a referência de variação para cada dispositivo. O experimento não possuía meios de medir a força do golpe, mas dentro da percepção tática humana o golpe fraco era sensivelmente menos potente que o golpe forte.

O áudio captado pelo AT2020 indica que o golpe forte era pelo menos 2 (duas) vezes mais potente que o golpe fraco. Essa primeira análise de volume já indica que os smartphones possuem uma normalização de volume de áudio que deve ocorrer via software.

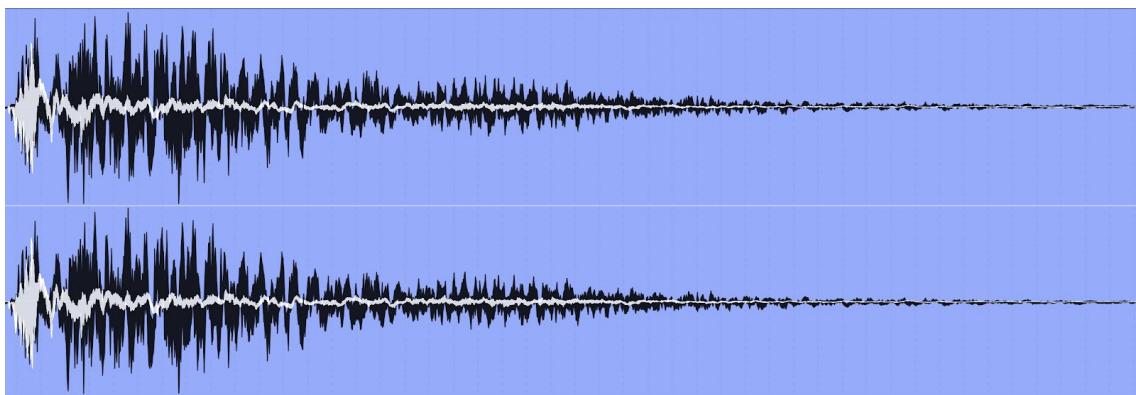
Porém um fato desanimador é que essa variação entre golpes mesmo no caso do AT2020 atingiu extremos que vão de 2,9 dB até 12,4 dB. Seria necessário realizar o experimento inúmeras vezes em vários ambientes diferentes para tentar identificar um padrão.

#### 4.2.2 Comparação do envelope ADSR gravado por cada dispositivo

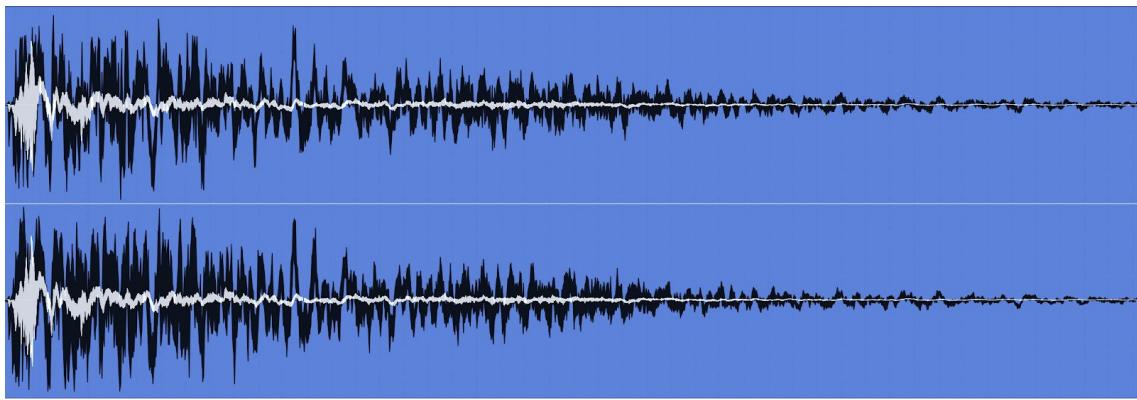
Analizando o envelope ADSR das gravações feitas por cada um dos dispositivos fica ainda mais evidente o processamento interno feito no áudio captado. As figuras a seguir comparam o sinal de áudio captado pelo AT2020 (sinal de referência, em branco) com cada um dos sinais captados pelos smartphones (curva escura no gráfico) vindos do mesmo som, um chute forte aplicado a 4 metros dos instrumentos de captação. As amostras de som possuem duração total de 350 ms e o limitante superior do eixo vertical de cada um dos gráficos é de 0 dB.



*Figura 4.4 - Sinal gravado pelo Apple iPhone 8 (preto) comparado com a referência, amostra com duração de 350 ms.*



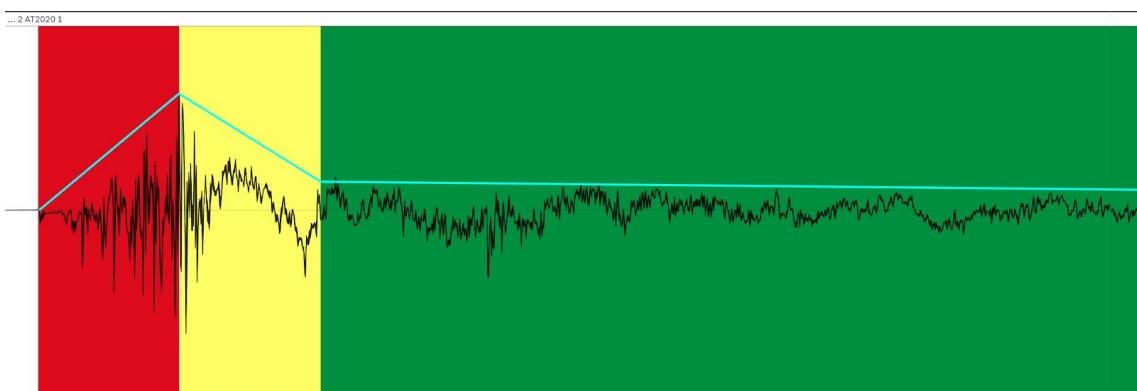
*Figura 4.5 - Sinal gravado pelo Samsung J8 (preto) comparado com a referência, amostra com duração de 350 ms.*



*Figura 4.6 - Sinal gravado pelo Xiaomi Mi 8 (preto) comparado com a referência, amostra com duração de 350 ms.*

Nas figuras acima é possível perceber que o ataque inicial da onda sonora é similar nas quatro gravações, porém, na referência ele está com um nível de volume consideravelmente maior que o restante da própria amostra. Esse comportamento da onda sugere que o impacto do golpe ocorre nesses primeiros 10 ms e que a "cauda" de aproximadamente 340 ms é apenas o amortecimento do sistema e volta para o estado de repouso.

Na Figura 4.7 observa-se em detalhe um trecho de 53 ms da amostra gravada pela referência (AT2020), a região em vermelho constitui o ataque do som, na região em amarelo é possível notar o decaimento da onda sonora, o período de sustentação é infinitesimal pois na região verde já atinge o período de relaxamento do envelope ADSR.



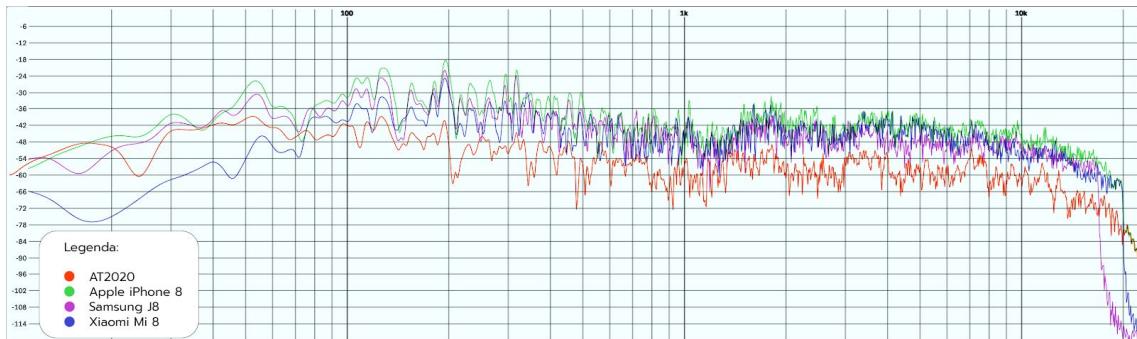
*Figura 4.7 - Detalhamento do envelope ADSR da referência em uma amostra de 53 ms.*

O que ocorre nas ondas captadas pelos smartphones é que elas sofrem uma compressão na dinâmica do sinal que diminui o volume do ataque inicial e nivela ele com o

relaxamento da onda sonora, aplicando um ganho posterior ao sinal após esse processo de normalização. O uso de compressores é largamente difundido na indústria musical [15] e faz sentido ser aplicado na captação de áudio de smartphones pois ao normalizar o volume toda a amostra de som captada fica mais audível o que é desejável em aplicações cotidianas. O compressor funciona como um amortecedor do sinal que diminui o ganho dos picos para que todo o sinal possa receber um ganho global sem que ocorra a clipagem digital, aumentando o volume RMS do áudio previamente comprimido [15] [25] [4]. Esse ganho do sinal não precisa ser necessariamente uniforme ou proporcional entre todas as frequências do espectro, na próxima sessão é possível notar que cada um dos smartphones captou não apenas diferentes envelopes de volume como também diferentes espectros de frequência sobre o mesmo som.

#### 4.2.3 Comparação do espectro de frequência dos áudios captados

Tratando-se do espectro de frequência de cada uma das amostras, não foram observadas diferenças significativas entre elas como pode-se observar na Figura 4.8. De forma sutil é possível perceber que o áudio captado pelos celulares possui um pequeno ganho nas regiões de frequência dos 200 Hz até os 10 kHz, que compreende também a faixa mais audível para o ser humano [15]. Os espectros apresentados a seguir são do mesmo áudio utilizado para analisar os envelopes, um chute forte dado a 4 metros da captação.



*Figura 4.8 - Comparativo entre o espectro de frequência das amostras captadas.*

A correspondência dos espectros da Figura 4.8 é dada a seguir:

- Laranja - Referência (AT2020);
- Verde - Apple iPhone 8;
- Rosa - Samsung J8;
- Azul - Xiaomi Mi 8.

Percebe-se que os smartphones Apple iPhone 8 e Xiaomi Mi 8 possuem filtros passa-baixa com frequência de corte igual a 20 kHz, no Samsung J8 existe também um filtro passa-baixa mas com frequência de corte de 17 kHz. O aparelho da Samsung foi o que apresentou pior resposta em frequência na região dos graves e subgraves.

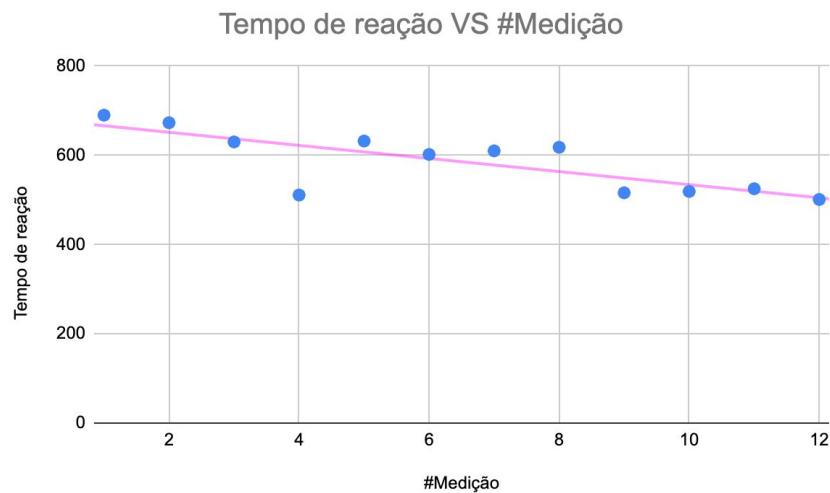
#### 4.2.4 Velocidade de reação do atleta

O último aspecto a ser considerado é o que está mais intimamente relacionado com a gamificação dos esportes de combate abordada em [1]. Durante o experimento o atleta precisava aguardar um sinal sonoro para golpear o aparador. A diferença entre o sinal de aviso e a onda sonora do golpe gravadas por cada dispositivo permitiu ver qual era a velocidade de reação do atleta. Os valores divergiram muito pouco entre os dispositivos, diferenças de até 3 ms sendo que a resolução máxima do Ableton Live é de 1 ms. As médias dos resultados estão listadas na Tabela 4.2.

*Tabela 4.2 - Média do tempo de reação do atleta em cada dispositivo.*

	Distância: 1 metro		Distância: 2 metros		Distância: 4 metros	
	FRACO	FORTE	FRACO	FORTE	FRACO	FORTE
	REAÇÃO (ms)	REAÇÃO (ms)	REAÇÃO (ms)	REAÇÃO (ms)	REAÇÃO (ms)	REAÇÃO (ms)
AT2020	785	585	796	589	857	585
Apple iPhone 8	786	587	797	589	858	585
Samsung J8	787	587	798	590	859	588
Xiaomi Mi 8	786	588	798	589	860	586

Além disso, considerando-se as múltiplas medições é possível notar que o atleta foi melhorando seu tempo de reação ao longo do experimento, como nota-se no Gráfico 4.1 o desempenho do atleta em chutes fortes a 4 metros da captação.



*Gráfico 4.1 - Variação do tempo de reação do atleta em 12 medições.*

## **5. Conclusão**

O projeto comprovou que não é possível medir a força de um golpe em um aparador utilizando o áudio captado por um smartphone dado o processamento que cada aparelho realiza internamente no áudio via software. Os experimentos mostram que cada um dos smartphones manipula tanto o envelope de amplitude quanto o espectro de frequências das amostras captadas de uma mesma fonte sonora. A hipótese talvez poderia ser validada se os fabricantes dos smartphones disponibilizassem informações sobre o processamento realizado e permitissem o desenvolvimento de softwares e aplicativos que trabalhassem apenas com o áudio puro captado pelo microfone interno do aparelho.

No entanto foi observada uma melhoria da velocidade de reação do atleta em um espaço amostral pequeno (Gráfico 4.1) mas que abre a possibilidade de utilizar os aparelhos para desenvolver esse fundamento da velocidade do atleta através de aplicativos que medissem essa reação. Isso poderia permitir que atletas em diferentes regiões geográficas treinassem juntos e competindo entre si.

A idéia de utilizar os smartphones para desenvolver a prática de esportes de combate continua sendo uma possibilidade pois estes aparelhos possuem poderosos sistemas embarcados e estão totalmente presentes no cotidiano do povo brasileiro.

Este projeto permitiu que o aluno desenvolvesse seu conhecimento sobre o método científico como a determinação de uma hipótese, o levantamento da fundamentação teórica e a realização de experimentos para validá-la.

Também contribuiu na compreensão aprofundada de conceitos de Sinais e Sistemas, Processamento Digital de Sinais e principalmente conceitos da engenharia de áudio como síntese sonora, história dos sintetizadores analógicos, história da Ableton, história e princípio de funcionamento de vários dos tipos de microfones existentes, princípios do áudio digital aplicados na produção musical e como eles se aplicam dentro do Ableton Live.

Após realizar um trabalho envolvendo aspectos da engenharia, o aluno também conclui este projeto tendo evoluído como produtor musical e artista.

### **5.1 Considerações sobre o Curso de Graduação**

A graduação em Engenharia Elétrica oferecida pela USP São Carlos é um curso que aprofunda cada um dos alicerces da física, da matemática e das ciências em geral que compõe o futuro profissional formado. Oferece conteúdo tanto teórico quanto prático, além de disciplinas que complementam os conhecimentos de gestão, planejamento e ciências humanas dos alunos. É um curso de formação técnica mas que desenvolve um grande senso de proatividade no processo de aprendizagem do aluno e futuro engenheiro, permitindo que o mesmo atue em várias áreas do conhecimento não apenas dentro de sua formação acadêmica.

# Referências bibliográficas

- [1] Cabral, M.P. Gamificação para o ensino de Tae-Kwon-Do; EESC-USP, 2020.
- [2] Zawawi, S.A.; Hamzah, A.A.; Majlis, B.Y.; A Review of MEMS Capacitive Microphones; Universiti Putra Malaysia, 2020. Disponível em:  
<[https://www.researchgate.net/publication/341301603\\_A\\_Review\\_of\\_MEMS\\_Capacitive\\_Microphones](https://www.researchgate.net/publication/341301603_A_Review_of_MEMS_Capacitive_Microphones)>. Acesso em 25 de novembro de 2020.
- [3] Dixon-Warren, S. Overview of MEMS Microphone Technologies for Consumer Applications; Mems Journal Inc. Southfield, MI, USA, 2011. Disponível em:  
<<https://www.memsjournal.com/2011/03/overview-of-mems-microphone-technologies-for-consumer-applications.html>>. Acesso em 25 de novembro de 2020.
- [4] Ableton Reference Manual Version 10. Disponível em:  
<[https://cdn-resources.ableton.com/resources/0b/c1/0bc1007e-bd0b-4d6f-b52d-ee0054f3a6f8/110manual\\_en.pdf](https://cdn-resources.ableton.com/resources/0b/c1/0bc1007e-bd0b-4d6f-b52d-ee0054f3a6f8/110manual_en.pdf)>. Acesso em 25 de novembro de 2020.
- [5] Colete LED SulSport eletrônico para treinamento em academia. Disponível em:  
<<https://www.sulsport.com.br/produtos/colete-led-sulsport-colete-eletronico-para-treinamento-e-m-academia/?variant=43601214>>. Acesso em 25 de novembro de 2020.
- [6] Haykin, S. Sinais e Sistemas.
- [7] Oppenheim, A. V.; Willsky, A. S.; Hamid, S.; Nawab, S. H. – Sinais e Sistemas Capítulo 7 - Amostragem.
- [8] Profundidade de bit de áudio. Disponível em:  
<[https://pt.wikipedia.org/wiki/Profundidade\\_de\\_bit\\_de\\_%C3%A1udio](https://pt.wikipedia.org/wiki/Profundidade_de_bit_de_%C3%A1udio)>. Acesso em 25 de novembro de 2020.
- [9] Halliday; Resnick; Walker. Fundamentals of Physics - 7th edition Chapter 17 - Waves - II.
- [10] Portal FGV - Brasil tem 424 milhões de dispositivos digitais em uso, revela a 31ª Pesquisa Anual do FGVCia. Disponível em:  
<<https://portal.fgv.br/noticias/brasil-tem-424-milhoes-dispositivos-digitais-uso-revela-31a-pesquisa-anual-fgvcia>>. Acesso em 25 de novembro de 2020.
- [11] MEMS (Micro Electro Mechanical Systems). Disponível em:  
<<https://www.embarcados.com.br/mems/>>. Acesso em 25 de novembro de 2020.
- [12] Datasheet microfone condensador cardióide fabricante Audio-Technica modelo AT2020. Disponível em: <[https://www.audio-technica.com.hk/templates/index/file/106\\_1\\_AT2020.pdf](https://www.audio-technica.com.hk/templates/index/file/106_1_AT2020.pdf)>. Acesso em 25 de novembro de 2020.

- [13] Focusrite iTrack Solo user guide. Disponível em:  
<<https://fael-downloads-prod.focusrite.com/customer/prod/s3fs-public/focusrite/downloads/8183/ittrack-solo-user-guide-03.pdf>>. Acesso em 25 de novembro de 2020.
- [14] Página do produto: Focusrite itrack Solo. Disponível em:  
<<https://focusrite.com/en/ios-audio-interface/ittrack/ittrack-solo>>. Acesso em 25 de novembro de 2020.
- [15] Hass, J. Introduction to Computer Music. Indiana University, 2019. Disponível em:  
<<https://cmtext.indiana.edu/toc.php>>. Acesso em 25 de novembro de 2020.
- [16] Portal FIFA - VAR at the 2018 FIFA World Cup. Disponível em:  
<<https://football-technology.fifa.com/en/innovations/VAR-at-the-World-Cup/>>. Acesso em 25 de novembro de 2020.
- [17] System Plus Consulting - iPhone X MEMS Microphones sample. Disponível em:  
<[https://www.systemplus.fr/wp-content/uploads/2018/03/SP18384-MEMS-iPhone-X-MEMS-Microphones\\_sample.pdf](https://www.systemplus.fr/wp-content/uploads/2018/03/SP18384-MEMS-iPhone-X-MEMS-Microphones_sample.pdf)>. Acesso em 25 de novembro de 2020.
- [18] Página do produto: Logic Pro 10.6. Disponível em: <<https://www.apple.com/br/logic-pro/>>. Acesso em 25 de novembro de 2020.
- [19] Cubase Pro 11 Operation Manual. Disponível em:  
<[https://steinberg.help/cubase\\_pro/v11/en/Cubase\\_Pro\\_11\\_Operation\\_Manual\\_en.pdf](https://steinberg.help/cubase_pro/v11/en/Cubase_Pro_11_Operation_Manual_en.pdf)>. Acesso em 25 de novembro de 2020.
- [20] Página do produto: Pro Tools. Disponível em: <<https://www.avid.com/pro-tools>>. Acesso em 25 de novembro de 2020.
- [21] About Ableton. Disponível em: <<https://www.ableton.com/en/about/>>. Acesso em 25 de novembro de 2020.
- [22] Sweetwater - The History of SSL. Disponível em:  
<<https://www.sweetwater.com/insync/the-history-of-ssl/>>. Acesso em 25 de novembro de 2020.
- [23] Wikipedia - Envelope (music). Disponível em:  
<[https://en.wikipedia.org/wiki/Envelope\\_\(music\)](https://en.wikipedia.org/wiki/Envelope_(music))>. Acesso em 25 de novembro de 2020.
- [24] Wikipedia - Bass Drum. Disponível em: <[https://en.wikipedia.org/wiki/Bass\\_drum](https://en.wikipedia.org/wiki/Bass_drum)>. Acesso em 25 de novembro de 2020.
- [25] Wikipedia - Clipping (audio). Disponível em  
<[https://en.wikipedia.org/wiki/Clipping\\_\(audio\)](https://en.wikipedia.org/wiki/Clipping_(audio))>. Acesso em 25 de novembro de 2020.