


Estudo de Timbres e Envelopes Sonoros



Por: Gustavo Henrique Siviero Soares.

Sob a orientação de: Prof. Gonzalo Travieso.

Agradecimentos ao Prof. Luciano da Fontoura Costa.

Trabalho de Conclusão de Curso – Monografia apresentada ao curso de Bacharelado em Física Computacional no Instituto de Física de São Carlos da Universidade de São Paulo.

São Carlos. Dezembro de 2021



Resumo.

Neste trabalho foi desenvolvido um estudo dos formatos de sinais sonoros, chamados envelopes, e mostramos as relações destes com o timbre e como modificam os espectros de frequências de um sinal. Em particular, como os envelopes adicionam novas frequências ao som ouvido, tornando um som sintético mais rico e semelhante aos sons naturais. Também buscamos compreender a natureza de certas sutilezas sonoras que enriquecem o som ouvido, denominadas dicções. Mostramos como adicionar certos efeitos de dicções simuladas via envelope, e como alguns envelopes naturais, obtidos de sons gravados, podem conter determinados efeitos de dicção. Discutimos os limites de descrição de um som gravado e suas relações com as informações de espectro, envelope e eventualmente a presença de dicções. Permitindo-se assim uma análise dos efeitos do componente envoltória nos sinais, sob o ponto de vista de estrutura espectral, dinâmica de volume e outras sutilezas apresentadas pelos sons, tendo a finalidade de compreender melhor a natureza de cada elemento sonoro, podendo auxiliar no ganho de qualidade de sons sintéticos ou mesmo da melhor compreensão dos fenômenos acústicos relacionados.

Palavras-chave: (1) Análise de Som. (2) Processamento digital. (3) Reconhecimento de padrões. (4) Modelagem Computacional

“A música é um exercício velado de aritmética sem que o espírito saiba que está lidando com números.”

Gottfried Wilhelm von Leibniz (1646-1716)

1. Introdução

Este excerto de Leibniz pode inspirar a forma como a música está intimamente relacionada a matemática e as concepções físicas modernas que utilizam desta para descrever o mundo. A frase do filósofo e grande contribuidor do cálculo diferencial e integral está em completa consonância com as descobertas de Pitágoras, que no século VI A.C. realizou experimentos com um monocórdio e observou que o som produzido por um estímulo na corda tensionada tem uma relação de proporcionalidade com o seu comprimento. Durante todo o período clássico a música foi firmada como parte substancial da matemática, assim como a geometria e aritmética. Na idade média esta disciplina compunha o *quadrivium*, etapa mais sofisticada do *curriculum* medieval, juntamente com aritmética, geometria e astronomia, todas disciplinas contemporaneamente associadas às ciências exatas.

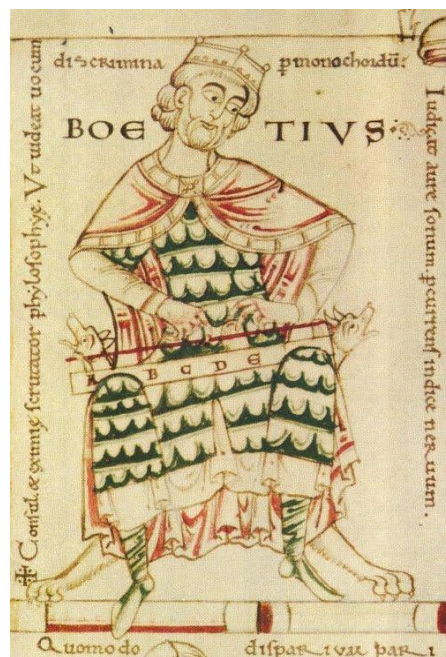


Imagem 1: Ilustração mostrando Boécio, filósofo romano autor de *De Institutione Musica*, com um monocórdio. Fonte: Boethius: *De arithmetica, De musica*. Early 12th century. Cambridge University Library, MS li.3.12, fol. 61v.

Sabemos que a proporção observada por Pitágoras tem a seguinte relação física: A frequência (f), propriedade responsável pela nota do som que ouvimos, varia inversamente com o comprimento da corda (l). Como sugere a fórmula:

$$v = \lambda \cdot f \rightarrow f = \frac{v}{\lambda}; \text{ onde } v = \sqrt{\frac{T}{\mu}} = cte$$
$$\therefore f = \frac{1}{2l} \sqrt{\frac{T}{\mu}}$$

No monocórdio a velocidade de propagação da onda na corda (v) está relacionada à tensão aplicada (T) e à densidade linear da própria corda (μ), características que permanecem constantes ao longo do experimento e pretendidamente também devem permanecer constantes em um

instrumento de cordas durante a execução de uma música. A característica modificável pelo experimentador, no monocórdio, e pelo músico, no instrumento de cordas é somente o comprimento desta, comumente prendendo-a com os dedos em um comprimento correspondente a uma determinada nota musical.

O modo fundamental é a vibração de uma corda quando só existem dois nodos fixos nas extremidades e nenhum ao longo do comprimento, ou seja, o comprimento todo da corda forma uma barriga, sendo assim metade de um comprimento de onda (λ). Quando uma corda é tocada ela vibra em uma frequência associada ao modo fundamental que confere a tonalidade do som ouvido. Portanto, se calcularmos a frequência de vibração de uma corda (onde a velocidade é conhecida), utilizando como comprimento de onda duas vezes o comprimento da corda, ou seja o modo fundamental, estaremos calculando a frequência fundamental da nota ouvida.

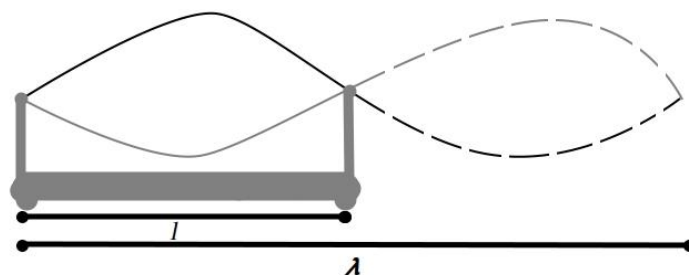


Imagem 2: ilustrando o modo fundamental e o comprimento da corda

No entanto, se produzirmos um sinal sonoro através de uma função senoidal na mesma frequência fundamental calculada de uma corda e compararmos com um som gravado da própria corda tocada, mesmo notando que ambos são a mesma nota, perceberemos também que são muito diferentes entre si. O som produzido seria limpo, constante e passaria a sensação de artificialidade, seria até mesmo desinteressante ou simplório. Enquanto que o som original proveniente da corda tocada seria perceptivelmente mais rico, discernindo as características do instrumento tocado e com uma dinâmica própria que dificilmente seria constante, e contaria com variações e decaimentos da amplitude ao longo do soar. Pode-se dizer que comumente os ouvidos humanos têm grande preferência por este segundo som.

Quais seriam então as diferenças fundamentais entre estes dois sons? O primeiro é apenas uma função seno pura, de amplitude constante enquanto dure. O segundo som, porém, tem bem mais frequências que o compõem do que somente da frequência fundamental. Podemos entender que a corda não vibra somente no modo fundamental, mas também em várias outras frequências múltiplas da frequência fundamental, frequências que chamamos de harmônicos. Desta forma o som

ouvido é uma composição da frequência fundamental com os seus harmônicos em diferentes proporções. Esta composição é responsável pelo timbre característico de cada instrumento, que segundo seu mecanismo de funcionamento, estrutura, materiais, geometria e também a maneira como se aplica a excitação, forma diferentes composições de harmônicos.

Então a maioria dos sons são uma mistura de diversas frequências, mas que também dispõem de outras características relacionadas à dinâmica de como este evolui ao longo do tempo e outras sutilezas. Conseguimos auditivamente perceber estalos, ataques, decaimentos, batimentos e diversos outros efeitos sonoros. Veremos que estudar como diferentes elementos físicos contribuem para formação destes efeitos possibilita uma compreensão da natureza física de cada elemento dos sons e sobretudo das relações que estes elementos guardam entre si.

2. Conceitos e Metodologia

Chamamos **espectro** a composição de frequências de um sinal com suas devidas proporções. Pode ser determinado e visualizado através do método da transformação de Fourier.

Convencionamos utilizar o módulo da função transformada para determinar as frequências e proporções. **Timbre** é a característica que discerne auditivamente os sons e possibilita reconhecer diferentes instrumentos musicais, o som de um automóvel ou a voz de uma pessoa. Fisicamente ele compreende o conjunto dos espectros que cada um destes referentes podem produzir. **Envelope ou envoltória** é a forma geral do sinal desconsiderando a oscilação, a função que expressa amplitude do sinal ao longo do tempo.

Definimos o termo **dicção** para referirmos às sutilezas sonoras intrínsecas da interação com o instrumento, pequenos ruídos e estalos que não são propriamente característica do timbre, mas acidente de um determinado sinal gravado.

2.1 WAV, Python e Métodos de Programação

WAV, abreviação de *Wave Form Audio Format* (Formato de áudio em forma de onda) é uma espécie de arquivo de áudio que guarda um vetor representando uma função discreta de acordo com uma determinada taxa de amostragem. Portanto, se intencionamos lidar matematicamente com um sinal numérico este é um formato adequado para armazenamento de som. Outra vantagem é a facilidade de leitura e manipulação deste tipo de formato utilizando softwares ou bibliotecas da linguagem Python.

Para análise e processamento de sinais utilizamos de métodos matemáticos que permitem identificar características, isolar, visualizar ou operar de diversas maneiras um sinal de áudio. A linguagem Python possibilita facilidades de implementação se recorremos a bibliotecas com rotinas muito robustas para a execução destes métodos, e também uma sintaxe muito simples utilizando da orientação ao objeto. A utilização de bibliotecas para a execução dos métodos faz-se necessária devido à grande quantidade de dados que é preciso processar para a realização dos experimentos. Uma única gravação de poucos segundos em uma taxa de amostragem de 44100 quadros por segundo tem uma quantidade expressiva de dados que poderia comprometer o tempo de execução de um programa implementado em python, já que esta é uma linguagem interpretada e não dispõe de mecanismos de otimização de alocação de memória. Recorrendo à biblioteca numpy, por exemplo, as operações são realizadas em rotinas otimizadas implementadas em C++, tornando a execução bem mais eficiente.

Parte importante deste trabalho foi desenvolver o código que proporciona um arcabouço de métodos aplicáveis aos objetos estudados e a gravação, pelo autor, de uma base de dados que auxiliaria nos experimentos. Intenciona-se que este código possa ser utilizado para os experimentos aqui presentes, mas que também seja adequado para diversos experimentos com envelopes e timbres, podendo ser usado por terceiros para além dos experimentos aqui presentes. A seguir serão apresentados alguns dos métodos mais importantes para processamento de sinais e sobre os quais o código está baseado.

2.2 Transformação de Fourier (\mathcal{F})

No estudo de fenômenos periódicos a frequência é uma componente essencial de um sinal oscilatório, é o valor que descreve a periodicidade da oscilação. Uma função senoidal constante tem uma determinada frequência intrínseca. Suponhamos compor um sinal com duas funções senoidais de frequências diferentes, desta forma também obteremos um sinal periódico, porém com dois valores que descrevem sua periodicidade. Se a partir deste sinal quisermos descobrir quais são estes valores, precisaremos de um método capaz de decompor este sinal repetitivo obtendo as características desta periodicidade, ou seja as frequências que o compõe.

A transformação de Fourier é um método matemático bastante conhecido e amplamente utilizado no processamento de sinais. O procedimento designa uma integração da função temporal (sinal) e permite obter, através de análise complexa, o espectro das frequências que compõem este

sinal, no espaço amplitude por frequência. Tal que o valor do módulo da função transformada significa as proporções de cada frequência na composição do sinal original.

A definição matemática da transformada (\mathcal{F}) requer a integração, em todo o seu domínio, do sinal $s(t)$ com relação ao tempo.

$$\mathcal{F}\{s(t)\} = \int_{-\infty}^{\infty} s(t) \cdot e^{-i\omega t} dt$$

Porém como o sinal digital está na forma discreta, utilizamos métodos de cálculo numérico implementados computacionalmente para obter os pontos transformados para o espaço das frequências. (Brigham. The Fast Fourier Transform). Para otimização do processamento utilizamos a rotina `fft` (fast fourier transform) da biblioteca numérica padrão da linguagem python: `numpy`, que opera de acordo com a integração numérica na forma da série:

$$\mathcal{F}\{s(t)\} = \sum_{n=0}^N s(t) \cdot e^{-\frac{2\pi i}{N} n}$$

2.3 Envelope e Transformação de Hilbert (\mathcal{H})

Outro objeto de grande importância para este estudo de sinais é o envelope. Podemos dizê-lo em termos de sua significação física: a função da amplitude com relação ao tempo. Sabemos que há uma equivalência na percepção humana entre a amplitude da onda e a variação da intensidade do som ouvido. Sinais obtidos de gravações de instrumentos musicais costumam apresentar envelopes com características muito particulares. Para obter a função capaz de expressar as informações que compõem o envelope de um som gravado utilizamos o método da transformação de Hilbert.

$$\mathcal{H}\{s(t)\} = h(t) * f(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(u)}{u - t} dt$$

Quando aplicamos a transformação de Hilbert temos uma representação da função transformada como o produto de duas funções conjugadas, onde uma contém as altas e outra as baixas frequências que compõem a função original. O que permite utilizar este método como ferramenta para demodular um sinal, obtendo seu envelope.

Na forma analítica a transformação de Hilbert pode ser obtida utilizando um processo sobre a sobre a transformação de Fourier. (Costa. CDT-40 Instantaneous Signal Analysis. p.7). Convenientemente podemos tomar apenas a parte que nos interessa das frequências de um sinal,

convoluindo o espectro com uma função que anula as frequências indesejáveis, depois tomamos a transformação de Fourier inversa. Criamos um sinal analítico através deste novo mecanismo. A amplitude instantânea é obtida tomando o módulo deste sinal analítico.

$$s_a(t) = \mathcal{F}_{inv}[\mathcal{F}(s(t)) \cdot g(f)] : g(f) = \begin{cases} 0, & f < 0 \\ 1, & f = 0 \\ 2, & f > 0 \end{cases}$$

A função analítica obtida, que representa a envoltória, não é perfeita. Ela ainda carrega muitas oscilações indesejáveis. Podemos contornar este problema usando um filtro passa-baixa aplicado à esta função. Claro que adicionamos um componente subjetivo: É preciso definir o quanto queremos filtrar o envelope obtido. Desta forma faz-se necessária uma espécie de regulação manual na obtenção de um envelope. Para cada envelope que quisermos obter precisamos visualizar o resultado e ajustar o parâmetro do filtro. Como o filtro suaviza a forma da envoltória, se filtrarmos demais arriscamos perder informações relevantes.

```
sd = sound(database.dataset('viola_caipira')['F#3'])
env = envelope().get_hilbert(sd).lowpass(0.002)
env.plot(size=(20,6))
```

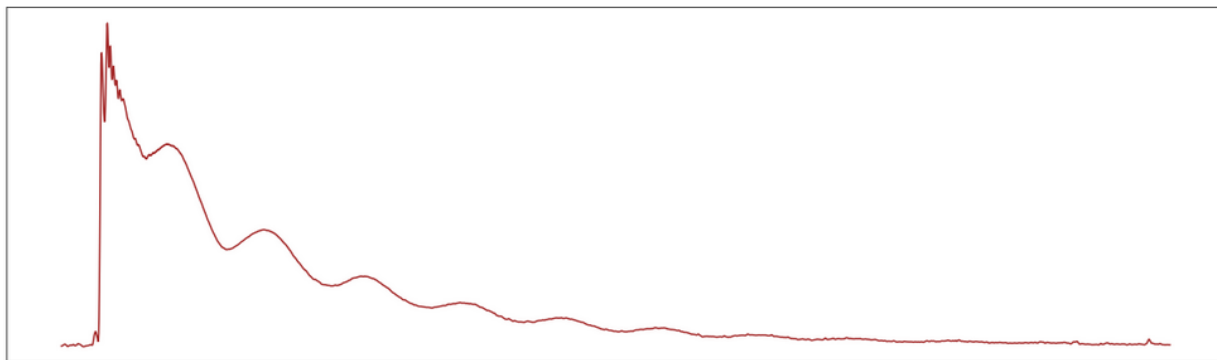


Imagem 3: Ilustrando a obtenção do envelope e filtro ajustado.

2.4 Predicados do Envelope

As dinâmicas da intensidade de um pulso sonoro em um instrumento musical, presentes no envelope, costumam apresentar um ciclo básico composto de quatro etapas. De acordo com estas etapas divide-se as partes temporais do envelope em **Ataque**, **Decaimento**, **Sustentação** e **Relaxamento**, que constituem a modelagem de uma forma geral dos envelopes sonoros chamados de envelopes **ADSR**, sigla que referencia estas etapas.

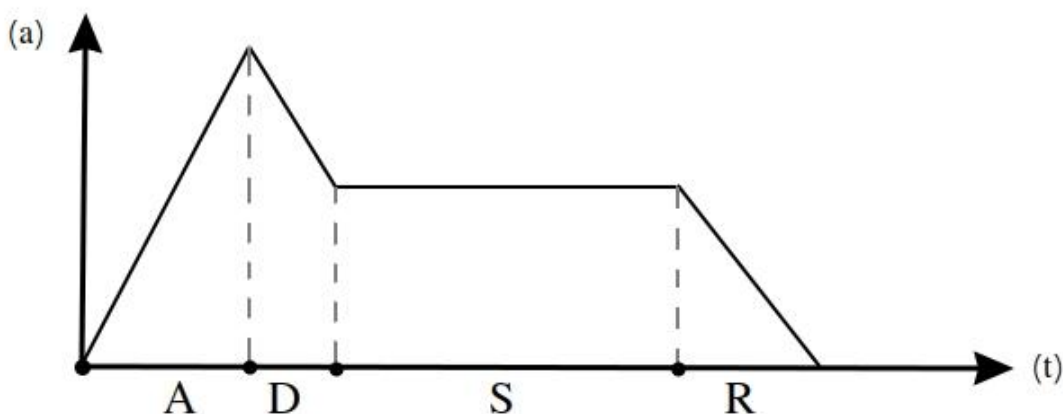


Imagem 4: Ilustração de um envelope ADSR típico.

A primeira etapa, chamada ataque, designa o tempo que o som leva para atingir sua intensidade máxima, desde o silêncio, quando uma corda é pinçada, por exemplo. É comum que após o ataque o som sofra uma queda de intensidade, o que designa a segunda etapa, chamada decaimento. Em seguida a sustentação é o período de soar do som, aqui representada simbolicamente como uma região de amplitude constante, porém é mais natural que haja algum decaimento. Trata-se na verdade da etapa de menor variação da intensidade. Por último, o relaxamento caracteriza o final, quando a intensidade diminui até o som cessar. (Costa. CDT-42 Single note, single frequency)

Estas etapas podem ter tempos de duração e variações de amplitudes de diversos valores diferentes de acordo com o emissor do som. Instrumentos de percussão comumente tem um tempo de ataque muito rápido, um decaimento bastante evidente, praticamente nenhuma sustentação e um período de relaxamento muito dependente da reverberação do lugar onde o som foi gravado. Em contrapartida, instrumentos de sopro ou de arco podem apresentar um tempo de ataque muito longo, nenhum decaimento, sustentação com a duração que pretende o executor e relaxamento a depender de como o próprio executor cessa o som. Os instrumentos de cordas e o piano são os que têm o ciclo com as etapas mais bem definidas e se aproximam mais das proporções da ilustração.

3. O Programa e a Base de Dados

O código do programa implementado, juntamente com a base de dados gravada e manual de uso, encontra-se disponível na plataforma GitHub, lançado sob domínio público.

Link de acesso: [Inserir o link](#)

A base de dados contém sons de diferentes notas musicais provenientes de diversos instrumentos e foram organizados pelo autor em uma estrutura de arquivos em formato WAV. Cada arquivo é um pulso sonoro com duração padronizada de 7 segundos, todos foram gravados com um microfone condensador cardióide a uma taxa de amostragem de 44100 quadros por segundo e normalizados para amplitude de pico de -1 dB. Cada som pode ser acessado facilmente através do código implementado.

3.1 Classes e Funcionamento

O programa possibilita a criação de objetos das seguintes classes:

sound: Objetos desta classe são os pulsos de áudio. Utilizando de seus métodos é possível tocar o som, visualizar o sinal e visualizar o espectro de frequências através da transformação de Fourier.

envelope: Obter o envelope de um som através da transformação de Hilbert, criar um envelope geométrico, adicionar um decaimento exponencial e aplicar um filtro passa-baixa ou passa-alta neste envelope.

signal: Utilizando esta classe é possível criar um sinal estabelecendo sua frequência principal, adicionar novas frequências ao sinal, também visualizar o sinal e seu espectro de frequências, envelopar o sinal passando um objeto do tipo envelope, tocar o som produzido e salvar o sinal em um arquivo WAV.

spectrum: Podemos obter os valores das frequências mais presentes no espectro ordenados pela proporção na composição do sinal.

E a classe própria da base de dados **pulses_database** que possibilita acessar facilmente os sons da base criando objetos do tipo sound.

```
database = pulses_database()
database.show()

Database Available Instruments and Their Respective Keys

violin_pulse: ['G3', 'D4', 'E5', 'A4']
drums: ['snare', 'ground', 'crash16', 'open_hihat', 'crash_bulbe', 'closed_hihat', 'tom']
steel_guitar: ['B3', 'D3', 'E4', 'A2', 'G3', 'E2']
electric_guitar: ['D3', 'B3', 'E2', 'A2', 'E4', 'G3']
nylon_guitar: ['D3', 'A2', 'E4', 'E2', 'G3', 'B3']
distortion_guitar: ['G3', 'E2', 'D3', 'B3', 'E4', 'A2']
harmonica: ['C6', 'E5', 'E4', 'C5', 'C4', 'Cmajor', 'G4']
viola_caipira: ['A4', 'F#3', 'A2', 'D4', 'D3']
bass: ['E1', 'G2', 'D2', 'A1']
violin_pick: ['A4', 'D4', 'G3', 'E5']
```

Imagem 5: Sons disponíveis na base de dados

4. Experimentos

Buscamos compreender de que modo o envelope interfere nas características do som. A resposta mais imediata e dedutível através do seu significado físico é: Criando uma dinâmica variável na intensidade do som ao longo de sua evolução no tempo. É realmente perceptível que este efeito é expresso no formato da onda de maneira condizente com o som que ouvimos. Porém, realizando os experimentos aqui presentes, utilizando os métodos descritos, podemos averiguar a completude desta resposta e aferir se existem outros elementos que são modificados pelo envelope.

4.1 Contribuições do Envelope no Espectro do Som

Primariamente criamos um sinal de amplitude constante, que oscila na frequência de 440 Hz, correspondente à nota Lá da quarta oitava. Vemos que o espectro deste sinal mostra uma linha vertical sobre a marca de 440 Hz no eixo da frequência e tem o valor 0 em todos os outros pontos, afinal este sinal é composto unicamente pela frequência citada.

```
sig = signal().oscillation(440, 4)
sig.plot(size=(20,2.5))
sig.plotfft(size=(20,2.5), xzoom=800, yzoom=4000)
```

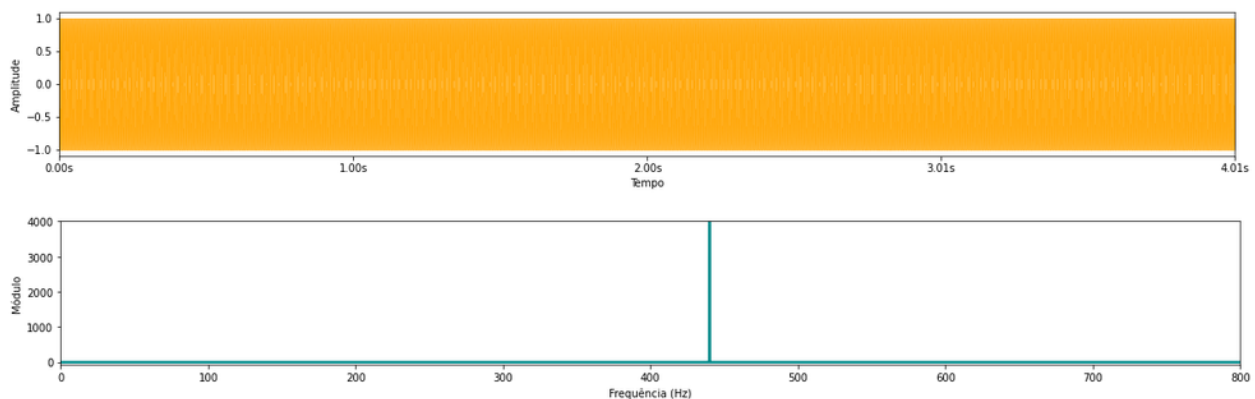


Imagem 6: Sinal de 440 Hz constante e sua transformação de Fourier.

Link para ouvir o sinal seno puro: [Inserir Link](#)

Porém, se aplicarmos um envoltório que faz decair a amplitude desta oscilação na forma exponencial, vemos que o espectro de frequências muda. A linha vertical se abre no formato de uma gaussiana centralizada na mesma frequência, é conhecido que a função obtida da transformação de Fourier de uma função exponencial declinante é uma função gaussiana (Wikibooks. Engineering Tables/Fourier Transform Table 2, signal 14). Isto significa que o sinal agora contém um conjunto de frequências em pequenas proporções ao redor da frequência principal. A abertura desta gama de frequências é proporcional ao decaimento da função exponencial na razão inversa: quanto mais

rápido este decaimento, maior a abertura. De fato, podemos notar auditivamente que este novo sinal é um som mais rico do que a oscilação pura de amplitude constante.

```
sig.exp_decay(2.0)
sig.plot(size=(20,2.5))
sig.plotfft(size=(20,2.5), xzoom=800)
```

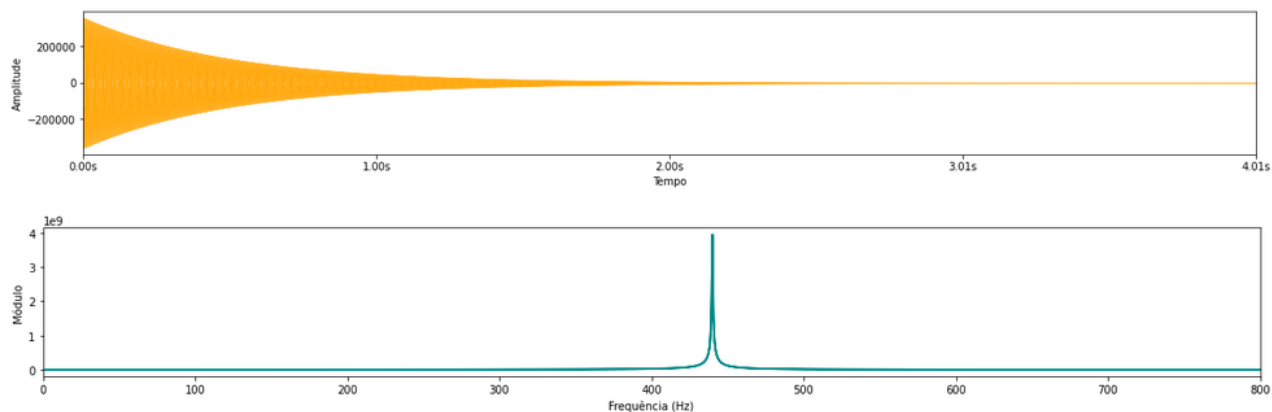


Imagem 7: Espectro do sinal com envelope de decaimento exponencial.
Link para ouvir o sinal com envelope exponencial declinante: [Inserir Link](#)

Se fizermos o mesmo procedimento, porém agora aplicando um envelope geométrico do tipo ADSR, notamos que também existe uma abertura no espectro de frequências. Ainda que esta nova abertura não seja propriamente uma gaussiana, tem um formato substancialmente parecido.

```
sig = signal().oscillation(440, 4)
env = envelope().geometric([(0,0),(0.2,0),(0.2001, 1),(0.35,0.4),(2,0.2),(3,0.01),(4,0)], 4)
sig.envelopment(env)
sig.plot(size=(20,3))
sig.plotfft(size=(20,4), xzoom=800, yzoom=4000)
```

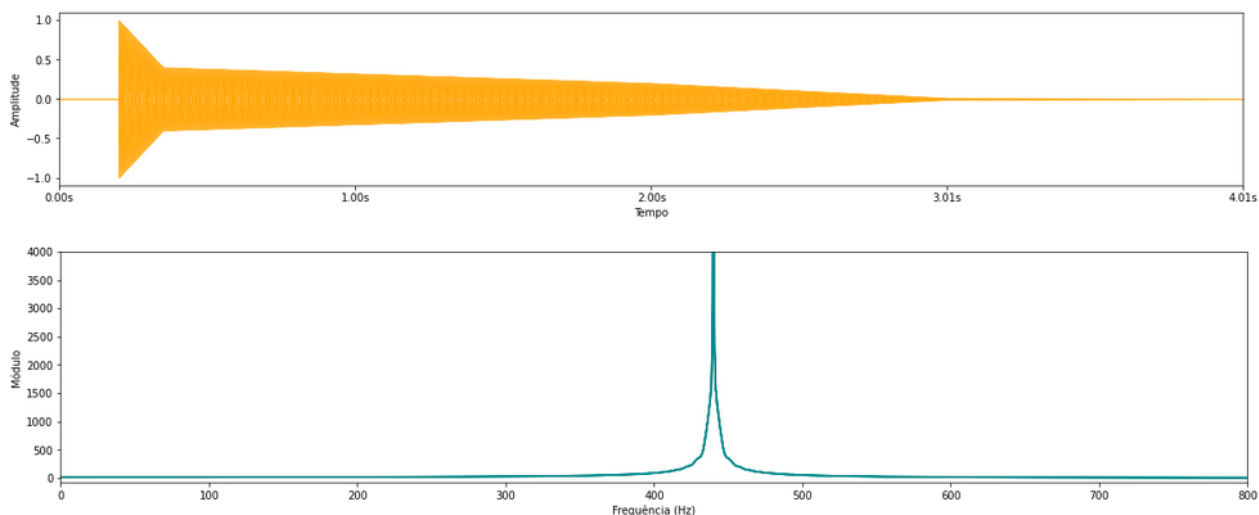


Imagem 8: Envelope geométrico ADSR sobre uma oscilação pura e o espectro do sinal

Note que o tempo de ataque deste envelope é extremamente rápido, pelo código é possível ver que esta etapa tem uma duração de apenas 0.0001 segundo. Se aumentássemos este tempo de

ataque não haveria nenhuma grande diferença visível na maneira como abre o espectro. Porém auditivamente é possível notar que quando o tempo de ataque torna-se muito pequeno percebemos que o som deixa de apresentar uma suavidade no início e surge uma dicção parecida com um estalo de quando uma corda é pinçada. Note a diferença do som com tempo de ataque de 0.05 segundo e com tempo de ataque de 0.0001 segundo.

Ataque 0.05 s: [Inserir link 0.05](#)

Ataque 0.0001 s: [Inserir link 0.0001](#)

Aferimos que, apesar do tempo de ataque rápido em um envelope ADSR adicionar uma característica notável ao som ouvido, esta dicção apresentada não causa uma deformação expressiva no espectro de frequências do sinal. Demonstrando que esta dicção tem uma presença extremamente sutil no espectro de frequências.

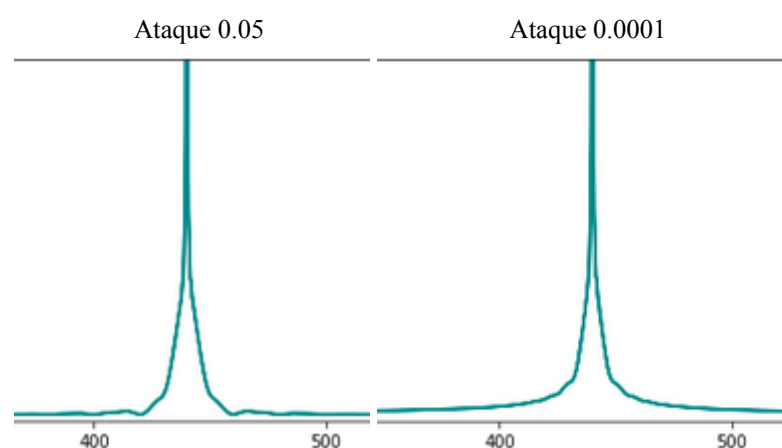
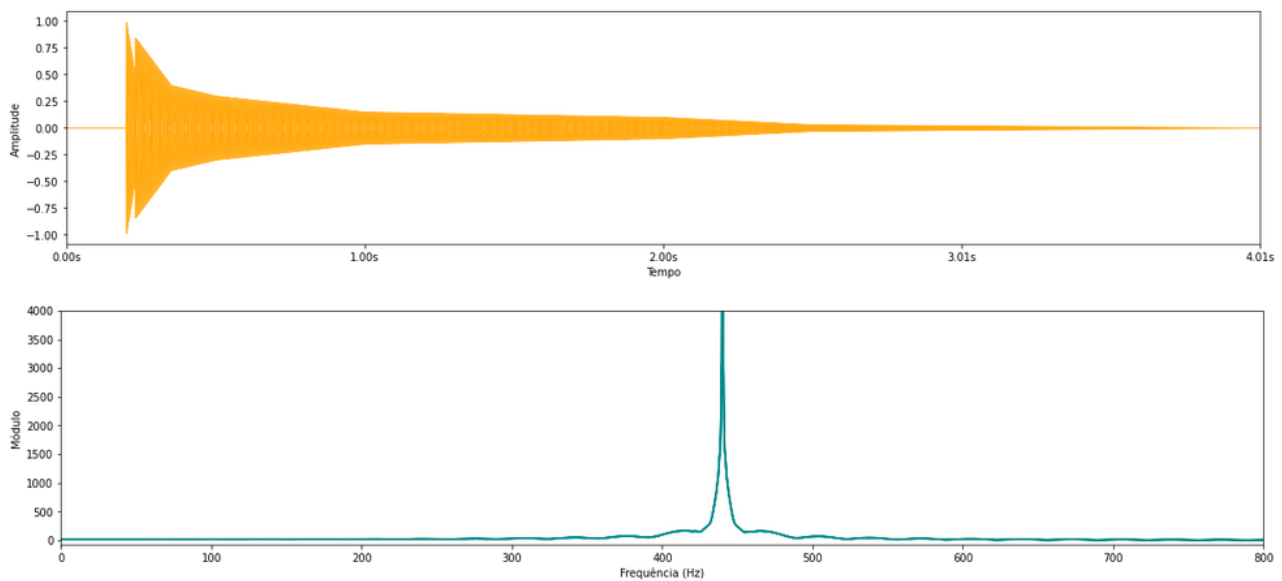


Imagem 9: Comparação dos espectros dos pulsos de tempo de ataque diferentes

4.2 Etapas de Ataque-decaimento e Consequências no Espectro.

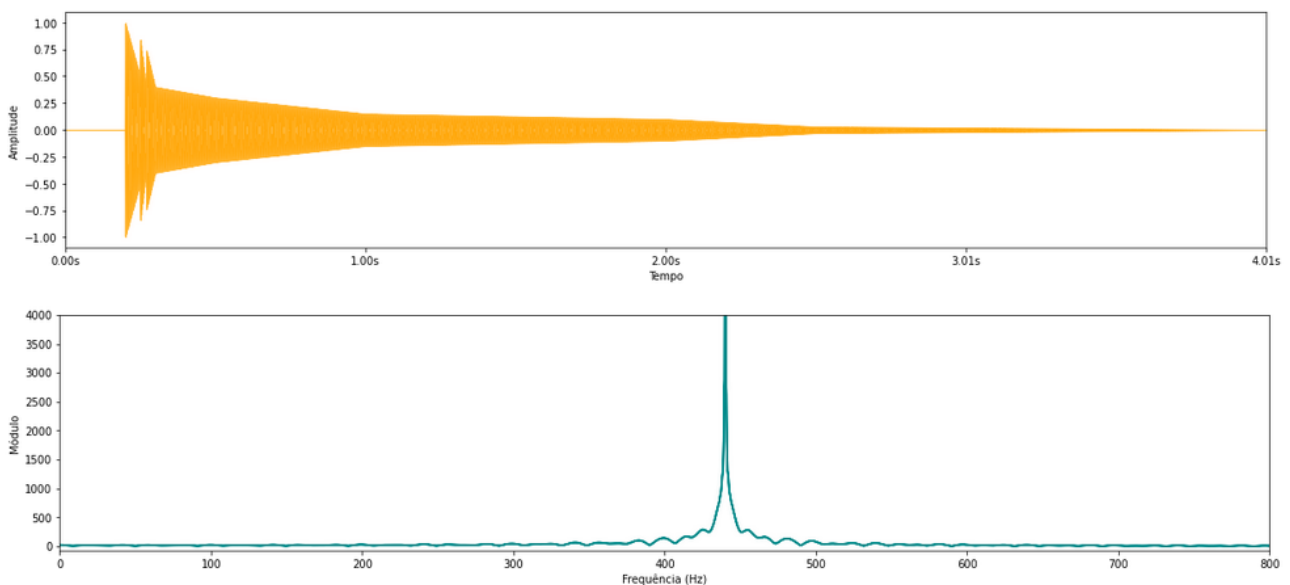
Podemos extrapolar a modelagem do envelope ADSR para uma estrutura de envelope um pouco mais sofisticada, onde adicionamos mais uma etapa de ataque e decaimento, tornando um envelope 2(AD)SR.



*Imagem 10: Sinal e espectro de um pulso com 2 etapas de ataque e decaimento.
Link para ouvir o som deste sinal: [Inserir Link](#)*

O som produzido agora apresenta uma nova dicção mais complexa. É possível também perceber que a forma como este envelope abre o espectro confere uma ondulação simétrica no entorno da frequência principal, portanto adiciona um conjunto muito peculiar de frequências ao sinal.

Com a intenção de observar as formas de abertura do espectro conforme o sinal apresenta novas dicções, podemos criar um envelope ainda mais extravagante: 3(AD)SR, com 3 etapas de ataque e decaimento.



*Imagem 11: Sinal e espectro de um pulso com 3 etapas de ataque e decaimento
Link para ouvir o som deste sinal: [Inserir Link](#)*

É notório que esta dicção tripla gera deformações mais visíveis no entorno da frequência principal, mantém-se a simetria e a também as baixas proporções entre estas frequências surgentes e a frequência principal. No exemplo mostrado o valor do pico da frequência principal foi de 9600 enquanto que toda a ondulação adicionada não excedeu o valor de 300. Portanto, todas as novas frequências adicionadas pela ondulação na abertura têm valores abaixo dos 3.2% de proporção na composição do espectro.

4.3 Envelopes naturais

Através da base de dados gravada temos a disponibilidade de utilizar diversos envelopes obtidos de sons naturais para inserir nesta análise. Utilizando o método da transformação de Hilbert capturamos o envelope de um sinal e inserimos uma oscilação com a frequência desejada. Então tomamos o espectro.

```
database = pulses_database()
sd = sound(database.dataset('drums')['tom'])
env = envelope().get_hilbert(sd).lowpass(0.0009)
env.plot(size=(20,3))
sig = signal().oscillation(440, 4)
sig.envelopment(env)
sig.plot(size=(20,2))
sig.plotfft(size=(20,3), xzoom=800, yzoom=0.5e8)
```

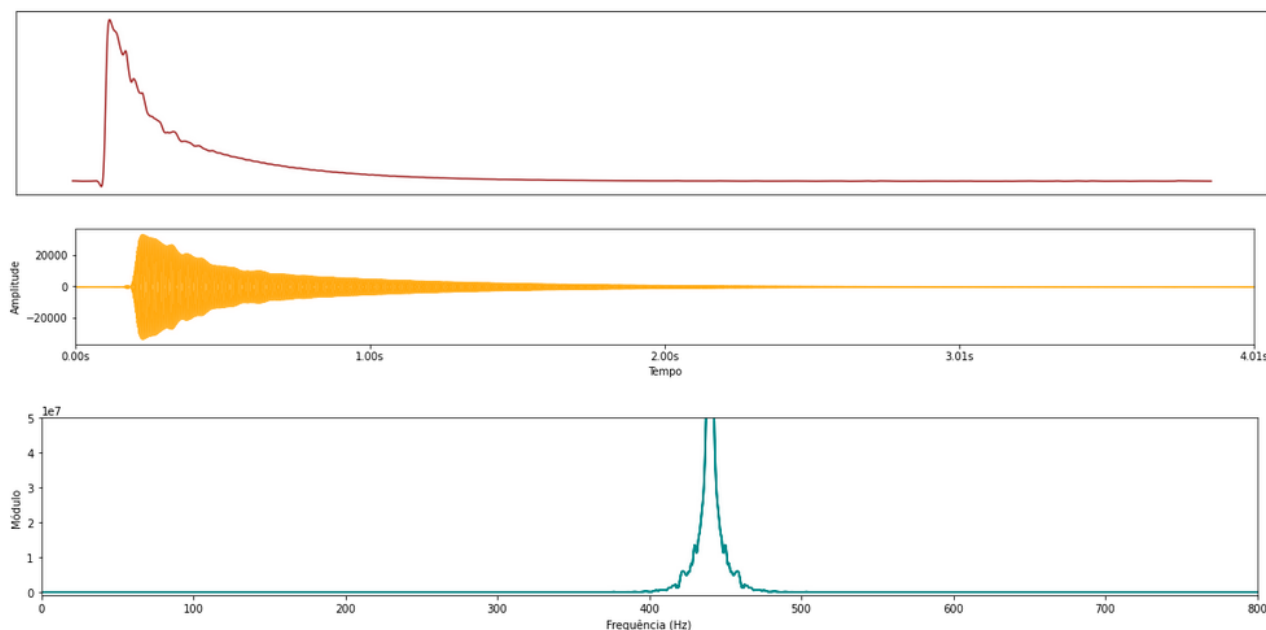


Imagem 12: Obtenção do envelope de um sinal gravado, sua inserção em um sinal de 440 Hz e seu espectro.

Link para ouvir o som deste sinal: [Inserir Link](#)

Vemos que envelopes naturais também causam ondulações simétricas na abertura do espectro e mesmo assim conservam-se em pequenas proporções em relação ao valor da frequência principal.

4.4 Informações do Som e Reconstrução Sintética.

Uma possibilidade que as classes do código permite, além de obter o envelope de um som da base de dados, é obter também os pontos de pico mais expressivos do espectro de frequência de um som. Criando um objeto da classe *spectrum*, Obtemos através do método *get()*, uma lista de todos os pontos de pico maiores que uma porcentagem da frequência principal (O valor desta porcentagem é passado como parâmetro ao método *get()*), esta lista contém os respectivos valores de frequência e proporção na composição do espectro. Com estas informações podemos construir um novo sinal, através da classe *signal*, adicionando ao sinal gerado as frequências desta lista com suas devidas proporções. Estaríamos criando um som que simula o timbre do som original. Se obtivermos os pontos de pico maiores que uma certa porcentagem mínima na composição do sinal, podemos considerar os maiores que 3%, por exemplo, teremos uma lista de pontos que descrevem o timbre do som ouvido, com maior ou menor precisão conforme a porcentagem escolhida.

```
database = pulses_database()
spec = spectrum().get(sound(database.dataset('harmonica')['E4']), 3)
spec.plot(lw=1)
```

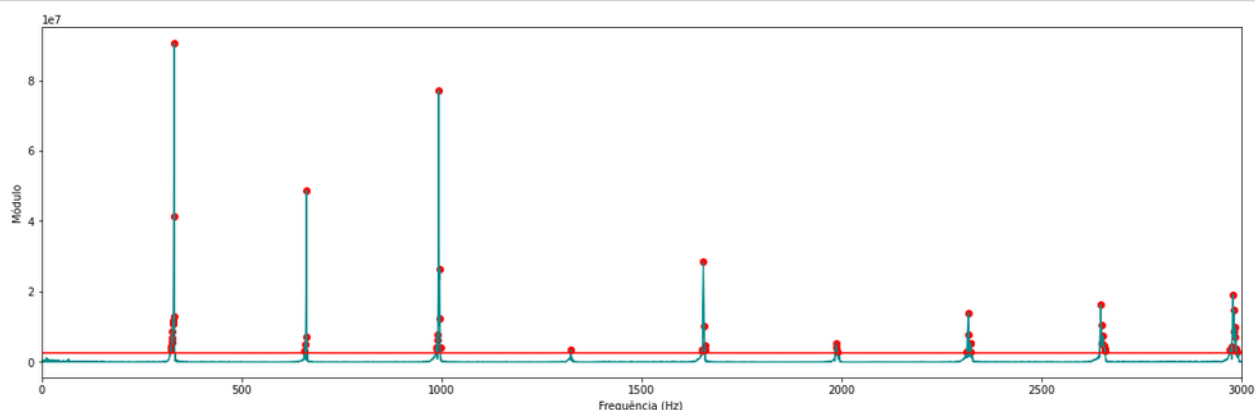


Imagem 13: Pontos de pico do espectro da nota E4 de uma gaita diatônica.

Utilizando este método obtivemos os valores de pico do espectro de uma nota Mi da quarta oitava de uma gaita diatônica. Com a porcentagem escolhida de 3% foram obtidos 174 pontos de frequência e proporção. Podemos criar um novo sinal utilizando estas informações, e em seguida aplicar um envelope geométrico do tipo ADSR construído manualmente, que se assemelha ao envelope original.


```
geo = envelope().geometric([(0,0),(0.15,0.55),(0.35,1),(0.42,0.98),(0.65,0.45),(1,0.2),(1.8,0.01),(4,0)], 4)
```

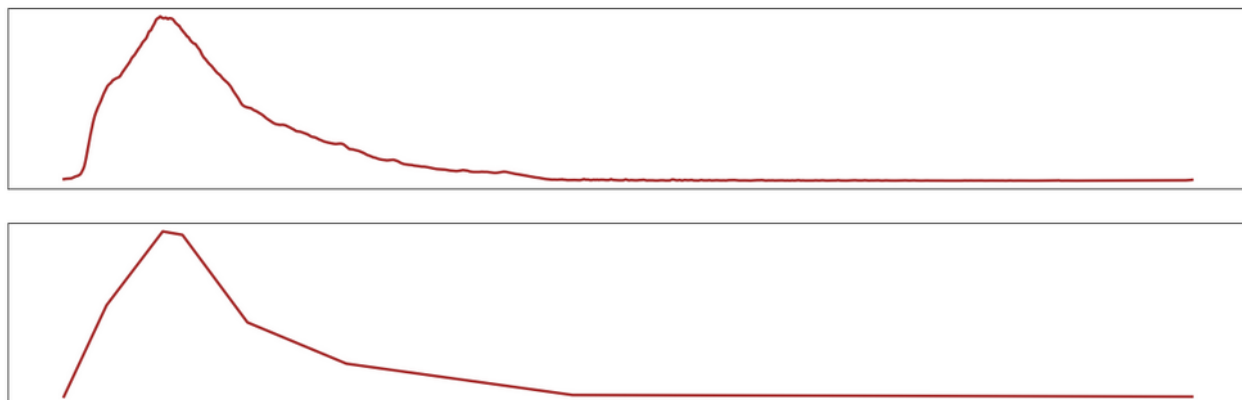


Imagem 14: Envelope natural do som da gaita diatônica e um ADSR semelhante.
Link para ouvir a comparação entre o som natural e o sintético: [Inserir Link](#)

Nota-se que o som gerado é muito semelhante ao original, simulando bem o timbre e a dinâmica de evolução no tempo. Considerando o vetor de 174 frequências e proporções, que correspondem a 348 informações numéricas, mais as 16 informações que compõem o envelope e o tempo de duração, somam-se 365 informações, no total, capazes de descrever os aspectos de timbre e dinâmica de evolução temporal do som. É claro que este número pode ser maior ou menor de acordo com a porcentagem escolhida para obter os picos, com a possibilidade de o próprio som apresentar mais frequências acima desta porcentagem e com a complexidade do envelope produzido. No entanto, ainda assim teríamos da ordem de centenas de informações capazes de descrever o som. Uma diminuição significativa frente às centenas de milhares de informações que compõem o sinal gravado.

Porém percebemos também que este novo som sintetizado não é capaz de se assemelhar inteiramente com o som original. O principal elemento do som original que não está presente no som sintético são as dicções. No exemplo da gaita diatônica, ouve-se um ruído proveniente da efusão do ar pelas palhetas, este elemento não é preservado na reconstrução do som, o que significa que estas 365 informações não guardam esta sutileza. Fica evidente a possibilidade de identificar auditivamente qual é o som da gaita em comparação com o som sintético.

Repetimos o mesmo experimento, mas desta vez sintetizando o som de um instrumento de cordas pinçadas: Gravado da nota Mi da quarta oitava de uma guitarra elétrica, que corresponde à tocar a corda mais aguda solta. Desta vez usamos uma porcentagem de consideração ainda menor, de 1.3% e obtivemos 63 pontos de pico, o novo envelope também contém 16 informações, totalizando 143 informações numéricas. Mostramos anteriormente que ajustando o tempo de ataque de um envelope ADSR, podemos adicionar uma dicção parecida com o ataque de uma corda

pinçada. Então compomos um sinal com o espectro obtido e este envelope construído para conter um dicção.

[Inserir Links para ouvir a comparação](#)

Percebemos que agora o som original e simulado são muito parecidos, e adicionando uma dicção de ataque tornamos o som sintético bem mais natural aos ouvidos.

4.5 Composição de Espectros e Envelope

Dispomos de vários sons na base de dados, cada um deles com distintas características de espectro e envelope. Podemos utilizar dos métodos mostrados anteriormente para criar novos sons compondo o espectro de um som escolhido com o envelope de outro.

Tomamos como exemplo obter as características do espectro da corda Sol do violino e compor um novo som com um envelope obtido de um prato do tipo *crash* de bateria.

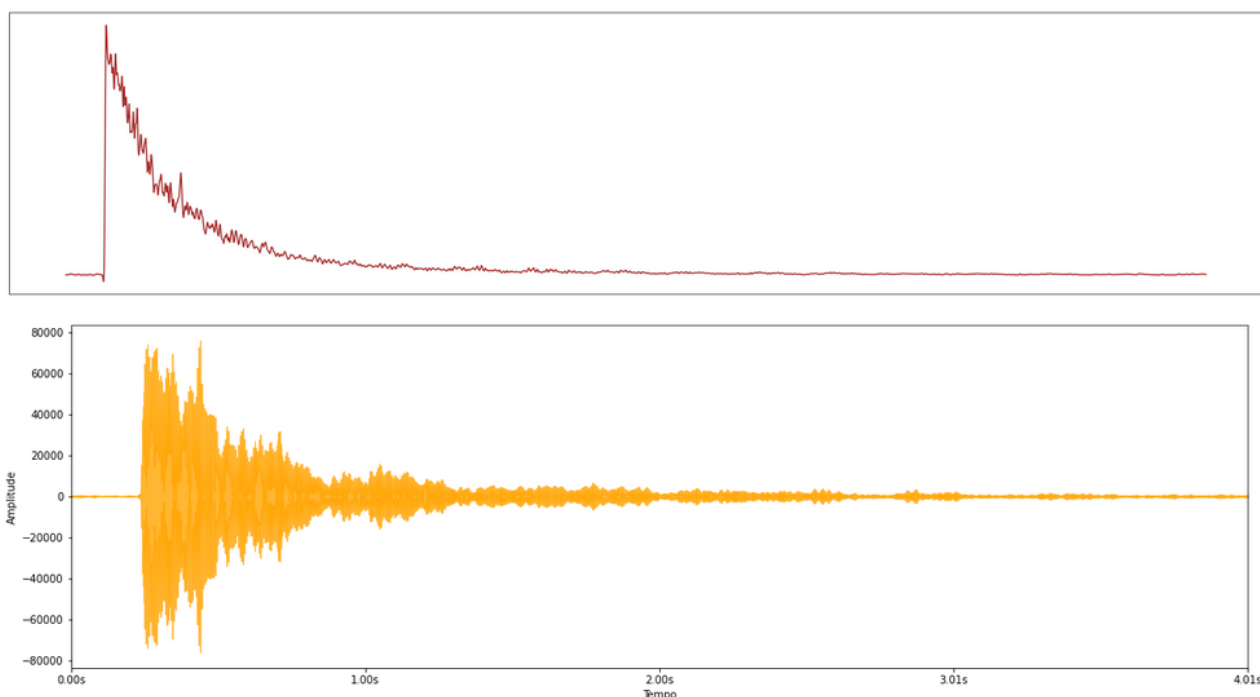


Imagem 15: Envelope do crash e sinal composto com espectro do violino.

Link para ouvir: [Inserir Link](#)

Nota-se que o envelope utilizado tem um tempo de ataque bastante rápido, um decaimento semelhante ao exponencial e ao longo deste decaimento há uma ondulação rápida e íngreme, como é característico o efeito chuveiro do som de um prato de bateria soando. Esta dicção referida como

efeito chuva também está presente no som sintético e confere uma sonoridade metálica ao timbre proveniente do violino, de modo que o som resultante se assemelha a um sino de igrejas barrocas. Vemos aqui que certas dicções podem ser preservadas pelo envelope, como é o exemplo deste efeito metálico que advém de uma espécie de ruído branco ao longo do decaimento da envoltória. Então nos ajuda a determinar os elementos que originam este timbre. Um sino tem uma característica harmônica, assim como tem o espectro do violino com seus harmônicos nas frequências de Sol, e também uma característica ruidosa no envelope, que confere o som metálico. Compondo estes elementos provenientes de sons que têm cada uma das características temos a receita para simular um som deste tipo.

Este é um exemplo de como podemos utilizar os elementos descritos para definir características de timbres e também simular diferentes sons em sintetizadores digitais.

5. Conclusões

Estudando as características dos sinais sonoros e categorizando as naturezas matemáticas de diversas delas, pudemos entender melhor como se compõem os sons em seus aspectos de timbre, dinâmica e dicções. Mostramos que o timbre, expresso pelo espectro de frequências, e a dinâmica de evolução da intensidade, expressa pelo envelope, se relacionam de maneira necessária: As características do envelope definem sutilezas importantes no timbre, e adicionando ou alterando um envelope, alteramos também as características do timbre, de maneira sutil ou expressiva à depender do próprio envelope. O simples decaimento da amplitude é capaz de incrementar uma gama de frequências, no formato gaussiano, ao redor da frequência principal, enriquecendo o som ouvido. Outras características no envelope podem resultar em formatos de abertura singulares, conferindo um espectro mais sofisticado ao som.

Muitos instrumentos musicais carregam atributos importantes no seu envelope, como por exemplo as dicções de ataque nos instrumentos de corda pinçada. Esta dicção, particularmente, está relacionada a variações súbitas da intensidade do som. Usando da modelagem de envelope ADSR, controlando as etapas de ataque e decaimento, é possível simular esta dicção, conferindo ao som produzido uma semelhança maior com sons naturais.

Esta análise traz a compreensão que a variação de amplitude da onda não somente se relaciona com a dinâmica da intensidade do som, como diretamente sugere seu significado físico, mas também com a variedade de frequências ouvida. E que, para descrever um som, precisamos coletar informações relativas às frequências principais do espectro e ao formato do envelope. Desta

forma, coletando centenas de informações, conseguimos simular com boa precisão um som gravado que precisa de centenas de milhares de informações para ser representado.

A deficiência desta análise reside na dificuldade de conservar dicções do som original. Como vimos, estas dicções são expressas de maneira muito ínfima nas características do espectro e mesmo do envelope, portanto são difíceis de serem coletadas através destes. No entanto elas têm uma importância muito grande para a naturalidade do som ouvido. Descobrimos a natureza de uma delas, o tempo de ataque no efeito da corda pinçada, conseguimos reinserir esta dicção no som sintético, melhorando-o significativamente.

Também mostramos que certos envelopes naturais são capazes de portar dicções mais expressivas, como é o exemplo do efeito chuveiro no prato de bateria. Compondo envelopes com espectros diferentes conseguimos observar a origem de certos timbres, possibilitando descrevê-los em termos de fenômenos físicos. Este entendimento pode auxiliar, por exemplo, na construção ou melhoramento de sintetizadores digitais de áudio, não somente com a intenção de torná-los mais próximos de sons naturais, mas também lançando possibilidades de composições de elementos inesperados, criando timbres complexos e incomuns, mas que soem com maior naturalidade. A música moderna tem explorado cada vez mais sons artificiais criados digitalmente. Conforme esta tendência, poderia se beneficiar de novidades auditivas proporcionadas por equipamentos ou softwares que trouxessem novas tecnologias, possibilitadas por uma compreensão mais categórica dos efeitos sonoros, seus elementos matemáticos e relações com os fenômenos físicos.

O código produzido e a base de dados gravada, ainda podem ser explorados em seu potencial para experimentos envolvendo os elementos aqui trabalhados, possibilitando novas descobertas. Conceituações sobre ruídos ao longo da sustentação, ruídos brancos inseridos no som, efeitos de tremolo, ataques mais sofisticados em instrumentos de cordas duplas, pausas abruptas ou quaisquer efeitos característicos de outros instrumentos musicais ainda estão para ser feitas com as análises mostradas.

6. Referências

1. Brigham, E. Oran, and R. E. Morrow. "The fast Fourier transform." *IEEE spectrum* 4.12 (1967): 63-70.

2. S. Orphanidis, Introduction to Signal Processing, Prentice-Hall, Upper Saddle, NJ., 1996.

3. A. V. Oppenheim and R. Schaffer. Discrete-Time Signal Processing. Pearson, 2009.

4. Costa, Luciano da F. *Instantaneous Signal Analysis (CDT-40)*. 2020

5. Costa, Luciano da F. *On Sound Synthesis I: Single Note, Single Frequency (CDT-42)*. 2020.

6. Costa, Luciano da F. *On Sound Synthesis II: Single Note, Varying Frequency (CDT-43)*. 2020.
7. Costa, Luciano da F. *On Sound Synthesis III: Single Note, Multiple Frequencies (CDT-44)*. 2020.
8. B. Boashash. Estimating and interpreting the instantaneous frequency of a signal: Part 1. *Proc. IEEE*, 80:520–538, 1992.
9. B. Boashash. *Time Frequency signal analysis and processing: A comprehensive reference*. Academic Press, 2015.
10. Jaffe, David A. "Spectrum analysis tutorial, part 1: the discrete Fourier transform." *Computer Music Journal* 11.2 (1987): 9-24.
11. M. Alonso and E.J. Finn, *Física: Um Curso Universitário*, v. 2 - *Campos e Ondas* (Editora Edgard Blücher Ltda., São Paulo, 1995), 1ª ed.
12. H.L. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (Cambridge University Press, Cambridge, 2009), 3ª ed.
13. L.L. Henrique, *Acústica Musical* (Fundação Calouste Gulbenkian, Lisboa, 2002), 2ª ed.
14. M. Hewitt, *Music Theory for Computer Musicians* (Cengage Learning, Stamford 2008), 1ª ed.
15. A.V. Oppenheim, R.W. Schaffer and J.R. Buck, *Discrete-Time Signal Processing Vol. 2* (Prentice-Hall, Englewood Cliffs, 1989), 2ª ed.
16. Ulrich, T. J. "Envelope calculation from the Hilbert transform." *Los Alamos Nat. Lab., Los Alamos, NM, USA, Tech. Rep* (2006).
17. J. Dugundji, "Envelopes and pre-envelopes of real waveforms," in *IRE Transactions on Information Theory*, vol. 4, no. 1, pp. 53-57, March 1958, doi: 10.1109/TIT.1958.1057435.
18. Purves, Steve. "Phase and the Hilbert transform." *The leading edge* 33.10 (2014): 1164-1166.
18. Williams, Earl G. *Fourier acoustics: sound radiation and nearfield acoustical holography*. Academic press, 1999.
20. Cochran, William T., et al. "What is the fast Fourier transform?." *Proceedings of the IEEE* 55.10 (1967): 1664-1674.
21. Laurenti, Nicola, and Giovanni De Poli. "A method for spectrum separation and envelope estimation of the residual in spectrum modeling of musical sound." *Proc. of the COST G-6 Conference on Digital Audio Effects (DAFX-00)*, Verona, Italy. 2000..
22. S. Goldman. *Frequency Analysis, Modulation and Noise*. Academic Press, 2011.
23. C. Phillips, J. Parr, and E. Riskin. *Signals, Systems and Transforms*. Pearson, 2013.
24. R. Carmona. *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms with an Implementation in S*. Mc Graw-Hill, 1948.
25. Provencher, S. W. "A Fourier method for the analysis of exponential decay curves." *Biophysical journal* 16.1 (1976): 27-41.
26. Wikibooks. *Engineering Tables/Fourier Transform Table 2*, signal 14. https://en.wikibooks.org/wiki/Engineering_Tables/Fourier_Transform_Table_2