

TEXTO DAS ANÁLISES DE FLAUTA

(FLUTE ANALYSIS)

PARTE INTEGRANTE/ANEXO À DISSERTAÇÃO DE MESTRADO

(ATTACHMENT OF THE MASTER DISSERTATION)

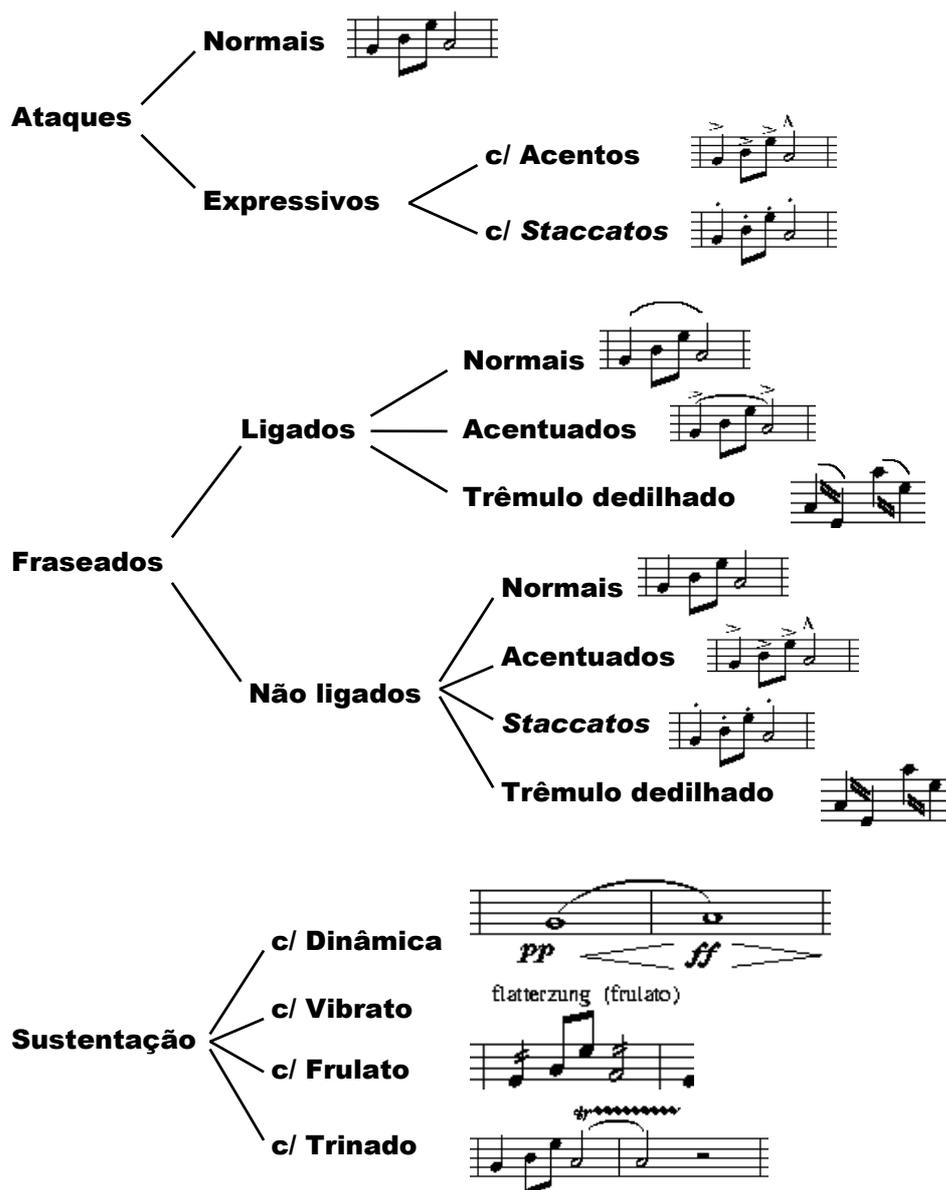
Faria, Regis Rossi Alves. **Aplicação de wavelets na análise de gestos musicais em timbres de instrumentos acústicos tradicionais**. Dissertação (Mestrado). Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Eletrônica. São Paulo 1997. 182p.

Faria, Regis Rossi Alves. **Application of wavelets on the analysis of musical gestures in traditional acoustic instruments timbres**. Master Dissertation in Electronic Engineering. Polytechnic School of the University of São Paulo. São Paulo 1997. 182p.

ANÁLISES DE FLAUTA

A Figura 1 abaixo resume todos os tipos de eventos musicais amostrados e analisados para a flauta. São ao todo 14 eventos analisados.

Figura 1 - Eventos musicais amostrados e analisados para a flauta



Os ataques abordam a análise dos períodos de ataques de notas individuais. Os fraseados estudam as transições entre notas embutidas num contexto melódico. As sustentações analisam um conjunto de efeitos sobre notas sustentadas.

As partituras ao lado de cada evento meramente ilustram a simbologia adotada em música para representá-los, isto é, a notação escrita para guiar a interpretação, não consistindo propriamente dos trechos musicais tocados.

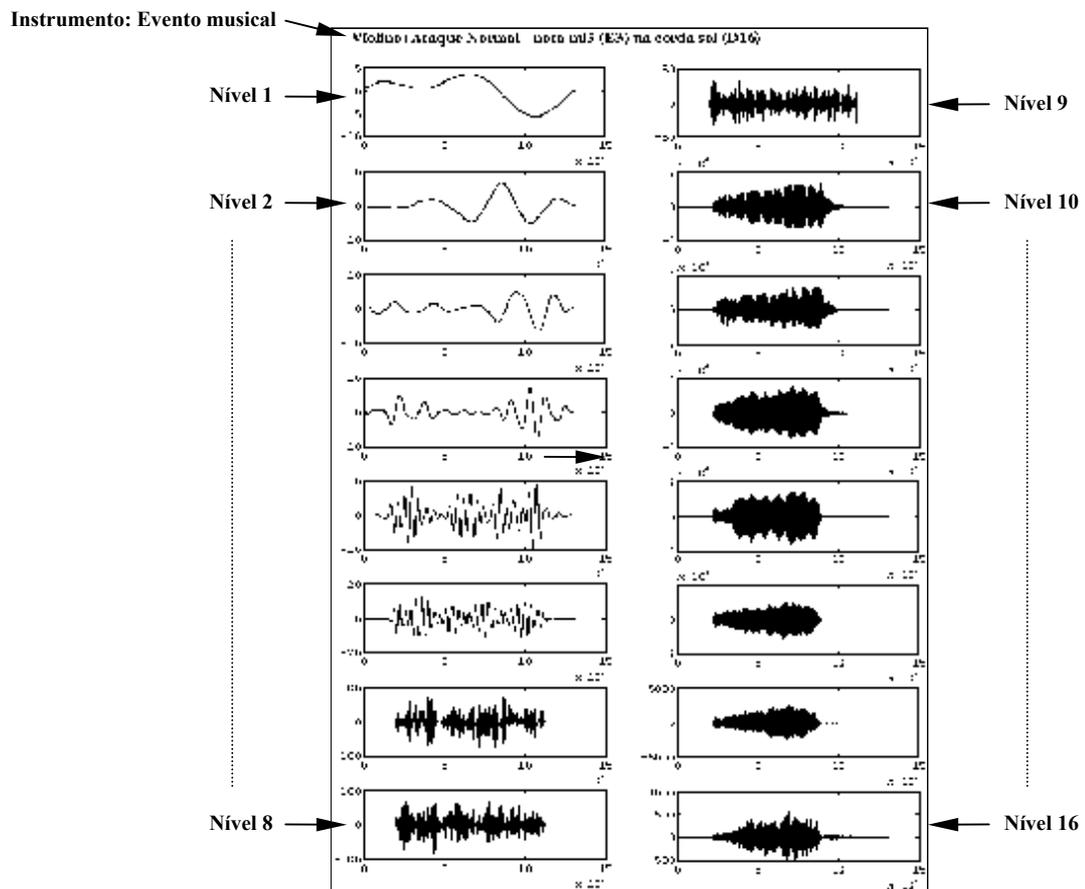
A análise *Wavelet* em multiresolução (AWMR) de cada evento musical compreende:

- os gráficos dos **níveis *Wavelet*** da análise,
- uma descrição do **evento** (uma descrição da interpretação e/ou sobre o trecho tocado),
- uma ilustração da **partitura** representativa do trecho tocado (quando aplicável),
- a representação gráfica do **som original**,
- a **extensão da amostra** (sinal musical),
- a ***Wavelet*** de análise,
- os **resultados individuais** por níveis (pontos mais relevantes e revelações em cada nível *Wavelet*),
- gráficos mostrando a **distribuição de energia** pelos níveis,
- e eventuais **espectrogramas** realizados.

A página inicial de cada análise consta de um **título**, que especifica o **instrumento** e o **evento musical** abordado, e dos **gráficos dos níveis *wavelets*** da análise do evento, isto é, a projeção do sinal em cada um dos subespaços da AWMR, do **segundo nível *Wavelet* mais grosso** (nível 1) ao **mais fino** (nível $N-1$, onde N é o \log_2 da extensão do sinal musical). Por exemplo, para sinais de extensão 2^{17} pontos amostrais, o nível 16 será o mais fino (W_{16}) e o nível 1 (W_1), o mais grosso (ou menos fino). O nível 0 não é mostrado, e corresponde à informação encerrada nos subespaços W_0 e V_0 da AWMR. O sinal nestes níveis são desprezíveis em termos de energia, e consistem basicamente na forma de onda da *Wavelet* de análise, em nada contribuindo à descrição ou formação do sinal.

A Figura 2 que segue ilustra uma página inicial de uma análise de evento musical, indicando a disposição e ordenação em que os níveis *wavelets* são plotados. O primeiro nível acima à esquerda é o nível *Wavelet* 1. O último nível (abaixo à direita) é o nível mais fino (que pode ser o 17, 16 ou 15 para sinais de extensão 2^{18} , 2^{17} ou 2^{16} , respectivamente).

Figura 2 - Exemplo de página inicial da análise AWMR de um evento musical com extensão 2¹⁷



O eixo horizontal de cada gráfico mostra a extensão do sinal (número de pontos amostrais ou amostras discretas), e corresponde diretamente à dimensão temporal (44100 pontos = 1 segundo). O eixo vertical mostra a intensidade (amplitude ou magnitude) em unidades (adimensionais) em escala linear.

Da segunda página em diante apresentam-se os outros itens da análise (textos, partituras e gráficos). Usualmente o item mais longo é aquele que discorre sobre os resultados e revelações individuais de cada nível: a **análise dos níveis**. O texto é detalhado, abordando tanto aspectos genéricos - *muitas vezes comuns a outras análises, como descrições do envelope, de padrões periódicos e estruturas produzidas por gestos de interpretação, avaliações sobre a suavidade e regularidade da forma de onda, da textura, taxa de variação de subida ou descida, qualidade percebida do timbre, pitches percebidos via audição, textura sonora, quantidade e características de ruído presente* - quanto aspectos exclusivos - *como análises espectrais suplementares, caracterização de série harmônica associada, análise da evolução da forma de onda acompanhando sua regularidade e/ou variabilidade gráfica e sônica, correlações com gestos de interpretação, relação de transientes e objetos sônicos relevantes e comparações com outros níveis*.

Ao final de cada análise, dois gráficos mostram a **distribuição de energia** do sinal pelos níveis *wavelets*; o **primeiro**, traçado em escala linear, evidencia os níveis mais energéticos do sinal musical; no **segundo** gráfico, o eixo vertical (normalizado) mostra o logaritmo da energia de cada nível sobre o logaritmo da energia total, desta forma comprimindo grandes variações de intensidade no topo do gráfico, permitindo que os níveis menos energéticos apareçam e possa-se

inferir sobre a distribuição relativa e a contribuição dos níveis que não aparecem sob o escalamento linear. O escalamento neste último é portanto logarítmico.

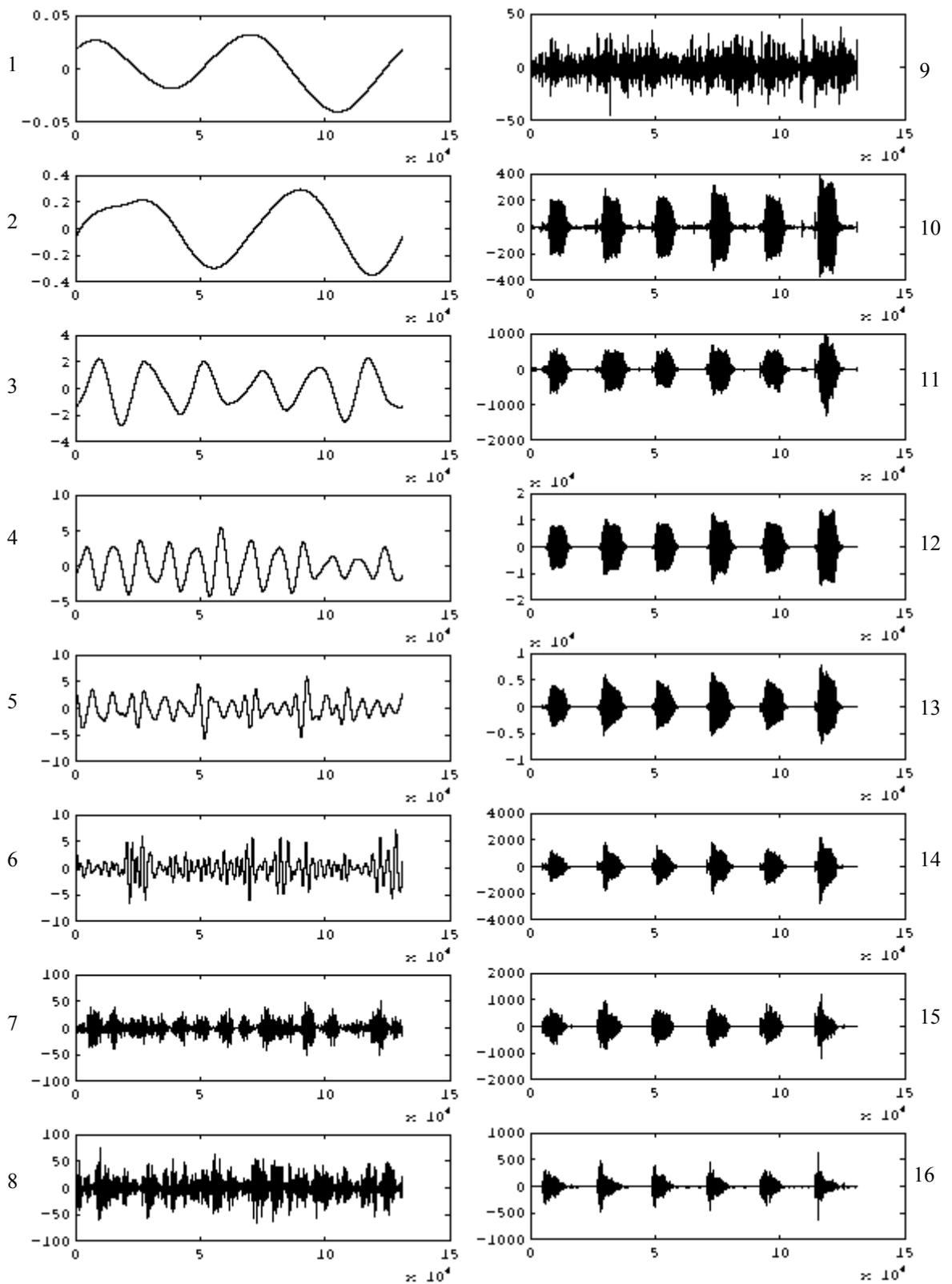
A quantidade de informação gerada é relativamente extensa, e optou-se pela exclusão da sua íntegra dentro do corpo do texto, por motivos de clareza, espaço e concisão. O material integral, todavia, está disponibilizado para consulta em meio digital anexo (CD ROM e disquete, em formato Microsoft Word for Windows 95 - versão 7.0).

As análises de flauta estão neste arquivo ("**an_flauta.doc**"), e as de violino, no arquivo "**an_violin.doc**". Estes arquivos consistem num amplo material de referência para musicólogos, engenheiros de som/áudio, projetistas e desenvolvedores de tecnologia musical interessados em conhecer os detalhes da análise *wavelet* de sinais musicais, avaliar sua aplicabilidade em projetos comerciais e/ou científicos específicos, e elaborar um roteiro de especificações iniciais para nortear explorações mais profundas, focalizadas, e orientadas à concepção de produtos que utilizem a tecnologia.

No item "Apreciação e discussão dos resultados " da dissertação é feita uma avaliação crítica e uma consolidação dos resultados das análises de todos os eventos musicais, e um levantamento das relações importantes comuns entre elas.

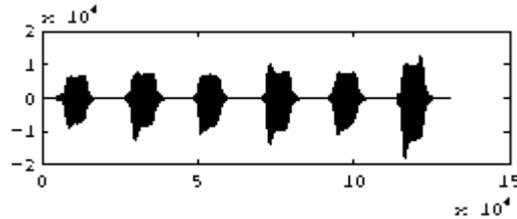
Nas conclusões finais elabora-se uma *tabela de distribuição de gestos/eventos musicais por níveis wavelets* que relaciona os objetos sônicos, eventos e gestos musicais aos níveis onde se distribuem e mais se concentram. Apresentam-se as estruturas sônicas que aparecem destacadas em cada nível, a percepção gráfica e sônica do nível, e uma lista de sensibilidade de cada nível, sumariando os resultados mais relevantes deste trabalho.

Flauta: Ataques, Normais - notas si4 (B4) na segunda oitava da flauta



Evento: Uma rajada de 6 notas caracterizando ataques normais com sopro tipo "tu", à nota si4 (B4) na segunda oitava da flauta.

Gráfico do som original:



Extensão da amostra: 2^{17} pontos (131072). **Wavelet de análise:** D16

Análise dos níveis:

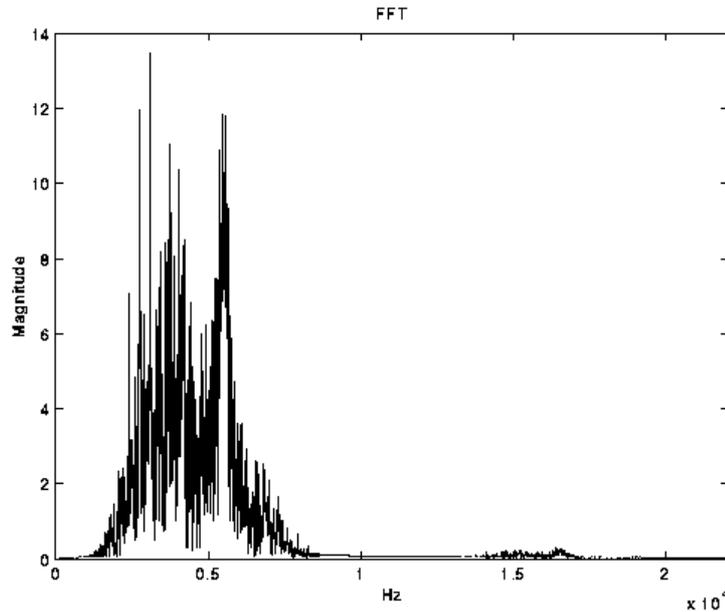
Nível 16: Envelope do nível acompanha envelope natural do sinal. Há explosões (*bursts*) na segunda e na última notas, causando uma sonoridade característica de *clicks*, coincidindo com as notas que mais apresentam uma qualidade de ataque súbito, com rápida variação da pressão de sopro e o engajamento da tonalidade correta imediatamente a seguir. A qualidade do som é ruidosa, retendo a maior parte do ruído do sopro característico, especialmente nos ataques.

Nível 15: O nível de ruído de fundo é baixo. O envelope geral segue o da nota original, porém com diferenças na textura, que deixa de ser tão suave e apresenta mais pontas e transientes, com amplas variações de amplitude, o que no mundo acústico é percebido como fortes variações de pressão acústica, um som mais áspero, ruidoso. Basicamente este nível "é o anterior mais suavizado" em termos de textura. Também característico neste nível é sua maior atenuação dos ruídos de fundo, o de sopro permanecendo "visível".

Nível 14: A última nota apresenta as maiores amplitudes (~3000 unidades). Neste nível já aparecem melhor caracterizadas as regiões de ataque, apresentando intensidades menores que as porções centrais das notas - *a amplitude máxima está em 266 unidades para a primeira região do ataque (medido para a primeira nota entre 4000:7100 pontos)* - em comparação com a máxima amplitude na região central - *1180 unidades, aproximadamente 4,43 vezes mais intensa em amplitude e cerca de 22 vezes mais energética que a região inicial da primeira nota.*

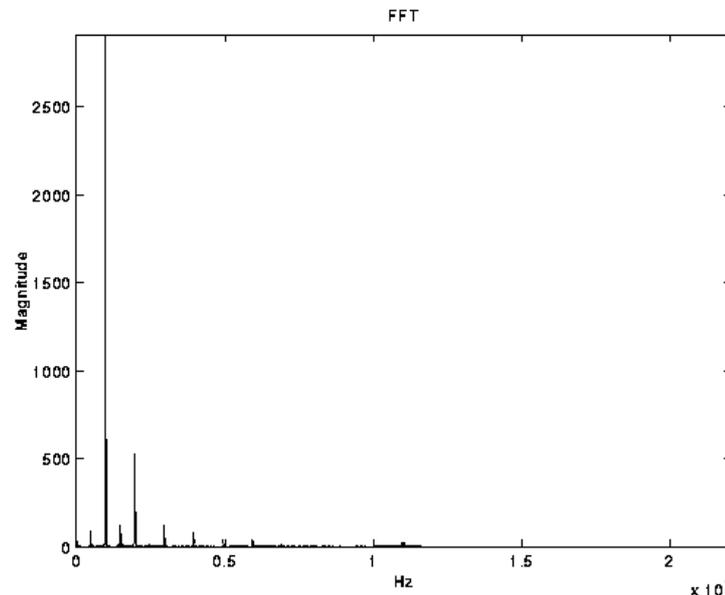
O espectro desta região inicial está na Figura 3 e apresenta picos máximos em 3100 Hz, 2756 Hz, e 5480 Hz. Na nota original, na mesma faixa de 4000:7100 pontos, o pico máximo de frequência está em 1001 Hz, seguido por 10971 Hz (quase 11^o harmônico). O espectro de toda a primeira nota no nível 14 traz a série harmônica associada a outras frequências como 3536,8 Hz (pico máximo) que vem seguido pelo 6^o harmônico (2966,2 Hz, desviado de 2,8 Hz do esperado), e também existe uma densa população de frequências com magnitudes inferiores a 10% do valor do pico máximo, distribuídas principalmente na faixa de 2000 a 8000 Hz.

Figura 3 - FFT da região inicial (ataque) da primeira nota no nível 14 (entre 4000 e 7100 pontos)



O espectro da nota inteira está na Figura 4 e apresenta uma constituição bem mais harmônica e limpa comparado ao espectro da porção inicial somente (não mostrado, e que apresenta uma densa população de frequências até 12500 Hz, caracterizando o período em que ainda há muita turbulência e as ondas estacionárias ainda não se firmaram).

Figura 4 - FFT da primeira nota do exemplo (som original)



A nota si4 por este espectro apresenta portanto *pitch* central em 987,8 Hz (segundo harmônico do si3). O segundo pico do espectro está em 1975,7 Hz, exatamente o segundo harmônico para o *pitch* central (quarto para si3, que pela estimativa deste espectro está em 493,9 Hz, visível na primeira raia à esquerda no gráfico). A composição harmônica é portanto muito forte, comprovando a característica harmônica do instrumento. É importante notar que, apesar da nota possuir *pitch* em 987,8 Hz este não é o primeiro harmônico da série de si: ele é o segundo, sendo 493,9 Hz o primeiro harmônico ou fundamental, presente com uma amplitude bem reduzida.

Nota-se também que as frequências cobertas pelo nível 14 aumentam de intensidade somente após o período inicial (ataque) onde esta é baixa e regular, passando para uma segunda fase com patamares de maior amplitude e apresentando regularidade ligeiramente menor que na fase inicial, diferença existente pela sensibilidade maior nesta fase à intensidade variável do sopro.

Nível 13: Percebe-se duas notas neste nível: o si₄ (confundindo-se com sua oitava fundamental si₃) e um sol₃, ou seja surgem 2 vozes. Pelo espectro do nível vemos que 1977 Hz (4^o harmônico) é o pico máximo, seguido por 1768,4 Hz (que não faz parte da série harmônica de si₃) que deve ser parte da série harmônica de sol₃ (G₃). A presença da nota sol foi confirmada via audição comparativa com a mesma nota tocada na flauta.

Nível 12: Também neste nível é perceptível, desta vez com mais "pureza" das vozes, a série harmônica do si₃ (com a fundamental se confundindo com sua oitava) e a presença de um tom sol₃ (G₃). O som é mais puro, ausente de ruídos, com os dois tons bem relevantes. Trata-se do nível mais energético (9,5 vezes mais energético que o anterior), o que mostra serem os níveis mais energéticos aqueles que possuem a melhor textura de som, descartando ruídos.

Tal "pureza sonora" é revelada pelo espectro de todo o nível (para todas as notas, vale lembrar): pico máximo de frequência em 987,8 Hz (fundamental de si₄, e segunda harmônica de si₃), e um segundo pico em 1768,4 Hz, que não faz parte da série harmônica de si₃, mas de sol₃.

As operações de decomposição e reconstrução utilizadas para gerar o sinal de cada nível abrem espaço para a ocorrência de um efeito de replicação de bandas, gerando, por assim dizer, "novas bandas" de frequência, especulares às originais. A regra de formação destas novas bandas, entretanto, leva geralmente à obtenção de vozes ou tons inscritos na escala cromática, isto é, gera tons afinados, geralmente em harmonia com os tons originais. No presente exemplo, surge a presença da série harmônica de sol, ou seja, a nota percebida além do si é um sol, e não um tom microtonal ou atonal, ou pertencente a uma escala diferente da cromática ocidental.

Nível 11: A textura é um pouco mais áspera, o sinal, mais irregular. Dois *pitchs* perceptíveis: um sol₃ e a presença de ré₃ (no início principalmente) e seu segundo harmônico, ré₄, mais fraco que sol₃ mas simultâneo. As vozes se projetam num registro mais grave, como era de se esperar para este nível. Partes iniciais dos ataques revelam características de um envelope diferente para esta região das notas neste nível. O nível é sensível a qualquer excitação provocada por deslocamentos de ar (sopro e respiração), e não somente sensível aos tons desejados e suas regiões (ataque, colapso, etc.). Um indício que comprova essa sensibilidade são os objetos existentes antes dos ataques, que correspondem a sutis respirações e gestos de sopro imperceptíveis no som original.

Nível 10: À primeira vista as formas de onda neste nível sugerem um "ressurgimento dos ruídos", haja visto a presença de um som de fundo que revela gestos associados à respiração. Ouvindo-se a primeira nota original a 11025 Hz ¹ percebemos claramente a existência de um sopro, e então o estabelecimento do tom vibrante, as ondas estacionárias. Ouvindo-se o nível 10 também lento nota-se a presença de um som peculiarmente grave, que se espalha por todo o sinal. Em vista da nota si₄ ser uma nota de registro mais agudo, portanto colapsando mais rapidamente que notas no primeiro registro da flauta, sabe-se que existem pausas entre as notas não preenchidas por nenhum som gerado pelo instrumento. O nível 10 entretanto mostra um

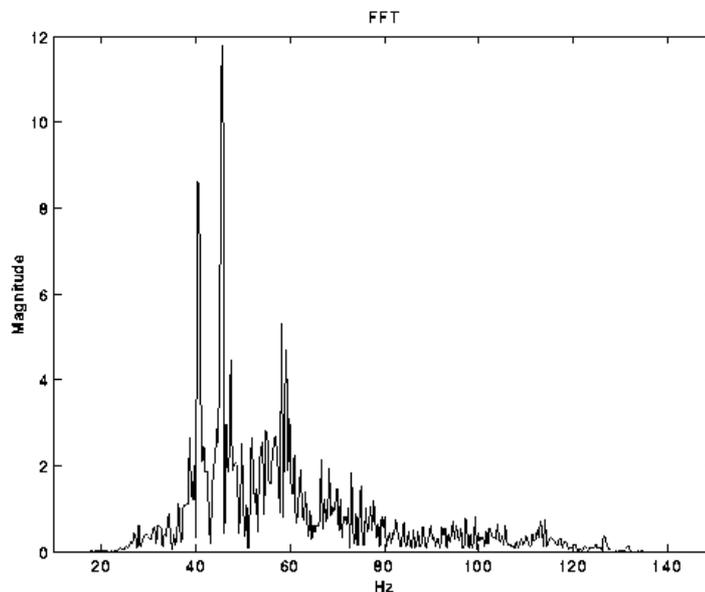
¹ Normalmente a frequência de reprodução é a mesma da gravação (44100 Hz), isto é, a frequência da taxa de amostragem. Entretanto, em algumas situações torna-se útil realizar a reprodução com uma taxa de reconstrução menor que a original, aumentando a duração do evento reproduzido.

ruído grave de fundo, cuja associação a eventos acústicos específicos não consiste em dedução óbvia. No entanto, eventos de respiração são bem localizados neste nível pela sua sensibilidade a transientes, além da sensibilidade usual aos padrões vibratórios acústicos.

Nível 9: Ruído grave de fundo. Não há um envelope regular neste nível: trata-se de um som grave, continuamente distribuído por toda a duração do exemplo, de textura áspera (muitas "descontinuidades"), com variações grandes de amplitudes em intervalos muito curtos (pelo que se pode visualizar do gráfico do nível). Eliminá-lo em nada altera na reconstrução do som original, como pode-se verificar através de uma reconstrução com todos os níveis abaixo do 9 (inclusive) zerados.

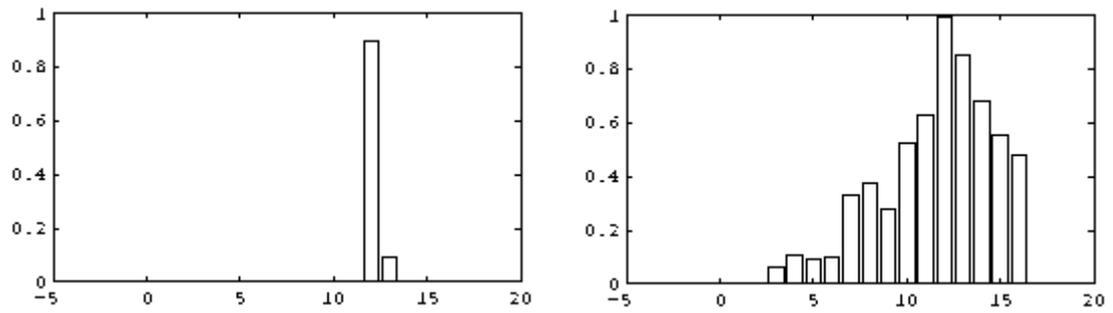
Nível 8: Ruído grave, existente por todo exemplo, não localiza nenhum evento musical. A Figura 5 mostra a FFT deste nível, com frequência de pico em 45,75 Hz, seguido por 40,37 Hz.

Figura 5 - FFT do nível 8 da análise



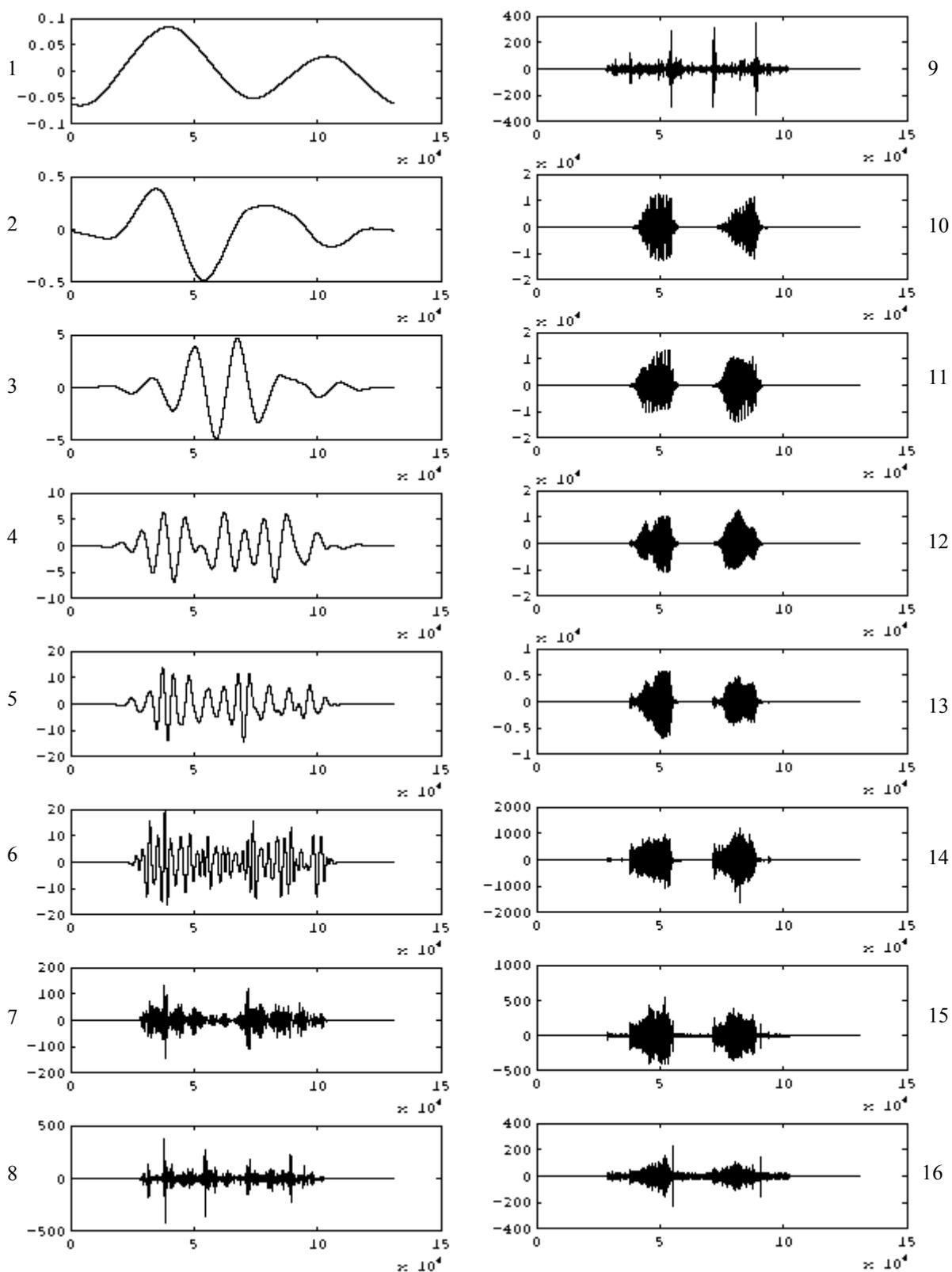
Uma reconstrução foi feita aumentando-se a amplitude dos coeficientes neste nível **em 12 vezes**, e mostra somente que ele contribui para um ruído de fundo grave, não sendo importante para os eventos musicais do exemplo.

Nível 7: Já apresenta alguma diferença em termos de segmentação (de notas) do que o nível anterior, que era um corpo contínuo sem diversidade. Neste nível aparece uma modulação em amplitude, que gera uma sequência de objetos ondulatórios. Entretanto tal modulação não presta-se para identificar objetos musicais neste nível, nem mesmo o número de notas presente. Uma reconstrução intensificando em 12 vezes os coeficientes para este nível não provocou nenhuma alteração perceptualmente significativa no som, comparado ao original.

Distribuição de energia:

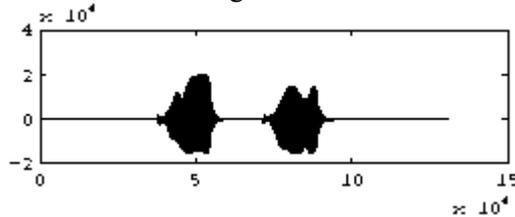
Mais de 99% de toda a energia do exemplo se concentra nos níveis 12 e 13. O gráfico à direita mostra a contribuição de cada nível numa escala logarítmica. Os níveis inferiores ao nível 10 não contribuem significativamente para a formação do som original.

Flauta: Ataques, Expressivos, com Acentos - notas mi3 (E3) na primeira oitava



Evento: Duas notas mi³ (E₃), da primeira oitava da flauta, tocadas acentuadamente em sequência. Acentos com dinâmica de piano, quase normais. No som original há um corte súbito das notas, realizado com um gesto labial. Um estalido característico do contato dos lábios durante o corte é perceptível nas notas.

Gráfico do som original:



Extensão da amostra: 2^{17} pontos (131072). **Wavelet de análise:** D16

Análise dos níveis:

Nível 16: Oferece meios para se rastrear os seguintes eventos musicais e acústicos:

1. Ruído: existe entre as notas um nível de ruído apreciável, e que pela sua amplitude relativamente alta mascara aspectos importantes para caracterização do acento, objetos que podem estar abaixo de seu limiar de magnitude. As fontes deste ruído já foram abordadas anteriormente.
2. É possível verificar que o acento em notas no primeiro registro da flauta consistem em fraseados difíceis, uma vez que precisa-se vencer subitamente a inércia da coluna de ar e estabelecer as ondas estacionárias enquanto se projeta um fluxo de ar mais vigoroso, de forma a caracterizar o acento. O efeito desejado só aparecerá de forma clara não no ataque, mas na região de sustentação, central da nota, como sugere o envelope das notas neste nível.
3. O gesto labial de corte de fluxo de ar revela aqui seu estalido característico, sentido neste nível como um transiente com qualidade próximo de um *click*, uma "descontinuidade" súbita e intensa.

Nível 15: o nível de ruído já é menor, mas aparece ainda com certa intensidade, podendo estar mascarando alguns objetos musicais relacionados ao ataque. No entanto, pode-se estimar com boa margem de acerto, que os ataques acentuados não possuem inícios abruptos como os observados no caso do violino. Estando a nota está num registro de difícil estabelecimento rápido de ondas estacionárias, e também pelo próprio princípio de produção tonal na flauta, pode-se compreender que há um atraso no estabelecimento do acento, que é produzido por um ataque com um fluxo intenso de sopro

Não há a percepção de um timbre característico de flauta neste nível: o som é bastante ruidoso. Nota-se, acima do limiar do nível de ruído, que há um ataque súbito, seguido de um decaimento rápido e em seguida uma valorização na intensidade, em ambas as notas. Na primeira, entretanto, há um tipo de crescendo, que evolui até o início da fase de colapso. No colapso há uma queda também vertiginosa dos contribuintes desse nível (de frequências mais altas).

Nível 14: O envelope sugerido para estas notas nesse nível apresenta um formato tendente ao quadrilátero, possuindo bordas quase quadradas, mas com pontos de alta amplitude na região central das notas. Este tipo de formato caracteriza o fato de que, sob acento, as componentes contribuintes deste nível crescem de intensidade abruptamente durante o ataque, mantêm-se sobre a sustentação e morrem depressa durante o colapso. Entre o ataque e o decaimento existe uma fase em que a energia trafega do ruído do sopro para o tom mi³ estável, e isto fica evidente

pela caracterização do estrangulamento que existe entre o início do ataque e a região central da nota (mais visível na segunda nota).

Nível 13: A primeira nota apresenta uma intensidade relativa maior na sua fase central, em comparação com a mesma fase na segunda nota. Em ambas, entretanto, a sustentação tende a apresentar intensidades constantes e uniformes. O ataque e o decaimento, visíveis no nível anterior, agora ficam melhor caracterizados (uma magnificação desta região propicia uma melhor visualização deste fato). Um estrangulamento observado na região de sustentação, tanto neste nível quanto no próximo, é verificado também no sinal original: trata-se de uma modulação de sopro sobre a região de sustentação produzida naturalmente pelo intérprete, uma particularidade desta interpretação que a individualiza, e está dissociada de transientes do ataque ou colapso. Isto ocorre de forma não determinística, e constitui elemento de variabilidade em música.

A região de colapso pode ser aproximada por duas fases: uma em que a amplitude relativa cai abruptamente até um valor mínimo, e uma segunda fase que consiste na gradual queda da amplitude até a extinção, mas numa taxa de decrescimento menor. O nível de ruído neste nível se encontra nos mesmos níveis visíveis no sinal original.

Nível 12: O timbre de flauta já é perceptível neste nível. A parte mais alta da série harmônica de E3 aparece, e o tom percebido é algo como uma mistura das 2 oitavas (E3 e E4). As intensidades relativas nas regiões centrais ainda exibem uniformidade, e sugerem um envelope aproximadamente quadrilátero, principalmente para a primeira nota. Note que a medida que o nível é mais energético ele concentra também mais da qualidade vibratória do timbre, o que é aferido pela audição do nível e pela verificação de que a região de sustentação, onde as vibrações estacionárias já estão estabelecidas, são as mais destacadas.

Nível 11: Envelope também "quadrilátero", com períodos de ataque e colapso pouco energéticos neste nível. Há também uma modulação de amplitude mais evidente nas partes centrais, onde as amplitudes relativas são maiores. Note que uma boa aproximação de envelope para as componentes deste nível utilizaria um *fade in* e *fade out* triangular, uma vez que a amplitude surge do zero e sobre gradualmente no ataque, e faz o mesmo caminho, mas inverso, durante o colapso, não havendo "descontinuidades" destacadas.

Nível 10: Os envelopes neste nível possuem contornos suaves, e há também uma modulação de amplitude por toda a duração das notas. As transições são graduais, não havendo descontinuidades expressivas. As bordas na primeira fase de colapso entretanto são íngremes, mas não possuem transientes significativos, o que se evidencia pelo som do nível, suave, de *pitch* único (não há a percepção de dois *pitchs* diferentes).

Nível 9: O envelope natural das notas não é observado neste nível. No entanto, há picos de amplitude locais que correspondem a gestos labiais, e eventos correlacionados com a respiração.

Nível 8: Revela pulsos periódicos, associados aos gestos bucais do início e final das notas; não apresenta um envelope regular, e exibe um ruído grave regularmente distribuído, também presente em todos os silêncios. Os pulsos apresentam periodicidade regular, estando ritmados.

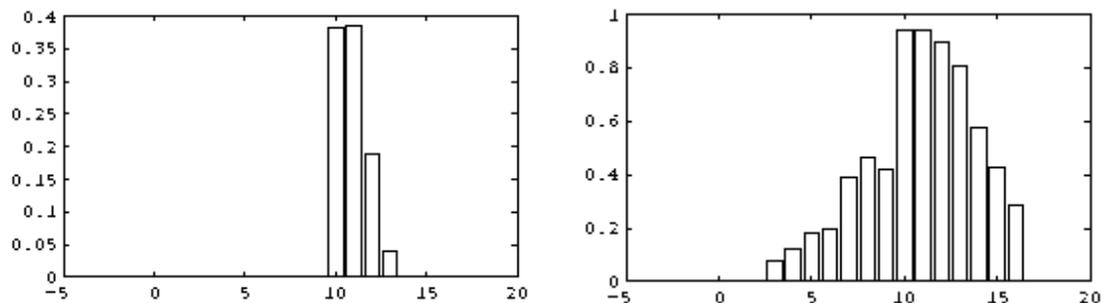
Nível 7: Neste nível volta a aparecer o formato do envelope natural das notas, podendo aproximar as regiões onde há notas, como também locais de ataques, os acentos estando associados a intensos picos locais nas proximidades da região de decaimento, imediatamente após o ataque. A pausa entre as notas não existe: há um preenchimento contínuo. Isto nos mostra que as componentes das notas neste nível só vêm a intensificar as componentes já

existentes na ausência dos tons, mas a intensificação ocorre de maneira organizada, proporcionalmente ao aumento das amplitudes das frequências tonais. Caracterizado por um som grave modulado em amplitude.

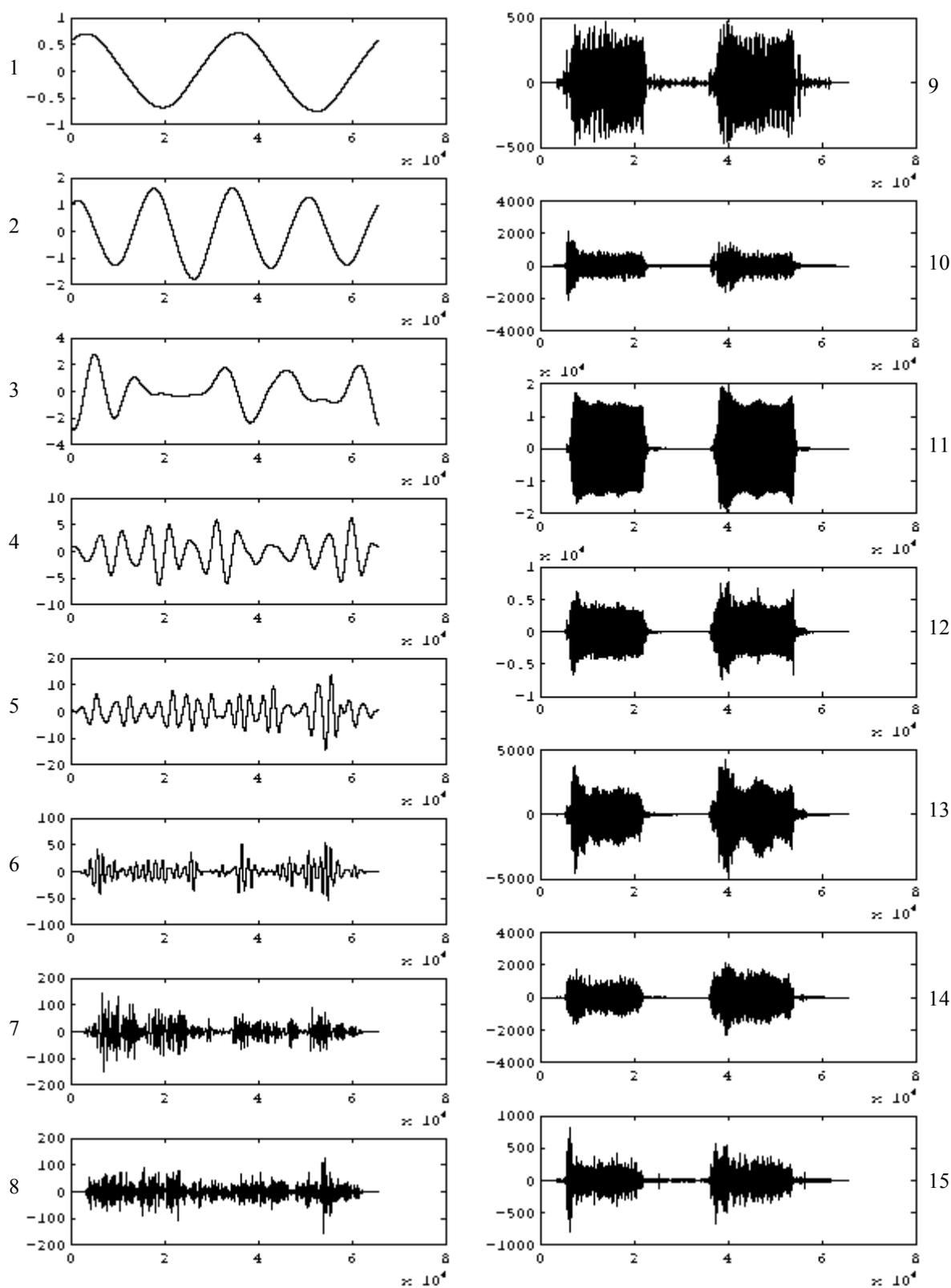
Nível 6: Perda do envelope. Não há estruturas, formas ou objetos que indiquem algum evento musical atuante neste nível. Praticamente inaudível.

Nível 5: Inaudível. Sem indicações de eventos acústicos.

Distribuição de energia:

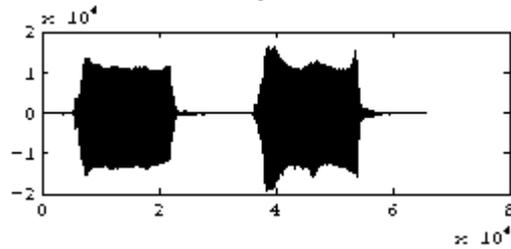


Mais de 99% de toda a energia do sinal musical se concentra nos níveis de 10 a 13. O segundo gráfico mostra que os níveis inferiores ao 10 exibem uma participação inexpressiva no conteúdo energético do sinal original.

Flauta: Ataques, Expressivos, com acentos - notas si4 (B4) na segunda oitava

Evento: duas notas si4 (B4), na segunda oitava da flauta, tocadas com acentos. Os ataques na interpretação foram súbitos e intensos. Esta análise complementa a análise anterior de ataques acentuados em flauta, mostrando o comportamento do gesto musical e som produzido quando realizado na segunda oitava deste instrumento.

Gráfico do sinal original:

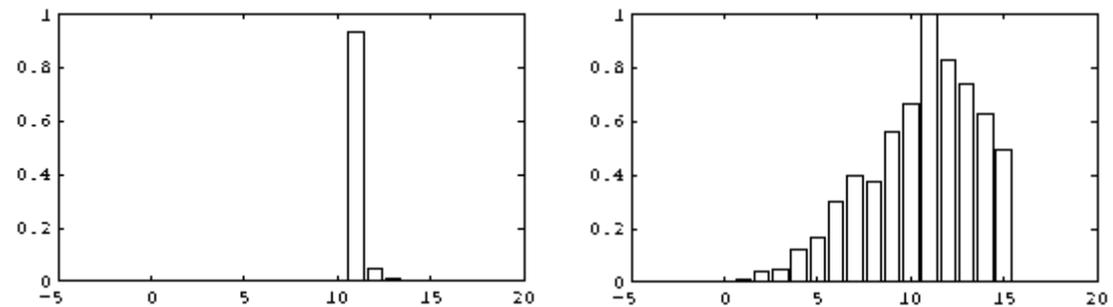


Extensão da amostra: 2^{16} pontos (65536). **Wavelet de análise:** D16

Análise de níveis:

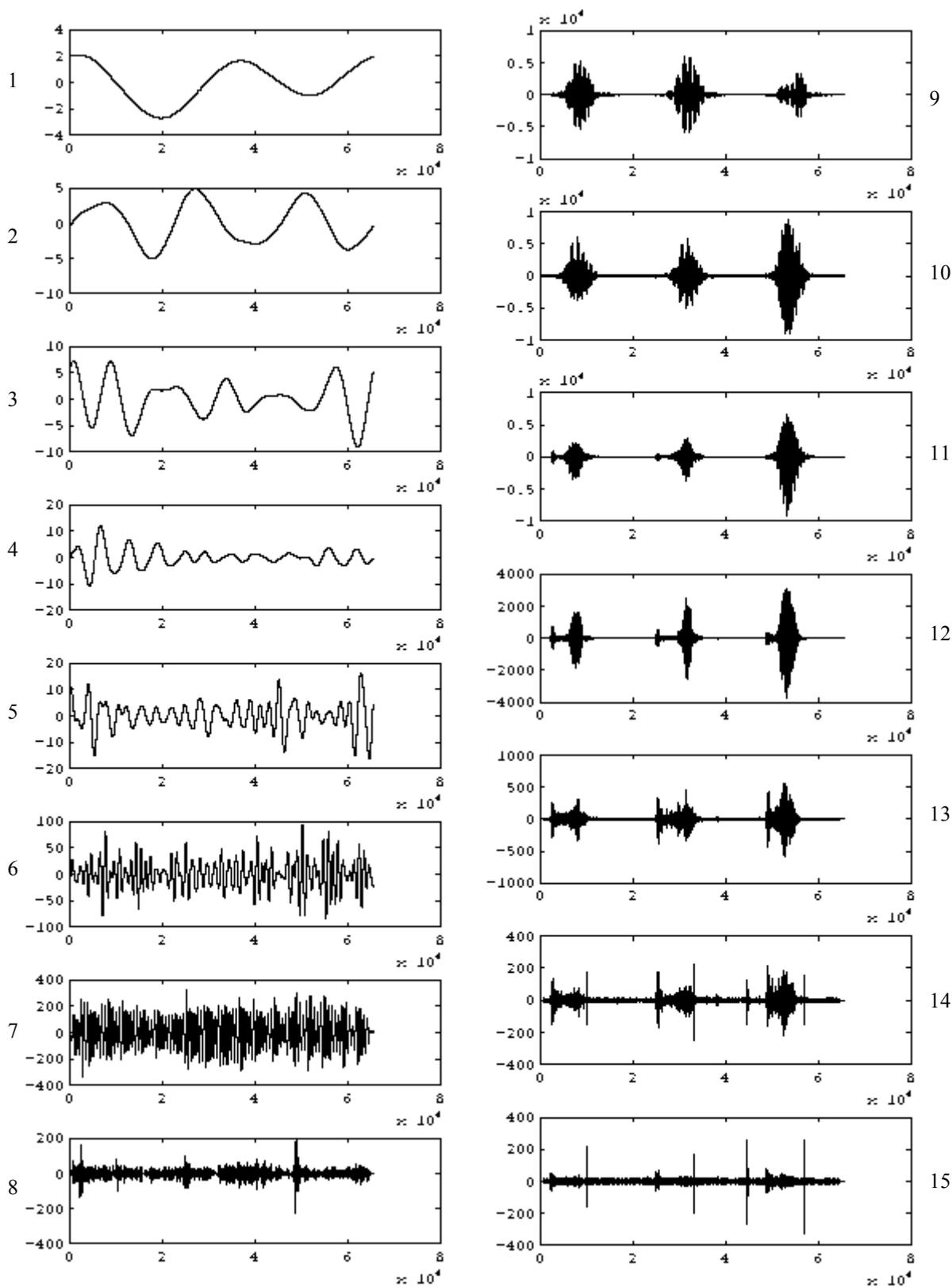
Nível 7: o nono nível mais fino. Mostra sensibilidade aos transientes súbitos do ataque acentuado, assim como aos associados ao colapso, permitindo a localização destes gestos.

Distribuição de energia:



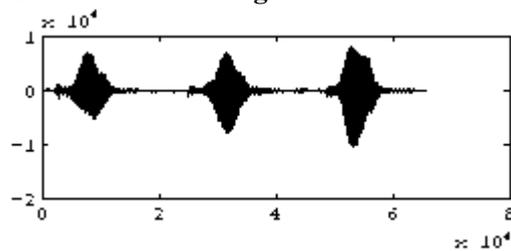
Mais de 99% de toda a energia do exemplo se concentra nos níveis de 11 a 13, sendo o nível 11 o mais energético, com mais de 90% de toda energia.

Flauta: Ataques, Expressivos, com *staccatos* - nota mi3 (E3) na primeira oitava



Evento: O exemplo consiste de 3 notas (mi3) em *staccato* tocadas em sequência e com ataque tipo "tu". O ruído de gesto bucal, efeito da movimentação da língua no sentido de modular o jato de ar, foi destacadamente captado pelo microfone, e na maioria das vezes tal ruído é inaudível nas audições ao vivo, pela própria natureza da dispersão e atenuação das ondas sonoras em auditórios.

Gráfico do sinal original:



O ruído do gesto bucal também fica realçado neste exemplo porque a nota escolhida para amostrar o *staccato* foi o primeiro mi (E3) da flauta transversa, uma nota no registro mais baixo, que requer a excitação de uma longa coluna de ar, apresentando portanto maior inércia acústica para ser produzida, e também maior dificuldade ao intérprete para sua execução. A lentidão de se produzir uma nota súbita como o *staccato* no mi3 pode ser útil na análise dos momentos transitórios entre o início do jato e o estabelecimento do tom, motivo pelo qual observamos este fraseado para esta nota. A análise qualitativa visual e auditiva dos níveis é o principal objeto de interesse no momento.

Extensão da amostra: 2^{16} pontos (65536). **Wavelet de análise:** D16

Análise dos níveis:

Como auxiliar para análises subsequentes, a FFT do exemplo (incluindo todas as 3 notas) indica a presença das seguintes raias (em Hz) componentes da série harmônica de mi3 (E3):

57,87	(terceiro pico, não faz parte da série harmônica)
325,0168	(fundamental de E3 para a flauta transversal, segundo pico)
654,071	(segundo harmônico, pico mais alto)
978,4	(terceiro harmônico, quinto pico)
1306,1	(quarto harmônico, quarto pico)
1630,5	(terceiro pico, quinto harmônico, praticamente empatado com o quarto harmônico)

Pela análise harmônica, vemos que os parciais pares possuem maior amplitude que os ímpares, com a exceção do 5º harmônico que está pouco mais alto que o quarto. Na verdade o primeiro componente é uma contaminação. Uma análise de Fourier da região de pausa entre a primeira e segunda notas (pontos entre 17000:25000) mostra uma raia de frequência com pico em 59,21 Hz, bastante próxima da frequência da rede elétrica de 60 Hz.

Nível 15: O ruído de fundo é a componente mais realçada neste nível, estando os ruídos dos gestos bucais em segundo lugar.

O ruído de fundo que se apresenta neste nível tem origem externa aos eventos sonoros associados à flauta, e possivelmente foi gerado pelo processamento dos arquivos sonoros em alguma etapa desde a gravação até a finalização deste exemplo em arquivo de som, conforme já abordamos anteriormente. Este ruído é indesejado, e inutiliza de certa forma o nível para

observação dos componentes de ruído oriundos do sopro, uma vez que eles se associam. *Wavelet-packets* podem ser úteis para separar estes componentes, haja visto que uma inspeção visual mais acurada da região de pausa (entre notas) e o início das notas mostra evidências de que as formas de onda apresentam espectros diferentes nestas regiões. Apesar do ruído de sopro ser uma importante componente na qualidade do timbre percebido da flauta, declinamos no momento de perfazer uma análise deste tipo em função de nossos propósitos objetivarem a caracterização de gestos musicais transitórios para este exemplo utilizando análise *Wavelet* simples.

Já os gestos bucais, estes são facilmente rastreáveis no nível pela sua qualidade de pulso, um transitório que apresenta amplitudes altas relativamente ao restante do sinal. Ao mesmo tempo, é bastante provável que não somente os gestos bucais no interior da terceira nota sejam responsáveis pelas altas intensidades naquela região, mas também que componentes do sopro contribuam para elevar a energia nos níveis mais altos. Evidências a favor dessa hipótese estão no nível 8 desta análise, onde há intensidade para os gestos bucais no início da nota, mas sem contribuição energética no gesto bucal interno à nota.

Nível 14: O timbre de flauta não é perceptível. Os gestos labiais, bem localizados, agem no sentido de mostrar os *onsets* (início do ataque) de cada nota e os instantes de colapso forçado. Em *staccatos* em flautas, a forma mais eficiente para se forçar o colapso das notas é cortando-se subitamente o jato de ar imprimindo a língua contra o céu da boca. Em circunstâncias ideais o gesto bucal associado não deveria produzir estalidos, mas deveria ser seco. A presença de ruídos bucais em fraseados velozes, entretanto, é frequente, e mostra-se que este nível é particularmente sensível a tais gestos.

Pela análise das notas neste nível, percebemos que a terceira nota foi a que mais apresentou transientes associados aos gestos da língua. Em três momentos houve a produção de "estalidos bucais":

1. no início do ataque, quando o jato de ar, sob pressão, é liberado pela retração da língua
2. no meio da nota, devido a uma movimentação da língua indesejada, mas plenamente natural
3. no final, caracterizando o momento do colapso forçado, novamente marcado por um movimento da língua.

O envelope das notas já se projeta além do ruído de fundo associado ao exemplo, que neste nível é menos energético. Nas redondezas dos pulsos que marcam os estalidos bucais, as frequências que contribuem nesta faixa estão variando suas intensidades dinamicamente, sugerindo um envelope englobando-os. Grande parte da energia neste nível se deve a frequências presentes nas notas em si, como pode-se verificar pelas intensidades relativas maiores sobre as notas, o que aproxima mais o envelope deste nível em direção à forma do envelope do sinal original.

Nível 13: Os gestos bucais aparecem como delimitadores do *staccato* da nota mi_3 (E3), uma nota difícil de se articular. O timbre característico de flauta não é perceptível. Uma análise FFT mostra que temos presentes os parciais 12 (pico mais elevado), 11, 10, 9 (segundo em elevação), 8 (já bem menor em amplitude comparado aos outros) e traços de outros parciais mais baixos e mais altos que estes. A presença dos parciais altos sem no entanto a percepção do timbre

característico de flauta sugere que as intensidades relativas entre os parciais (i.é, o arranjo de harmônicos neste nível) não se fundem de forma a produzir uma sonoridade de flauta.

Nível 12: O som característico lembra o de uma corneta. Os gestos bucais ainda bastante visíveis, funcionam como delimitadores das notas, o que é bem visível para as primeiras duas notas. A terceira, recordando a existência já detectada de um estalido bucal no interior da nota, observa-se que este apresenta uma correlação com os níveis mais altos de amplitude nesta nota. Outro fato que deve ser levado em conta para esta terceira nota é a sua qualidade soprada e o rápido estabelecimento de uma faixa estacionária do tom, exatamente pela transitoriedade do efeito de *staccato*, é exacerbado neste nível, como sugere o envelope mais intenso para esta nota.

A grosso modo, o envelope neste nível apresenta três regiões distintas:

uma **primeira**, com amplitudes baixas (máximos em torno de 100 unidades) correspondentes aos períodos que vão do ataque da nota até o engajamento do tom correto, e uma **segunda** região com amplitudes mais intensas, correspondentes ao corpo da nota ou região de sustentação (máximos na faixa de 1000-2000 unidades, cerca de 15 vezes mais intenso que a região anterior, entre 30000 e 33000 pontos). O colapso forçado, localizado pelo gesto bucal que interrompe o jato, é a **terceira** região que vem em seguida com um curto segmento com amplitude decrescente, em média cerca de 20 x mais baixa que na sustentação (máximos de 300 unidades).

Auditivamente a qualidade do sopro da flauta reside em parte neste nível. A região de ataque de cada nota possui neste nível uma qualidade soprada, que sugere ao ouvinte um indício de que o instrumento é de sopro.

Nível 11: De novo as amplitudes relativas na última nota são as mais altas. Um outro detalhe sobre a última nota é que o ataque súbito nela excitou harmônicos mais altos, e é possível ouvir-se o som de uma oitava acima de mi^3 (E3). A terceira nota possui portanto duas características de interpretação que a tornam mais evidente nos níveis mais altos: (1), a qualidade soprada da nota, certamente realçada por uma pressão de jato de ar mais intensa que nas outras duas anteriores, e (2), o fato de que harmônicos mais altos foram excitados.

A terceira nota sem dúvida é a mais energética, e isso fica evidenciado pelas amplitudes altas que apresenta neste nível. Também nesse nível o timbre característico flauta não é perceptível. Há também uma certa modulação em amplitude nas notas, o que confere o contorno vibratório visíveis nas notas. Sobre esta vibração, todavia, é perceptível que o envelope geral das notas apresenta um contorno mais suave.

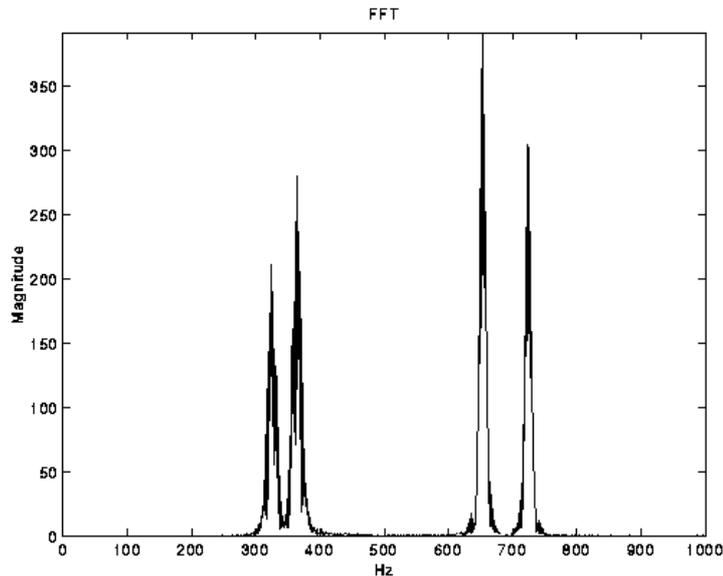
Nível 10: Nível mais energético. Modulação em amplitude notável pelo formato oscilante das notas. Os envelopes das notas apresentam contornos mais suaves que no nível anterior. A frequência de vibração visível nas duas primeiras notas é a metade da vista para terceira (onde harmônicos oitavados são excitados), diferença também sensível via audição. Novamente observamos um fenômeno de modulação de anel², com raias espelhadas, simetricamente dispostas ao redor de frequências que são exatos divisores (frações racionais) da frequência de amostragem.

A Figura 6 mostra as 4 raias mais intensas do sinal neste nível. As duas primeiras são raias espelhadas, ou bandas laterais simétricas em relação a 344,53125 Hz (1/128 da frequência de

² Espécie de modulação de amplitude onde a portadora é suprimida e somente as bandas laterais existem.

amostragem). As duas segundas também são espelhadas (com exceção às amplitudes relativas que são diferentes) e simétricas em relação a $1/64$ da frequência de amostragem (689,0625 Hz).

Figura 6 - FFT do nível 10



A primeira banda da figura corresponde à fundamental da nota (em 325,0168 Hz), sendo a segunda (a banda lateral superior) gerada pelo processamento da análise *Wavelet*. Para as duas primeiras notas há uma preponderância do primeiro par de raias em relação ao segundo, e a frequência de batimento para aquelas é de 39,03 Hz, a frequência que identificamos visualmente modulando a amplitude das duas primeiras notas.

Já para a última nota, o segundo par de raias (em 654,07 - o segundo harmônico - e 724,05 Hz, a banda lateral superior criada na "modulação de anel") é mais intenso, e a frequência de batimento de 69,98 Hz é a que verificamos visualmente modulando a nota no gráfico.

Nível 9: Contorno do envelope é suave. Há modulação de anel, com a mesma frequência de batimento para todas as três notas, de 39,03 Hz. As maiores amplitudes estão localizadas nas regiões centrais das notas, mais especificamente nas regiões onde os tons estacionários se estabelecem.

Nível 8: Perda do formato do envelope para as notas. Há um sinal contínuo, estreitando-se nas regiões de pausa, e alargando-se nas regiões dos gestos bucais, mas se estendendo por todo o exemplo. Os picos de amplitude no gráfico do nível correspondem aos gestos bucais iniciais, associados à liberação do jato de ar pela língua.

O gesto bucal central observado no interior da 3ª nota não possui caracterização neste nível, e pode-se portanto afirmar que ele possui qualidades tonais diferentes dos outros gestos. Pode-se mesmo extrapolar admitindo-se que não se trata propriamente de um gesto bucal no interior da nota, mas sim de um evento acústico causado pela variação intensa do jato de sopro liberado na 3ª nota. Realmente, atento à audição da nota original, verifica-se que não há gesto bucal no seu interior, sendo na verdade um gesto de sopro, um jato intenso com qualidade soprada, provavelmente obtido com alguma modulação na embocadura, e que não é perceptível neste nível (8).

Nível 7: Bem menos energético, e com uma forma de onda contínua e de amplitude regular por todo o exemplo, com leve estrangulamento nas regiões correspondentes às pausas. Uma FFT do

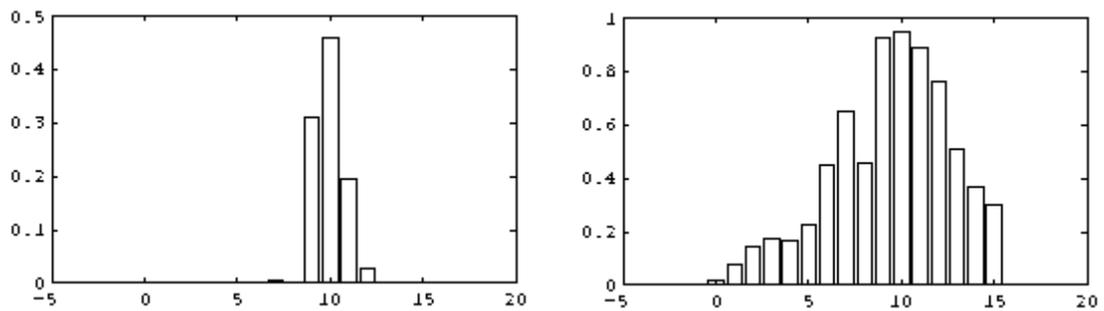
nível mostra que ele apresenta pico máximo de frequência em 57,87 Hz, sendo esta a raia mais energética do espectro do nível, e muito próxima à frequência oficial da rede elétrica, de 60 Hz, o que sugere alguma forma de contaminação. 45,7 Hz e 114,4 Hz são outros picos presentes, cerca de 6 vezes menores que o pico máximo (i.é, com menos de 15% de sua magnitude). O nível não mostra nenhum evento musical ou gesto relevante.

Nível 6: Não é possível extrair características musicais relevantes desse nível. Ele também varia em amplitude mais que o nível anterior, e é ainda audível.

Nível 5: Inaudível.

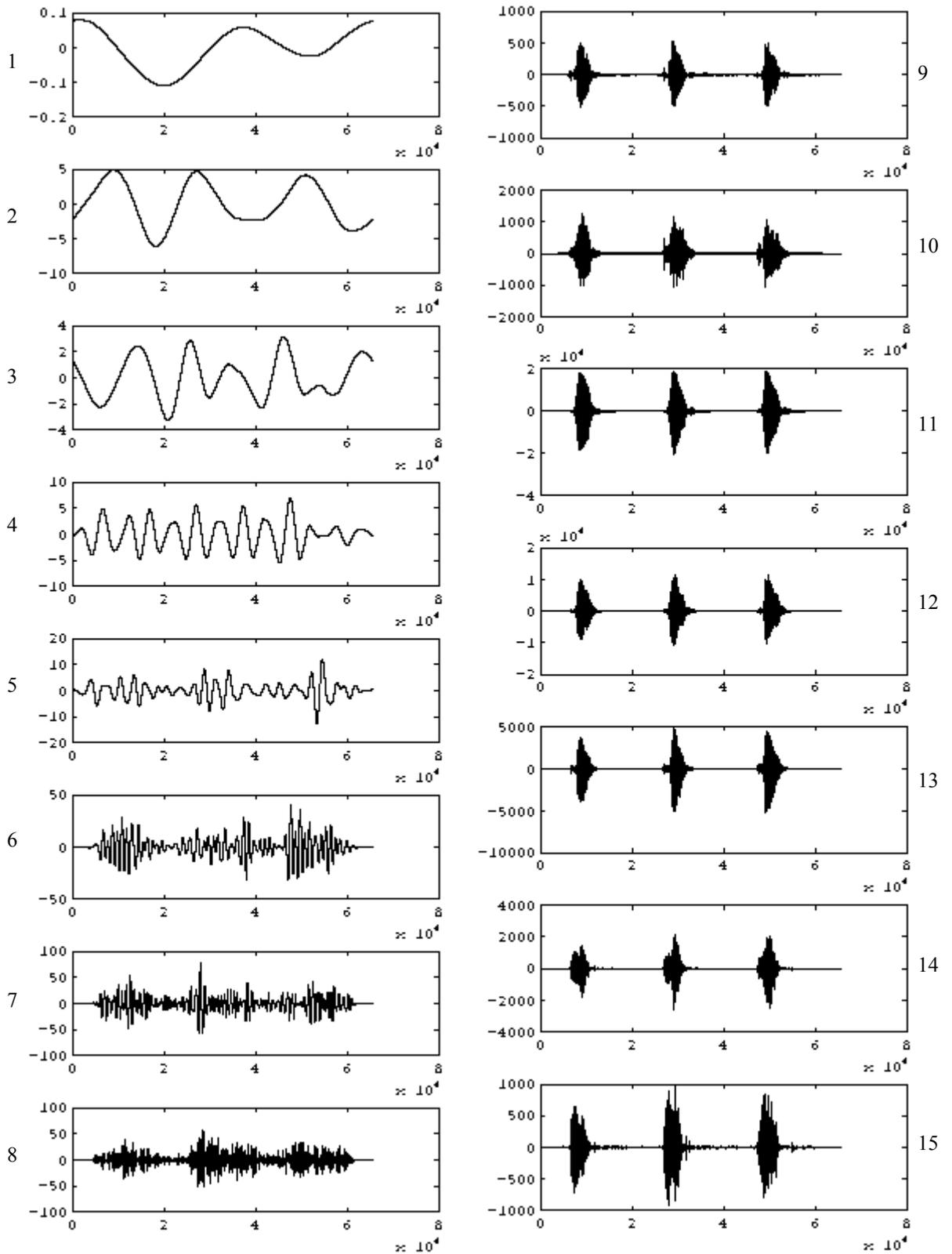
Nível 3: Nível inaudível. Revela patamares quadrados nas cristas que correspondem às pausas.

Distribuição de energia:



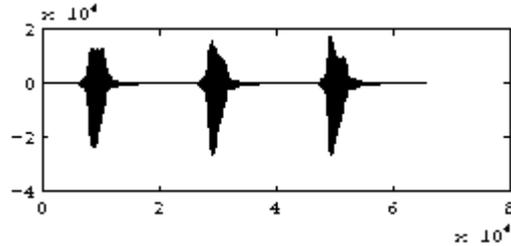
Mais de 99% de toda a energia do exemplo se concentram nos níveis de 9 a 12. O segundo gráfico ilustra a distribuição de energia pelos níveis numa escala logarítmica, com grandes variações de energia no topo do gráfico comprimidas em pequenos deslocamentos.

Flauta: Ataques, Expressivos, com *staccatos* - nota si4 (B4) na segunda oitava



Evento: três notas si4 (B4), da segunda oitava da flauta, executadas com acentos *staccatos*. Esta análise complementa a análise anterior de ataques com *staccatos*, desta vez analisando *staccatos* executados na segunda oitava da flauta.

Gráfico do sinal Original:

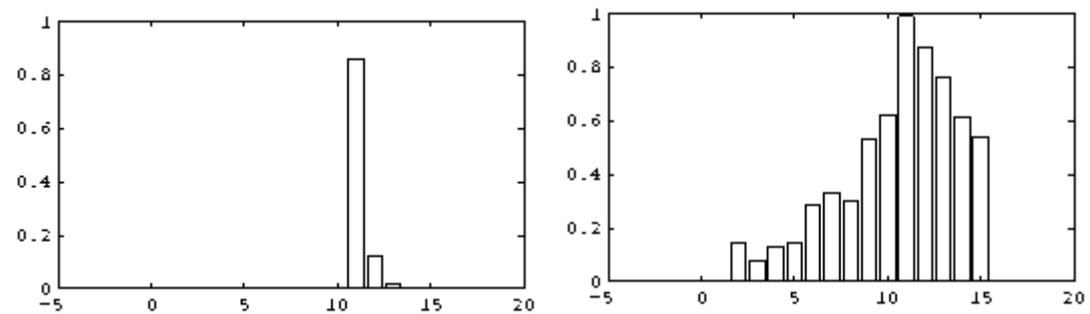


Extensão da amostra: 2^{16} (65536). **Wavelet de análise:** D16

Análise de níveis:

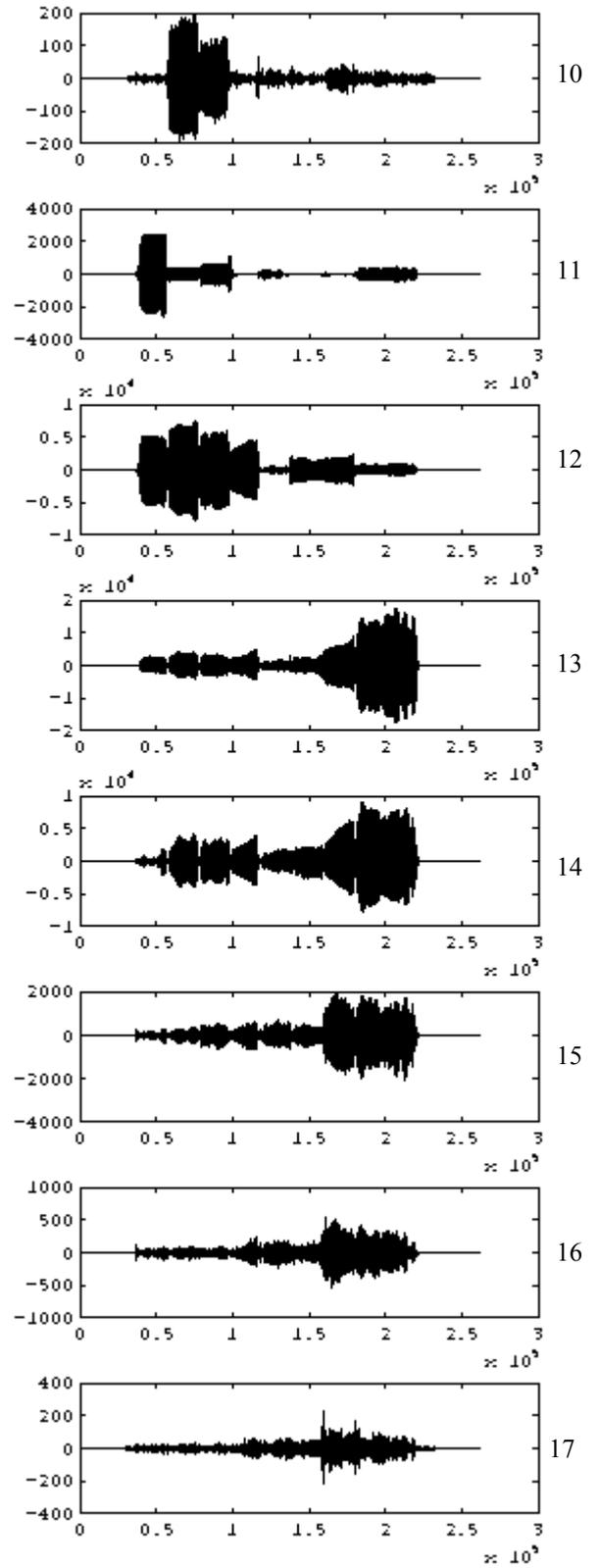
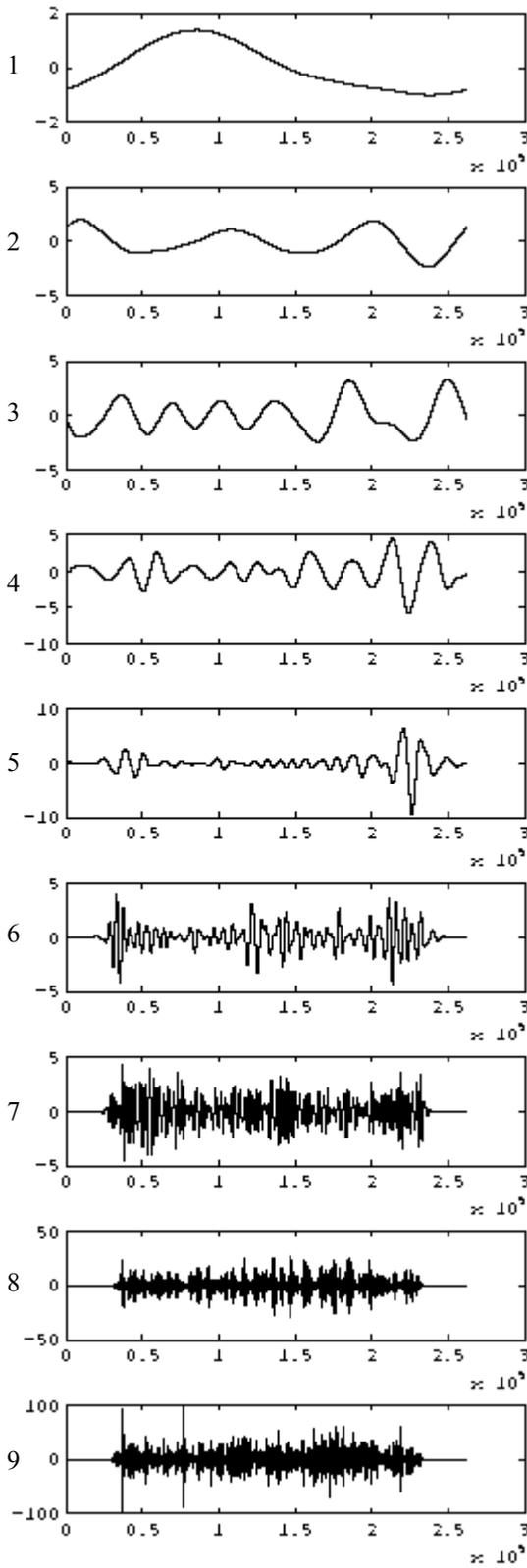
Não foram detectados nos níveis desta análise artefactos relevantes diferentes da análise anterior.

Distribuição de energia:



Mais de 99% de toda a energia do exemplo se concentra nos níveis de 11 a 13, sendo o nível 11 o mais energético. O gráfico à direita mostra o logaritmo da energia de cada nível pelo logaritmo da energia total.

Flauta: Fraseado, Ligado, Normal

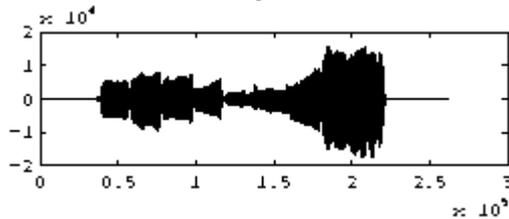


Evento: Fraseado ligado normal. Execução de um motivo simples, tema com fraseado ligado, notas executadas sem efeitos, mas com expressividade. Leve vibrato na última nota. O tema em questão foi executado realizando-se leves modulações ou gestos de sopro nas transições, de forma a destacar cada nota sem no entanto separá-las. A partitura abaixo ilustra o fraseado tocado.



O efeito pode ser conseguido utilizando-se uma modulação no sopro pelo movimento sutil da língua contra a face posterior dos incisivos superiores, como na pronúncia da sílaba "du" (ataque tipo "du"), e enriquece de expressividade a interpretação ligada, motivo pelo qual preferiu-se analisar este exemplo com esta interpretação específica.

Gráfico do sinal original:



Extensão da amostra: 2^{18} pontos (262144). **Wavelet de análise:** D16

Análises dos níveis:

Nível 17: Nível mais fino contém a maior parte do ruído de sopro. A sétima nota apresenta grande nível de ruído de sopro.

Nível 16: A forma de onda do nível já apresenta um melhor quantidade de segmentação (separação de notas). Ainda presença potente do ruído de sopro associado principalmente às sétima e oitava notas. O ataque da primeira nota também é bem localizado, consistindo num objeto de formato quadrilátero, rampa de subida alta e decaimento para a região de sustentação bem demarcados. As notas seguintes realizadas em fraseado ligado não apresentam tal característica de ataque. As duas últimas notas apresentam as maiores amplitudes e intensidades energéticas do exemplo.

A transição da **sexta para a sétima** nota (Eb4 para Ab4) consiste num passo de 5 semitons. Para transpor intervalos desta grandeza neste registro da flauta o intérprete tem de ajustar a embocadura de forma a estreitar levemente o orifício labial por onde flui o jato de sopro e concentrar o feixe de ar numa área menor (menor diâmetro) na direção do bucal da flauta. Isto propicia um jato de maior velocidade, pressão e concentração, necessário para excitar modos mais altos.

Tal ajuste de embocadura consiste num gesto que exige bastante habilidade a fim de se obter melhor foco de sopro sem aumentar o seu nível de ruído, o que naturalmente tende a ocorrer em vista da rápida alteração que é realizada na disposição dos lábios e na largura do orifício soprador. Naturalmente também o intérprete, ciente das exigências para excitar a nova nota mais alta em frequência, aumenta o fluxo de ar levemente para evitar desvios tonais, como por exemplo excitar uma oitava abaixo do desejado. O aumento do ruído de sopro portanto nas 2

últimas notas se dá em resposta a esses gestos musicais envolvendo a embocadura e um leve aumento na pressão de sopro.

Nível 15: A segmentação das notas é melhor que no nível anterior. Não se nota pela análise gráfica nenhuma estrutura especial nos ataques das notas intermediárias, somente a primeira possui um ataque diferenciado, caracterizado por uma fase de sopro com amplitudes altas e um decaimento na transição para a sustentação onde os tons harmônicos soam.

As duas últimas notas apresentam de novo a maior amplitude, apesar de também não exibirem nenhuma forma de ataque diferenciado. O colapso na última nota é rápido e não ocorre gradualmente.

Analisando-se as transições mais proximamente verificamos que existe um estrangulamento entre as notas, **antecedido por um colapso rápido mas suave e sucedido às vezes por um breve gesto de sopro**. Isto ocorre com maior realce entre a **segunda e a terceira** notas, e entre a **quarta e quinta** notas: a transição realizada não foi um *legato* perfeito, mas um *quasi-legato*, expressivamente marcado por uma leve modulação no fluxo de sopro, algo entre o ataque não ligado e o ligado (ataque do tipo "*du*"), produzindo um rápido colapso no tom precedente pela queda no fluxo de ar mas não interrompendo-o completamente.

Nas transições da **sexta para sétima**, e **sétima para a oitava** notas não ocorrem tal tipo de modulação expressiva no fluxo de ar, mas um estrangulamento entre elas localiza o ponto da saída de um padrão harmônico para outro. Importante notar que este nível já denota diferenças tonais entre as notas. Os níveis anteriores não realçavam o fato de que se trata de uma pequena melodia, com notas diferentes ascendendo a escala.

Nível 14: Amplitude cresce bastante para as duas últimas notas, aproximando sua dinâmica do *mezzoforte*. Estas últimas notas também estão localizadas num registro mais alto, e portanto suas séries harmônicas distribuem-se em níveis mais altos, logo contribuindo mais (em energia) nos níveis mais altos em comparação às outras notas.

As diferenças tonais entre as notas possibilita uma melhor estimativa das notas originais e por conseguinte do tema em si, apesar do timbre não lembrar timbres de flauta. O vibrato na última nota é em essência um modulação de amplitude, como pode ser verificado por análise gráfica da última nota e pela audição da mesma.

Nível 13: Nível mais energético. As duas últimas notas (em especial a última) apresentam as maiores intensidades. Revela o envelope mais suave e os tons mais puros, sem *clicks* ou estalos. A geração de bandas laterais espelhadas ao redor de divisores (frações racionais) da frequência de amostragem faz surgir *pitchs* ou tonalidades inexistentes no som original. Neste exemplo específico temos um exemplo importante dos efeitos que esse comportamento da análise *Wavelet* pode provocar. A audição do tema neste nível revela dois aspectos novos importantes:

1. Surgem novas vozes para algumas notas, arranjadas em acorde ou tendendo a compor uma função harmônica.
2. A melodia no nível é diferente da melodia original, pois que algumas notas estão em tons diferentes, havendo migrado intervalos variando de um semitom a vários tons.

A Figura 7 abaixo mostra uma boa aproximação para a melodia (harmonizada) percebida pela audição do nível.³

³ Diz-se aproximação porque as notas geradas nos níveis *wavelets*, usualmente produtos da integração das linhas espectrais presentes, apresentam *pitch* percebido (altura tonal) passível de interpretações diversas,

Figura 7 - Melodia percebida no nível 13 (sem indicação da dinâmica associada)



Um outro detalhe notável é a existência de um vibrato bem caracterizado para a terceira nota do tema, na forma de uma modulação de amplitude. A quinta nota também apresenta leve vibrato.

Nível 12: As 4 primeiras notas agora são a que apresentam as maiores intensidades. Por estarem num registro mais grave é natural que se concentrem mais nos níveis audíveis menos finos, que compreendem faixas de frequências abaixo das cobertas por níveis mais finos. A melodia neste nível também é diferente da melodia original, algumas notas estando deslocadas na escala. A escala, como no nível 13, também não se comporta propriamente como uma escala temperada. A Figura 8 abaixo mostra a melodia (harmonizada) percebida neste nível 12 transcrita em partitura:

Figura 8 - Melodia percebida no nível 12 (sem indicação da dinâmica associada)



A transição da **quinta para a sexta** nota localiza um gesto de sopro devido a alteração da embocadura para vencer o intervalo entre F4 e Eb4 (ajuste de foco de sopro). O efeito é o de uma elevação de intensidade na região de colapso da nota, que está correlacionado a uma proporcional elevação nas componentes harmônicas do tom, já que o foco do sopro estando ajustado não produzirá um suplemento de energia considerável no ruído. Isto explica, em parte, a não localização deste objeto em níveis onde o ruído de sopro é melhor caracterizado, mas nos níveis energéticos, onde a energia do sopro é convertida em vibrações harmônicas mais intensas.

Nível 11: A nota mais baixa (G3) apresenta as maiores amplitudes como era de se esperar. Este nível também revela aspectos interessantes novos: as quarta (D4), sexta (Eb4) e sétima (Ab4) notas estão praticamente ausentes neste nível, isto é, elas não apresentam contribuição relevante neste nível, e tornam-se praticamente pausas. A Figura 9 abaixo mostra a melodia percebida neste nível, transcrita em partitura:

Figura 9 - Melodia percebida no nível 11 (sem indicação da dinâmica associada)



A região central da quinta nota (F4) e a região de transição entre a sexta e a sétima notas apresentam componentes neste nível visualizáveis e audíveis.

Nível 10: A textura áspera (e grave) do som, o envelope descontínuo e sem suavidade apresentados neste nível traçam um perfil típico de ruído. Evidência de que gestos de sopro

que pode variar de pessoa para pessoa. A percepção de microtons, da presença de mais vozes em dinâmicas diferentes e mesmo da existência de outra escala (não temperada) também pode ocorrer.

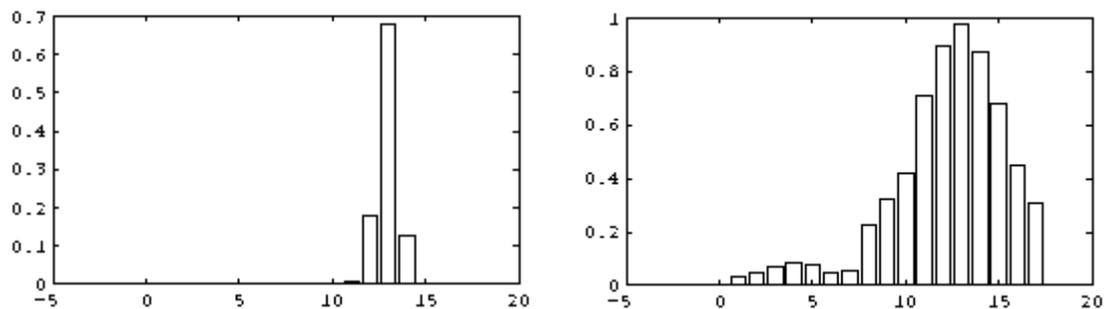
podem ser localizados neste nível é confirmada pela indicação de um objeto com amplitudes elevadas em relação a sua vizinhança na transição entre a quarta e a quinta notas. Estruturas nesta transição só foram localizadas antes no nível 15, onde o ruído de sopro e gestos de sopro são mais visíveis. A forma de onda do sinal neste nível liga todas as notas, não havendo localização explícitas de estrangulamentos ou outro objeto empregável para efeito de segmentação. A segunda e a terceira notas aparecem intensas neste nível, as restantes não tendo participação expressiva no nível.

Nível 9: Envelope das notas não é suave, e está pontilhado de transientes intensos localizando gestos musicais pontuais, intensos, periódicos ou não. A segmentação das notas é inexistente, e o envelope não guarda familiaridade alguma com o envelope original. Os picos mais intensos do nível localizam os pontos de ataque da primeira nota e a transição entre a segunda e a terceira notas.

Nível 8: Sem segmentação de notas. Sem evidência ou localização de gestos musicais. Presença de uma modulação em amplitude distribuída por todo o sinal do nível. Formato do envelope pontilhado de transientes, semelhante ao nível anterior. Pouco energético, e quase inaudível.

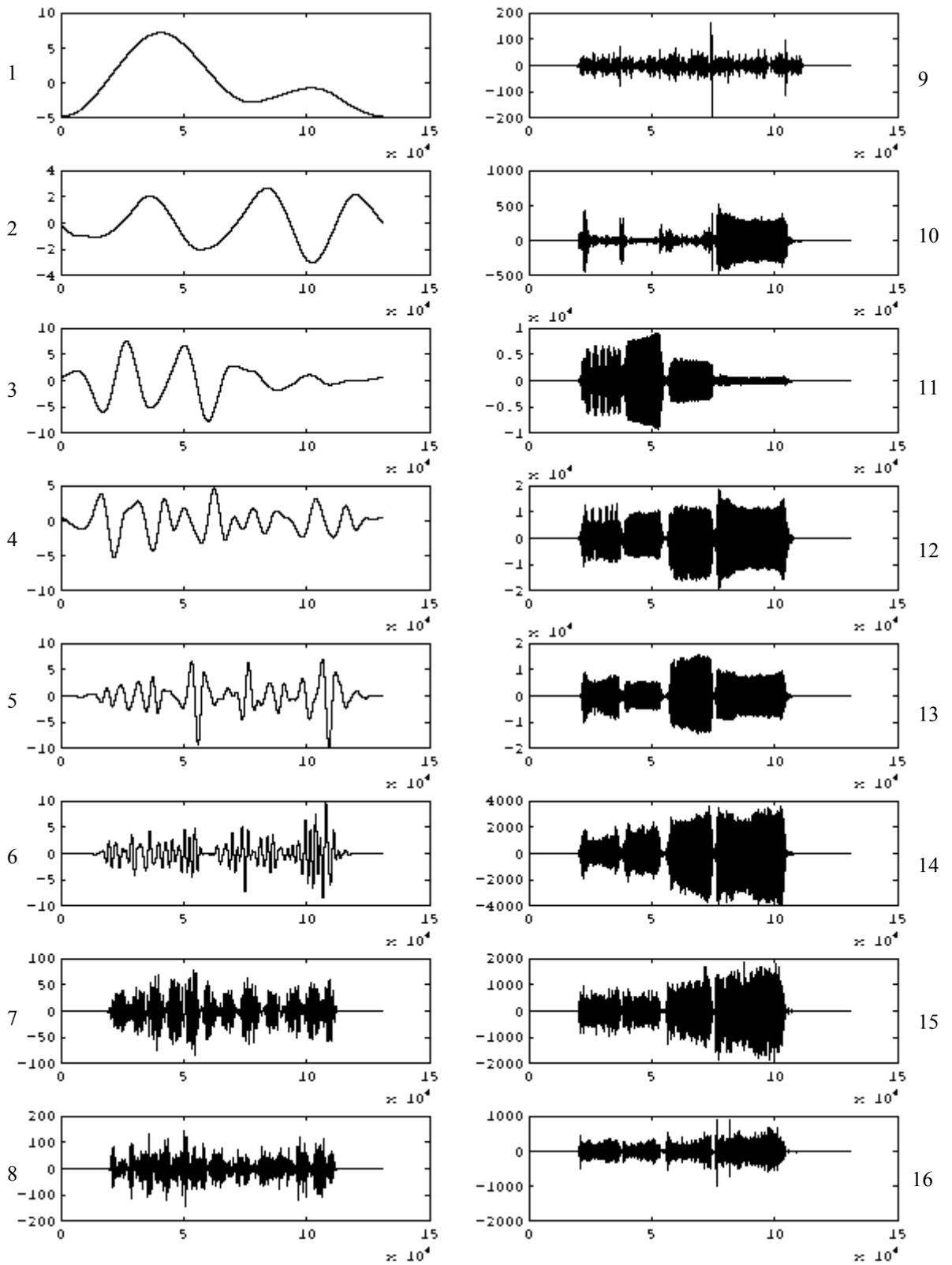
Nível 7: Inaudível. Envelope pontilhado de transientes, como nos níveis anteriores.

Distribuição de energia:



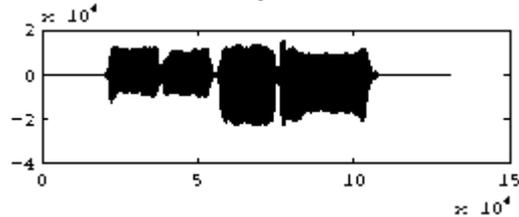
Mais de 99% de toda a energia do fraseado do exemplo se concentra nos níveis de 11 a 14, sendo o nível 13 o mais energético. O gráfico à direita mostra que a contribuição dos níveis inferiores ao 9 é bastante insignificante em termos energéticos.

Flauta: Fraseado, Ligado, Acentuado



Evento: As quatro últimas notas do tema melódico do exemplo anterior (fraseado ligado normal), agora interpretado aplicando uma leve acentuação às notas e mantendo-se a ligadura articulada por sutis ataques tipo "du" (como se pronunciando a sílaba "du").

Gráfico do sinal original:

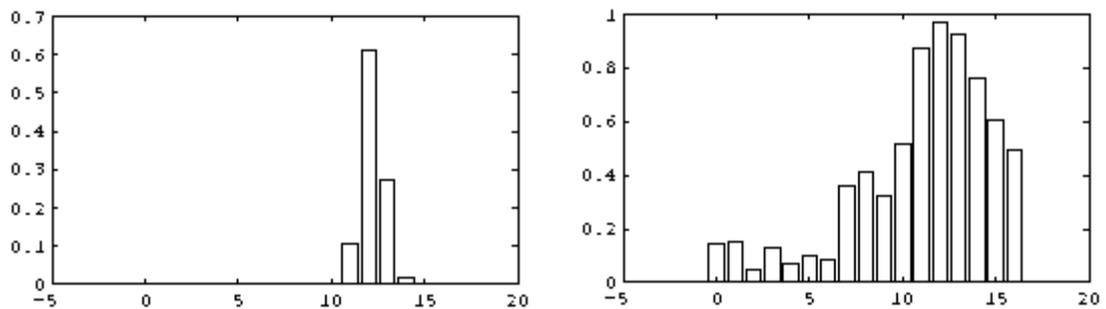


Extensão da amostra: 2^{17} pontos (131072). **Wavelet de análise:** D16

Análise dos níveis:

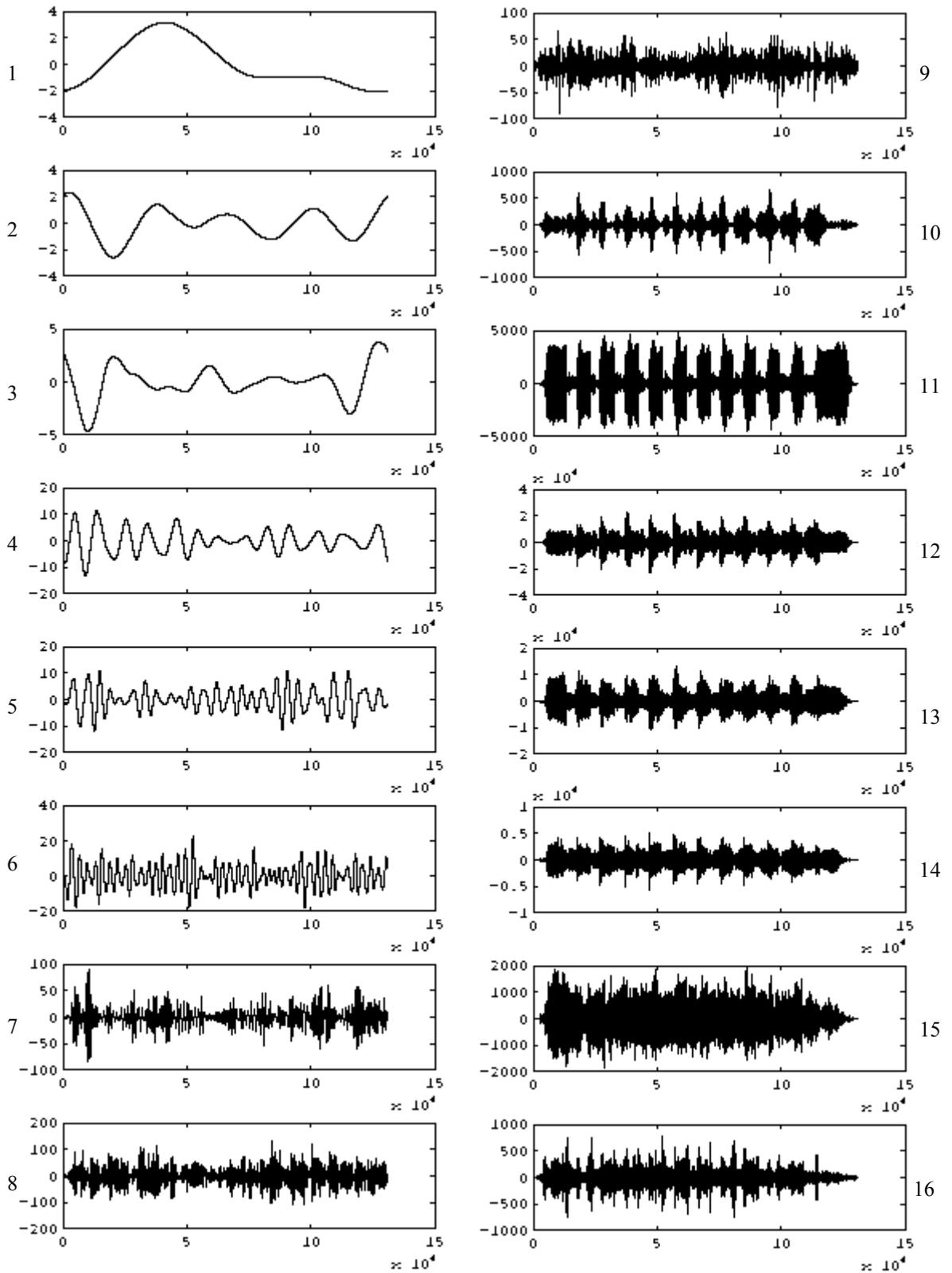
Nível 10: transientes intensos localizam neste nível os acentos. Observa-se que acentos implicam num aumento brusco de intensidade neste nível nas redondezas do leve ataque que foi dado às notas, mais precisamente após a estabilização dos tons (início da região de sustentação).

Distribuição de energia:



Mais de 99% de toda a energia do exemplo se concentra nos níveis de 11 a 14, sendo o nível 12 o mais energético. Verifica-se que a acentuação altera o padrão de distribuição de energia, como pode-se deduzir pela simples comparação dos gráficos acima com os do exemplo anterior (fraseado ligado normal). Naquele exemplo a distribuição no gráfico à direita foi mais regular, e relativamente simétrica em relação ao nível mais energético. Neste exemplo, verifica-se que o acento implica num aumento na quantidade de energia nos níveis adjacentes ao nível mais energético.

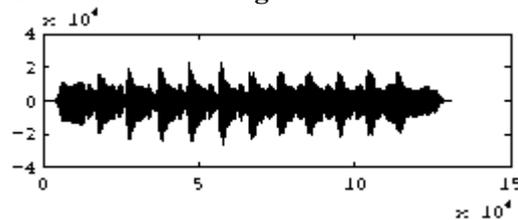
Flauta: Fraseado, Ligado, Trêmulo dedilhado - notas sol4 (G4) e si4 (B4)



Evento: Trêmulo em fraseado ligado das notas sol4 (G4) e si4 (B4) na segunda oitava da flauta transversa. Pela capacidade de captação direcional do microfone utilizado, o gestual do dedilhado foi capturado, mas este é um evento acústico indesejado porque não foi intencional nesta técnica de interpretação destacar o ruído do chaveamento (em algumas técnicas pode ser intencional a captura destes ruídos).

O trêmulo é cíclico apresentando um período onde as duas notas soam em sequência. O ciclo ativo é de aproximadamente 60% do período ocupado pelo sol4 e 40% ocupado pelo si4. O si, nota mais alta, é emitido quando o dedilhado é aliviado e a coluna de ar reduz seu comprimento. O sol é novamente emitido quando o dedilhado retorna e força o instrumento a vibrar uma coluna de ar maior. Pela dinâmica física envolvida na excitação de diferentes comprimentos de ar em sequência e rapidamente, é natural esperar que o si seja mais rápido e fácil de se estabelecer, enquanto o sol, por excitar uma coluna de ar mais longa, demora um pouco mais para estabelecer as vibrações correspondentes, devido a própria inércia acústica do instrumento. Estas peculiaridades agem influenciando o ciclo ativo do período do trêmulo.

Gráfico do sinal original:



Extensão da amostra: 2^{17} pontos (131072). **Wavelet de análise:** D16

Nível 16: Ruído de sopro do trêmulo. Há variações amplas e periódicas, com a intermitência periódica de transientes com alta declividade. Estes podem estar captando o chaveamento do dedilhado, o que pode ser ou não mais facilmente confirmado ouvindo-se o nível e comparando-o com o nível mais grosso bem abaixo onde também aparecem componentes ruidosos, mas numa menor frequência.

Os transientes com alta declividade periódicos marcam os instantes em que há novo ataque na alternância das notas envolvidas no trêmulo. Notar que além do dedilhado mais difícil, o trêmulo também envolve um ajuste mais crítico e também periódico da embocadura, que pode ficar bem difícil para alguns intervalos e notas específicas.

Nível 15: Forma de onda não é suave: as variações de amplitude são rápidas e intensas indicando um padrão ruidoso, de envelope irregular e pontilhado de transientes. A textura sônica é ruidosa, de um ruído de alta frequência, associado a componentes harmônicas que poderiam ser melhor extraídas com uma decomposição com *Wavelet packets*.

Analisada mais proximamente, em um período do trêmulo, observa-se que a nota sol4 apresenta forma de onda com menor suavidade e maior teor de ruído. As transições não são segmentadas ou de localização óbvia, mas acredita-se poder melhorar a definição dos instantes de transição com o uso de ferramentas *Wavelets* que discriminem ou dividam melhor o espectro de níveis muito finos, como esse.

Nível 14: O sinal apresenta objetos (transientes) periódicos, que localizam os instantes de emissão inicial ou excitação das notas sol4 do trêmulo. A súbita alteração do comprimento da coluna de ar na transição do si4 para o sol4 produz um transiente intenso neste nível, dando uma conformação peculiar para o trêmulo neste nível, e possibilitando uma melhor estimativa do

período de trêmulo. As transições do sol4 para si4, que implicam numa súbita diminuição da coluna de ar vibrante, forçam o colapso das vibrações do sol antes que as vibrações harmônicas do si4 se fortaleçam. O efeito disso é uma leve depressão das amplitudes neste nível nas regiões de transição do sol para o si.

O transiente intenso que marca o "ataque" das notas sol4 pode também estar reforçado pelo ruído produzido no ato de se fechar as teclas correspondentes para o sol. No entanto, não há evidências da captação relevante de ruído de chaveamento, uma vez que a liberação do dedilhado para emitir a nota si4 também produz um ruído característico o qual não está sendo discriminado no nível.

Nível 12: Nível mais energético. O envelope, entretanto, não é tão regular quanto observado nos exemplos anteriores nos níveis mais energéticos. O conteúdo porém é essencialmente o proveniente das séries harmônicas, e a forma da onda apresenta contornos mais suaves que nos níveis ruidosos. Todavia, o nível não deixa de exibir picos intensos nas regiões de transição do si4 para o sol4.

Nível 11: A nota sol apresenta amplitudes maiores que o si - cerca de 6,17 em média maiores (calculado com a função **amp** do **Wavesynt** nas regiões do si (226) e do sol (1400)). Evidentemente, o sol é mais baixo (em frequência), e seus níveis mais energéticos são portanto mais baixos que os níveis mais energéticos para si4. O nível portanto serve também a função de segmentador ou separador de notas, uma vez que as notas envolvidas, devido à distância tonal entre elas, apresentam diferentes níveis de amplitude. Supondo que haja interesse em separar a nota sol4 da si4 poder-se-ia utilizar da capacidade do nível de separá-las (até certa extensão). Vale dizer também que tal efeito de separação sempre será mais acentuado nos exemplos de trêmulo, uma vez que os intervalos entre as notas é maior que por exemplo no trinado (que é de 1 semitom).

A análise sonora do som do nível mostra que as notas do trêmulo apresentam tonalidades muito próximas, sendo possível diferenciá-las mas mais difícil inferir as tonalidades e mesmo o intervalo tonal entre elas.

Nível 10: Os picos de máximo (e de mínimo) se situam **sobre as regiões centrais das notas**, e não sobre o ataque. Isso mostra que o nível é sensível aos componentes harmônicos que só são realçados ou alcançam máxima excitação nas regiões de sustentação de cada nota. Nas regiões de transição, observa-se um estrangulamento substancial nas amplitudes, marcando os pontos onde as vibrações estacionárias são débeis porque o padrão vibratório está variando. O som do nível já é bem grave, mas ainda revela ser dominado pelos padrões harmônicos de cada nota.

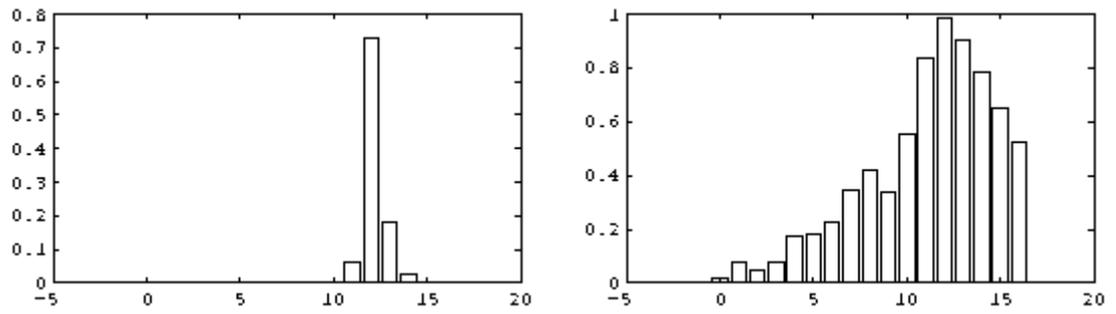
Cumprir notar que apesar de apresentar uma ótima segmentação das notas, evidenciando as posições centrais destas com as maiores amplitudes, as notas soam semelhantes, não sendo possível inferir sobre as reais tonalidades (alturas) das notas originais.

Nível 9: Nível pouco energético. Não demonstra segmentação de notas, possui textura irregular e contornos não suaves. Também exibe uma textura sônica ruidosa, realçando o ruído de fundo, não gerado a partir dos eventos musicais.

Nível 8: Quase inaudível, muito semelhante ao nível 9 anterior.

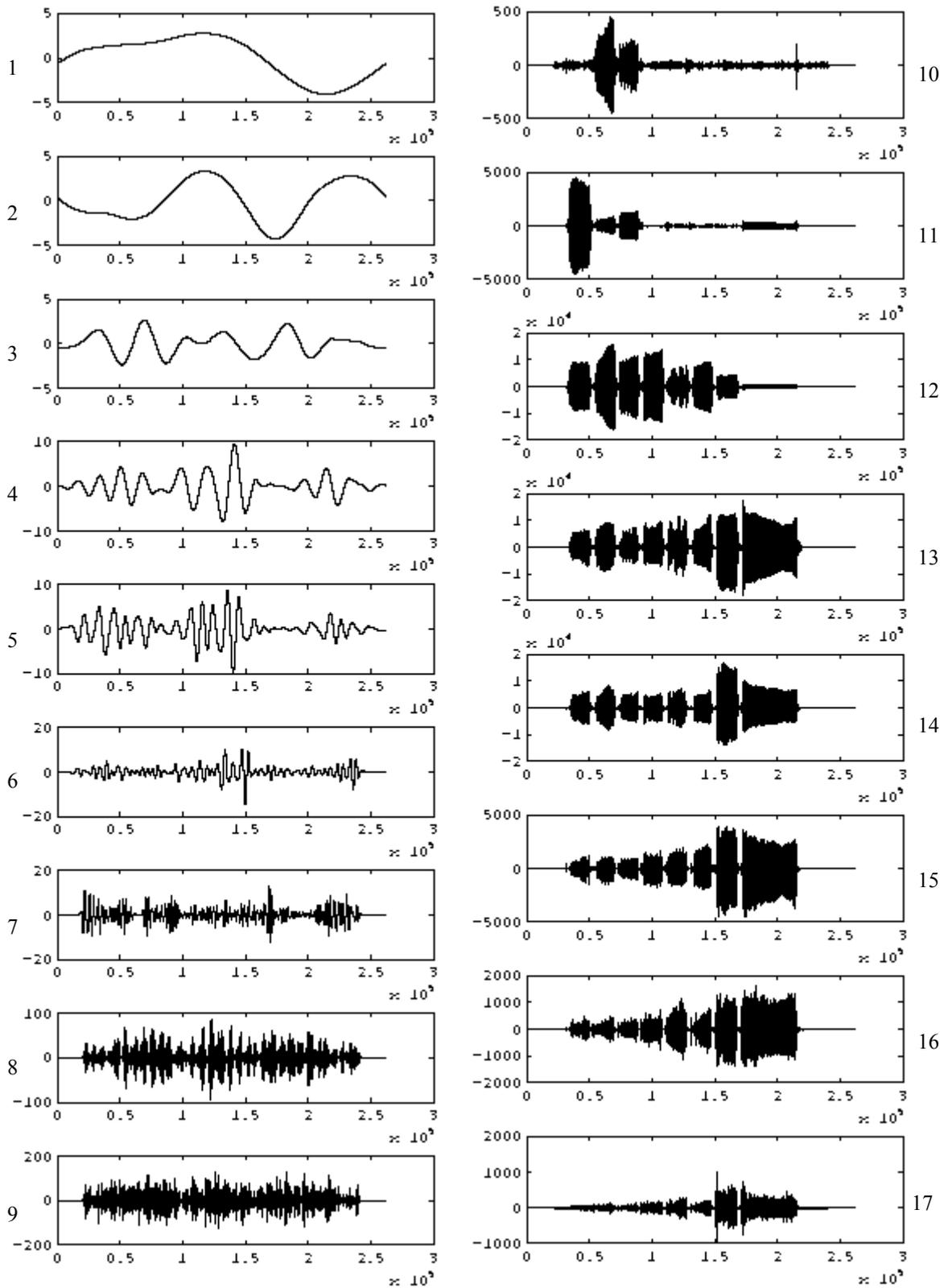
Nível 6: Inaudível, exceto pelos *clicks*.

Distribuição de energia:



Mais de 99% de toda a energia deste fraseado ligado (trêmulo dedilhado) se concentra nos níveis de 11 a 14, sendo o nível 12 o mais energético, com mais de 70% de toda a energia total. O gráfico ao lado, à direita, mostra a contribuição energética de cada nível, numa escala logarítmica.

Flauta: Fraseado, Não Ligado, Normal

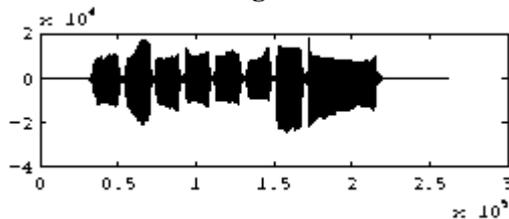


Evento: Fraseado não ligado normal. Execução de um tema melódico simples, o mesmo executado no exemplo com fraseado ligado, aqui porém interpretado com um ataque a cada nota separadamente. A partitura abaixo ilustra o fraseado tocado.



O exemplo destaca a segmentação das notas, isto é, o fato de que cada uma sofre um ataque específico, consistindo no que em música se denomina fraseado não ligado (*non legato*).

Gráfico do sinal original:



Extensão da amostra: 2^{18} pontos (262144). **Wavelet de análise:** D16

Análise dos níveis:

Nível 17: Ruído de sopro. Segmentação das notas é boa, sendo possível localizar com precisão os momentos de ataque, colapso e pausas. No entanto a audição não diferencia os tons das notas, isto é, não é possível inferir sobre as notas tocadas, o som é um ruído de alta frequência, e o envelope das notas varia em amplitude na medida em que a pressão de sopro é maior.

A avaliação tanto auditiva quanto gráfica do sinal no nível mostram que as duas últimas notas foram executadas com maior pressão de sopro. Os ataques dessas notas foram mais intensos, com uma súbita elevação nas amplitudes nesta região, seguidas por um decaimento e uma região de sustentação com amplitudes também mais elevadas que nas notas anteriores. Os ataques das notas anteriores não apresentam grandes variações de amplitude, mas estão bem caracterizados por rampas súbitas, uma elevação seguida por decaimento, e portanto os momentos de *onset* ficam bem localizados.

Uma explicação suplementar para o fato de as primeiras notas não apresentarem amplitudes tão altas quanto as últimas é que estas são notas mais altas em frequência, estão num registro mais alto, e aquelas são notas mais graves, e devem apresentar concentração de energia nos níveis inferiores.

Os colapsos de todas as notas apresentam conformações semelhantes: descidas relativamente íngremes, sendo mais curtas que os colapsos vistos no sinal original. Isso explica-se porque neste nível há grande contribuição do ruído de sopro, que é interrompido ao final de cada nota, e dos modos harmônicos mais altos, que na ausência do sopro morrem mais rápido. Quase todas as notas apresentam um leve crescendo em direção ao final, como indica a elevação gradual e sutil das amplitudes no decorrer das notas neste nível.

Nível 16: A análise gráfica do nível mostra em crescendo em direção ao final do tema, a primeira nota menos intensa e a última a mais intensa. Tal fato não está evidente no sinal original.

Confirmando o que foi verificado também no nível anterior, as notas individualmente mostram um crescendo em direção ao seus finais. A última nota possui em envelope que poderia ser aproximado por um quadrilátero, com extremidades em rampas íngremes. A textura das formas

de onda é pontilhada de descontinuidades e picos aleatoriamente distribuídos, não sendo portanto suave. A evidência de que se trata de um fraseado não ligado é óbvia pela verificação de estrangulamentos nas formas de onda entre as notas. Os ataques não são acentuados, e possuem descrição semelhantes aos do nível anterior, bem caracterizados com uma rampa de subida, picos locais elevados, e um decaimento seguido pela região de sustentação.

Os colapsos são curtos e íngremes, resultados de gestos de interrupção do sopro impostos pelo intérprete em cada nota, a fim de destacá-las. Aliás este destaque pode ser a estrutura mais relevante na impressão musical do tema tocado sob este fraseado, visivelmente menos expressivo e soando mais "mecânico" que o exemplo com fraseado *legato*. O som já fornece indícios sobre a tonalidade das notas, e ainda possui uma textura bastante ruidosa.

Nível 15: A audição do nível mostra os mesmos intervalos tonais entre as notas, isto é, é possível inferir com maior precisão acerca das notas originais e suas tonalidades. As duas últimas notas ainda são as mais intensas, com envelopes ainda aproximáveis por quadriláteros. A forma das ondas não possuem um contorno suave.

Os ataques ficam bem localizados e caracterizados neste nível, conforme pode-se visualizar e ouvir. A audição em 11025 Hz tornará evidente os gestos de sopros envolvidos nos ataques. A textura de cada nota é diferente. As segunda e quinta notas por exemplo apresentam um "vibrato" ao nível 15, sendo que a frequência de modulação (de amplitude) é maior na segunda nota e mais lento na última.

Outra característica extraída da análise do nível se refere à modulação de amplitude observada, semelhante a um vibrato (de amplitude, para flauta), e onde a frequência modulante varia dentro de uma única nota. Na modulação na quinta nota, por exemplo, a frequência de batimento inicial é mais lenta, e acelera em direção ao final. Isto sugere a existência de uma **correlação com o aumento de intensidade no decorrer da nota** (*crescendo*) percebido nos níveis mais finos, em especial no nível 16. Nas regiões onde o sopro é menos intenso o "vibrato" é mais lento, e cresce em velocidade nas regiões onde há maior energia de sopro aplicada.

Note entretanto que a correlação proposta se aplica a "vibratos" verificados nos sinais de níveis específicos, e não ao sinal original. A audição da quinta nota original não revela nenhum vibrato intencional relevante.

Nível 14: A sétima nota saiu mais destacada neste nível, mostrando as maiores magnitudes. Pela análise gráfica verifica-se que o nível permite uma boa caracterização da região do colapso das notas, evidenciando a forma de onda nestas regiões diferenciadamente dos outros níveis. A região de ataque apesar de bem localizada também, não o é tanto quanto as regiões de colapsos.

Modulações de amplitude ("vibratos") aparecem em quase todas as notas, em especial na terceira ("vibrato" mais rápido), quarta e quinta ("vibrato" lento). A sétima nota apresenta um envelope quase quadrilátero, e possui o envelope de contornos menos suave do nível. A oitava nota apresenta também um envelope muito pouco suave. A característica mais evidente do nível é a presença de modulações de amplitude em algumas notas, frutos do processamento da análise *Wavelet*, gerando bandas laterais espelhadas equidistantes de frequências que são frações racionais de números inteiros da frequência de amostragem.

Em exemplos de fraseados as análises sempre procuram abordar aspectos mais qualitativos e identificar estruturas que diferenciam as notas entre si, não objetivando análises profundas sobre o espectro dos níveis.

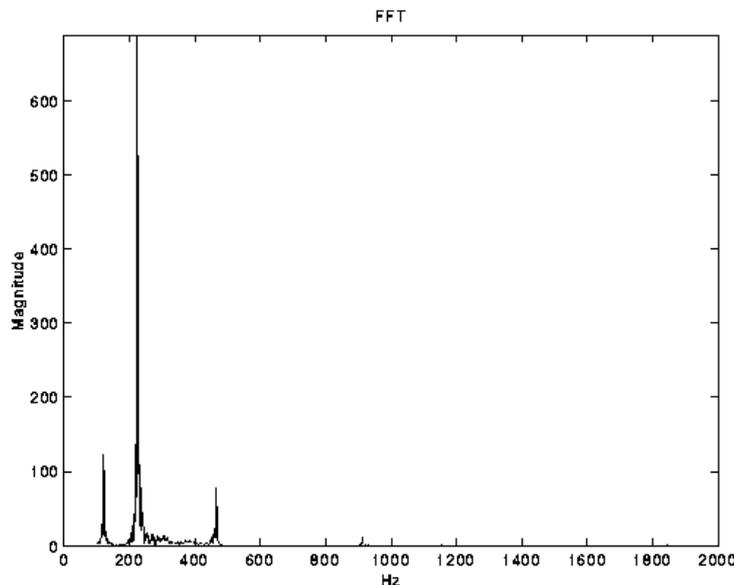
Nível 13: É o nível mais energético.

Nível 11: Nível relativamente suave. As três primeiras notas são mais intensas, a segunda entre elas a de menor amplitude. A última nota também aparece.

Característica do nível é a ocorrência de modulação de amplitude ("vibrato" de baixa frequência, como um batimento) em todas as notas visíveis e audíveis. Dada a inexistência de vibratos intencionais no sinal original, procurou-se verificar a origem desta modulação em amplitude que ocorre no nível.

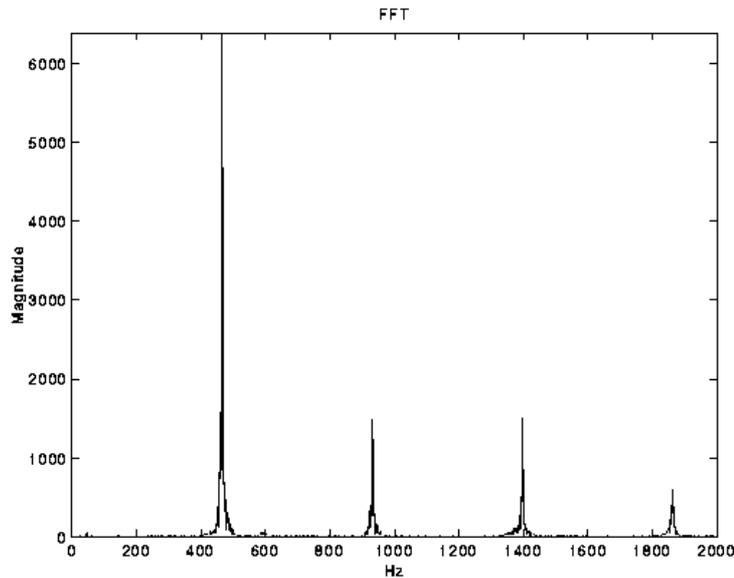
A terceira nota, que apresenta graficamente a mais clara flutuação característica de modulações, foi analisada para este propósito. Sua FFT revelou três bandas com elevações destacadas: uma central em 223,407 Hz, exibindo a maior magnitude (de 5 a 6 vezes maior que as outras duas), uma à esquerda em 121,1243 Hz, e a terceira à direita, em 465,65 Hz, como na figura abaixo.

Figura 10 - FFT da terceira nota no nível 11



As bandas laterais mostraram ser na verdade "raias espelhadas" da banda central, respectivamente espelhadas em relação a 172,2656 Hz (1/256 da frequência de amostragem) para a banda esquerda, e em relação a 344,5313 Hz (1/128 da frequência de amostragem) para a banda direita. A FFT da nota original (figura abaixo) não revela a existência de componentes relevante em 223,4 Hz (a banda esquerda identificada no nível acima) ou em 121,1243 Hz. No entanto, a banda lateral direita identificada acima (465,6555 Hz) corresponde à fundamental da nota Bb3 tocada, o que nos mostra que ela é a raia original, sendo as outras duas espelhos desta criados pelo processamento da análise *Wavelet*.

Figura 11 - FFT da terceira nota (sinal original)



A frequência modulante, visível no gráfico do nível, corresponde à metade do intervalo de frequência entre as bandas mais intensas da FFT na Figura 10 (a esquerda e a central).

Nível 10: As únicas notas relevantes são a segunda e a terceira. O som do nível já é bem grave⁴. A textura sônica e também a textura das formas de onda, e logicamente o espectro denso na faixa coberta pelo nível, são indícios da composição basicamente ruidosa existente por todo o sinal no exemplo. O envelope não é suave, não segmenta as notas, nem serve à localização de estruturas vibrantes, harmônicos, entre outros.

Um aspecto interessante entretanto está na capacidade do nível em **localizar instantes de interrupção do sopro**, como bem o faz na última nota. O pico local que sobressai ao final da última nota localiza exatamente o ponto onde o fluxo de ar é interrompido, ou seja o início do colapso da última nota. Note que a localização do ponto de interrupção de ar só foi relevante para a última nota, ou porque ela foi definitivamente o final, ou talvez porque fosse a mais alta em frequência (*pitch*) e o bloqueio de ar para colapsar notas agudas requer um gesto de interrupção total de sopro com características diferentes dos gestos interruptores anteriormente usados para colapsar as notas anteriores mais baixas em frequência. Mais pesquisa experimental sobre esse fato é necessário para clarear estes aspectos e direcionar a hipótese para uma conclusão correta.

Nível 9: Não possui segmentação de notas, não apresenta envelope suave, e o ruído grave se distribui por todo o exemplo. Quase inaudível.

Nível 8: Como o nível anterior, mas com uma modulação em amplitude mais visível e ligeiramente mais suave que o nível anterior. Não revela nenhum aspecto musical importante, e é pouco energético.

Nível 7: Envelope do nível se estende a regiões além do envelope do som original, mostrando que ele é um péssimo localizador de estruturas acústico-musicais. As regiões de cristas das

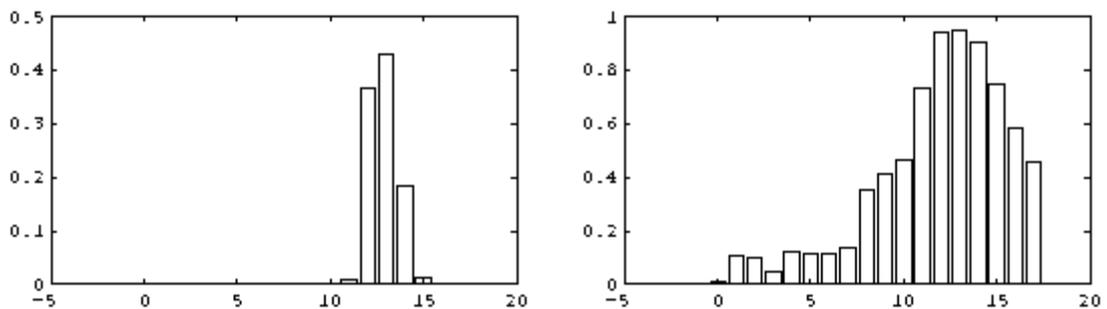
⁴ Para o leitor que aprecia o gênero ficção científica, o som deste nível (o 8º nível mais fino) assemelha-se ao ruído grave do "gerador warp" da nave *Enterprise*, do seriado televisivo *Jornada nas Estrelas, Nova Geração*.

formas de onda, isto é, onde elas apresentam maiores amplitudes, coincidem com as regiões no sinal original onde existem **pausas ou baixas amplitudes**, como sobre as transições, e nas regiões correspondentes às sustentações das notas, o sinal neste nível apresenta baixas amplitudes. Isto sugere que o nível seja sensível aos ruídos de fundo, ruídos ambiente ou gerados pelo processamento do sinal digital, mostrando maiores amplitudes nas regiões onde estes componentes estão em evidência. Nas regiões de sustentação das notas, onde a composição harmônica prevalece sobre os sinais mais fracos ruidosos, o nível revela baixas amplitudes.

Nível 6: Mais ou menos como o anterior, inclusive com a propriedade de ser o negativo do sinal original, i.é, revelar picos de amplitude onde existem pausas no sinal original. No entanto esta propriedade é evidente para algumas notas, como por exemplo a quinta e a sexta notas, e pouco evidente para outras, como por exemplo para as terceira e quarta notas, não consistindo em uma regra geral.

Nível 5: Apresenta maiores amplitudes para notas médias, ao meio do exemplo, e as menores para as notas mais altas. Não revela qualquer correlação facilmente identificável com os gestos musicais ocorrendo no exemplo.

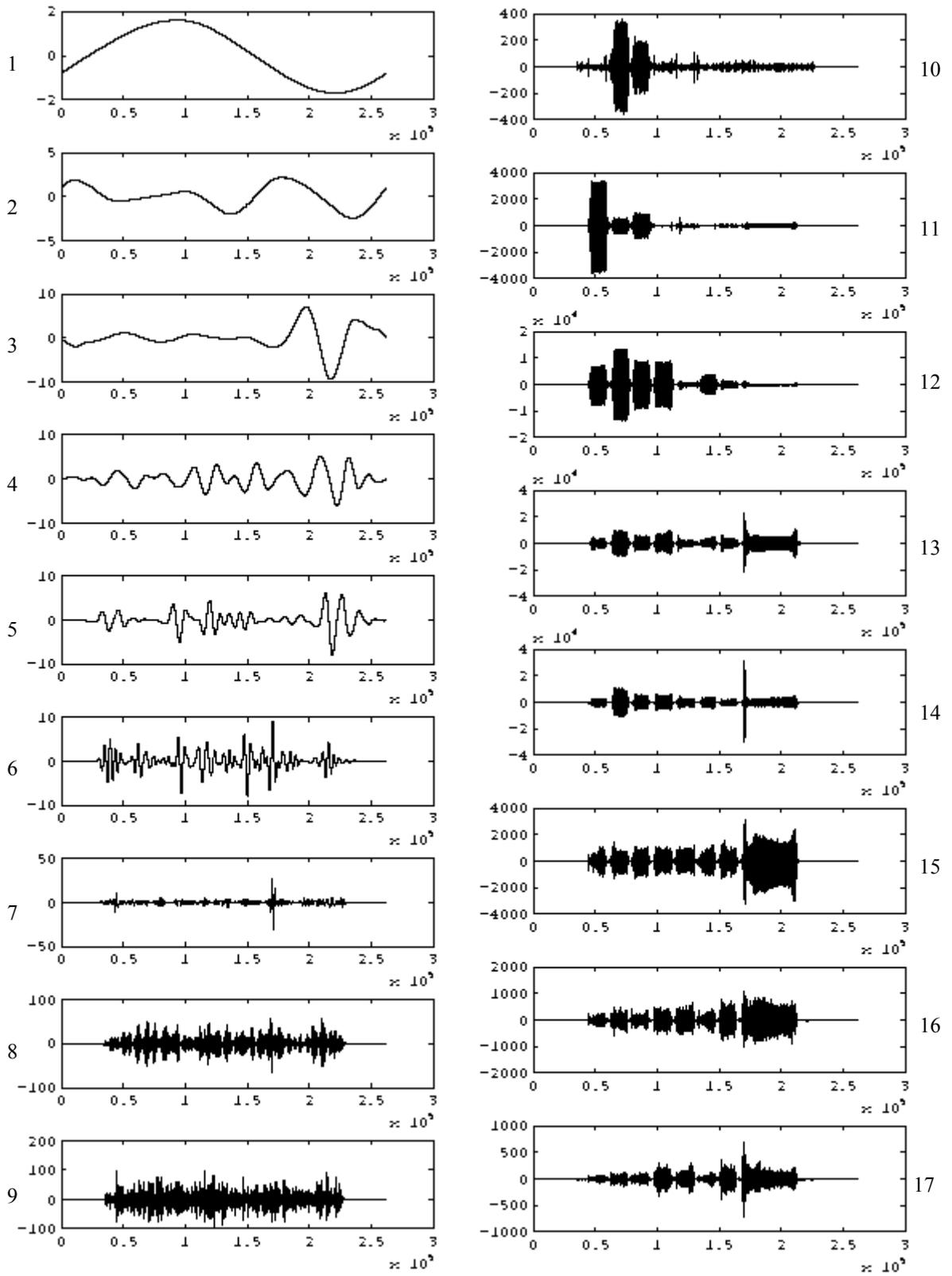
Distribuição de energia:



Mais de 99% de toda a energia do exemplo concentra-se nos níveis de 11 a 15, sendo o nível 13 o mais energético.

O gráfico à direita mostra o logaritmo da energia de cada nível pelo logaritmo da energia total. Os níveis inferiores ao nível 9 (inclusive) não contribuem significativamente em termos energéticos, e reconstruções desprezando-se os coeficientes destes níveis não exibem perda de qualidade significativa no som final.

Flauta: Fraseado, Não ligado, Acentuado

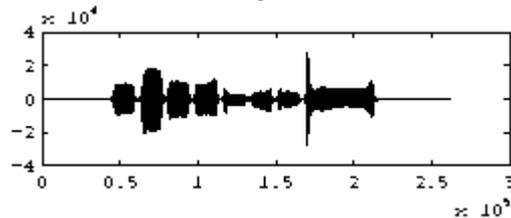


Evento: Fraseado não ligado acentuado. O exemplo é a execução do tema melódico introduzido nos exemplos anteriores (fraseado normal ligado e não ligado), executado agora em fraseado não ligados com as notas acentuadas, com ataques do tipo "tu".



A interpretação foi realizada de modo natural, e apresentou um realce no nível de ruído de sopro. A última nota expõe o maior realce no ataque. Ouvida a 11025 Hz ela mostra claramente que devido ao ataque súbito e intenso (acentuado) harmônicos mais altos foram excitados no ataque por um curto intervalo de tempo, estabilizando em seguida a tonalidade correta a caminho da região central de sustentação.

Gráfico do sinal original:



Extensão da amostra: 2^{18} pontos (262144). *Wavelet* de amostra: D16

Análise dos níveis:

Nível 17: Nível característico de ruído de sopro. Última nota mais destacada, com maiores intensidades relativas, por dois motivos:

Por ser a última de uma sequência acentuada, e a final do motivo temático, sua interpretação foi a mais intensa e o ataque mais realçado para fechar o tema com expressividade

Ela também é a nota mais alta em *pitch* (tom) da sequência, logo sua concentração energética se desloca em direção aos níveis mais altos.

Impossível de reconhecer os tons (as notas) nem o instrumento (timbre) pelo som do nível. O envelope não é suave, existem pulsos e transientes. Da análise gráfica vê-se que para as últimas duas notas os ataques caracterizam uma rampa abrupta no início das notas, marcando o ataque com as mais altas amplitudes nas notas.⁵ As amplitudes na região central das notas tendem a atingir uma dinâmica homogênea. As notas mais baixas (em frequência) também caracterizam os ataques de modos semelhantes, porém em menores amplitudes.

Os colapsos ficam bem marcados por uma abrupta queda nas amplitudes, e uma análise mais próxima (magnificando esta região) revela que formam uma sequência de curtíssimos intervalos com amplitudes sucessivamente menores. Um ruído de fundo de baixa amplitude povoa todo o exemplo.

Subida e descida das notas íngremes sugerem que uma boa aproximação para o envelope no nível seja na forma de um quadrilátero, no início e no fim. Gesto bucal associado com

⁵ Um exame da região do ataque magnificada mostra que as transições da pausa para o ataque não constituem uma "descontinuidade" tão abrupta como uma rampa, e sim constituem transientes com uma rápida variação crescente porém gradual de amplitude em um curto espaço de tempo, levando a uma sobrelevação e a um decaimento em seguida.

movimentos da língua e à respiração são localizados entre a penúltima e última notas. **O nível é bastante sensível a objetos e gestos associados ao sopro e suas variações.**

Nível 16: Uma visão panorâmica dos envelopes das notas sugere que uma boa aproximação para eles seria na forma de um quadrilátero, com a região inicial do ataque apresentando uma sobrelevação, caracterizando o acento neste nível. Uma visualização mais próxima destas regiões todavia mostram que não há realmente degraus descontínuos presentes no sinal, mas sim rampas, com variações de amplitude intensas. As últimas notas, mais altas em frequência, se encaixam melhor nesta descrição acima.

Os ataques ficam bem caracterizados neste nível, assim como os gestos bucais entre as notas, oriundos do chaveamento do ar realizado pela língua. A movimentação da língua previamente à liberação do jato de ar produz um sutil e breve estalido que ocorre em média a 28 ms antes do jato de sopro. Tal gesto bucal não é sensível no sinal original, mas fica bem visível neste segundo nível mais fino.

Em várias notas as maiores amplitudes coincidem com a região de sobrelevação do ataque. Em algumas, entretanto, as maiores amplitudes ocorrem após o ataque inicial acentuado, já na região de sustentação, como por exemplo na segunda, terceira e sexta notas. Verifica-se portanto que estas notas neste nível não aderem ou satisfazem ao modelo tradicional para envelopes, apresentando uma conformação diferenciada para as regiões de ataque, sobrelevação e decaimento.

Em verdade, pela análise gráfica da forma de onda, por exemplo, para a segunda nota, verifica-se, uma leve sobrelevação local cerca de 8 ms após o início do ataque, e um pequeno decaimento logo após a subida do ataque, cerca de 17ms após o início do mesmo. Mas a sobrelevação mais relevante, que usualmente está associada perceptualmente ao ataque, ocorre cerca de 70 ms após o início do mesmo, já portanto numa região que melhor corresponde à sustentação da nota.

A oitava nota apresenta um período de ataque mais longo, exibindo um aumento gradual que levará a uma sobrelevação máxima relacionada ao ataque cerca de 40 ms após o seu início. O intervalo entre o início do ataque e sua sobrelevação máxima associada varia entre 3 e 8 ms para as outras notas, com exceção da sétima, que é cerca 30 ms. A tabela abaixo procura sumariar estes dados para todas as notas do exemplo.

Tempo entre o início à sobrelevação do ataque, e tempo do início ao decaimento do ataque

Nota	Tempo do início do ataque à sobrelevação	Tempo do início do ataque ao decaimento
nota 1	7 ms	70 ms
nota 2	6 ms para primeira sobrelevação, 70 ms para a mais relevante após o ataque	25 ms
nota 3	2 a 3 ms (ataque não muito caracterizado como acentuado)	15 ms
nota 4	7 a 8 ms (ataque não muito bem caracterizado como acentuado)	inconclusivo
nota 5	5 ms	34 ms
nota 6	3 a 4 ms	27 ms
nota 7	23 a 30 ms	40 ms
nota 8	37 a 40 ms	62 ms

Início é a primeira subida significativa na amplitude, e *decaimento*, a maior depressão entre o ataque e a região de sustentação.⁶

O envelope não é suave, mas pontilhado de transientes. A textura do som é ruidosa, áspera, e já é possível inferir sobre as diferenças tonais (altura em frequência) das notas. Ouvido a 22050 Hz percebe-se a existência de uma modulação em amplitude que provoca um batimento nas notas, em especial na última nota onde este efeito é melhor percebido.

A última nota foi fortemente atacada e excitou harmônicos superiores por breve intervalo de tempo, o que por si só pode justificar uma sobrelevação nas amplitudes atrasada.

Nível 15: Um envelope aproximadamente “quadrado” especialmente para a última nota neste nível é o que sugere a imagem gráfica do som neste nível, visto numa escala panorâmica. Como viu-se, o envelope pode ser aproximado por quadriláteros, mas na verdade não mostra rampas descontínuas quando visualizado mais proximamente, numa escala menor, ou seja quando magnificamos uma região do exemplo (uma nota). O quadrado no caso seria na prática implementado por rampas íngremes, com uma taxa de variação alta.

A última nota é bem mais energética do que as outras neste nível, como sugerem as altas amplitudes relativas que apresenta no gráfico. O envelope para as outras notas não é suave, apresenta muitos transientes e um aspecto amorfo.

Os ataques descrevem altas taxas de elevação, e há várias sobrelevações posteriores ao início do ataque para a maioria das notas, o comportamento das variações de amplitude ocorrendo de forma aparentemente aleatória, descontrolada, não sujeita a regras determinísticas simples.

Já é possível estimar as alturas das notas originais pois que a tonalidade das notas é claramente perceptível, o que denota a presença da composição harmônica neste nível.

Nível 14: As notas estão melhor segmentadas neste nível. O ataque na última nota mostra-se o melhor caracterizado como acentuado neste nível, com uma sobrelevação intensa. Uma visão magnificada do sinal revela a existência de modulação em frequência agindo sobre as notas, manifestando-se como um batimento visível na forma de flutuações de baixa frequência. Como em exemplos anteriores, a ocorrência de tal modulação no nível deve-se ao espelhamento de bandas relevantes do sinal original neste nível em torno de frequências divisoras da frequência de amostragem.

Os ataques em todas as notas apresentam uma subida em rampa com inclinações menores que nos níveis anteriores mais finos, não havendo degraus no início do sinal. Isso mostra que as frequências relevantes neste nível não apresentam variações tão intensas e rápidas verificadas nos níveis mais agudos/finos. A última nota é a que, entretanto, melhor caracteriza o acento, com uma sobrelevação quase 3 vezes maior que a maior sobrelevação nas notas anteriores. A amplitude das regiões de sustentação tendem a ser semelhantes para todas as notas. O nível apresenta um envelope com maior suavidade, e uma visão panorâmica revela subidas e descidas íngremes. Estruturas caracterizadoras de gestos bucais entre as notas não contribuem relevantemente neste nível.

⁶ Os valores acima devem servir somente como uma estimativa aproximada para efeito ilustrativo, uma vez que foram obtidos com *softwares* editores de som para PC de uso genérico que não dispõem de ferramentas precisas e específicas para esta tarefa.

Há uma clara diferença tonal entre as notas, as alturas correspondendo às alturas do sinal original, como ocorreu nos exemplos anteriores do mesmo tema. A percepção de harmônicos mais altos excitados no ataque acentuado da última nota é clara.

Nível 13: Os ataques são mais bem caracterizados que no nível anterior, com sobrelevações mais destacadas. Anteriormente ao colapso da última nota ocorre uma elevação de amplitude, bem marcada neste nível.

A região de sustentação das notas apresenta uma flutuação de baixa frequência indicando a presença de um tipo de modulação em amplitude, que se manifesta imediatamente após a sobrelevação inicial associada ao ataque e continua até a região do colapso. A envelope da quinta nota (F4) melhor caracteriza o ataque, a sobrelevação, decaimento e a região de sustentação modulada sugerida neste nível. Os colapsos apresentam uma descida aproximadamente exponencial, decaindo rapidamente, não apresentando a modulação em amplitude verificada na região de sustentação, fato que sugere que já nesta região certas frequências (que produziam imageamento de bandas e conseqüentemente a modulação de amplitude verificada) tenham desaparecido ou reduzidas significativamente.

Nível 12: O envelope neste nível dá destaque às 4 primeiras notas do exemplo, na verdade as notas mais baixas em *pitch*, e que apresentam neste nível as maiores amplitudes relativas. Uma visão panorâmica sugere envelopes quadriláteros, especialmente para as 4 primeiras notas, mais intensas, com as regiões de sustentação apresentando amplitudes máximas aproximadamente constantes.

As 4 primeiras notas apresentam envelopes com ataques e colapsos aproximadamente semelhantes, com rampas de subida e descida íngremes com inclinações ligeiramente maior (em média) nos ataques. A região de sustentação exibe um envelope aproximadamente constante, o sinal flutuando sob uma modulação de amplitude de baixo índice de modulação. O sinal em todo o nível é suave, não apresentando descontinuidades ou transientes abruptos, como estalidos e *clicks*, e sim uma forma de onda contínua e suave.

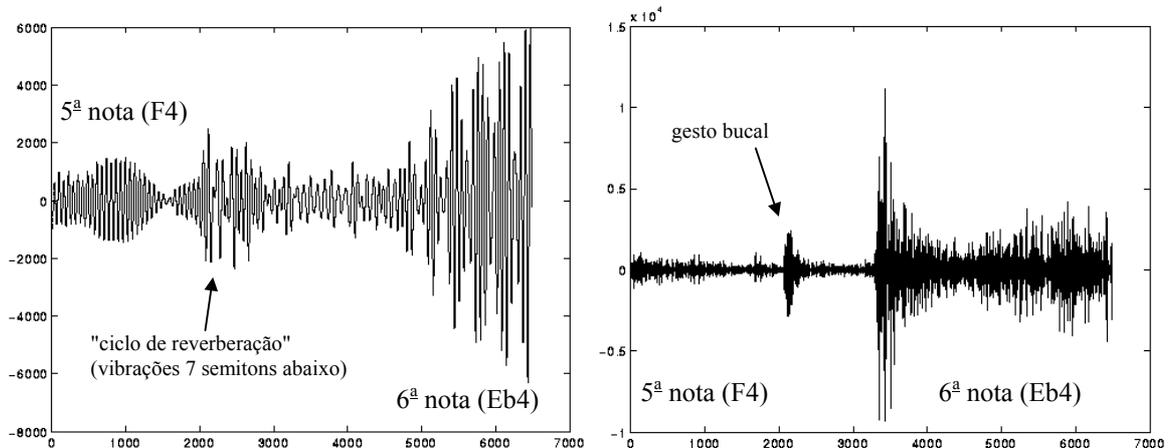
A região de ataque fica bem demarcada para as 4 últimas notas, com uma sobrelevação destacada da região de sustentação. A quinta e sexta notas apresentam um envelope neste nível semelhante ao visto no sinal original, com exceção da modulação de amplitude que não é verificada no som original.

O colapso da quinta nota (F4), ouvido numa velocidade mais baixa (11025 Hz), apresenta um "ciclo de reverberação" ou "eco" das vibrações harmônicas no tubo da flauta, e, antes do sopro que excitará os harmônicos de Eb4, soam vibrações percebidas 7 semitons abaixo, em si bemol 3 (Bb3). Isso é um fenômeno esperado, uma vez que no colapso das notas não mais há a excitação do sopro e os modos mais altos extinguem-se mais cedo que os mais graves.

Verificou-se que as vibrações - 7 semitons abaixo de F4 - iniciam tão logo a excitação das vibrações de F4 terminam, o que coincide com a localização do gesto bucal existente entre as quinta e sexta notas no nível 16, e vão terminar após iniciado o ataque da sexta nota (Eb4), mais especificamente após o decaimento na sexta nota.

As Figura 12, abaixo, mostra a região de transição da quinta para a sexta nota, cobrindo o mesmo intervalo de tempo nos níveis 12 e 16 respectivamente:

Figura 12 - Transição da 5ª para a 6ª notas - nível 12 (esquerda) e 16 (à direita)



O nível 12 focaliza melhor as componentes harmônicas do som, localizando melhor as regiões onde estas são mais energéticas. O nível 16 permite localizar os gestos bucais e os momentos onde há transições e chaveamentos no sopro. Note que as vibrações harmônicas da sexta nota (Eb4) no nível 12 ganham energia a partir do final do decaimento verificado no nível 16 após o ataque inicial. As vibrações entre o gesto bucal e o final do decaimento do ataque da sexta nota não apresentam os parciais harmônicos esperados para a nota Eb4, que portanto só se estabelecem na região de sustentação desta nota.

Nível 11: A primeira nota, mais baixa em frequência (G3), é a que apresenta maior intensidade neste nível, seguida pelas duas seguintes, de 2 a 3 vezes entretanto menos intensas que a primeira. A última nota, também a mais alta em frequência (B4), apresenta mais componentes neste nível que as quarta, quinta, sexta e sétima notas, por originalmente ser também mais intensa que estas.

As transições entre as três primeiras notas, vistas mais proximamente, constituem em regiões de estrangulamentos (baixas amplitudes) entre estas notas. Entretanto, a forma de onda é contínua e não desaparece completamente nas pausas, evidenciando a presença de uma reverberação ou um colapso mais lento para os modos mais baixos (baixas frequências) das notas mais graves.

Nível 10: Somente a segunda e terceira notas apresentam vibrações harmônicas neste nível, as outras notas exibidas sem segmentação ou destacando identidades auditivas ou representações gráficas individuais. O som é o de um ruído grave para todas as notas, com exceção da 2ª e 3ª. Os envelopes não são suaves, e as formas de onda apresentam picos distribuídos de forma aleatória.

Várias faixas curtas e intensas, com picos em amplitude destacados do corpo do sinal, localizam os instantes de ataques das notas, marcando assim gestos musicais associados de baixa frequência. O instante que marcaria o ataque da última nota, que se mostrou a mais energética em vários níveis anteriores mais finos, não aparece em destaque.

Nível 8: A visão panorâmica da representação gráfica do nível mostra a existência de uma modulação de amplitude agindo sobre o sinal, produzindo 4 regiões de máximos (cristas). O sinal neste nível extrapola os limites, contornos naturais reais das notas, e se espalha por toda a duração do exemplo, isto é, onde não foi realizado acolchoamento com zeros. Isso impossibilita a localização de instantes relevantes da interpretação, e por conseguinte a localização de

estruturas e gestos musicais na grade de tempo correta. Uma investigação mais criteriosa ainda se faz necessário para verificar a utilidade deste nível em aplicações musicais.

A frequência modulante que produz o efeito de 4 cristas acima é de aproximadamente 0,9 Hz, calculado tomando como período a quarta parte da extensão do som original (antes do *zero-padding*) que é de 4,327 segundos. Magnificando regiões do sinal verifica-se a ocorrência de outras frequências modulantes sobre o sinal. O sinal é quase inaudível.

Pela análise do espectro do nível notamos um pico de máximo em 40,37 Hz e um segundo cerca de 0,84 Hz abaixo, em 39,53 Hz, e também o espelho destas bandas em torno de 1/1024 da frequência de amostragem logo acima, com picos em 45,76 e 46,60 Hz.

A diferença entre os picos mais altos do espectro é cerca de 5,38 Hz, uma frequência de batimento verificada magnificando o sinal acima. Este batimento entretanto é gerado pela análise *Wavelet*, uma vez que consiste na diferença entre as bandas espelhadas acima, uma delas somente presente no sinal original. A segunda diferença encontrada, de 0,84 Hz entre os picos de uma das bandas, é bastante próxima da frequência estimada acima de 0,9 Hz que modula a visão panorâmica do sinal.

Nível 7: Um sinal muito pouco energético, revelando dois picos de amplitude, um sobre a primeira nota no ataque, e outro sobre o ataque da última nota.

Apresenta amplitudes altas em regiões onde ocorrem pausas ou silêncios musicais no sinal original. Este fenômeno no entanto não pode ser generalizado como uma propriedade intrínseca do nível, porque esta correspondência não é precisa, isto é, as pausas não são localizadas com precisão. Por exemplo, os picos máximos no presente exemplo ocorrem localizando os ataques da primeira e da última notas, portanto em regiões onde há sinal, posteriores às pausas. Também sobre as regiões correspondentes aos colapsos de várias notas há máximos de amplitude verificados neste nível.

Todavia a proximidade entre as regiões de maior amplitude e as regiões de pausas (mesmo sem uma localização precisa) é um comportamento recorrente se admitirmos um grau de desvio, uma vez que também os ataques e os colapsos de notas são eventos vizinhos a pausas. O nível tende a localizar estas regiões apresentando máximos de amplitude nas regiões circunvizinhas às pausas e transições entre notas. **Como pausas são silêncios, vazios de eventos musicais, supõe-se que este nível seja sensível a sinais de baixas frequência e intensidade, como ruídos de fundo, que se tornam pronunciados na ausência de sinais musicais.**⁷

O sinal é inaudível, possui uma textura ruidosa e não é suave. Não permite a segmentação das notas, e a forma de onda é contínua por todo o exemplo, ocorrendo inclusive nas pausas entre notas. Sobre as regiões onde se situam as notas, entretanto, há um reforço nas amplitudes do sinal, mostrando que o nível é sensível à presença de eventos ou vibrações musicais.

Nível 6: Sinal não suave, não audível, não localiza as notas, forma de onda contígua por todo o exemplo. A segmentação verificada não corresponde à segmentação do sinal original. O sinal no nível apresenta propriedades semelhantes ao nível anterior, localizando com seus pontos de

⁷ Na ausência de sinais intensos durante a gravação com equipamentos digitais, principalmente naqueles onde empregam-se compressões e expansões de amplitude, a codificação com PCM formará palavras digitais onde a informação está nos bits menos significativos que cobrem faixas dinâmicas de baixas amplitudes, e que são mais propensos a ruídos de quantização.

máximos (ou mínimos) de amplitude os instantes de ataque ou colapso da maioria das notas. O sinal apresenta **transientes** em vários instantes, pontos onde a amplitude é máxima.

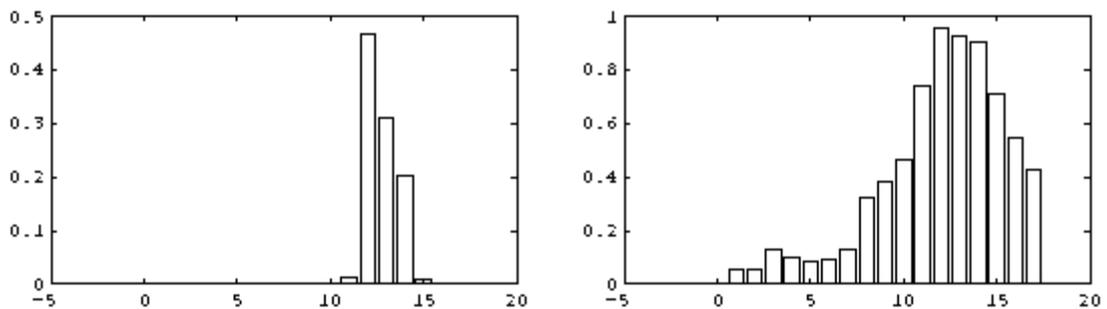
Nível 5: Inaudível, exceto pelo *click* dos picos sobre a última nota, que não localizam nem o ataque nem o colapso desta com precisão. O nível pode ser visto como uma média do sinal original, uma medida da oscilação associada a eventos ou gestos em baixa frequência (subsônica). O emprego do nível para analisar aspectos musicológicos, como estimar a intensidade de determinado trecho tocado em relação a outros, pode ser plausível, mas somente a análise de partes orquestrais mais longas poderão mostrar sua eficácia neste sentido.

Nível 4: Uma forma de onda que poderia ser passível de emprego para análises de cunho musicológico. Note que as amplitudes do sinal neste nível são crescentes, culminando com os máximos sobre a nota mais intensa e acentuada do motivo, a última. De certa forma, há uma leve progressão na direção do desfecho acentuado na última nota neste motivo, e tal informação está traduzida nas medidas de magnitude exibida pelo sinal do nível.

Nível 3: Uma versão mais suave e mediana do nível anterior, uma medida da progressão e desfecho na última nota.

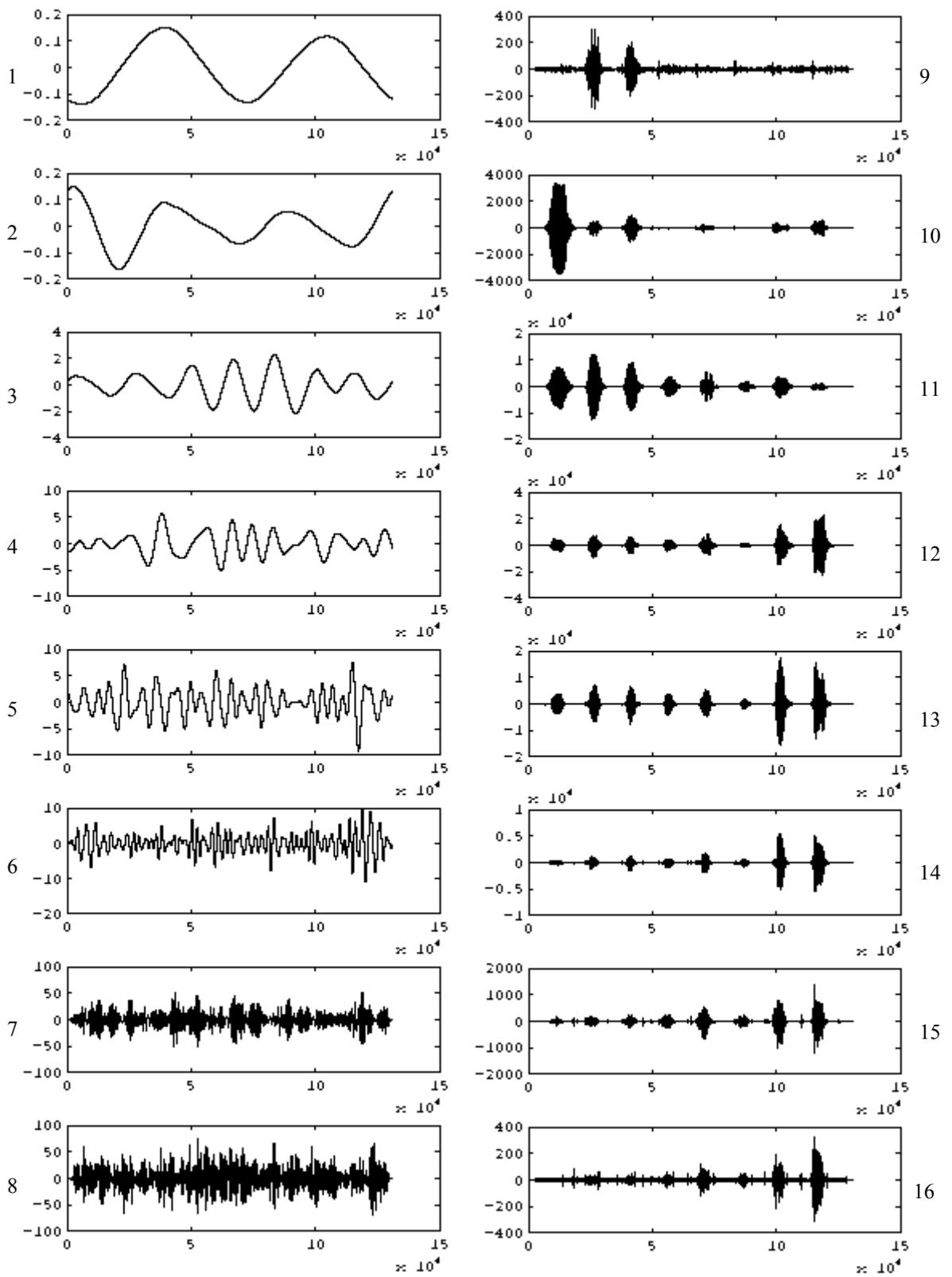
Nível 2: O nível não fornece indícios para o presente exemplo de uma utilidade significativa para análise musicológica ou dos eventos e gestos musicais existentes na execução do exemplo.

Distribuição de energia:



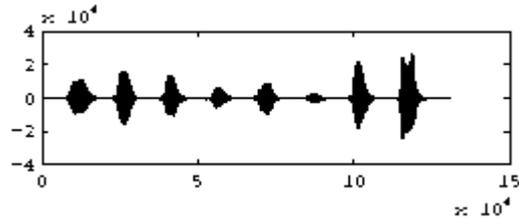
Mais de 99% de toda a energia do fraseado se concentra nos níveis de 11 a 15, sendo o nível 12 o mais energético. Níveis abaixo do sétimo nível da análise são virtualmente insignificantes em termos de energia. A maior contribuição energética é originada das vibrações acústicas do instrumento, e portanto os níveis mais energéticos são aqueles que cobrem as frequências relevantes das séries harmônicas.

Flauta: Fraseado, Não ligado, com *staccatos*



Evento: Fraseado, não ligado, com *staccatos*. Interpretação do mesmo tema melódico do exemplo anterior (fraseado não ligado normal) executando-se as notas agora sob *staccato*.

Gráfico do sinal original:



Extensão da amostra: 2^{17} pontos (131072). **Wavelet de análise:** D16

Análise dos níveis:

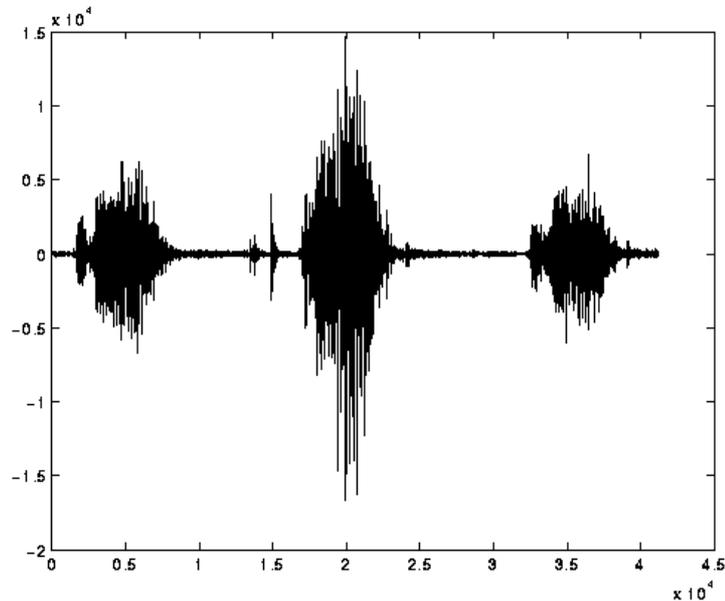
Nível 16: Nível que caracteriza o ruído de sopro. A última nota apresenta as maiores amplitudes. Ruído de fundo espalhado por todo o exemplo. Regiões de sustentação das notas apresentam amplitudes mais pronunciadas.

Ataques e colapsos das 6 primeiras notas podem ser localizados, mas as formas de onda associadas a esses eventos exibem amplitudes no mesmo nível de intensidade que o ruído de fundo que se distribui regularmente sobre todo o exemplo, sendo portanto indiscerníveis deste. O nível localiza com eficácia os eventos acústicos e gestos bucais, respiração, movimentos de língua e lábios, que no ato de chaveamento do sopro produziram estalidos.

Nível 15: Estalidos bem localizados nas transições, e apresentando elevações comparáveis às amplitudes das notas adjacentes (veja nas próximas figuras). Para as notas mais intensas os ataques *staccatos* apresentam subidas íngremes, mas sem caracterizar grandes declividades e sobrelevações. Já nas notas menos intensas, fica melhor caracterizado o envelope típico de ataques súbitos, com a região de ataque, sobrelevação associada e respectivo decaimento. O decaimento neste nível leva a um estrangulamento que é mais acentuado para os exemplos de *staccatos*, constituindo regiões de ataque bem separadas das regiões de sustentação nas notas menos intensas.

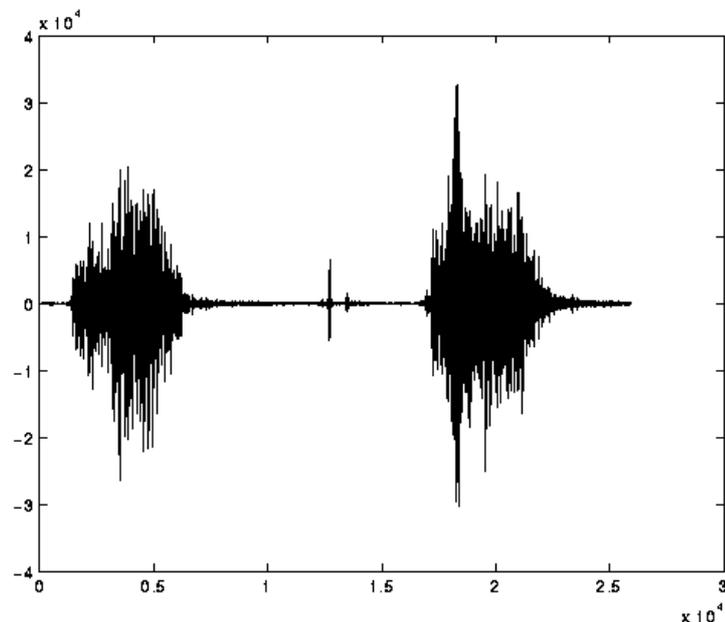
A Figura 13 mostra as quarta, quinta e sexta notas no nível 15 com mais detalhes, exatamente as notas menos intensas do exemplo. Note a caracterização do ataque com uma região de decaimento bem destacada.

Figura 13 - Quarta, quinta e sexta notas do exemplo no nível 15 (segundo nível mais fino)



Na Figura 14 mostra-se as duas últimas notas, as mais intensas do exemplo, com maiores detalhes (magnificadas). Note a presença de uma subida inicial íngreme, porém menos abrupta que na figura anterior, e com a região de decaimento apresentando maiores amplitudes do que para as notas menos intensas da figura anterior.

Figura 14 - As duas últimas notas do exemplo no nível 15



Isso ocorre porque, sendo mais intensas uma maior pressão e fluxo de sopro excita os modos vibratórios da flauta, causando maiores amplitudes em todas as regiões, e não somente na sustentação. A sobrelevação no ataque, neste caso, pode até ser inferior em amplitude que a região de sustentação das notas.

A região de colapso para a maioria das notas exibe uma queda gradual, sem decréscimos súbitos e acentuados. Os estalidos nas transições das notas são objetos acústicos transientes provocados pela movimentação da língua contra o céu da boca nos instantes de chaveamento do sopro,

geralmente ocorrendo previamente à liberação de novo jato (veja as figuras anteriores). Na abertura de novo jato de sopro, quando a intenção do intérprete é de produzir um ataque intenso a pressão de ar também é maior, e as rampas de subida no ataque serão mais inclinadas.

Os contornos das notas não são suaves, a textura do som é ruidosa, áspera e aguda. No entanto é possível inferir sobre as tonalidades das notas.

Nível 14: A composição harmônica já é maior neste nível que nos anteriores, e tem-se uma maior percepção dos *pitchs* (alturas) das diferentes notas. Há uma clara percepção auditiva da qualidade do ataque *staccato* neste nível, diferenciável da textura e sonoridade do restante da nota (sustentação e colapso). As regiões de sustentação são curtas e não apresentam modulações extras como vibratos.

Nível 13: *Pitch* das notas soa como no som original. O nível é bem energético, e de contornos mais suaves, logo envelopes mais suaves também. O formato das notas numa visão panorâmica do nível sugere um envelope arredondado, com exceção para as duas últimas notas mais intensas, que apresentam sustentação mais irregular. As regiões de ataque e colapso formam estruturas discretas, de menores amplitudes mas contíguas à região central ovalada.

O nível apresenta sensibilidade em destacar as porções onde as vibrações harmônicas da flauta exibem maior contribuição nas notas, isso é, na região central. O formato ovalado de seu envelope sugere portanto que as vibrações estacionárias de cada tom cresce logo após o decaimento do ataque e diminui na direção do seu colapso, quando o fluxo de sopro está em redução. A ocorrência de uma súbita interrupção do sopro, marcada pelos gestos de sopro vistos nos níveis mais finos, provoca uma súbita diminuição nas amplitudes das vibrações neste nível.

O nível apresenta um som mais puro, sem a textura granulosa e áspera características dos níveis mais finos. As notas no nível apresentam algum índice de modulação na amplitude, fruto do espelhamento de certas bandas no espectro numa disposição tal que o efeito de batimentos (com baixo índice de modulação) se impõem à onda.

Nível 12: Nível mais energético. As tonalidades das notas (*pitchs*) diferem daquelas do som original.

Nível 11: Três primeiras notas são as mais intensas, a segunda é a mais (A3).

Na visão panorâmica percebe-se claramente uma modulação de amplitude provocando um batimento nas amplitudes das quinta e oitava notas. A oitava nota, a mais intensa do exemplo, aqui exhibe as menores amplitudes. Isso porque sua contribuição energética (das vibrações harmônicas) se dá nos níveis mais agudos, e esse é o mesmo motivo pelo qual as 3 primeiras são as mais intensas, isto é, concentram mais sua energia neste nível.

Vistas magnificadas, as notas neste nível exibem formatos ovalados, com os pontos de máxima amplitude ao centro. Isto é evidência de que o nível concentra primordialmente a contribuição das vibrações harmônicas estacionárias de cada nota, as quais usualmente se estabelecem nas regiões de sustentação. No caso de *staccatos*, esta região é mais curta que para notas normais ou acentuadas.

A audição do nível mostra que as tonalidades das notas correspondem às tonalidades originais, exceto pela última nota, percebida uma oitava abaixo.

Nível 10: Várias notas soam com dinâmica *pianíssimo*, isto é, não possuem contribuições de suas vibrações harmônicas concentradas neste nível. As tonalidades das notas não correspondem aos tons originais. Já bem menos energético que o nível anterior, ainda é observado um

contorno ovalado na primeira nota, a mais intensa do nível. Presença de modulação de amplitude.

Os ataques *staccatos* não ficam caracterizados neste nível: os contornos das notas são suaves, o envelope apresenta subidas graduais para todas as notas e descidas também em rampas pouco inclinadas. As declividades da subida dos ataques é ligeiramente maior que a observada nos colapsos. Modulações de amplitude são observadas em todas as notas. Os ataques são ligeiramente mais protuberantes que os colapsos.

Nível 9: Som do nível é o de um ruído grave, cuja forma de onda é contígua por todo o exemplo. Somente as segunda e terceira notas apresentam contribuição harmônica no nível, ainda audível.

Pulsos que se destacam sobre o ruído característico do sinal localizam os instantes de ataques das notas, evidência de que o nível é sensível a gestos musicais associados aos ataques. Os pulsos observados neste nível são mais evidentes que os pulsos verificados em nível semelhante para o exemplo de fraseado não ligado acentuado.

Nível 8: Ruído grave, textura áspera, característica de ruídos, presença de vários máximos e mínimos locais, envelope não é suave, a forma de onda apesar de contínua, é irregular apresentando transientes distribuídos por todo o exemplo. Segmentação das notas não há.

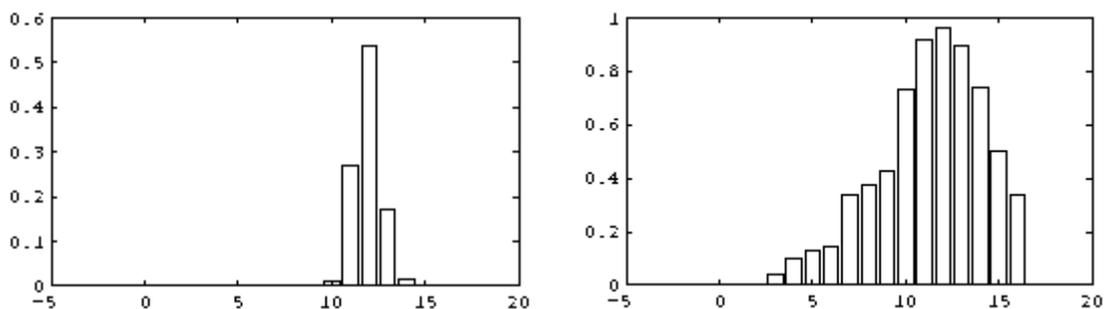
O emprego do nível, utilizando-se técnicas adicionais para limiarização e localização de limiares, para efeito de localização de eventos acústicos de baixa frequência, gestos musicais e transientes associados deve ser pesquisado em trabalhos futuros. O nível é audível.

Nível 7: Audível, ruído grave de fundo. Modulação de amplitude visível. Forma de onda não suave, contígua, sem segmentação de notas. Semelhante ao nível anterior, porém mais "limpo".

Nível 6: Forma de onda contígua, não suave, cheia de picos, leve ascendência dinâmica no final, onde estão as notas mais intensas. Efeito entretanto mais discreto do que aquele que ocorre nos exemplos com acentos.

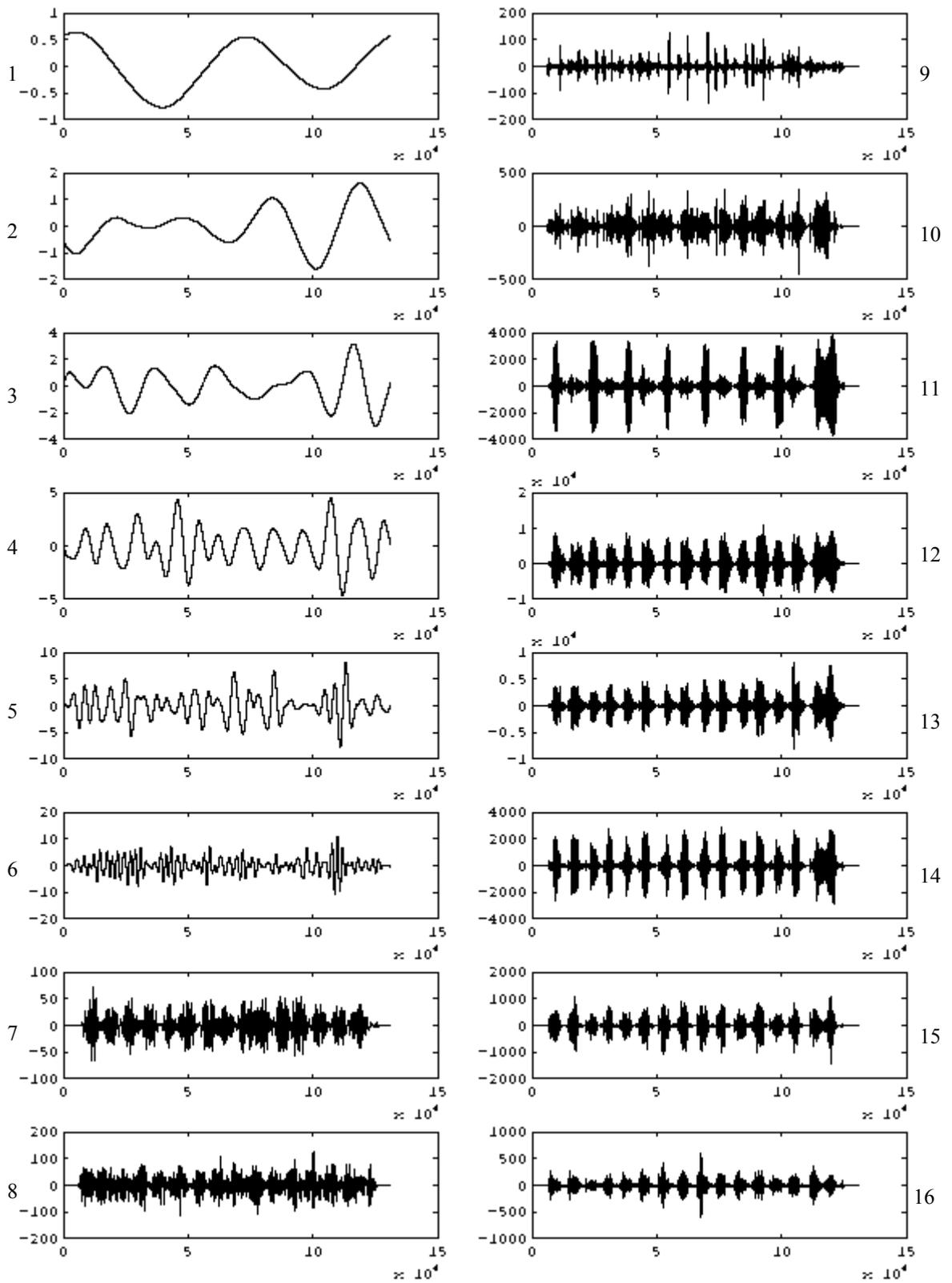
Nível 5: Inaudível, exceto pelos transientes no início e no fim. Picos de máximo estão sobre a segunda nota e a última nota.

Distribuição de energia:



Mais de 99% de toda a energia do exemplo se concentra nos níveis de 10 a 14, sendo o nível 12 o mais energético. A distribuição de energia se dá com certo grau de simetria em torno do nível 12, como pode-se verificar visualmente no gráfico de distribuição de energia à direita (acima).

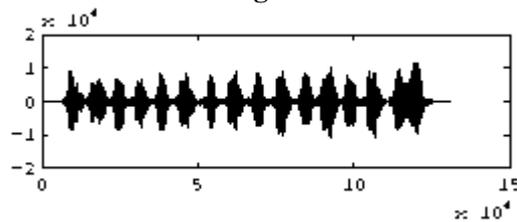
Flauta: Fraseado, Não ligado, Trêmulo dedilhado - notas sol4 (G4) e si4 (B4)



Evento: Fraseado não ligado: trêmulo dedilhado. Exemplo consiste na execução de um trêmulo com as notas sol (G4) e si (B4) da segunda oitava da flauta, executados em tempo rápido para se caracterizar a natureza das transições e da emissão desses tons num fraseado tremulado. O trêmulo é do tipo "trêmulo de dedilhado" consistindo na interpretação de duas notas com intervalo maior que 1 tom.

A última nota do exemplo possui uma identidade diferenciada, com uma leve modulação aplicada alongando-a, e produzindo um final dinamicamente mais relevante.

Gráfico do sinal original:



Extensão da amostra: 2^{17} pontos (131072). **Wavelet de análise:** D16

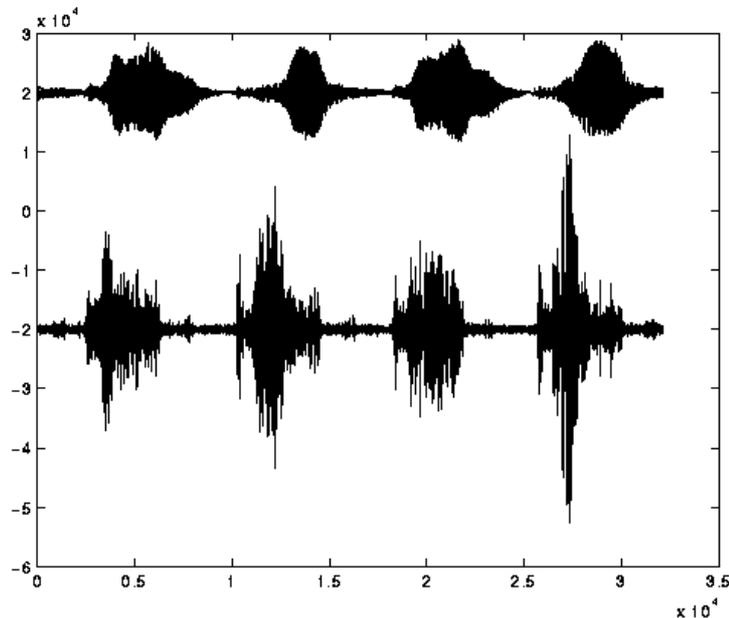
Análise dos níveis:

Nível 16: Percebe-se bastante modulação de sopro, isto é, o sopro é fundamental na formação do envelope do trêmulo não ligado. O ciclo do sopro envolve o ataque e o colapso forçado sucessivos. O nível mais fino portanto é importante para se caracterizar trêmulos não ligados, onde observamos a presença do ciclo completo do sopro, do ataque ao colapso forçado, em tempo rápido.

Uma inspeção visual mais próxima (magnificação) de algumas notas mostrará o envelope característico do ataque em fraseados não ligados, com uma rampa de subida íngreme, uma sobrelevação seguida de um decaimento.

Logo em seguida vem a região de sustentação com altas intensidades, superiores às da sobrelevação, e, em seguida, pela diminuição da pressão de sopro devido ao avanço da língua sobre o céu da boca reduzindo o sopro, uma queda progressiva nas amplitudes, entrando na região de colapso. A interrupção completa do sopro, que é cíclica no trêmulo não ligado, provoca o decaimento súbito ao final da região de colapso, conforme podemos observar na Figura 15, mostrando, para um mesmo intervalo de tempo (4 notas), o sinal original (sinal de cima) e o sinal no nível 16 (sinal de baixo).

Figura 15 - Parte do sinal original (topo) e do sinal ao nível 16 (abaixo)



Os "nódulos inter-notas", isto é, as pequenas elevações situadas entre as notas, correspondem aos gestos bucais que ocorrem quando temos um movimento de língua rápido avançando e retraindo em relação ao céu da boca, no sentido de se modular a passagem de ar e assim o fluxo de sopro. A pressão de sopro é máxima no centro das notas, onde a intensidade é mais alta.

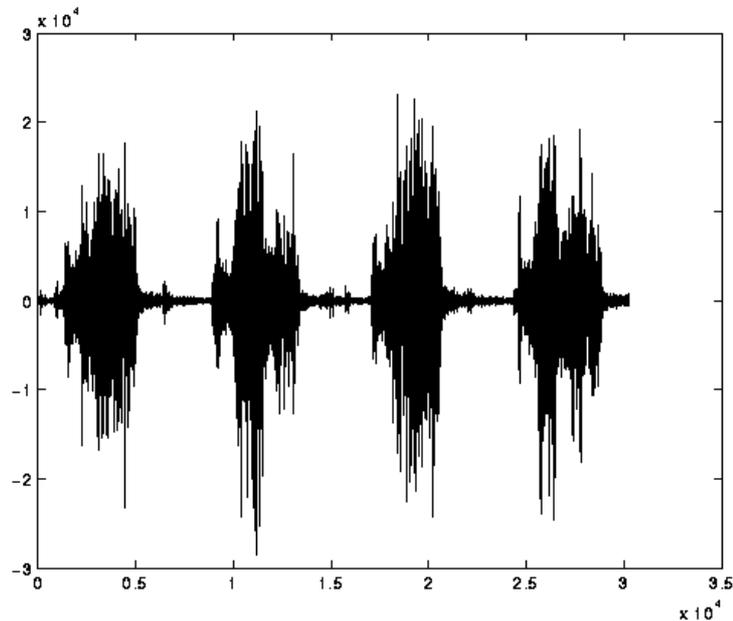
Nível 15: Verificamos que os envelopes das notas sol4 (G4) possui uma depressão de amplitudes ao centro da região de sustentação. Previamente ao colapso, as amplitudes crescem novamente e então caem subitamente em virtude do colapso forçado.

Estruturas intermediárias, isto é, pequenos nódulos ou elevações ainda são vistas entre as notas, embora bem mais sutilmente do que no nível anterior.

Para as notas si, observamos o ataque característico, com sobrelevação e decaimento, e em seguida um progressivo aumento de amplitudes em direção ao colapso, quando então caem abruptamente. Os contornos da forma de onda, como no nível anterior, não são suaves: são irregulares e com vários picos locais (de máxima amplitude).

A Figura 16 mostra uma parte magnificada do nível 15, entre as 6^a e 9^a notas inclusive, uma sequência de si-sol-si-sol (entre 0.944 e 1.630 segundos).

Figura 16 - Parte do sinal ao nível 15: intervalo entre 0,94 s e 1,63 s (da 6ª à 9ª notas)



Textura sônica ruidosa, como o nível anterior, mas apresentando já alguma contribuição das séries harmônicas das notas do trêmulo.

Nível 14: Primeiro nível que permite uma melhor percepção das tonalidades. Os envelopes das notas si4 (B4) diferem em forma dos envelopes das notas sol4 (G4), estes sendo ligeiramente mais suaves, com menos irregularidades e transientes, possuindo ataques discretos, e uma flutuação ou modulação de baixa frequência sobre as amplitudes na região de sustentação, com uma progressiva elevação que culmina nas maiores elevações para cada nota (sol4) do centro em diante, na direção do colapso.

Já as notas si4 apresentam ataques mais discretos, embora bem caracterizados, não apresentam flutuações de baixa frequência na amplitude após o decaimento, e sim as amplitudes crescem rapidamente e mantém uma média intensa por toda a sustentação, decaindo após no colapso.

O colapso para ambas as notas consiste numa queda de elevações menos abrupta do que nos níveis anteriores, e são estendidos por uma pequena reverberação ou eco, como num pulso após o corte da excitação.

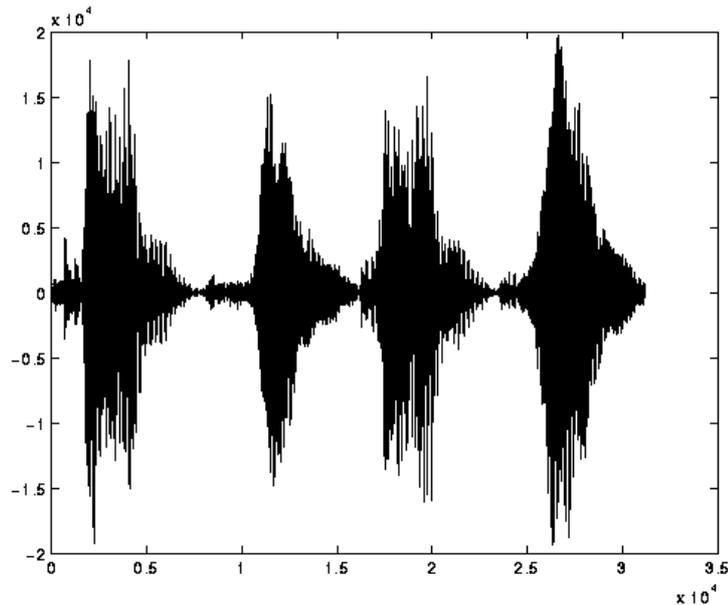
Ouvidas a 11025 Hz as duas notas apresentam diferenças nas regiões transitórias, do ataque, previamente à região estacionária. As notas sol4 apresentam um transitório mais complexo até estabelecerem o tom correto e soá-lo com regularidade. Esta complexidade no ataque implica num maior tempo necessário para atingir a zona de sustentação, e explica-se em parte pela dificuldade inercial de se excitar uma nota mais baixa (maior coluna de ar) no tubo vibrante. O fato de termos um ciclo de interrupção do sopro a todo momento, as vibrações que se processam no tubo estão sofrendo transitórios constantes, e as zonas de vibração estacionárias são mais curtas, especialmente se o tempo da interpretação for rápido.

Nível 13: As notas si apresentam uma modulação em frequência facilmente constatada por inspeção visual do gráfico, fruto do processamento da análise *Wavelet* em multiresolução. Também seus envelopes são mais irregulares, com ataques discretos, zonas de sustentação com envelopes quase quadrados, com subidas e descidas íngremes, e uma região de colapso mais longa, que decai gradualmente. Neste nível o colapso se mistura à região de sustentação, seu

início podendo ser considerado uma segunda parte daquela região onde as amplitudes são menores.

A Figura 17 ilustra o intervalo cobrindo da sexta à nona notas (si-sol-si-sol) mostrando os contornos e aspectos da forma dos envelopes. Os envelopes das notas sol apresentam ataques mais discretos, com pequenas elevações, e uma região de sustentação mais intensa, com as amplitudes gradualmente reduzindo-se em seguida, misturando-se à zona de colapso.

Figura 17 - Parte do sinal ao nível 13: da 6ª à 9ª notas (si-sol-si-sol)



Nível 12: Nível mais energético. Boa sonoridade, nível concentra energia dos parciais harmônicos. As sobrelevações nos ataques das notas si4 apresentam amplitudes destacadas no exemplo, contrastando com o envelope das notas sol4 (aproximadamente com um formato de gaussiana), com elevação e queda variando mais suavemente, de forma gradual.

Nível 11: As notas si4 apresentam intensidades bem menores que as notas sol4, em média 6 vezes menos intensas, assim como ocorreu no exemplo com fraseado ligado com trêmulo dedilhado.

Observando o gráfico do nível da sexta à oitava notas, percebe-se que as regiões centrais das notas sol4 (mais baixas em frequência) são as mais destacadas neste nível, e que as notas si4 apresentam formas de onda de envelopes amorfos, contornos não suaves.

Outro fato a ser observado, inclusive em níveis anteriores, é a não observância de silêncios completos entre as notas, em todos os níveis até o presente. Se o trêmulo fosse mais lento poderíamos visualizar momentos de pausa, uma vez que o tubo ressonante teria tempo o suficiente para colapsar todas as vibrações que nele trafegam.

Nível 10: Alguns dos picos de máximos no nível situam-se sobre as zonas centrais de ambas as notas do trêmulo. No entanto, existem também transientes associados às zonas de transição entre notas e ausência de máximos de amplitude pontiagudos sobre algumas notas. Essa irregularidade sugere que os picos estão rastreando eventos musicais e/ou gestos transitórios que não estão aparentes ou facilmente identificáveis no sinal original.

Os eventos ou gestos em questão seriam então aperiódicos, isto é, produzidos de forma aleatória, sua existência estando condicionada à natural manipulação do instrumento, a eventos

de respiração e movimentos bucais aperiódicos, e não aos processos vibratórios correlacionados aos modos ressonantes do tubo vibrante. Dado à textura já ruidosa e grave do som do nível (que cobre faixas de frequência mais baixas), é de se esperar que tais gestos impliquem em contribuições relevantes de frequências mais baixas.

Estudos mais específicos serão necessários para investigar a natureza dos picos de máximo associando-os a gestos específicos na interpretação do instrumento.

Nível 9: Sinal não suave e de contornos irregulares, mas apresentando certa periodicidade na ocorrência de transientes ou pulsos que se distribuem regularmente por todo o exemplo, com capacidade de localização (isto é, não extrapolam os limites do sinal original real invadindo a região de *zero padding*).

Os pulsos localizam as zonas centrais das transições entre notas e das regiões de sustentação, onde encontram-se pontos críticos associados à intermitência do sopro, quais seja os de máxima intensidade e pressão (quando a língua permite a passagem do maior fluxo de ar) e os de intensidade mínima ou nula, onde a língua interrompe por curtíssimo período de tempo o fluxo de ar.

Os estalidos e outros ruídos bucais associados aos gestos bucais da interpretação são os eventos mais prováveis de estarem visíveis neste nível, uma vez que são periódicos mas exibem frequência de periodicidade inconstante exatamente por serem processos sincronizados pelos ritmos naturais dos movimentos do intérprete. Este nível portanto revela-se mais útil que o anterior para rastrear tais gestos.

Cabe notar entretanto que em audições acústicas reais, onde vários instrumentos são tocados num ambiente adequado, tais gestos não produzem ruídos perceptíveis porque seriam fortemente atenuados pela impedância acústica do ambiente real, e também mascarados pela contribuição mais energética das vibrações harmônicas dos instrumentos. Só microfones com capacidade de captação direcional, como o que foi utilizado, podem revelar tais ruídos característicos da manipulação do instrumento ou ruídos de origem bucal.

Nível 8: O contorno do envelope do nível consiste numa aproximação ao envelope original: depressões ou estrangulamentos no sinal original que ocorrem nas zonas entre notas, e elevações correspondentes às zonas centrais das notas também ocorrem no sinal deste nível. Como o nível é baixo, já representa de certa forma uma média do sinal original projetada num nível de resolução específico onde as flutuações de larga escala do sinal original são mais visíveis. Em outras palavras, as frequências cobertas por este nível integram o sinal original e contribuem no delineamento ou conformação do envelope de larga escala do sinal original (mesmo tratando-se de nível pouco energético).

As notas entretanto neste nível não soam mais diferentes e seus envelopes são preenchidos por um conteúdo de frequências comum a todas as notas, isto é, um mesmo tom é modulado em amplitude pelo envelope do sinal original. Dois *pitchs* ou bandas de frequências preponderantes são encontradas no espectro do nível: 46 Hz e 40 Hz, a diferença entre estas raias de aproximadamente 6 Hz corresponde a grosso modo à periodicidade das notas, que ocorrem numa frequência de aproximadamente 6 por segundo.

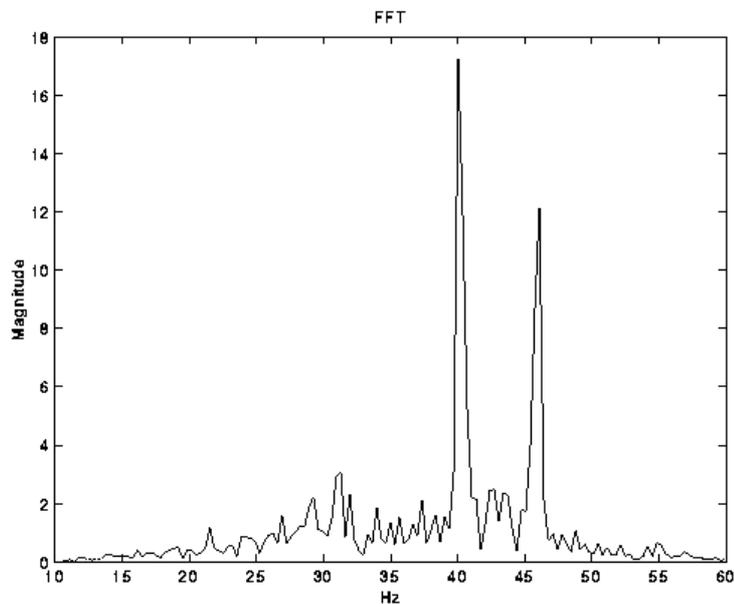
Portanto, quando temos uma interpretação bem periódica, com precisão no respeito à grade de tempos e de duração das notas - *especialmente se a partitura consiste em seqüências de notas de mesma duração* - a frequência de produção de notas aparecerá como simples diferença entre os raios do espectro deste nível. Se a frequência for inconstante, então a sua extração por

intermédio deste simples cálculo (o do batimento entre os *pitchs* preponderantes do nível) não será eficaz, e técnicas mais sofisticadas serão necessárias.

Somente com esta sequência de 15 notas é que primeiro pudemos descobrir a sensibilidade deste nível a periodicidades na interpretação de notas, agindo a grosso modo como um **contador de notas**.

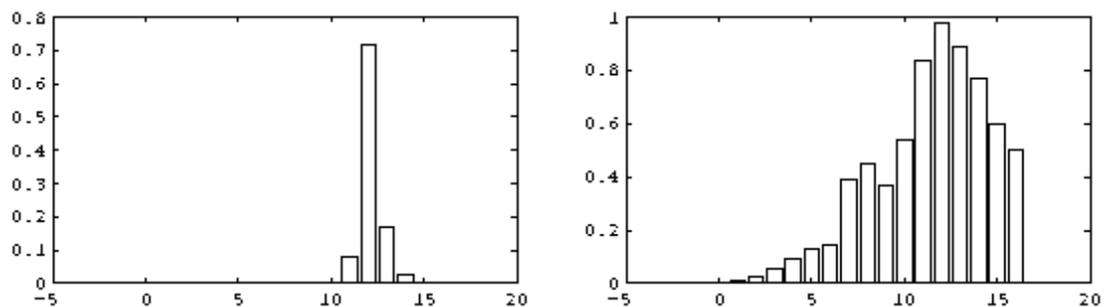
Nível 7: Também revela periodicidade, contando notas, com picos do espectro em 40 Hz (mais intenso) e 46 Hz, exatamente as mesmas frequências do nível anterior. O espectro do nível está na Figura 18 abaixo.

Figura 18 - FFT do nível 7 para a análise de fraseado não ligado, trêmulo dedilhado



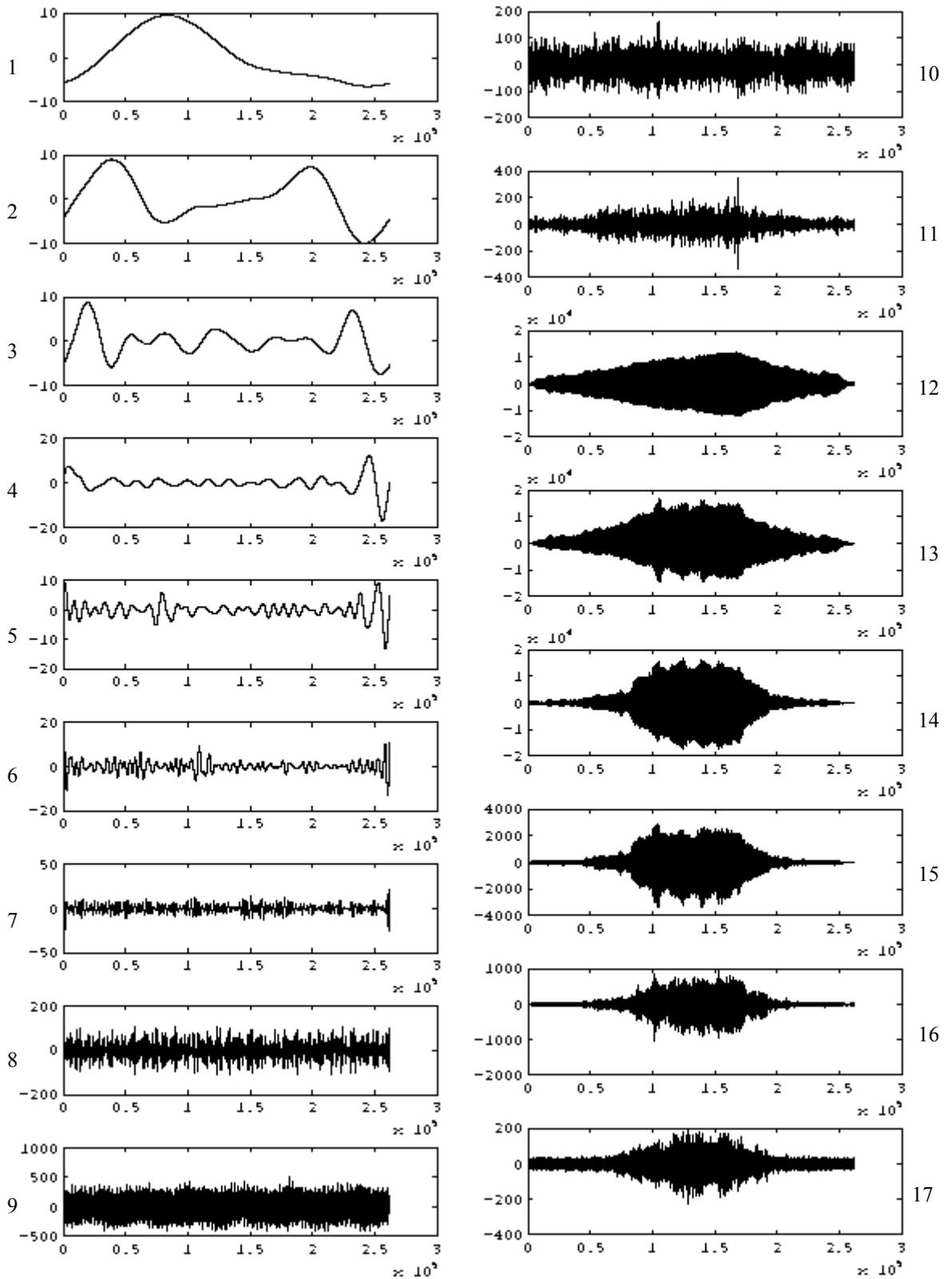
Nível 6: Inaudível (exceto pelos *clicks*), se espalha pela zona de acolchoamento, não permite localizar com precisão eventos em relação ao sinal original

Distribuição de energia:



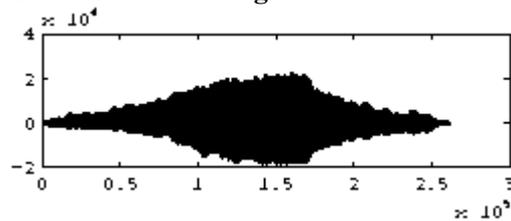
Mais de 99% de toda a energia do exemplo se concentra nos níveis de 11 a 14, sendo o nível 12 o mais energético.

Flauta: Sustentação com Dinâmica - nota ré4 (D4)



Evento: Sustentação da nota ré4 (D4) com variação contínua (ligada) da dinâmica (intensidade). Início com ataque soprado (sem a marcação com ataque de língua em "tu") em *pianíssimo*, seguido de um *crescendo* até *fortíssimo*, e um *decrecendo* até *pianíssimo*, sem vibrato. O exemplo apresenta uma textura "soprada", com o sopro fluindo externo ao foco da embocadura por toda duração da nota.

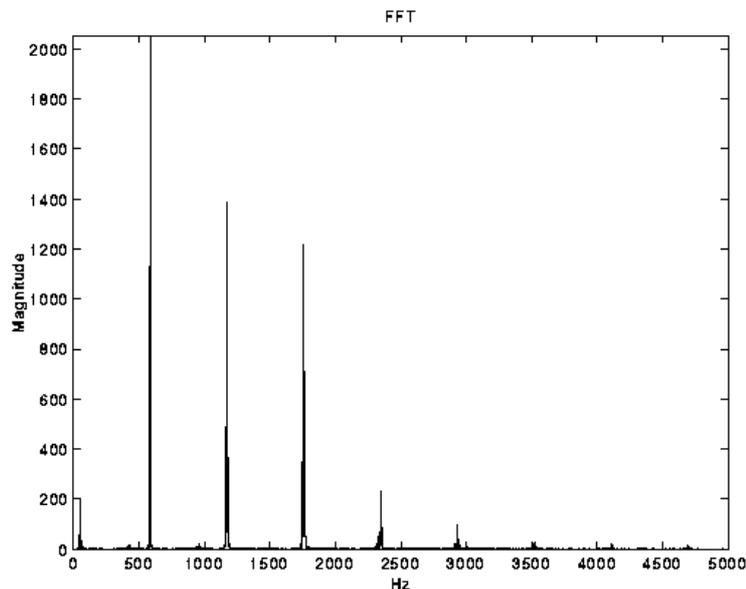
Gráfico do sinal original:



Na tabela abaixo listamos os harmônicos principais e suas localizações em frequência para o ré4 (D4) da flauta. O espectro da nota está ilustrado na Figura 19 a seguir.

1º harmônico	586,8 Hz - na verdade segundo harmônico para ré3 (D3)
2º harmônico	1173,6 Hz
3º harmônico	1759,5 Hz
4º harmônico	2346,3 Hz

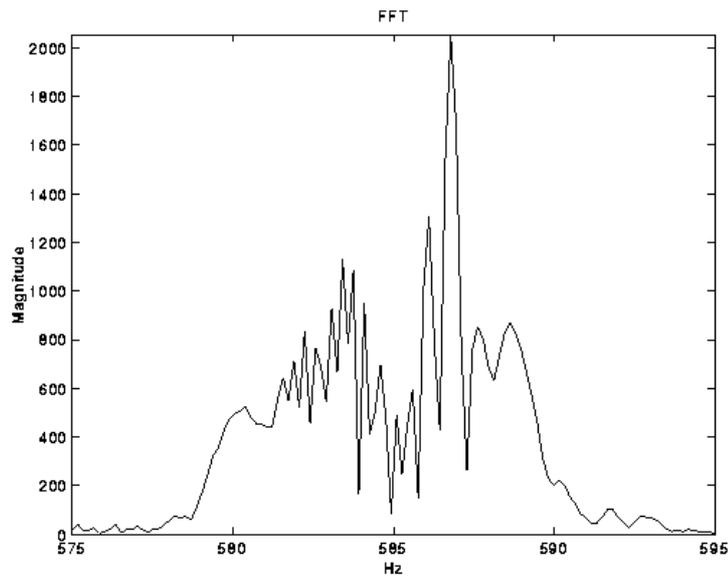
Figura 19 - FFT da nota ré4 tocada (exemplo de sustentação com dinâmica)



Um pico em 57,70 Hz aparece discretamente à esquerda do primeiro harmônico. A frequência está muito próxima de 60 Hz, a frequência da rede, e deve ser oriunda de efeitos espúrios devido à gravação ou processamento digital do sinal acústico. Não há evidências de que faça parte de qualquer série harmônica para esta nota. Que seja componente do ruído de sopro é uma hipótese improvável.

A Figura 20 (a seguir) mostra a região do espectro onde situa-se o primeiro harmônico deste sinal (a fundamental). Note que se trata de uma banda relativamente complexa, cujo centro está na frequência acima especificada para o primeiro harmônico.

Figura 20 - Região do primeiro harmônico (fundamental) para a nota ré4 (D4)



Extensão da amostra: 2^{18} pontos (262144). **Wavelet de análise:** D16

Análise dos níveis:

Nível 17: Por inspeção visual da representação gráfica do nível verifica-se a presença de um ruído de fundo por toda a extensão do exemplo. É provável que algum ruído de fundo exista neste exemplo independente do sinal acústico. Parte dele, entretanto, está associado ao ruído de sopro. De amplitude média regular em torno de 44 unidades, o ruído de fundo está mascarando o ruído do sopro abaixo desse limiar de magnitude, e as componentes do sopro só visíveis emergindo nas proximidades do centro do exemplo, onde as dinâmicas são mais fortes. Tal emergência é comprovação de que a dinâmica está fortemente correlacionada ao sopro.

Nível 14: Patamares sucessivos no ataque e no colapso, respectivamente crescentes e decrescentes.

Nível 13: Envelope exibe forma mais próxima do envelope original.

Nível 12: Nível mais suave, envelope aproximável por rampas. Contornos suaves da forma de onda.

Nível 11: É 10818 vezes menos energético que o nível anterior. Apresenta textura e contornos de ruído de baixa frequência, revelando certa semelhança (gráfica) com o nível 17. Também apresenta um sinal ruidoso de amplitude média regular por toda a extensão do exemplo, e que provavelmente mascara o componente acústico da interpretação cuja amplitude esteja abaixo desse limiar, sejam gestos musicais (que se traduzam em transientes) ou componentes vibratórios (harmônicos).

O nível, todavia, é sensível a variações de dinâmica, uma vez que apresenta amplitudes mais intensas onde esta é mais forte. Também o som apresenta diferenças, com o acréscimo ou intensificação de certas frequências.

Nível 10: Nível não apresenta sensibilidade alguma à variação dinâmica contínua. Consiste numa forma de onda não suave, de textura sônica ruidosa, contornos irregulares e pontilhado de pulsos e transientes aparentemente aleatórios, sem uma periodicidade definida. Ruído de baixa frequência associado. Superposto ao ruído grave, todavia, é perceptível um *pitch* estável e

estacionário por todo o exemplo, que centra-se em 114,56 Hz, a frequência de pico máximo mais expressiva no espectro do nível (não mostrado).

Nível 9: O nível apresenta um sinal de textura ruidosa, insensível a variações dinâmicas, se estendendo por todo o exemplo, e modulado por uma baixa frequência (com baixo índice de modulação) como pode ser aferido por simples inspeção visual das flutuações de amplitude periódicas. Este nível é mais energético que o nível anterior mais fino (10), e isto sugere que ruídos e vibrações de baixa frequência, sem correlação alguma com o material musical sob análise, encontrem-se presentes no som original, e aqui neste nível concentrem sua maior contribuição. Pode-se eliminá-los do som desprezando os coeficientes deste nível, uma vez que a sua ausência não prejudica a composição harmônica do som musical. Uma reconstrução assim foi realizada (`rec1.mat`) e, no sistema de áudio utilizado, não foi perceptível qualquer diferença com o sinal original. A comprovação de diferenças, porém, podem ficar evidentes se um sistema mais sofisticado de áudio digital, em estúdio apropriado, for utilizado. Cabe a um trabalho futuro dar cabo a este experimento.

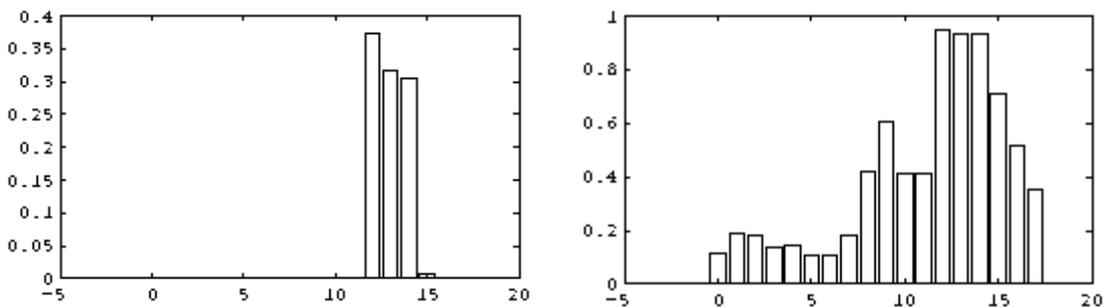
Nível 7: Inaudível, com exceção dos *clicks* nas bordas (efeitos de bordas)

Nível 6: Maiores amplitudes exibidas nas bordas do sinal.

Nível 5: Bordas mal localizadas. Inaudível, não revela aspectos musicalmente relevantes.

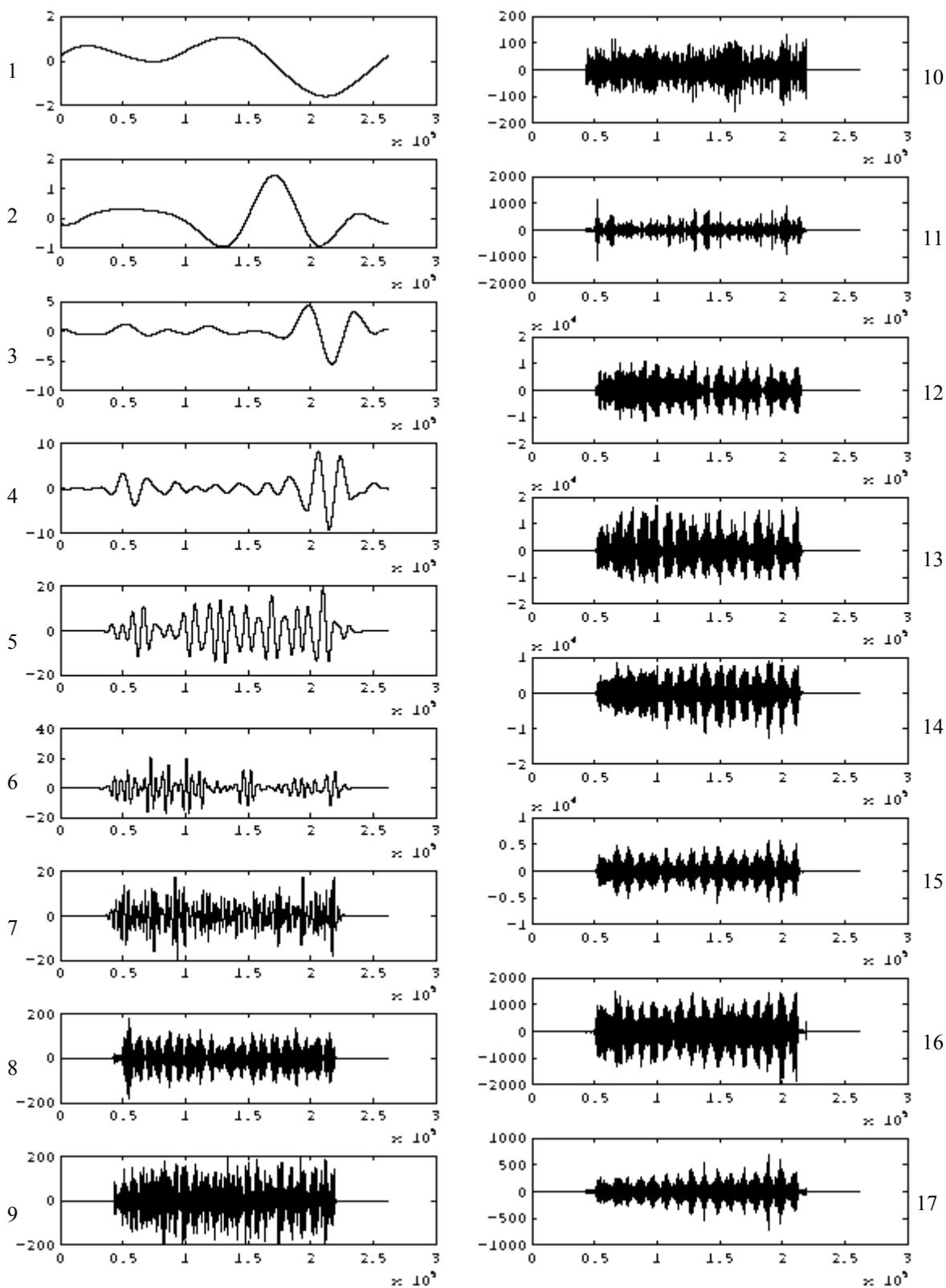
Nível 1: Revela um aumento de energia nos níveis mais grossos, diferente dos outros exemplos.

Distribuição de energia:



Mais de 99% de toda a energia da nota tocada sob variação dinâmica se concentra nos níveis de 12 a 15, sendo o nível 12 o mais energético. A distribuição evidenciada no gráfico à direita mostra que os níveis mais baixos participam ligeiramente mais do conteúdo energético do sinal em comparação aos exemplos (análises de outros eventos) anteriores. Todavia, sua participação é desprezível.

Flauta: Sustentação com Vibrato - nota fá4 (F4)



Evento: Sustentação da nota fá4 (F4) na segunda oitava da flauta com aplicação de vibrato. O espectro de fá4 (F4) neste exemplo com vibrato é mostrado na Figura 21 a seguir. A tabela em seguida identifica as raia com os harmônicos do tom.

Gráfico do sinal original:

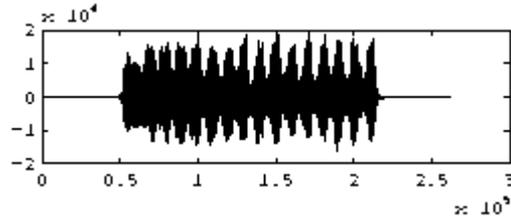


Figura 21 - FFT da nota fá4 (F4), tocada sob aplicação de vibrato (linhas principais, de 0-3000 Hz)

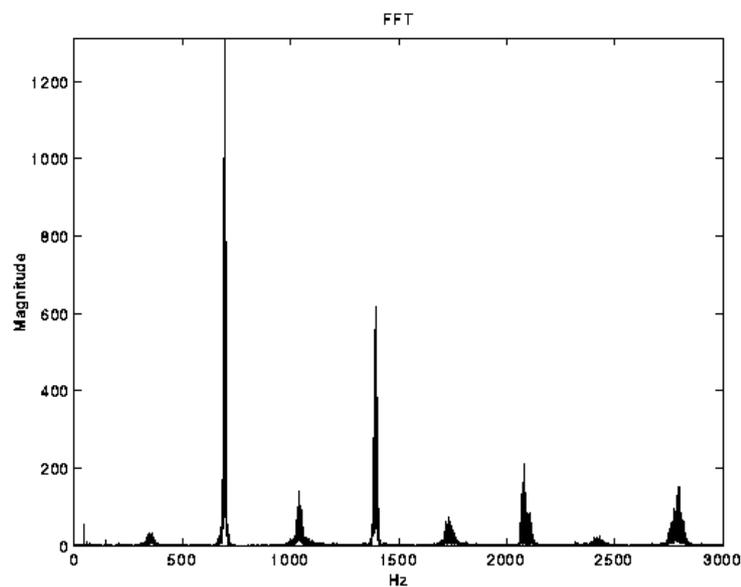


Tabela: Análise espectral do sinal original (raia em ordem crescente de frequência)

Frequência(s) central da raia	Harmônico correspondente de F4
696,9692 Hz (fundamental)	1 ^o (segundo harmônico aproximado de fá3 (F3), localizado em 346,55 Hz, a banda que aparece muito discreta à esquerda de 500 Hz)
1041,3 - 1039,5 Hz	(terceiro harmônico de F3)
1393,9 Hz	2 ^o (quarto harmônico de F3)
1734,4 e 1730,9 Hz	(quinto harmônico de F3)
2082,5 Hz	3 ^o (sexto harmônico de F3)
2426 Hz	(sétimo harmônico de F3)
2793,3 Hz	4 ^o (oitavo harmônico de F3)

Havendo uma modulação de amplitude devido ao vibrato, as raia dos modos harmônicos originais da nota ficam ligeiramente desviadas para cima e para baixo, gerando duas raia novas cuja diferença (intervalo de frequência entre elas) corresponde à flutuação visível e audível do vibrato. As frequências reais do tom não devem, portanto, ser atribuídas diretamente aos valores

acima, porque eles são bandas laterais de uma *modulação de anel*⁸. No entanto, os valores acima de picos das raias devem corresponder aos valores das frequências harmônicas desviadas para cima ou para baixo da metade da frequência de vibrato.

Uma estimativa para esta frequência pode ser feita calculando o período entre máximos na onda flutuante do sinal original, o que fornece um valor entre 4 e 4,2 Hz para o vibrato. Uma estimativa alternativa pode ser feita verificando a diferença entre picos dentro de uma mesma raia do espectro do sinal original, na esperança de que eles sejam bandas laterais geradas pelo vibrato. Essa interpretação no entanto é suscetível a uma larga margem de erro, porque na existência de várias frequências que são moduladas por uma frequência muito menor (caso do vibrato) as bandas laterais que surgem se combinam e o contorno original do sinal sem vibrato não é preservado nas bandas laterais, dificultando a identificação destas. Uma aproximação sugerida pela análise das bandas presentes na raia espectral mais forte para esta nota - *em torno de 696,96 Hz (segundo harmônico de F_{á3}, primeiro harmônico para F_{á4})* - revela uma diferença de 4,03 Hz em média para várias linhas espectrais, valor bem próximo daquele obtido pela estimativa anterior.

Extensão da amostra: 2¹⁸ pontos (262144). **Wavelet de análise:** D16

Análise dos níveis:

Nível 17: O sopro acompanha o vibrato para a flauta. Na verdade, ele é a origem do vibrato. A modulação de amplitude agora, exagerada, é intencional e não fruto do processamento da análise *Wavelet* em multiresolução, gerando bandas especulares. O espectro do nível mostra densa população de frequências entre 5000 e 22050 Hz. A linha espectral mais intensa está em 11517 Hz. Há bandas intensas espelhadas ao redor de 11025 Hz.

Nível 16: Objetos (nódulos) de ataque e colapso. Rampas de subida e descida bem íngremes, envelope aproximado de um degrau sob uma "visão panorâmica" da nota.

Nível 15: Modulação de amplitude bem visível.

Nível 14: Região próxima do ataque com amplitudes reduzidas. Dois *pitchs* (vozes) existentes modulados em amplitude.

Análise FFT do nível revela 8 raias principais, abaixo de 5000 Hz. A mais alta, em 1393,9 Hz (segundo harmônico de f_{á4}, com linhas de máximo distando cerca de 4,5 Hz), depois uma banda lateral desta raia logo abaixo em 1362,3 Hz (identificando uma modulação de amplitude com batimento de 31,6 Hz, as bandas laterais situadas em torno de divisor da frequência de amostragem), depois o quarto harmônico em 2082,5 Hz (diferença entre os picos nesta raia é aproximadamente de 4,4 Hz), e a terceira raia mais intensa em 2795,6 Hz (oitavo harmônico) também espelhada em torno de uma frequência que é fração racional da frequência amostragem.

A frequência de batimento do vibrato, estimada pelo cálculo aproximado do período da flutuação no sinal do nível, está entre 4 e 4,3 Hz. Pelo método da distância entre picos da mesma banda no espectro, o valor médio estimado gira em torno de 4,4 Hz.

Dois aspectos devem ser levados em conta nestas estimativas:

⁸ Uma espécie de modulação de amplitude com portadora suprimida, abordada no capítulo 5, item 5.2 da dissertação.

1. Não é o objetivo no momento quantificar com precisão tais propriedades do vibrato, justificando a não utilização de métodos mais sofisticados de análise espectral para obter valores mais precisos.
2. A própria natureza da interpretação acústica e da fisiologia dos gestos de sopro admite uma diversidade constante, isto é, todos os fenômenos consistem em processos variantes no tempo, o que implica numa frequência de vibrato variando numa faixa de frequências.

De posse de mais dados estimados a partir do sinal deste nível, podemos pois admitir que a frequência de batimento do vibrato neste exemplo situa-se numa faixa entre 4 e 4,5 Hz.

O vibrato não modula significativamente o início do exemplo para as frequências cobertas neste nível, conforme pode ser facilmente verificado por simples inspeção visual do gráfico do sinal do nível na região de ataque, onde as flutuações não exibem o padrão mais regular observados na região de sustentação. Isso é uma característica esperada por dois motivos:

1. O intérprete só dará início ao vibrato uma vez que se certifique da correta entonação e estabelecimento do tom
2. No ataque, tentativas de se impor uma modulação no jato de sopro para produzir vibrato vão também produzir uma instabilidade maior e portanto uma dificuldade para os modos harmônicos se estabelecerem.

O motivo 1 para o momento é mais relevante, porque o vibrato é intencional e o flautista só o inicia depois de um certo tempo.

A quebra da integração dos harmônicos fundidos num timbre único de flauta se revela na percepção de *pitchs* segregados como vozes. Alguns *pitchs* soam a partir de parciais da série harmônica. Outros, a partir de imagens espelhadas de raias originais do sinal. Se o espelhamento ocorrer para uma raia muito intensa, então a raia espelhada também apresentará intensidades suficientemente altas, e um novo *pitch* intenso é ouvido desta.

Nível 11: Transiente ou pulso na sobrelevação do ataque (maior amplitude do nível). Ruído grave de fundo perceptível. Sinal não é suave, apresenta muitos transientes e picos, de contorno irregular e envelope amorfo (não segue nenhuma forma convencional em música).

Vibrações da série harmônica ainda é perceptível na forma de um tom, situado em faixa de frequências mais alta que o ruído de fundo.

Nível 10: Textura ruidosa, envelope irregular, não suave. Audível, ruído de baixa frequência contribuindo relevantemente neste nível.

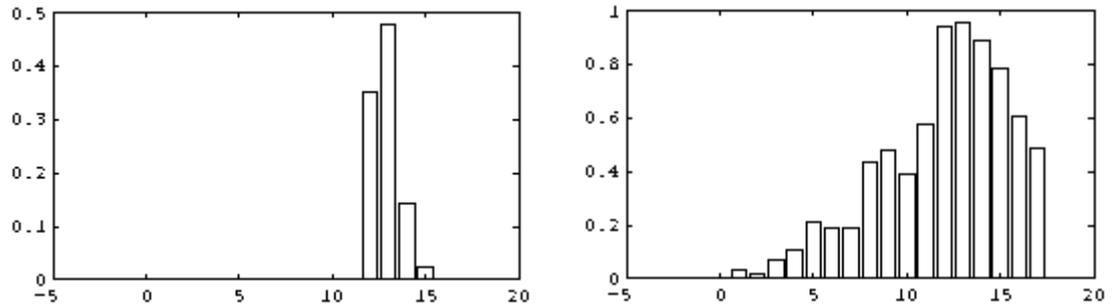
Nível 9: Formato semelhante ao nível anterior. Apresenta modulação de amplitude complexa, e fica difícil determinar uma frequência modulante como em níveis anteriores. O espectro também revela duas linhas intensas espelhadas em torno de submúltiplo da frequência de amostragem (45,758 e 40,375 Hz), duas outras em 58 e 69 Hz, mais uma densa população de frequências na faixa de 25 a 120 Hz, sem uma frequência de vibrato de destaque.

Nível 8: Espectro apresenta mesmas linhas espelhadas que o nível anterior. Ainda audível. A frequência do batimento mais evidente no sinal corresponde à diferença entre as linhas espelhadas, portanto é uma frequência gerada pelo algoritmo de análise *Wavelet* (cerca de 5,4 Hz).

Nível 7: No limite da audição. Só pulsos transientes mais intensos são ouvidos como *clicks*. Vibrações subsônicas estão presentes.

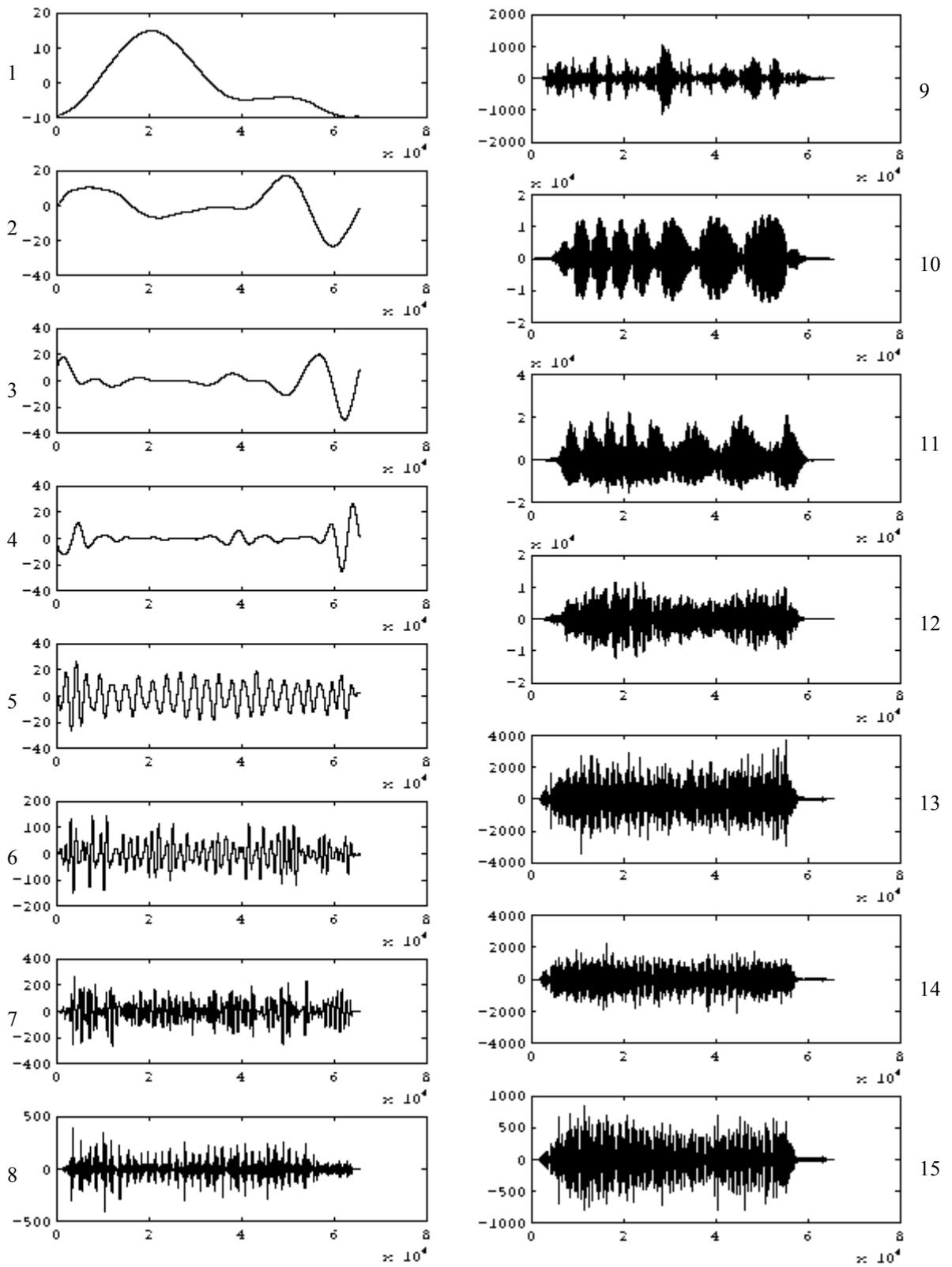
Nível 4: Início e final (mais amplo) são mais destacados. A forma de onda é suave. Inaudível. O sinal no nível extrapola os limites do sinal original, apresentando forma de onda nas regiões acolchoadas com zeros (*zero padding*). O nível anterior também faz isso, mas invade a região do acolchoamento com amplitudes bem mais baixas.

Distribuição de energia:



Mais de 99% de toda a energia se concentra nos níveis de 12 a 15, sendo o nível 13 o mais energético.

Flauta: Sustentação com Frulato - nota fá4 (F4)

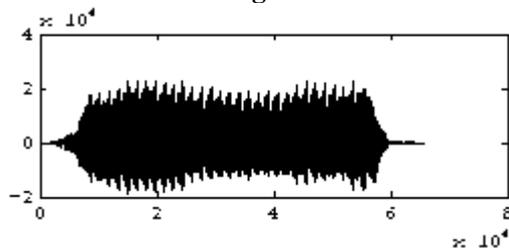


Evento: Sustentação com frulato. Exemplo de um nota executada com efeito de frulato, desde o ataque até o final da nota. Ataque em frulato do tipo "tu" à nota fá4 (F4) e sustentação do efeito.

Ouvido a 11025 é notável o efeito: inicialmente ouve-se o sopro e logo em seguida, quando quebra-se a inércia da coluna de ar, estabelecem-se os padrões estacionários, que são periodicamente modulados por um jato intermitente de sopro. O efeito de intermitência é produzido pela vibração da língua contra o céu da boca, interrompendo o fluxo de ar periodicamente. A cada ciclo as vibrações estacionárias tendem a colapsar, e o retorno do fluxo de ar ocorre antes que morram completamente. No entanto, a cada retorno do fluxo de ar o ruído de sopro é novamente introduzido, e temos um realce das componentes do sopro a cada ciclo.

Uma análise FFT da nota mostra a presença de uma série harmônica predominante, e obscurece o fato de que há um frulato executado na nota. Três raias se destacam: o **segundo harmônico do fá3** (F3) que na verdade é a fundamental desta nota F4 (em 691 Hz, amplitude 5456 unidades), o **quarto harmônico** (em 1387,5 Hz, amplitude 1616 unidades) e o **sexto harmônico** (em 2100,8 Hz, amplitude 472 unidades). Há portanto uma predominância dos parciais pares, no entanto as raias são relativamente mais largas comparadas às notas interpretadas sem o frulato, com picos locais distribuídos ao redor das frequências centrais discriminadoras da raia, acusando a presença de bandas laterais de uma modulação de anel, resultado dos desvios em frequência provocados pelo frulato. A frequência com que o frulato é executado não é entretanto de fácil verificação no espectro da nota - por exemplo, pelo menos duas frequências moduladoras diferentes podem ser verificadas na FFT entre 650 e 750 Hz.

Gráfico do sinal original:



Extensão da amostra: 2^{16} pontos (65536). **Wavelet de análise:** D16

Análise dos níveis:

Nível 15: O ataque consiste numa região de subida gradual até um ponto máximo do envelope. Na ponta inicial a presença do sopro é predominante. No decorrer da nota, no período de sustentação, observa-se uma textura pouco suave no envelope, com muitos picos localizados de amplitude e uma sequência de pulsos periodicamente distribuídos pela nota até a região do colapso, quando cessa a modulação do frulato e dá-se início ao colapso propriamente.

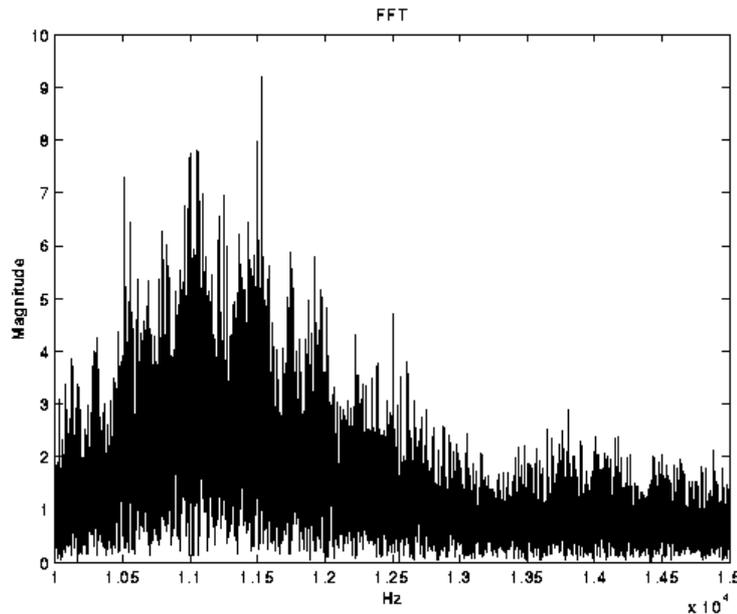
O colapso revela-se neste nível como um "apêndice" com amplitudes baixas, consistindo num ruído de fundo na sua maior parte.

O espectro do nível revela uma densa população de frequências acima de 5000 Hz até 22050 Hz, com picos duplos resultado de uma modulação de anel (modulação de amplitude sem portadora) gerada pelo efeito de frulato. O pico mais amplo está em 11535 Hz e sua banda lateral espelhada se encontra 42 Hz abaixo, em 11493 Hz, evidenciando uma modulação de amplitude de 21 Hz, a frequência estimada para o movimento lingual que oriunda o frulato.

A mesma frequência é extraível da segunda raia mais ampla, com pico principal em 11046 Hz, e sua banda lateral inferior espelhada, também a 42 Hz abaixo, em 11004 Hz, novamente

mostrando a frequência do ciclo intermitente de sopro do frulato. A FFT do nível está ilustrada na Figura 22, mostrando o espectro entre 10000 e 15000 Hz, onde se encontram as raias principais do som neste nível, e pode-se visualizar a modulação de anel implícita citada acima, com picos gêmeos separados de 42 Hz.

Figura 22 - FFT do nível 15 (sustentação com frulato)



A análise do espectro do som original não torna evidente a presença da frequência do frulato, e o sinal no nível 15, tipicamente um nível que realça as componentes do sopro, é capaz de clarear os aspectos cíclicos envolvidos no sopro intermitente do frulato.

O som do nível revela a periodicidade do sopro do frulato, tornando-se sua principal característica. Nota-se que a periodicidade é estabelecida desde o início da nota, caracterizando o ataque de frulato a partir do *onset* da nota. Somente na região do colapso o frulato é interrompido. O som lembra um chocalho de cascavel quando reproduzido em 11025 Hz.

Um nível mais elevado de energia para este nível mais fino revela que grande parte da energia do frulato fica esboçado no sopro.

Nível 14: O envelope do nível se assemelha ao do nível anterior, mostrando entretanto um nível de ruído de fundo menor (visível pelo apêndice do colapso). O som do nível é semelhante ao de um chocalho de cascavel.

A análise espectral estática do nível mostra que ele ainda representa uma significativa concentração de componentes do sopro, dada a ampla distribuição de frequências por uma faixa larga do espectro acima de 3000 Hz.

Nível 13: A presença do ruído de sopro ainda é forte como pode-se verificar pela FFT (espectro estático) do nível, e a periodicidade do ciclo intermitente do frulato é a mesma calculada nos níveis anteriores. O espectro (não mostrado) revela uma densa população de frequências entre 2000 e 8000 Hz principalmente, mas raias "mais comportadas" já aparecem claramente exibindo um padrão de série harmônica. O pico máximo do espectro se dá em 3411,7 Hz, e entre 3300e 3600 Hz aparecem várias raias onde a decifração visual da frequência do frulato associada se torna mais complexa.

Ao redor de $1/16$ da frequência de amostragem (2756,25 Hz) encontramos bandas laterais espelhadas, com ligeiras diferenças em amplitudes, um efeito de duplicação de bandas causado pelo algoritmo de análise *Wavelet* em multiresolução⁹.

A audição do nível sugere que há um *pitch* predominante e a frequência do frulato é aproximadamente a mesma dos níveis anteriores. Cálculos da distância entre picos, numa tentativa de estimar a frequência modulante associada ao frulato, revelam valores entre 18 e 27 Hz de frequência para o frulato. A estimativa da frequência de frulato pela simples análise visual dos picos e raias do espectro do nível é, entretanto, praticamente inviável, requerendo para tal o uso de técnicas mais complexas, como por exemplo o emprego de análise *pitch*-síncrona (*pitch synchronous*) sobre o sinal do nível, como Evangelista (1993) realizou em seus experimentos (veja citação de seu trabalho nas referências bibliográficas da dissertação).

A audição do nível a 11025 Hz revela que os componentes harmônicos do tom (vibrações estacionárias) surgem após 3 ou 4 ciclos do frulato, e que permanecem presentes até o fim da nota, flutuando sob a modulação do frulato, conforme já dito anteriormente.

Nível 12: O envelope deste nível revela duas características novas:

1. O ataque se processa numa ascensão gradual de amplitude
2. Existem duas regiões de máximos e uma de mínimo (amplitudes relativas) na fase de sustentação. O efeito entretanto sobre a amplitude percebida do som é desprezível, e até observa-se que a região de menores intensidades na sustentação é portadora de um som mais puro e regular.

A visualização da forma de onda revela que a frequência modulante do frulato no nível pode ser na verdade uma combinação de frequências modulantes, uma vez que pulsos perfeitamente distanciados não são verificados, e sim verifica-se a existência de uma modulação mais complexa. A audição do nível a 11025 Hz revela um início soprado seguido, após alguns ciclos, pela ascensão dos parciais harmônicos. Na região do colapso também verifica-se uma ressonância após o fim da excitação (sopro) produzindo uma espécie de eco ou reprodução menos intensa do último ciclo do som.

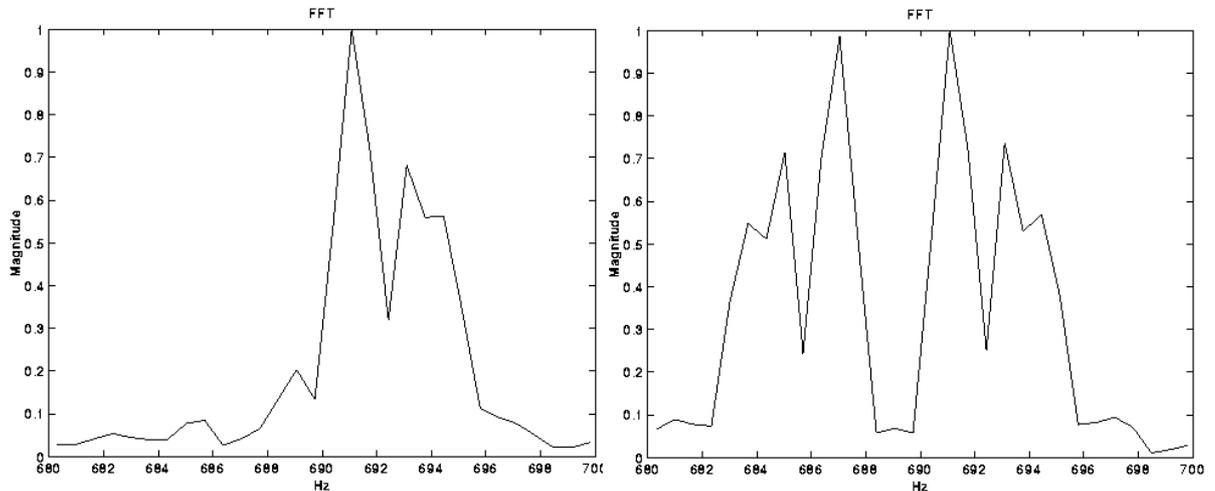
Nível 11: A inspeção gráfica da onda deste nível revela pelo menos duas regiões com diferentes modulações de amplitude, separadas aproximadamente ao meio da nota. De fato, o que se verifica é que a frequência modulante simples do frulato, calculada anteriormente em aproximadamente 21 Hz, não se aplica a todos os níveis, o que sugere a ocorrência de diferentes modulações de amplitude com diferentes frequências modulantes aplicadas a cada faixa do espectro audível. No nível 11 por exemplo nenhuma frequência modulante aparece com destaque, como 21 Hz o fez no nível 15. O espectro do nível basicamente revela as mesmas bandas existentes no espectro do sinal original, a única diferença está no aparecimento das raias espelhadas em torno de frequências que são razão de números inteiros da frequência de amostragem (f_a): $1/64.f_a$ (~689 Hz) e $1/32.f_a$ (~1378 Hz). Essa modulação, que tem origem no algoritmo usado para a análise *Wavelet*, é a causa da modulação em amplitude, tanto visualizada no gráfico da onda quanto percebida via audição.

A Figura 23 mostra a primeira banda espectral relevante do sinal original e do sinal ao nível 11, respectivamente (espectros normalizados). Note que o pico em 691 Hz (fundamental da nota)

⁹ Leia discussão sobre este efeito no capítulo 5, item 5.2 da dissertação.

aparece em ambas as figuras, mas ao nível 11 ela aparece espelhada em torno de $1/64 \cdot f_a$ (689,0625 Hz).

Figura 23 - Primeira banda relevante da FFT do sinal original (esquerda) e do nível 11 (direita)



A análise auditiva do sinal no nível 11 mostra um sinal mais puro e desprovido de ruídos, com uma só voz percebida, e uma modulação em amplitude percebida como um batimento com frequência em torno de 9 Hz para uma primeira região (aproximadamente de 0 a 680ms) e 4,5 Hz para uma segunda região (de 680ms ao final).

Nível 10: O nível apresenta características gerais semelhantes ao nível 11, como a fácil audição e visualização de diferentes frequências modulantes em forma de batimento; entretanto apresenta melhor suavidade, "pureza" de tom e o envelope possui contornos mais "contínuos". O batimento verificado no nível anterior fica muito mais explícito neste nível, onde as cristas e vales da onda ficam melhor caracterizados.

Uma análise mais cuidadosa da forma de onda apresentada no nível (e da modulação em amplitude em forma de batimento que molda seu envelope) mostra que existe uma variação contínua da frequência percebida de batimento. No início da nota esta frequência é máxima, e decresce a medida que a nota progride e o efeito do frulato se estabiliza. Isto sugere **(1)** que o regime permanente de um tom emitido sob frulato requer algum tempo para se estabilizar, ou **(2)** que o gesto musical de frulato evolua, alterando seu padrão modular no tempo. O batimento mais rápido ao início mostra que os modos harmônicos (representados neste nível) oscilam em amplitude respondendo ao ciclo de excitação e corte (sopro e ausência de sopro) do frulato. Esta frequência de oscilação tende a diminuir a medida que os modos vão ganhando energia em cada ciclo, e podem portanto manter suas oscilações por mais tempo. Esse comportamento pode ser vital para se produzir um frulato sintetizado com qualidade bem próxima ao natural.

Nível 9: Cerca de 600 vezes menos energético que o nível anterior, o nível 9 apresenta um envelope também diferenciado daqueles dos níveis anteriores, e seu som não guarda semelhança com os sons de níveis anteriores.

O som do nível não revela a preponderância dos modos harmônicos do tom, como nos níveis mais energéticos. O envelope não é suave, apresenta várias "descontinuidades", revela ataques e colapsos semelhantes aos dos níveis anteriores, com início e final graduais. A presença de uma modulação de amplitude em forma de batimento é entretanto uma característica impressa devida ao frulato, e inexistente de forma tão regular em notas interpretadas com outras técnicas.

Nível 8: A presença da frequência característica do frulato se faz novamente relevante agora no nível 8, muito menos energético que os anteriores. O envelope do nível revela inúmeras "descontinuidades", e o colapso, onde o frulato é cessado, não apresenta os picos localizados que se distribuem por toda a nota onde o frulato existe - inclusive no início, uma vez que a nota foi atacada sob frulato.

O som é característico dos níveis baixos menos energéticos, e realça os componentes ruidosos de baixa frequência que existem no som original, sejam eles oriundos de ruídos ambientes, das fases de gravação, conversão e preparo das amostras, seja oriundo de vibrações graves excitadas nos instrumentos durante a emissão dos tons, ou de gestos musicais devido a manipulação de arcos, sopros, respiração e dedilhados. O espectro do nível revela uma densa população de frequências entre 50 e 250 Hz, com picos em 109,0 e 134,6 Hz.

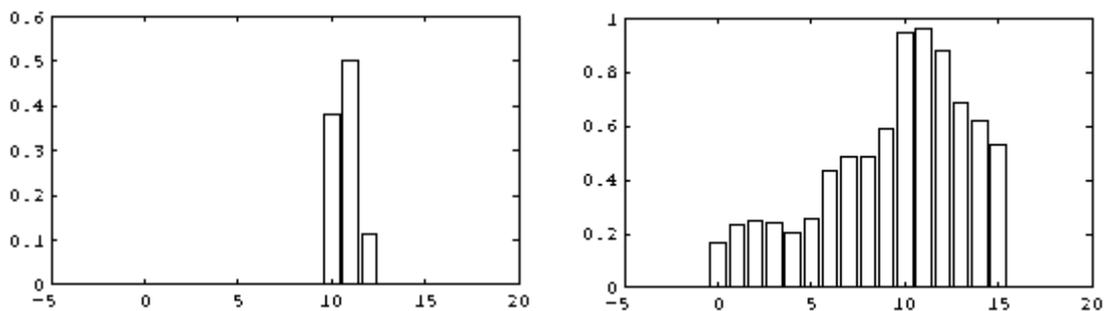
Nível 7: Não revela aspectos importantes do sinal musical, e guarda semelhança com os sinais de outras notas sem frulato.

Nível 6: Espectro revela pico de frequência em 27,59 Hz, seguido por outro em 40,37 Hz, frequências muito baixas. Não há contaminação de 60 Hz.

Nível 5: Inaudível (exceto pelo *click* do início e fim). Mostra um onda contínua, com frequência (*pitch*) central em 15,47 Hz, e um segundo pico local em torno de 20 Hz, um valor muito próximo à frequência de frulato estimada no nível 15. Não pode-se todavia afirmar que a frequência modulante do frulato fica evidente neste nível.

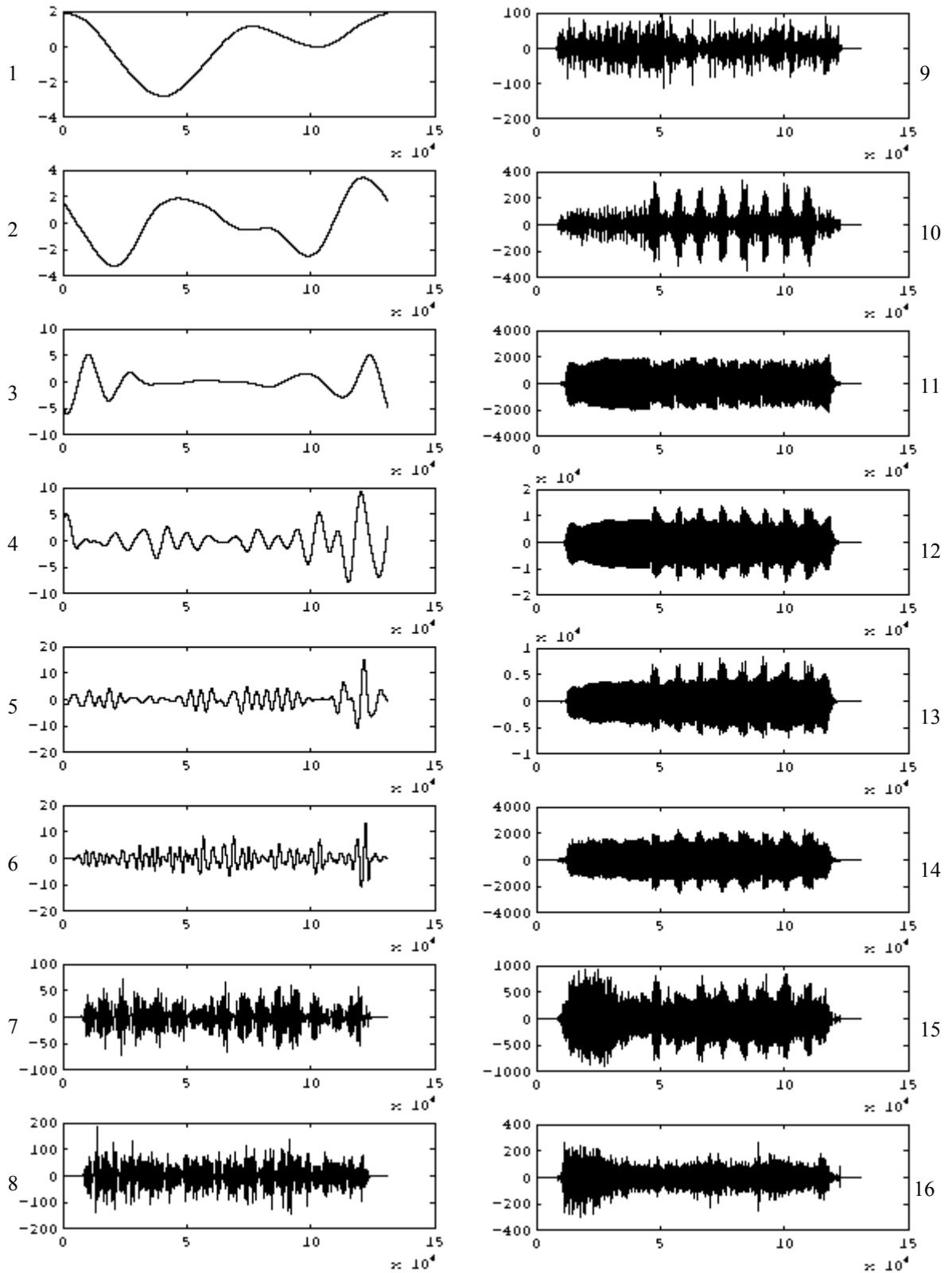
Do nível 4 para baixo as formas de onda não exibem as flutuações periódicas observadas nos níveis superiores, e não mostraram-se importantes para a verificação do frulato.

Distribuição de energia:



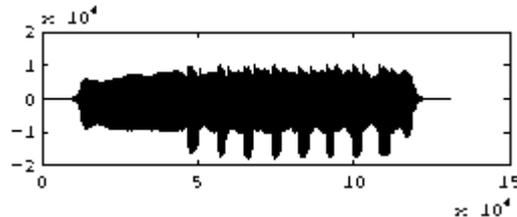
Quase toda a totalidade da energia do exemplo se concentra nos níveis de 10 a 12, sendo o nível 11 o mais energético (com aproximadamente 50% de toda a energia). Os níveis mais baixos -*em especial de 6 a 9*- revelam-se mais energéticos neste exemplo de frulato do que em outros eventos abordados neste trabalho, sugerindo que a modulação provocada por este adorno musical se imprime por um número maior de níveis *wavelets*.

Flauta: Sustentação com Trinado - notas lá4 (A4) e si bemol4 (Bb4)



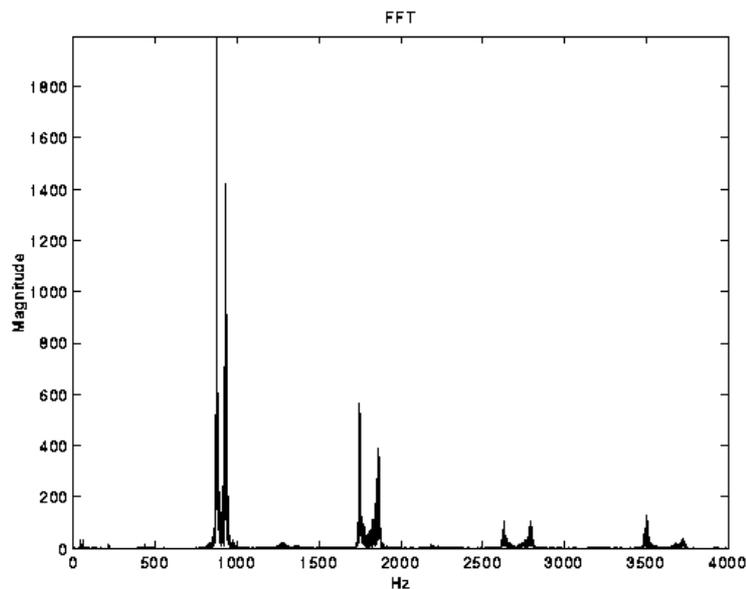
Evento: Sustentação com trinado em lá4 (A4) e si bemol4 (Bb4). A nota atacada é o segundo lá natural da flauta (A4, na segunda oitava), e o trinado, iniciado após estabelecido o lá, é executado com 1 semitom acima, em si bemol (Bb4). O objeto de interesse na análise são as transições entre o lá e o si bemol, as quais são periódicas.

Gráfico do sinal original:



O espectro do exemplo, exibindo as duas séries harmônicas (lá e si bemol) está ilustrado na Figura 24, entre 0 e 4000 Hz. Os picos de máximos, em 873,77 Hz e 930,974 Hz, são a fundamental respectivamente do lá e do si bemol da segunda oitava da flauta.

Figura 24 - FFT do sinal original para a faixa de 0-4000 Hz (sustentação com trinado)



Extensão da amostra: 2^{17} pontos (131072). **Wavelet de análise:** D16

Análise dos níveis:

Nível 16: O ataque do lá4 é a região que apresenta maior contribuição de ruído de sopro. A região do trinado não apresenta uma flutuação ou modulação em amplitude visível. O som também não demonstra que está havendo um trinado, o que sugere que o sopro mantém-se inalterado durante o trinado, sem modulações. Em verdade, o intervalo de um semitom do trinado é pequeno o suficiente para dispensar ajustes de embocadura regulares para focalizar alternadamente as notas envolvidas.

Pulsos transientes ocorrem aleatoriamente, ora sobre uma nota, ora sobre a outra. O som apresenta uma clara textura ruidosa, não conduzindo o ouvinte à compreensão das tonalidades envolvidas e nem mesmo à conclusão de que tipo de fraseado está ocorrendo.

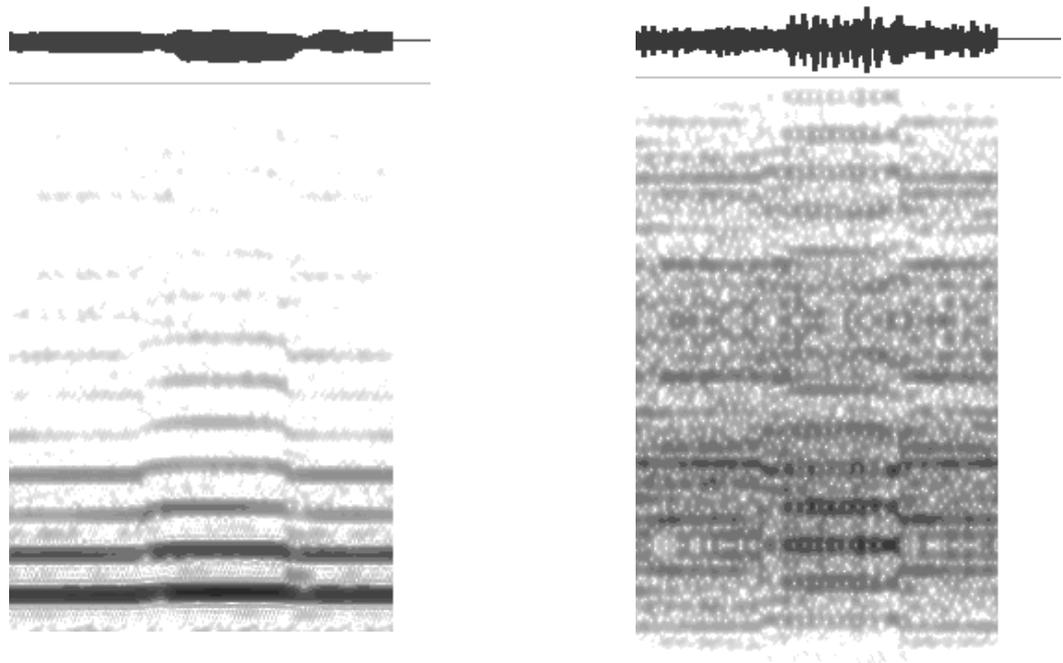
A ausência de flutuação ou qualquer diferenciação entre regiões do nível, acrescida da impossibilidade de se inferir sobre as tonalidades tocadas, revela a ausência da contribuição da série harmônica das notas tocadas.

Nível 15: O nível já revela uma flutuação ou modulação de amplitude de padrão semelhante ao observado no sinal original. O ataque apresenta intensa contribuição neste nível, provavelmente devido ao sopro, que ainda contribui neste nível. A estabilização do lá4 inicial é marcada pelo decaimento em amplitudes logo após o ataque. Em seguida inicia-se o trinado com si bemol4, as amplitudes maiores ocorrendo sobre o si bemol.

Os contornos da forma de onda são irregulares, com uma textura não suave e de padronagem ruidosa. O envelope do trinado é amorfo, e não apresenta claramente uma sequência de patamares de amplitude, como no sinal original. O som apresenta um "brilho metálico", realçado pelas altas frequências associadas às tonalidades harmônicas.

Em geral, sobre o si bemol4 as amplitudes são maiores, mas a localização do instante de transição não é definida. A Figura 25 mostra o espectrograma (de 1024 pontos, janela FFT de 8ms, 16 bits) de uma transição do trinado, uma sequência lá-sib-lá para o sinal original e para o mesmo trecho no nível 15, respectivamente. A transição no sinal original é contínua para os harmônicos mais baixos. Nas frequências mais altas, que não aparecem tão firmes e contínuas, mas em manchas ou nódulos, as transições sugerem uma curva contínua, mas ocorrem em saltos.

Figura 25 - Espectrograma de uma sequência "lá4-sib4-lá4" para o sinal original (esquerda) e para o nível 15 (direita)



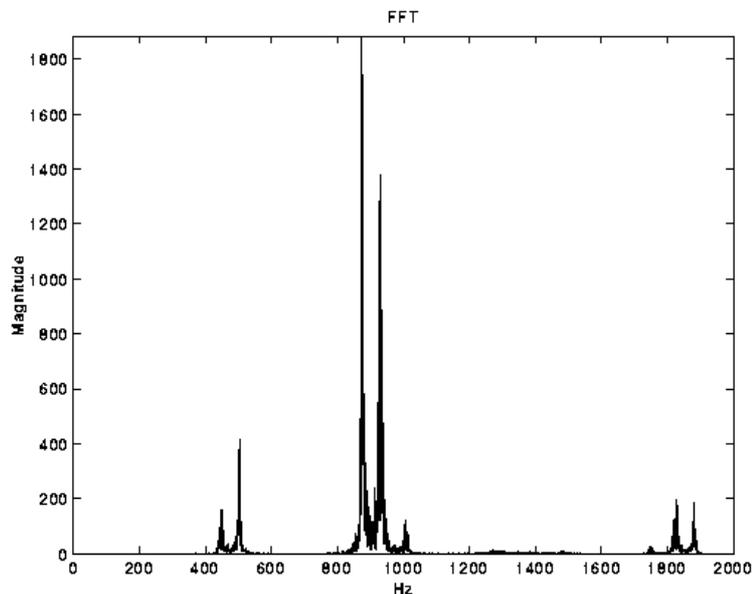
O espectrograma do mesmo trecho do nível 15 revela com mais detalhes a transição para as faixas de frequência cobertas neste nível, e vemos a ocorrência de faixas simétricas ou espelhadas, resultado da duplicação de certas bandas durante a decomposição *Wavelet*. Como o nível é muito ruidoso, vemos que a textura (gráfica) do espectrograma revela uma densa população de frequências, e as transições constituem saltos nas frequências (harmônicas) mais relevantes para os tons envolvidos.

Nível 14: Padrão flutuante de amplitudes é mais firme. Tonalidade dos sons também. Amplitudes das nota si bemol4 são maiores.

Nível 13: O sib4, de frequência mais alta, também mostra as maiores amplitudes. Este é um dos realces interessantes do trinado. Como já verificamos que o sopro se mantém praticamente regular durante o trinado, para o si bemol - que está meio tom acima de lá4 - a mesma pressão de sopro fornece maiores intensidades sonoras. Isto é uma característica da produção de tons na flauta, e deve variar dependendo das notas envolvidas no trinado e do registro onde se encontram. A síntese natural portanto de trinados não necessariamente obedece um padrão de diferenciação de amplitudes como vemos aqui, este padrão pode ser diferente dependendo das notas envolvidas no trinado.

Nível 12: A melhor localização das transições entre as notas envolvidas no trinado. Nível mais energético. O espectro entre 0 e 2000 Hz está na Figura 26 (abaixo).

Figura 26 - FFT do nível 12 do exemplo (plotado entre 0-2000 Hz)



A banda à esquerda, entre 400 e 500 hz é uma imagem (espelhada) da banda principal à direita, entre 800 e 1000 Hz. O pico máximo está em 873,77 Hz, o segundo em 930,974 Hz. O espectro entre 800 e 1000 Hz é idêntico ao espectro na mesma faixa do sinal original.

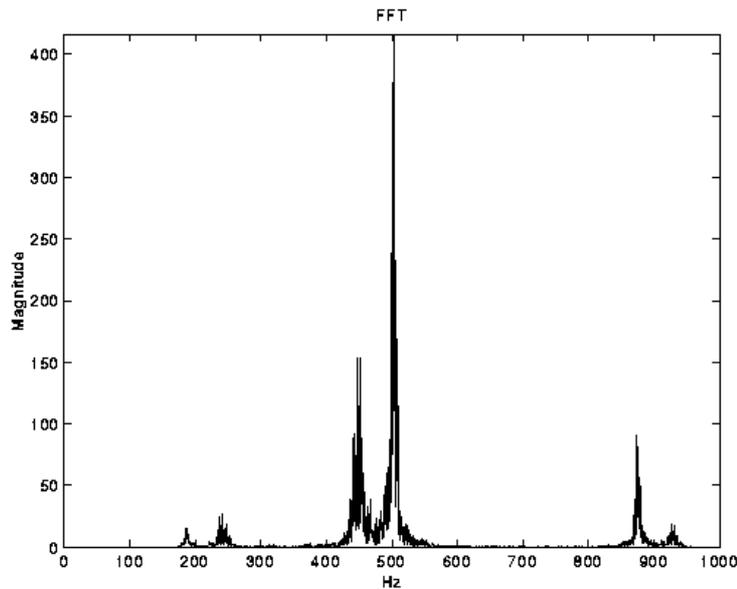
Nível 11: As amplitudes do si bemol4 são menores que do lá4. Isto porque o si bemol concentra maior energia em níveis mais altos que o lá, por estar localizado em uma tonalidade mais alta.

Os contornos do sinal não são muito suaves, e as transições de lá para si bemol são melhor localizadas por um estrangulamento (regiões de baixa amplitude) enquanto as de si bemol para lá não apresentam esta segmentação tão distinta, mas um estrangulamento muito mais discreto. De novo, estas são características que resultam diretamente da dinâmica da produção de sons em tubos vibrantes.

O *pitch* percebido das notas é diferenciado no nível, e uma sequência de *pitchs* alto-baixo-alto no sinal original soa aqui como uma sequência baixo-alto-baixo. A transição entre as notas é contínua, sem degraus ou descontinuidades. Examinadas de perto, duas notas apresentam aproximadamente a mesma modulação em amplitude, comum de se localizar nos níveis de análise *Wavelet*. O espectro do nível mostra uma maior presença do lá4 apresentando maiores amplitudes que o si bemol4, de novo um comportamento esperado, já que o lá é mais baixo (em frequência) que o si4, e portanto se concentra em níveis mais baixos.

O nível é baixo o suficiente para seu espectro realçar as bandas mais baixas do sinal harmônico presente, isto é, aquelas abaixo de 800 Hz que aparecem diminutas e fracas no espectro original. O espectro do nível está na Figura 27. As duas raias na banda central - entre 400 e 600 Hz - se espelham devido ao processamento *Wavelet* na banda à direita, entre 800 e 1000 Hz. As bandas à esquerda não são espelhos das raias principais. Os picos das raias principais estão em 504,348 Hz e 452,1973 Hz.

Figura 27 - Espectro do nível 11 (sustentação com trinado)



Nível 10: A diferença entre amplitudes exibidas pelo lá4 e pelo si bemol4 é maior neste nível. Há diferença tonal entre as regiões flutuantes no nível, indicando a presença das séries harmônicas; os tons ouvidos, entretanto, têm *pitchs* bem diferentes dos do sinal original.

Nível 9: Indiferente ao trinado. Sem segmentação das notas envolvidas, textura sônica ruidosa, forma de onda não suave, pontilhado de pulsos, caracterizando um padrão de ruído.

Nível 8: Quase inaudível. Ruído de baixa frequência.

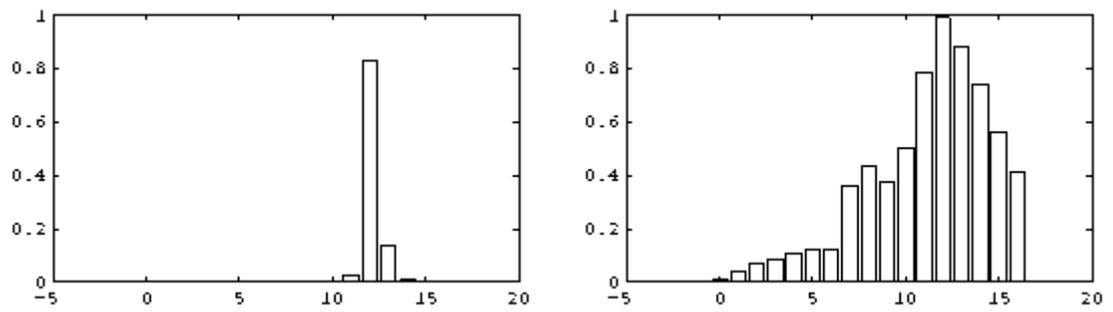
Nível 7: Idem nível anterior, menos audível ainda, só se ouvem os *clicks* dos pulsos mais amplos.

Nível 6: Amplitudes crescentes, nível de suavidade da forma de onda maior, isto porque este nível já exhibe uma "média" dos níveis anteriores, tendo como ponderação os filtros passa baixa e alta *Wavelet*. Inaudível.

Nível 5: Forma de onda extrapola limites do sinal original, invadindo a região acolchoada com zeros.

Nível 4: Bordas do sinal neste nível apresentam alta amplitude.

Nível 3: A amplitude no centro é baixa, e nas bordas altas. Sinal neste nível já é suave o bastante, aproximando-se da forma de onda básica da *Wavelet* D16.

Distribuição de energia:

Mais de 99% de toda a energia do trinado concentra-se nos níveis de 11 a 14, sendo o nível 12 o mais energético.