

TEXTO DAS ANÁLISES DE VIOLINO

(VIOLIN ANALYSIS)

PARTE INTEGRANTE/ANEXO À DISSERTAÇÃO DE MESTRADO

(ATTACHMENT OF THE MASTER DISSERTATION)

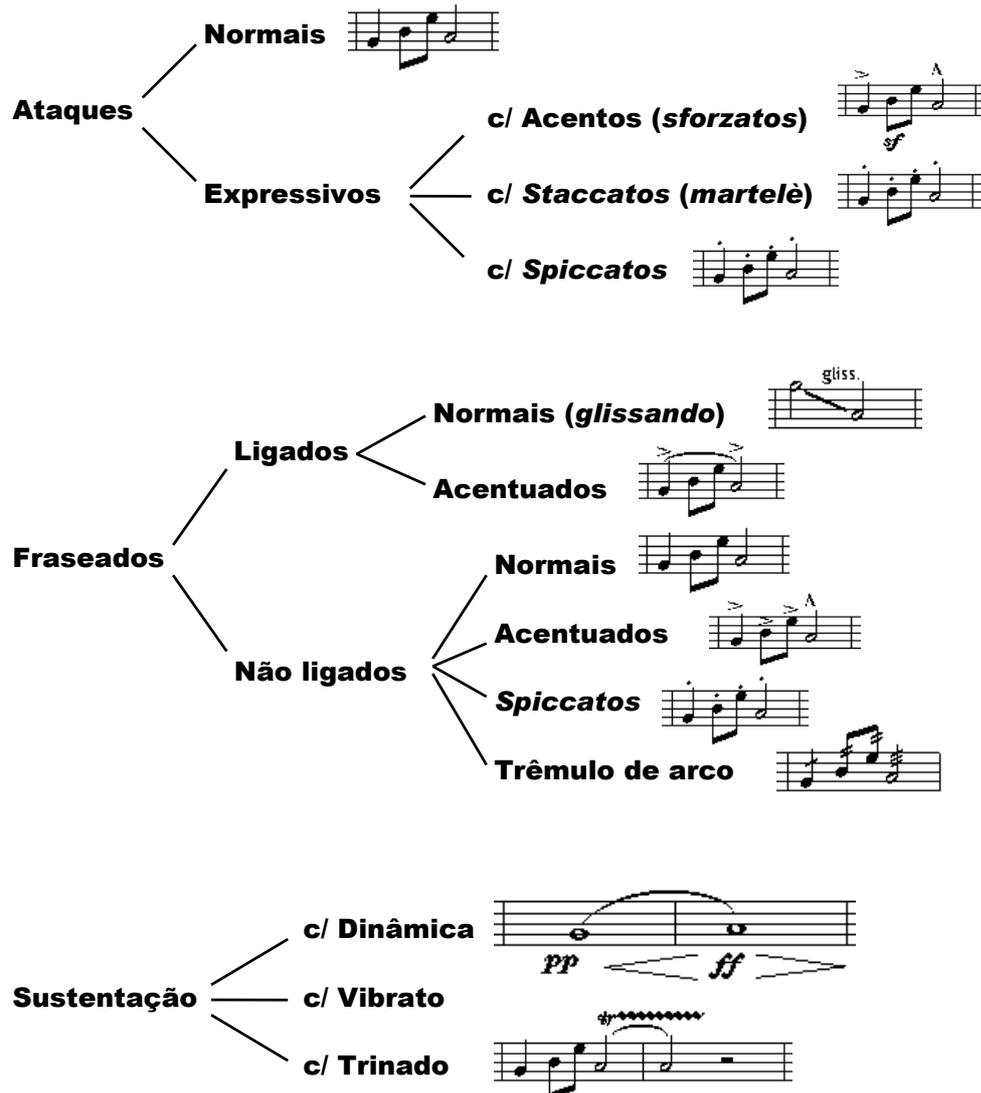
Faria, Regis Rossi Alves. **Aplicação de wavelets na análise de gestos musicais em timbres de instrumentos acústicos tradicionais**. Dissertação (Mestrado). Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Eletrônica. São Paulo 1997. 182p.

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 VIOLINO

A figura abaixo resume todos os tipos de eventos musicais amostrados e analisados para o violino. São ao todo 13 eventos analisados.

Figura 1 - Eventos musicais amostrados e analisados para o violino



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:

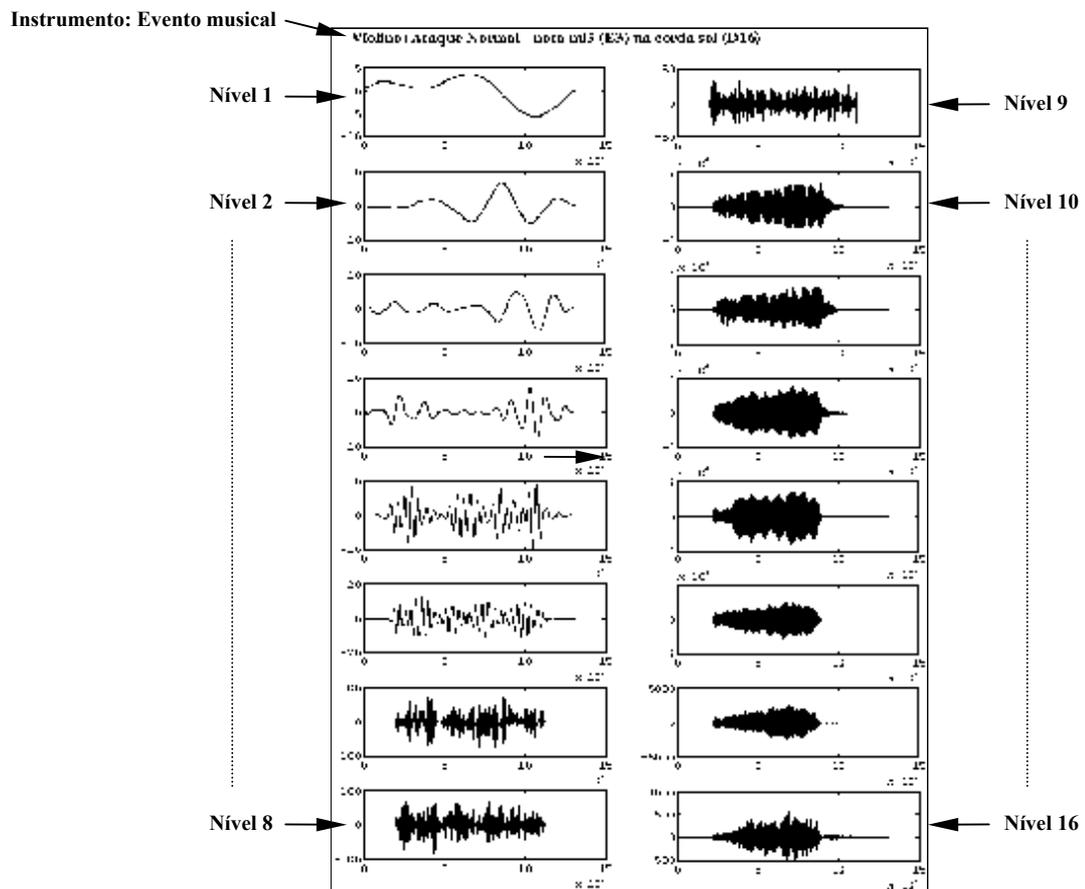
- os gráficos dos **níveis *wavelet*** da análise,
- uma descrição do **evento** (uma descrição da interpretação e/ou sobre o trecho tocado),

- 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.

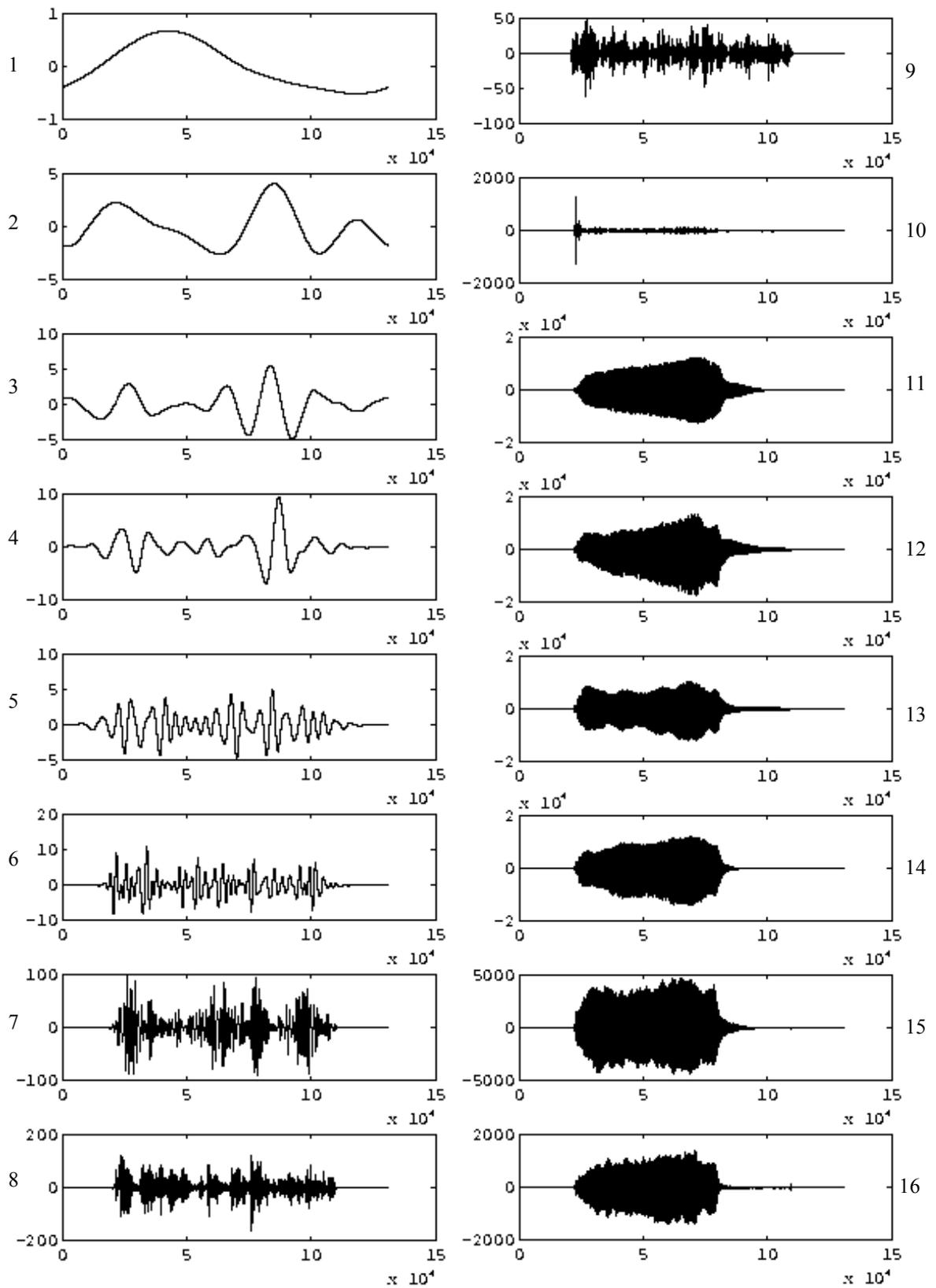
A quantidade de informação gerada é relativamente extensa, e optou-se pela exclusão da sua íntegra dentro do corpo do texto, por motivos de clareza, espaço e concisão. O material integral, todavia, está disponibilizado para consulta em meio digital anexo (CD ROM e disquete, em formato Microsoft Word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

As análises de flauta estão no arquivo "**an_flauta.doc**", e as de violino, neste 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.

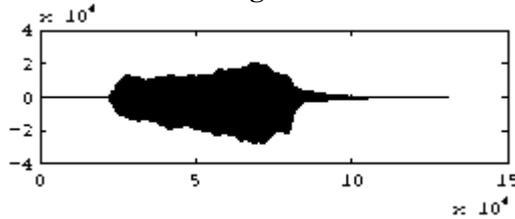
Nas conclusões finais elabora-se uma *tabela de distribuição de gestos/eventos musicais por níveis wavelets* que relaciona os objetos sônicos, eventos e gestos musicais aos níveis onde se distribuem e mais se concentram. Apresentam-se as estruturas sônicas que aparecem destacadas em cada nível, a percepção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

Violino: Ataque Normal - nota mi4 (E4) na corda mi (solta)



Evento: ataque normal, em arco *detachè*, à nota mi4 (E4) na corda mi (solta).

Gráfico do sinal original:



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

Análise dos níveis:

O nível mais fino (16) encerra a maior parte do ruído de arco característico.

O ataque normal à corda fica bem caracterizado nos níveis 15 e 10 principalmente. Ambos apresentam subida com alta taxa de variação, o instante do engate inicial do arco na corda bem localizado no nível 10.

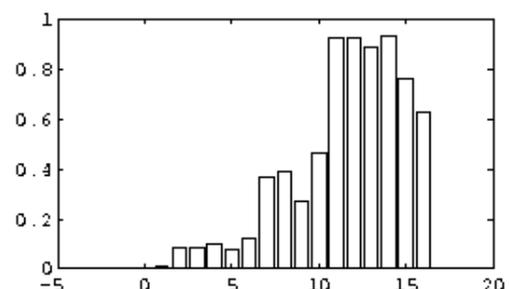
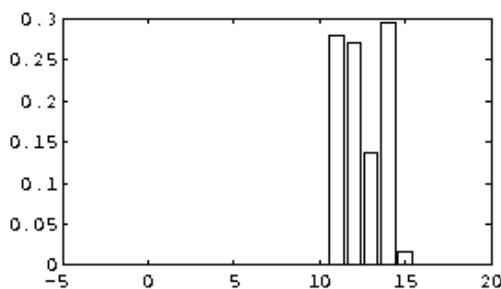
O nível 15 concentra uma expressiva contribuição da ressonância do tom durante o colapso, exibindo um envelope para esta região com contorno suave e amplitudes decrescendo gradualmente.

Os níveis de 11 a 14 concentram a maior parte da contribuição harmônica do tom, apresentam som "mais limpo", pouco ruído e envelopes mais regulares (suaves). O nível 11 mostra-se particularmente sensível ao final do tom, apresentando amplitudes crescentes e atinge um máximo na vizinhança (anterior) do início do colapso.

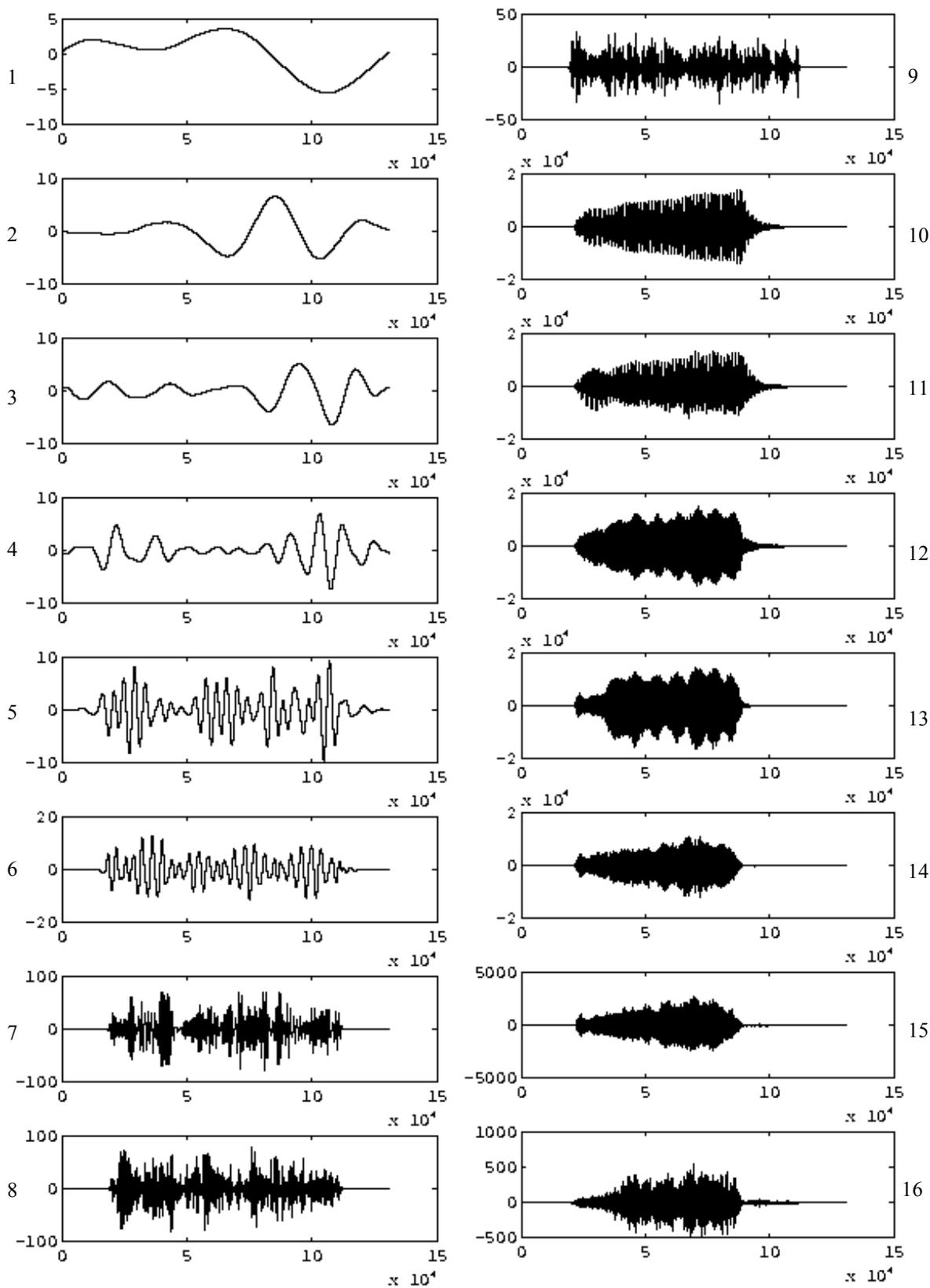
O nível 9 e inferiores não guardam a forma do envelope original, localizam alguns gestos e eventos transitórios, e não contribuem significativamente em energia.

Os níveis 4, 3 e 2 nos dizem sobre as regiões onde o sinal (nota musical) é mais intensa e sonora.

Distribuição de energia:

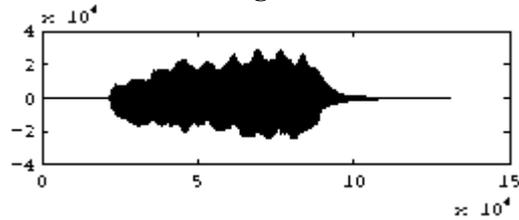


Mais de 99% de toda a energia deste ataque normal à nota mi na corda mi (solta) concentra-se nos níveis de 11 a 15, sendo o nível 14 o mais energético. Os níveis 11 e 12 exibem uma contribuição energética que soma-se a mais de 55% da energia total, e juntamente com o nível 14 concentram mais de 80% da energia total da nota. O gráfico à direita (acima) mostra o logaritmo da energia de cada nível pelo logaritmo da energia total. O eixo vertical, portanto, comprime maiores variações de energia no topo, enquanto perfaz o contrário na base.

Violino: Ataque Normal - nota mi3 (E3) na corda sol (análise com D16)

Evento: ataque normal, nota única, em arco *detachè*, à nota mi3 (E3) na corda sol (a mais baixa em frequência do violino) adornado com um leve vibrato na região de sustentação da nota.

Gráfico do sinal original:



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

Análise dos níveis:

O nível mais fino (16) mostra-se sensível à dinâmica do exemplo, as maiores amplitudes sobre a região de sustentação. O som é ruidoso, e concentra a maior parte do ruído característico do arraste do arco.

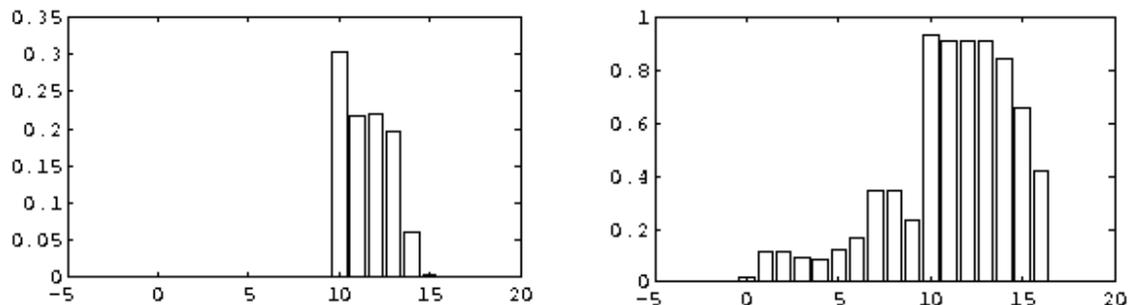
Uma melhor caracterização da região do ataque, decaimento e início da sustentação é observável nos níveis 15, 14, 13, 9 e 8.

Os níveis 10, 11, e 12 são os mais energéticos, apresentam contornos suaves, sons "limpos" (ausência de ruídos) e concentram a maior parte da contribuição harmônica do tom.

Flutuações de amplitude são visíveis nos níveis 12 e 13.

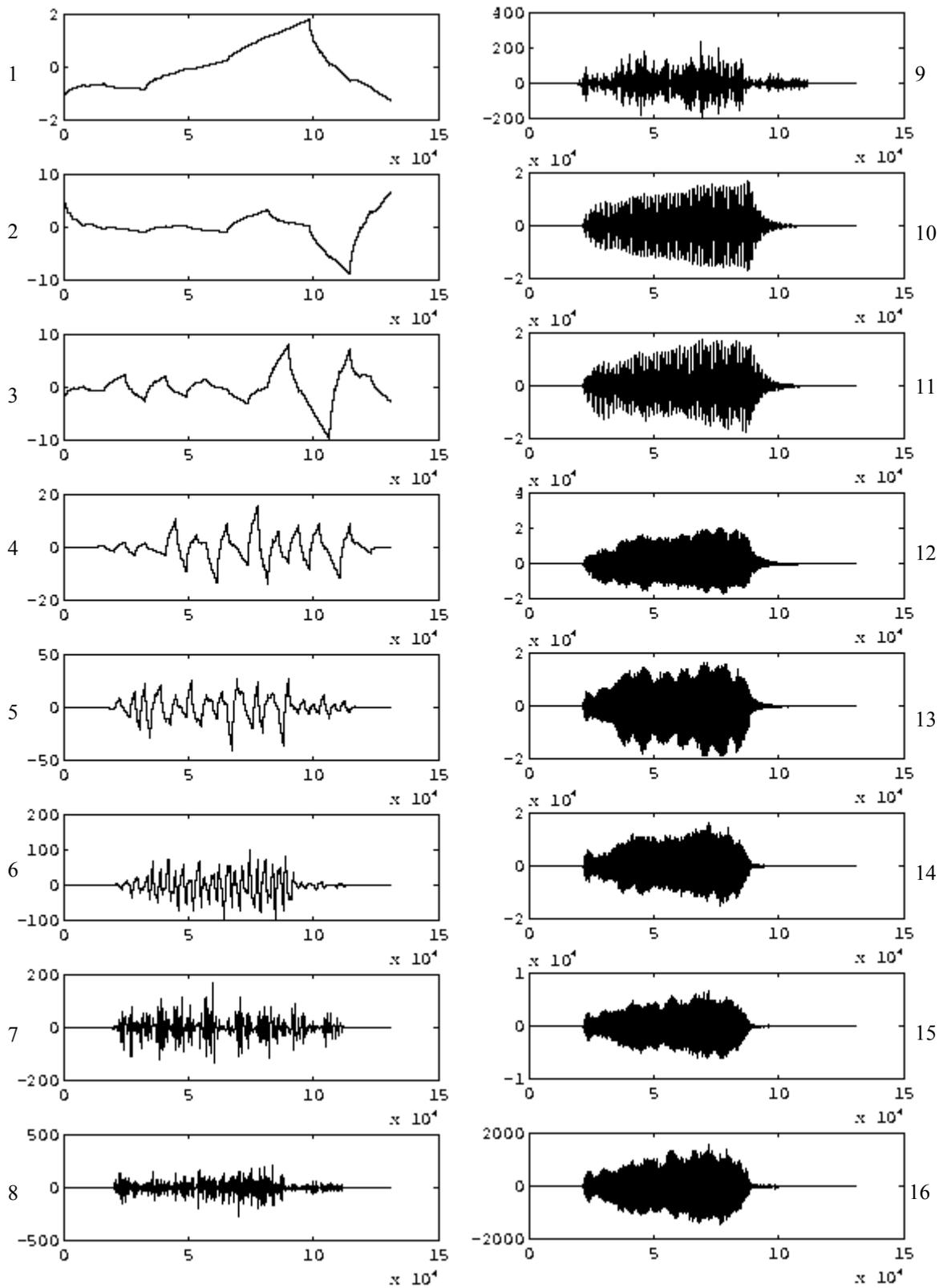
O nível 9 e inferiores não guardam o formato do envelope original, apresentam formas de onda irregulares com vários pontos de máximo (e mínimos) locais, apesar de exibirem regiões aproximadamente periódicas.

Distribuição de energia:



Mais de 99% de toda a energia se concentra nos níveis de 10 a 15, sendo o nível 10 o mais energético. O gráfico à direita (acima) mostra o logaritmo da energia de cada nível pelo logaritmo da energia total do exemplo. Vê-se que os níveis inferiores ao 10^0 exibem contribuição pouco significativa ao conteúdo energético do exemplo. Reconstruções do sinal original desprezando-se os coeficientes destes níveis mais baixos não implicam em significativa perda de qualidade ou informação relevante no sinal.

Violino: Ataque Normal - nota mi³ (E3) na corda sol (análise com D4)



Evento: ataque normal, nota única, em arco *detachè*, à nota mi3 (E3) na corda sol (a mais baixa do violino) adornado com um leve vibrato na região de sustentação da nota. Mesma nota da análise anterior, agora analisada com uma *Wavelet* menos regular.

Gráfico do sinal original: (veja análise anterior)

Extensão da amostra: 2^{17} pontos (131072). **Wavelet de análise:** D4

Análise dos níveis:

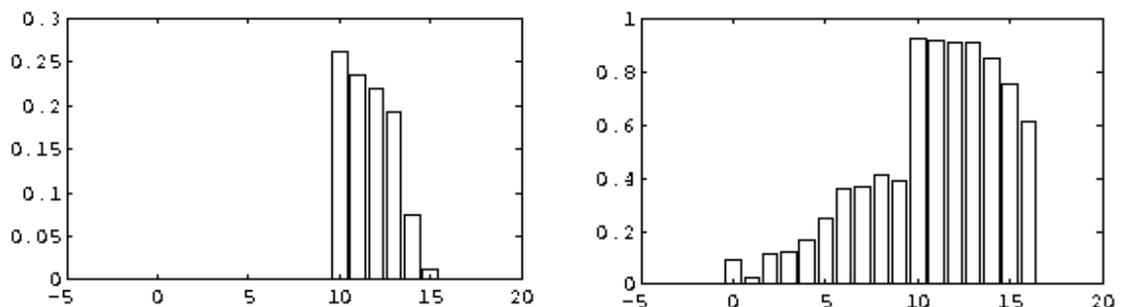
O nível mais fino (16) apresenta um ataque mais bem caracterizado nesta análise com D4 em relação à análise anterior, com D16. Há maior contribuição harmônica no nível, além do característico ruído associado ao arco, e o ataque fica bem definido até o nível 13 (quarto nível mais fino) inclusive.

A região do colapso fica melhor caracterizada nos níveis de 12 a 10, exatamente os níveis mais energéticos.

Os níveis de 9 a 16 são os mais relevantes em termos de concentração de quase a totalidade da informação sobre o evento musical e os gestos embutidos.

A partir do nível 5 para baixo o sinal começa a perder o poder de resolução e localização dos eventos existentes. O nível 4 e inferiores exibem uma representação do sinal em baixa resolução, e são inaudíveis.

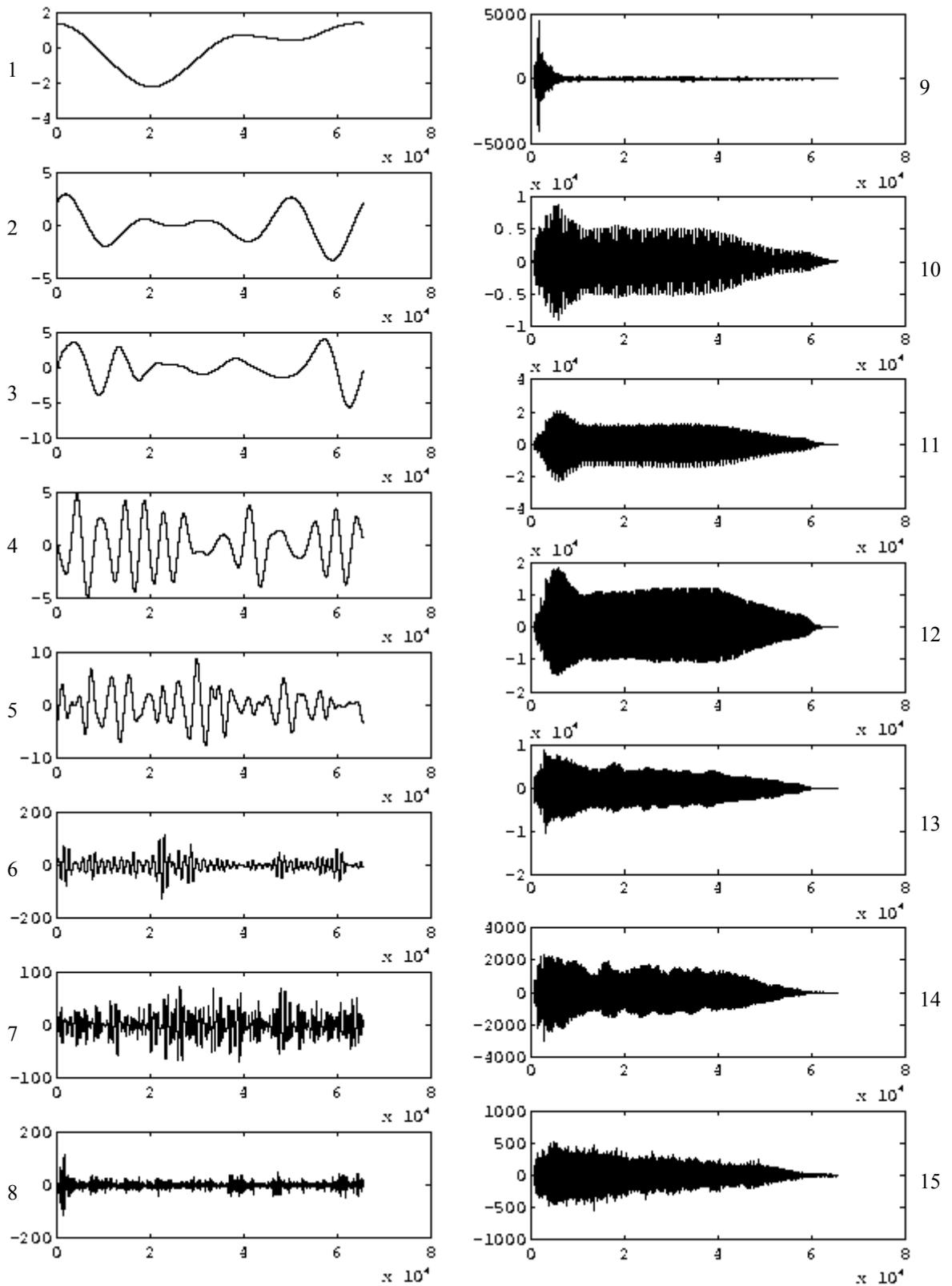
Distribuição de energia:



Mais de 99% de toda a energia se concentra nos níveis de 10 a 15, sendo o nível 10 o mais energético. O gráfico à direita (acima) mostra o logaritmo da energia de cada nível pelo logaritmo da energia total do exemplo. Vê-se que os níveis inferiores ao 10^o exibem contribuição pouco significativa ao conteúdo energético. Reconstruções do sinal original desprezando-se os coeficientes destes níveis mais baixos não implicam em significativa perda de qualidade ou informação no sinal.

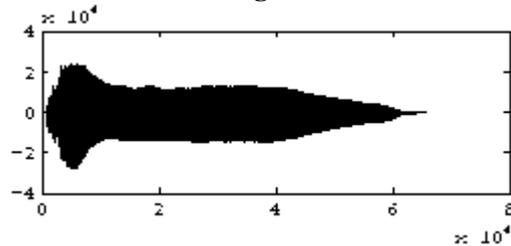
A avaliação do padrão exibido de distribuição de energia para esta amostra (ataque normal) é praticamente a mesma tanto para a *Wavelet* de análise D16 (análise anterior) quanto para a *Wavelet* de análise D4 (presente).

Violino: Ataque Expressivo com Acento (*Sforzato*) - nota mi4 (E4), corda lá



Evento: ataque acentuado (*sforzato*), nota única, ao primeiro mi (E4) na corda lá.

Gráfico do sinal original:



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

Análise dos níveis:

No ataque do nível mais fino (15) é maior a participação de parciais mais altos, efeito que se dilui da região de sustentação em diante. A maior sobrelevação neste nível ocorre entre 100 e 120 ms, e não no início, logo após a subida inicial. A textura sônica denota a presença do ruído do arco e a composição harmônica do tom, principalmente na região inicial onde soam mais intensamente os parciais mais altos. O formato do ataque, sobrelevação e decaimento no início não estão bem caracterizados, havendo em verdade uma sequência destes padrões na fase inicial da nota. Essa complexidade da forma de onda inicial reflete a complexidade das vibrações iniciais que podem ser estabelecidas pelo engate e arraste do arco na corda do violino.

Os níveis de 13 a 10 são os mais importantes para a formação desta nota acentuada. Pela audição (principalmente a 11025 Hz) e observação gráfica destes níveis, verifica-se que há um **período inicial de engate**, onde o arco "ataca" a corda e não são perceptíveis *pitchs* estáveis (padrões vibratórios com altura definida) mas sim padrões irregulares e ruidosos, e um período consecutivo onde *pitchs* definidos são estabelecidos, havendo a percepção de tonalidades.

No nível 13, ouvido a 11025 Hz, percebe-se a grosso modo 3 fases compondo a região inicial do tom: uma **primeira**, que cobre o momento do engate, uma **segunda** porção onde parciais mais altos são emitidos juntamente com 2 *pitchs* centrais, e uma **terceira** porção, onde parte do brilho sônico anterior se reduz e melhora-se a percepção de 2 *pitchs* centrais. Destes dois *pitchs*, o mais baixo fica levemente mascarado durante a segunda fase pelas frequências mais altas.

O nível 12 exibe um formato de envelope onde pode-se facilmente observar uma região de ataque, uma sobrelevação, um leve decaimento e as regiões de sustentação e colapso. Contudo, estas regiões, normalmente associadas aos processos transitórios que se desenvolvem antes do período de sustentação, ocorrem na verdade antes desta maior sobrelevação visível. Isto porque o nível 12 é um dos mais energéticos, exibindo a evolução dos padrões harmônicos e preterindo os componentes transitórios. Sendo natural que o surgimento e evolução dos padrões harmônicos se dê realmente após os transitórios iniciais da nota, a forma de onda observada ao nível 12 não dá informações sobre a "turbulenta" fase inicial, e não exibe os padrões característicos da região de ataque para o *sforzato*.

Uma visão mais detalhada dos primeiros 250ms da nota ao nível 12 revela uma fase inicial, com forma de onda aproximadamente irregular, não exibindo padrões de ataque, sobrelevações ou decaimentos, mas sim exibindo um sutil padrão periódico que cresce de amplitude em cada período, e ganha relevância e regularidade a partir de 188 ms do início.

O envelope visto no nível 12 é portanto uma versão escalada do envelope ideal. Esta interpretação em particular do *sforzato* contribuiu muito para moldar a forma de onda

observada: a intérprete caracterizou a qualidade de acento com um incremento na pressão e/ou velocidade do arraste de arco após o engate inicial do arco à corda, produzindo assim a sobrelevação máxima também após a região inicial, onde já soam modos estáveis, e que ocorre no momento em que a energia do arco encontra menos resistência à sua transferência à caixa ressonante.

A influência do intérprete sobre a caracterização dos acentos é portanto inegável, e as regiões de máxima intensidade de *sforzatos* podem ocorrer bem após a região de ataque, no início da nota. Isto mostra que a síntese de notas acentuadas nos violinos deve abordar uma fase de engate do arco e independentemente uma fase de incremento no arraste para caracterizar o *sforzato*, esta podendo ocorrer dissociadamente do ataque inicial.

O engate do arco à corda é bem perceptível pela audição do nível 10, onde dois *pitchs* centrais são perceptíveis. Também nos níveis 11 e 12 são ouvidos dois *pitchs* soando.

Os níveis 9 e 8 (respectivamente o 7º e o 8º níveis mais finos) localizam e caracterizam bem o ataque acentuado *sforzato*, especialmente o engate inicial do arco, apesar de serem muito menos energéticos. Parte da qualidade sonora do *sforzato* está portanto centrada em níveis pouco energéticos. O 7º nível mais fino nas análises de violino (9 nesta análise) usualmente revela os processos acústicos que ocorrem no corpo de madeira do instrumento.

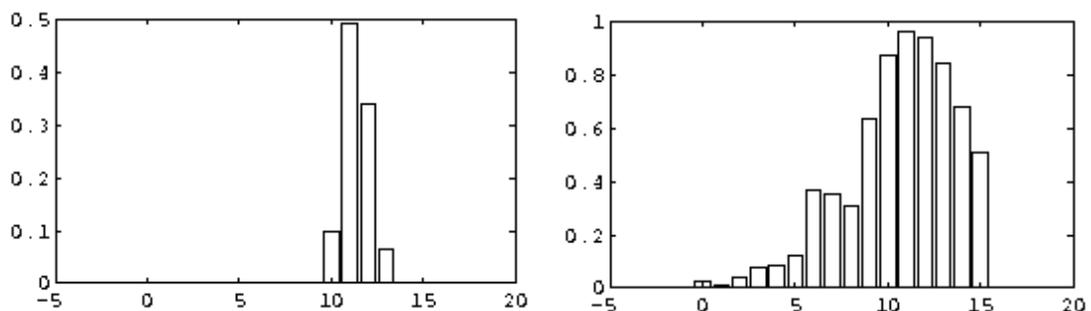
Violinistas são capazes de segregar as diversas sonoridades que emergem de pontos diferentes do instrumento, e identificar a parte do violino que está gerando aquele som, seja ele transiente ou estacionário. Os sons associados à caixa ressonante do violino, seu tampo e fundo, especialmente os transientes e pulsos, são enfatizados neste sétimo nível mais fino da análise *Wavelet*, podendo-se informalmente referenciá-lo pelo "**nível da madeira**", pela sua textura sônica, como se deixando passar somente a contribuição da caixa de madeira. Isto está de acordo com a percepção deste nível escutado por um violinista.

O nível 8 exhibe basicamente uma textura sônica de ruído de fundo.

O nível 7 da análise revela um ruído de fundo disperso por todo o exemplo, descorrelacionado com o envelope da nota tocada.

O nível 4 exhibe suas maiores amplitudes no início, exatamente sobre a região de sobrelevação do *sforzato*, e é inaudível.

Distribuição de energia:



O primeiro gráfico apresenta no eixo vertical a intensidade em escala linear. Mais de 99% de toda a energia da amostra analisada se concentra nos níveis de 10 a 13, sendo o nível 11 o mais energético.

No segundo gráfico o eixo vertical contém 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. Reconstruções do sinal original desprezando-se os coeficientes dos níveis 9 e inferiores (de 8 a 0) não implicam em perda de qualidade significativa no som final resintetizado (em comparação ao original).

Outras reconstruções relevantes foram realizadas para esse evento, inclusive algumas tomando-se alguns níveis (na verdade os coeficientes) da análise da mesma nota sob ataque normal, de forma a avaliar a importância sônica de alguns níveis para caracterizar o *sforzato*.

Uma delas foi feita zerando-se os níveis de 0 a 8 e também o nível mais fino (15); O sinal reconstruído foi praticamente idêntico à nota original, com diferença abaixo do limiar de percepção. Outra reconstrução zerando-se os níveis de 0 a 9 (inclusive) e também o nível mais fino (15) mostra que o nível 9 não é tão importante para a efetivação da percepção do *sforzato*.

O experimento mais significativo para se avaliar a concentração da qualidade do *sforzato* em alguns níveis utilizou a rotina CRUZAR para realizar reconstruções com níveis tomados de arquivos sonoros diferentes, implementando uma forma de **síntese cruzada** com *Wavelets*.

Dois arquivos sonoros, exemplificando eventos diferentes, foram empregados: **(A)** a nota mi na corda lá da presente análise, e **(B)** a nota mi também tocada na corda lá, sob interpretação normal¹. Os arquivos sonoros foram editados e devidamente acolchoados com zeros ao início e fim dos vetores, tais que o início de ambas as notas coincidisse no tempo, seus ataques comesçassem simultaneamente, e a duração dos vetores fosse igual em ambos. Em seguida, ambos os sinais foram decompostos com o sistema de AWMR, e a rotina CRUZAR foi empregada para se construir **sinais mistos**, contendo contribuições de ambos os eventos (ataque *sforzato*, A, e ataque normal, B).

Analisando-se comparativamente os níveis de A e B, nota-se que no nível 11 ambas apresentam mesmo *pitch*, mas exibem diferenças na região de ataque de A e na elevação ao final de B. O nível 10 difere bastante nos dois eventos: no ataque normal ele apresenta um envelope mais largo também no final, e varia de intensidade por toda a duração; no ataque com *sforzato* há um pico bem destacado no início. O nível 12 em ambos apresenta envelopes regulares e apresenta um *pitch* semelhante; o sinal com *sforzato* apresenta, contudo, um segundo *pitch* em destaque, associado à "mordida" do arco à corda.

Os sinais construídos em síntese cruzada foram em seguida ouvidos, interpretados e comparados com ambas as amostras originais de ataque normal e ataque acentuado com *sforzato*. A tabela abaixo mostra as mais importantes sínteses cruzadas realizadas e os arquivos onde se encontram armazenados seus sinais de áudio (disponíveis no CD ROM em anexo):

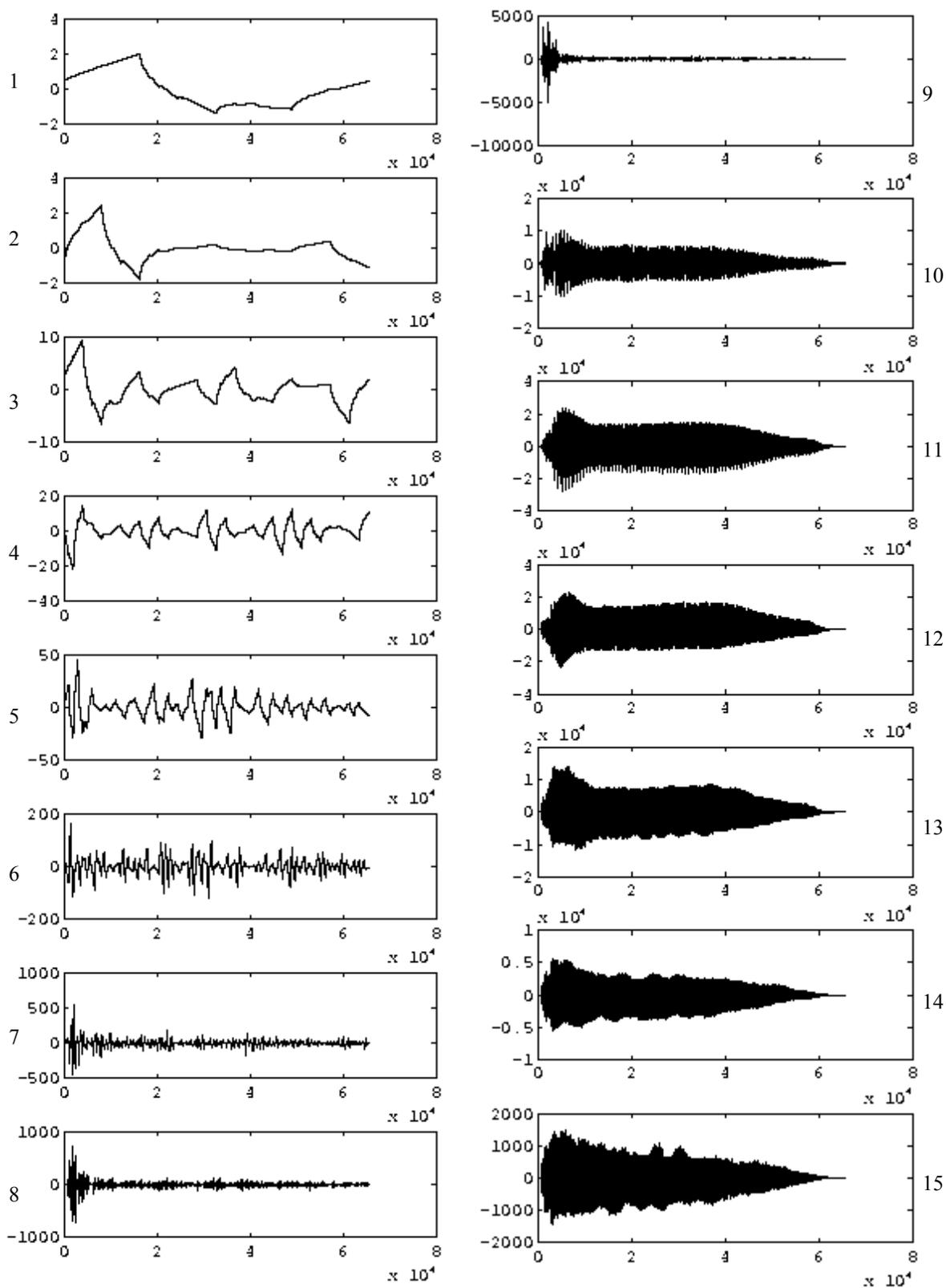
¹ Amostra não selecionada para análise neste trabalho. No entanto, é aqui utilizada por haver sido tocada também na corda lá, como a amostra de ataque com *sforzato* em análise o foi.

Arquivo	Níveis tomados de A (<i>sforzato</i>)	Níveis tomados de B (<i>normal</i>)
norm-sf1	10, com peso amplificado em 10 vezes (peso 10)	todos, exceto o 10
norm-sf2	10, 11 e 12, com peso 1	todos, exceto 10, 11, 12.
norm-sf3	11, 12, 13, com peso 1	todos, exceto 11, 12, 13.
norm-sf4	11, 12, com peso 1	todos, exceto 11 e 12
norm-sf5	10, com peso 3	todos, exceto o 10
norm-sf6	11, 12, com peso 1	13, com peso 0.225, 14 com peso 0.543, 15 com peso 0.552, os outros com peso 1
norm-sf7	8 e 9, com peso 3; 10, com peso 2,5	todos, exceto 8, 9 e 10

Somente com os níveis 11 e 12 da amostra A já é possível notar que uma qualidade de *sforzato* é adicionada à nota, mas existe uma degradação no som, evidenciada pela percepção de alguns *pitchs* (tons ou vozes adicionais) que não se integram (fundem num timbre ou *stream* único). A causa principal disso está no balanço deficiente dos pesos na reconstrução, o que é usualmente um motivo para se "dividir" uma nota em uma espécie de acorde onde outras vozes soam simultaneamente.

O *sforzato* ainda é perceptível se somente o nível 10 da amostra A for substituído com um peso 3 (3 vezes maior), como no arquivo **norm-sf5** acima na tabela. O timbre geral ainda degrada-se um pouco, distorcendo levemente a percepção de um violino, mas a qualidade do *sforzato*, em especial a região do ataque, é adicionada, mostrando a importância deste nível na formação da percepção do acento *sforzato*.

Violino: Ataque Expressivo com Acento (*Sforzato*) - nota mi4 (E4), corda lá (D4)



Evento: Ataque *sforzato*, nota única, ao primeiro mi (E4) na corda lá. Mesmo amostra sonora da análise anterior, desta vez analisada com a *Wavelet* D4.

Gráfico do sinal original: (veja análise anterior)

Extensão da amostra: 2^{16} pontos (65536). **Wavelet de análise:** D4

Análise dos níveis:

Os gestos observados e características gerais verificadas em cada nível são semelhantes às da análise anterior, da mesma nota acentuada com D16, especialmente para os níveis mais energéticos.

O nível mais fino encampa frequências mais baixas não cobertas pelo nível 15 da análise com D16, o que é perceptível pela audição comparativa dos mesmos. Seu som apresenta uma textura menos aguda que o som do mesmo nível para D16.

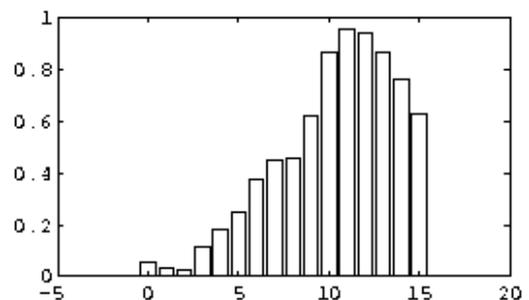
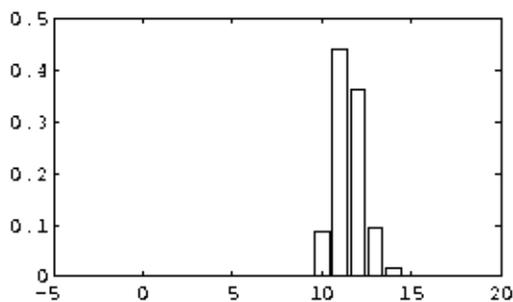
No nível 11 fica mais destacado o ataque de *sforzato*, sua audição apresentando a percepção de um tom ao início e estabilizando em outro: o ataque no sinal ao nível 11 apresenta 2 porções, cada qual com um *pitch* central dominante (mais pronunciado). O segundo tom permanece constante durante o exemplo, só reduzindo em intensidade no final.

O nível 10 caracteriza mais ainda o ataque *sforzato*. Comparando este nível com o nível correspondente da amostra da nota mi tocada na corda mi solta (da primeira análise de ataque normal, atrás) nota-se claramente a diferença provocada pela presença do *sforzato* neste nível. A intensidade no ataque acentuado (o presente) é mais forte, e além disso para aquela amostra foi utilizada a *Wavelet* D16.

O nível 7 com D4 permite uma localização do acento *sforzato*, guardando semelhança como nível imediatamente mais fino (8).

O nível 5 e inferiores também exibem formas de onda que já se assemelham à *Wavelet* D4, menos regular.

Distribuição de energia:



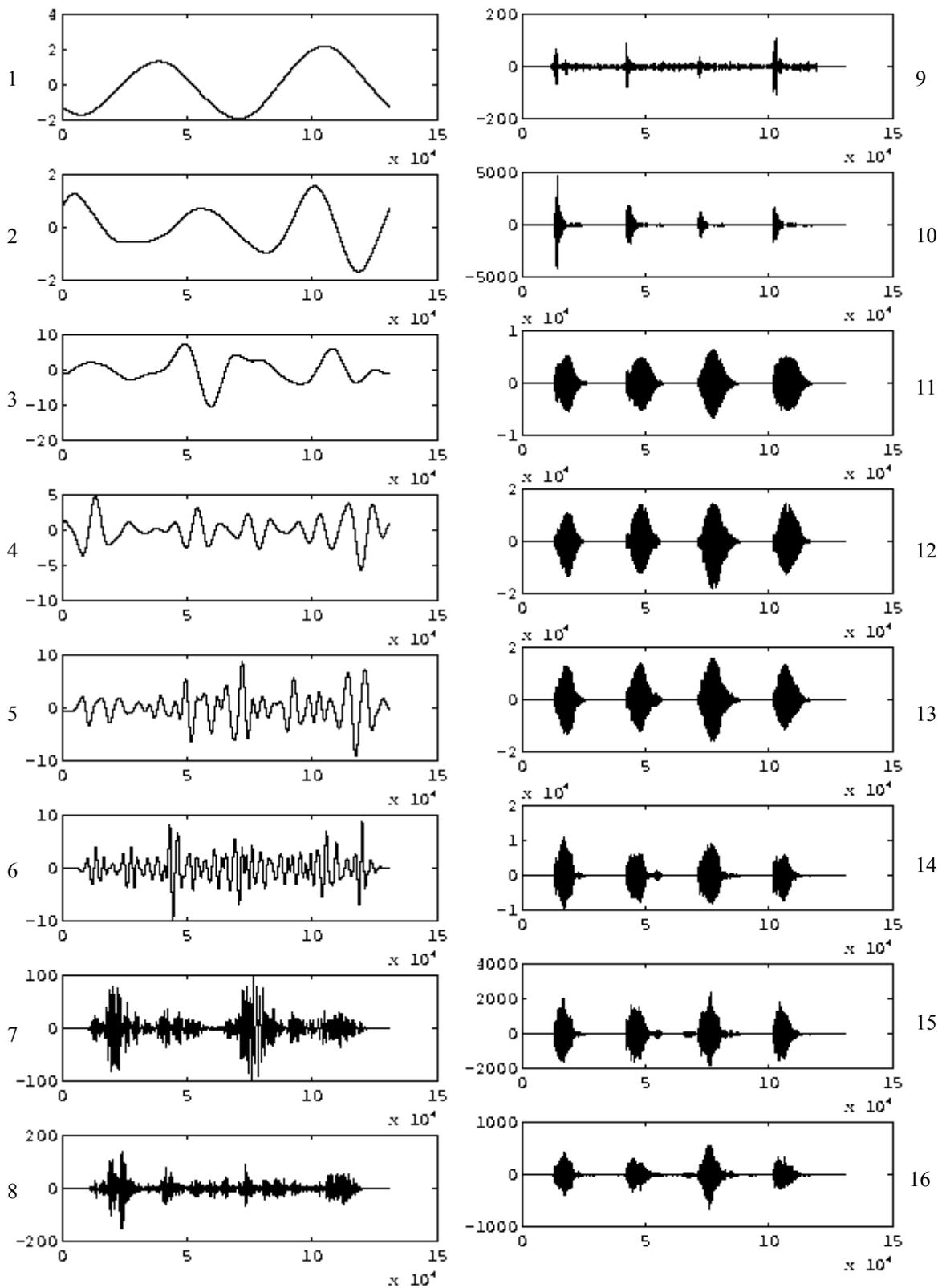
Mais de 99% de toda a energia do exemplo concentra-se nos níveis de 10 a 14, sendo o 11º o mais energético. Comparativamente à análise com D16, a energia se distribui mais uniformemente pelos níveis, numa gradação mais suave entre a energia de um nível e do nível adjacente (degrau menos alto), como pode-se verificar comparando a distribuição (em escala logarítmica) do segundo gráfico acima com seu correspondente na análise anterior (mesmo sinal analisado com D16).

Como na análise anterior, algumas reconstruções seletivas também foram realizadas com os coeficientes obtidos com D4, as quais relata-se abaixo:

1. reconstrução com níveis 9 e 10 com peso 0,5 e os níveis 11,12, 13 e 14 com peso 0,586: não elimina a qualidade percebida do acento *sforzato*.
2. com níveis 10 a 15 intactos, o restante zerado: leva a uma reconstrução praticamente idêntica à original, diferenças abaixo do limiar perceptivo.
3. níveis 10 a 14 intactos, o restante em zero: leva a uma reconstrução quase idêntica, havendo certa granulosidade na textura do som.
4. níveis 12, 13 e 14 intactos, o restante em zero.
5. níveis 13 e 14 intactos, o restante em zero: o tom fica muito degradado, falta muito do espectro harmônico, e o *sforzato* fica bastante atenuado, ainda que não eliminado.

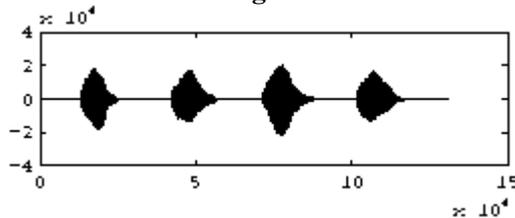
Várias reconstruções seletivas foram feitas eliminando-se os níveis 9 e 10 e deixando-se diferentes pesos para os níveis de 11 a 15, mas nenhuma delas eliminou o *sforzato*. Ele aparece em todos estes níveis, com características e composição diferentes.

Violino: Ataques Expressivos com *Staccatos* - nota mi4 (E4) na corda lá



Evento: Quatro notas executadas com ataques *staccato* (*martelè*) à nota mi (E4) na corda lá. Nesta interpretação em particular o *staccato* está acrescido de uma qualidade acentuada, do tipo *martelè*. A baixa velocidade na execução das notas potencializa esse efeito.

Gráfico do sinal original:



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

Nível 16: O nível mais fino (16) um objeto sonoro prévio ao ataque da nota aparece destacado. O ruído de arco é bastante significativo, especialmente sobre a terceira nota.

Nível 15: Rastreia todos os movimentos espúrios (ruidosos) indesejáveis na interpretação, como portamentos, rangidos e movimentos bruscos e não regulares do arco. Reproduzindo o som do nível com frequência de reprodução de 11025 Hz (4 vezes mais lento que o padrão CD) é possível ouvir estes movimentos e ruídos.

Nível 14: Revela, bem conformados, os envelopes das regiões de colapso das notas, e o efeito do arco sobre elas.

Nível 13: Não fornece detalhes expressivos. Evidencia o formato padrão do envelope das notas, e sonicamente consiste de um *pitch* predominante mais alto e um secundário, mais baixo, que se destaca no início dos ataques.

Nível 12: A terceira nota aparece mais intensa, sem o objeto sônico que lhe precedia bem destacado ao nível 16. É o nível mais energético, e revela predominantemente a composição harmônica (vibrações harmônicas) típicas do instrumento, não sendo muito sensível a artefatos e objetos sônicos transientes oriundos dos gestos de arco.

Nível 10: Revela com destaque o som dos ataques. O envelope caracteriza bem o ataque *staccato*. O ataque da primeira nota é mais intenso que o das demais neste nível.

Nível 9: Evidencia o ruído ambiente; preenche porções entre as notas (silêncios ou pausas); é sensível à rampa de subida inicial dos ataques (*onsets*); as intensidades maiores ocorrem sobre as regiões onde há notas, especialmente nos ataques.

Nível 8: Identifica com evidência as porções centrais das notas (sustentações).

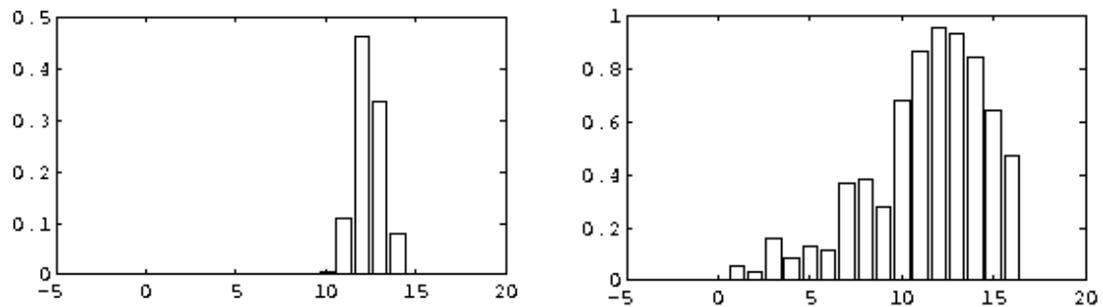
Nível 7: Como o nível anterior, identifica melhor as porções centrais. Quase inaudível. Este nível também fornece informações sobre o sentido do arco empregado na execução de cada nota. Note pelo gráfico que a segunda e a quarta (última) notas apresentam menores intensidades do que as primeira e terceira notas. Na execução de *staccatos*, o intérprete normalmente muda o sentido do arco para cada nota atacada, alternando arco para cima e arco para baixo. Geralmente o arco para baixo consiste num movimento mais enérgico e produz uma tonalidade mais destacada que pode ser tão diferente daquela obtida com o arco para cima quanto se desejar. O fato de que o ponto e o ângulo de contato arco-corda também é diferente em cada sentido influi na formação e textura sônica do tom. Este nível permite, pois, a identificação do padrão alternante de sentido de arco, característico das interpretações de *staccatos*. Também o nível 8 e 9 fornecem indícios sobre este padrão alternante de sentido de arco, embora menos evidente que neste nível.

Nível 6: Tende a preencher todos os pontos da extensão total do som, incluindo as regiões originalmente preenchidas com zeros (*zero padding*) do arquivo sonoro. A possibilidade de que seu envelope e os pontos de pico forneçam alguma informação relevante deve ser pesquisada em trabalhos futuros. Inaudível, exceto pelos *clicks* dos pulsos transientes).

Nível 5: Ruído de baixa frequência (soando com um chiado). Não identifica porções importantes.

Os níveis inferiores oferecem uma contribuição insignificante à energia total da amostra. A reconstrução do som original desprezando estes níveis não causa impacto diferencial algum.

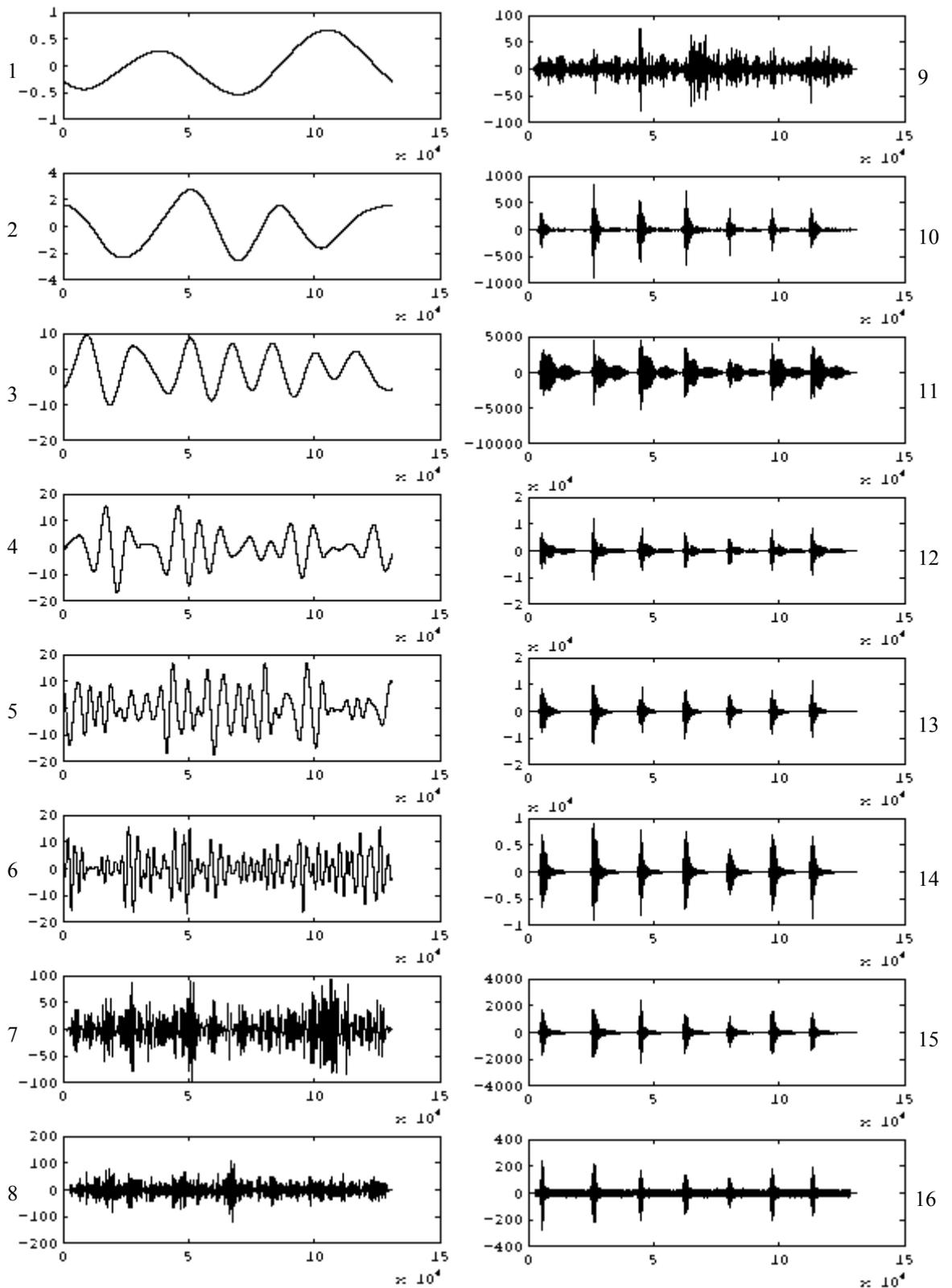
Distribuição de energia:



O primeiro gráfico apresenta no eixo vertical a intensidade em escala linear. Mais de 99% de toda a energia da amostra analisada se concentra nos níveis 11 a 14, sendo o nível 12 o mais energético.

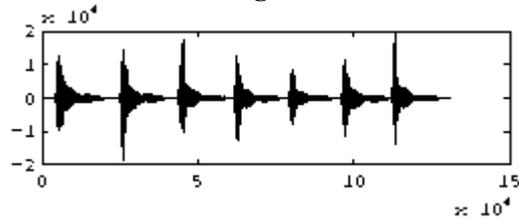
No exemplo anterior (ataques expressivos com *sforzato*) em que uma mesma nota mi4 (E4) é analisada verificou-se que a maioria de sua energia se concentrava nos níveis de 10 a 13, diferentemente do que aqui se verifica para uma mesma nota mi4. Todavia, cabe ressaltar que a ordenação numérica dos níveis difere quando a extensão original do sinal não é a mesma. Assim, o nível mais fino para aquele exemplo era o nível 15 (porque o sinal original possuía 2^{16} pontos) e aqui o nível mais fino é o 16 (porque o sinal original possui 2^{17} pontos). Observado este critério de ordenação vê-se que, na verdade, os níveis de 11 a 14 deste exemplo mapeiam as mesma faixas de frequência que os níveis de 10 a 13 do exemplo anterior e portanto são correspondentes. Assim, quase a totalidade da energia para as notas E4 de ambos os exemplos se concentram nos mesmos níveis *wavelet*: do 3^o ao 6^o níveis mais finos.

Violino: Ataques Expressivos com *Spiccatos* - primeiro mi (E4) na corda lá



Evento: Sequência de 7 ataques expressivos, em *staccato spiccato*, à primeira nota mi (E4) na corda lá.

Gráfico do sinal original:



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

Análise dos níveis:

Os níveis superiores (mais finos, do 13 ao 16) revelam grande sensibilidade às rampas de subida dos períodos do ataque *spiccato*. O formato do envelope típico para este tipo de ataque expressivo fica melhor caracterizado nestes níveis.

No nível 12 observa-se nitidamente a presença de uma "reverberação ou eco" seguindo às fases de sustentação, estas muito curtas. As reverberações constituem os "nódulos" proeminentes após as sustentações, ocorrendo tão logo o arco abandona a corda e a nota colapsa naturalmente.

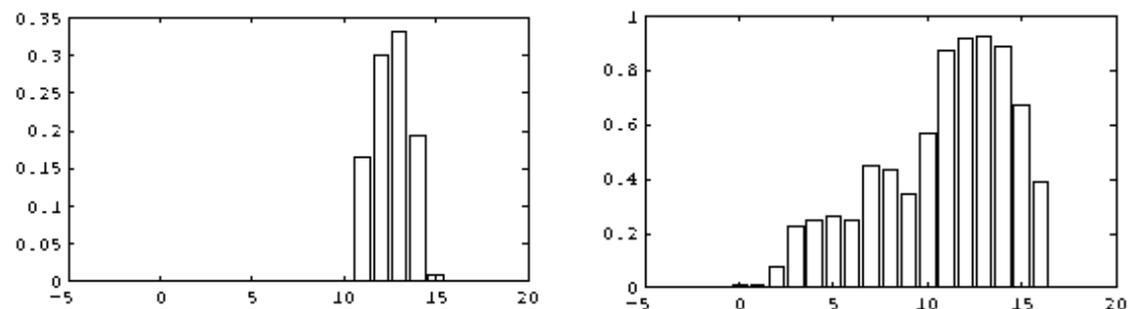
No nível 11 o mesmo efeito reverberante, seguido de um colapso natural dos modos vibrantes, é observado.

A técnica interpretativa de *spiccatos* ensina que o arco deve atacar súbito e sair da corda, deixando-a vibrar. O efeito imediato deste gesto musical está na produção de "nódulos reverberantes" principalmente nos níveis 12 e 11.

O nível 10 é muito sensível ao transiente provocado pela "mordida" do arco à corda em cada ataque, localizando com precisão o momento da sobrelevação em cada nota.

No nível 9 os pulsos (picos) observados não estão associados ao gesto do ataque, mas ao gesto da saída do arco, i.é, ao transiente causado pelo abandono do arco à corda.

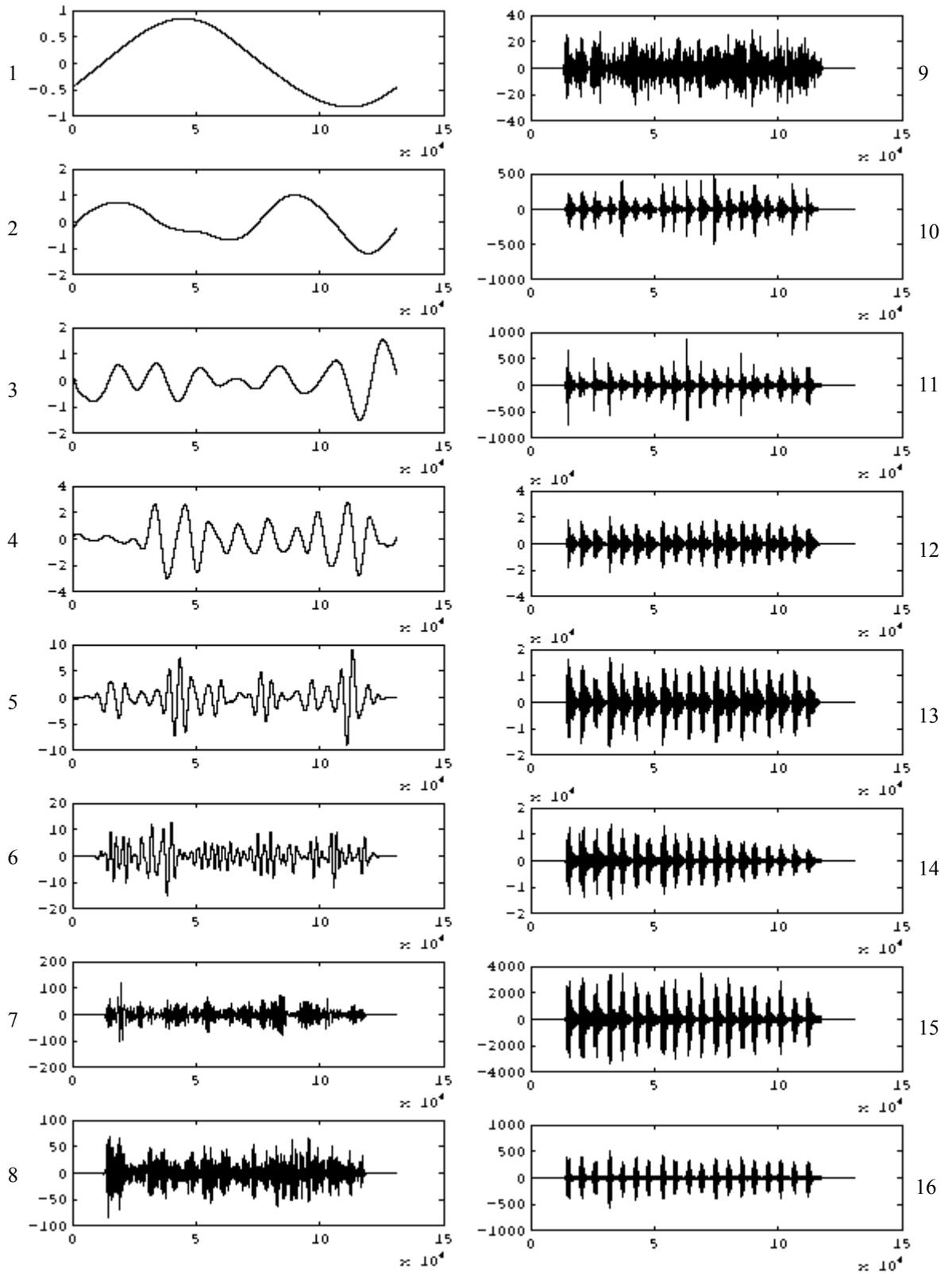
Distribuição de energia:



O primeiro gráfico (acima à esquerda) mostra no eixo vertical a intensidade energética de cada nível, plotada numa escala linear. Note que mais de 99% de toda a energia da amostra se concentra entre os níveis 11 e 15, sendo o nível 13 o mais energético para este exemplo de ataques *spiccatos*. Apesar de o padrão de distribuição nos 4 níveis mais energéticos (de 11 a 14) diferirem dos exemplos anteriores analisando notas mi4 (E4), estes níveis ainda comportam quase toda a totalidade da energia para estas notas, em concordância com conclusões anteriores.

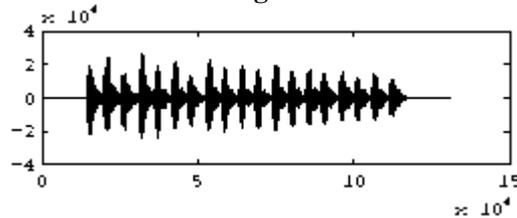
Note no segundo gráfico (à direita) que os níveis inferiores ao 9 contribuem muito pouco no conteúdo energético do evento musical analisado, sua participação sendo desprezível, sua ausência causando um impacto reduzido sobre a qualidade do som reconstruído.

Violino: Ataques Expressivos com *Spiccato*, naturalmente rápido



Evento: sequência de ataques *spiccatos* ao segundo mi (E5) da corda lá, em velocidade naturalmente rápida, como sugere a boa técnica de interpretação para este tipo de gesto expressivo.

Gráfico do sinal original:



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

Análise dos níveis:

O nível *wavelet* mais fino (16) é sensível ao *spiccato*, localizando e possibilitando a avaliação da intensidade e periodicidade das regiões onde há o arraste do arco.

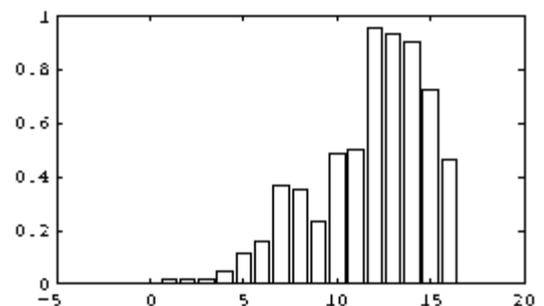
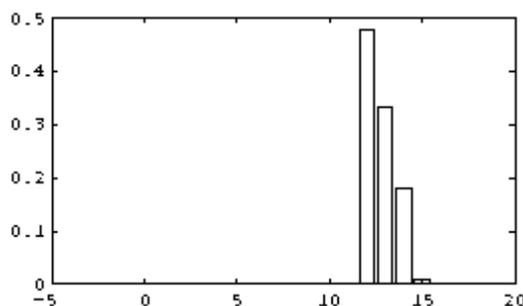
O nível 15 localiza com precisão os pontos de *onset* da nota (ataque inicial), e mostra-se sensível à "reverberação" da caixa ressonante do instrumento, que ocorre em função do abandono do arco à corda e sucessivo colapso natural de cada nota, com as vibrações decaindo de amplitude. O padrão gráfico exibido revela-se característico dos gestos musicais envolvidos na interpretação de *spiccatos*. A ocorrência da "reverberação" neste nível fino se deve à tonalidade mais alta da nota, inscrita na 5ª oitava. Também o nível 14 exibe regiões de colapso natural e ressonância da caixa, como no nível anterior mais fino.

O nível 11 é sensível a gestos de arco (variações na velocidade, pressão, ângulo de ataque) e empunhadura, e localiza bem as sobrelevações após os *onsets* das notas.

Os níveis abaixo do 10^o não oferecem muita informação sobre os gestos musicais.

Claramente se nota que para tons mais altos (oitavas altas, acima da 4ª) os eventos e processos vibratórios se projetam em níveis mais altos/finos, exatamente aqueles de bandas cobertas largas, onde a segregação de estruturas e observação de gestos musicais é menos evidente.

Distribuição de energia:



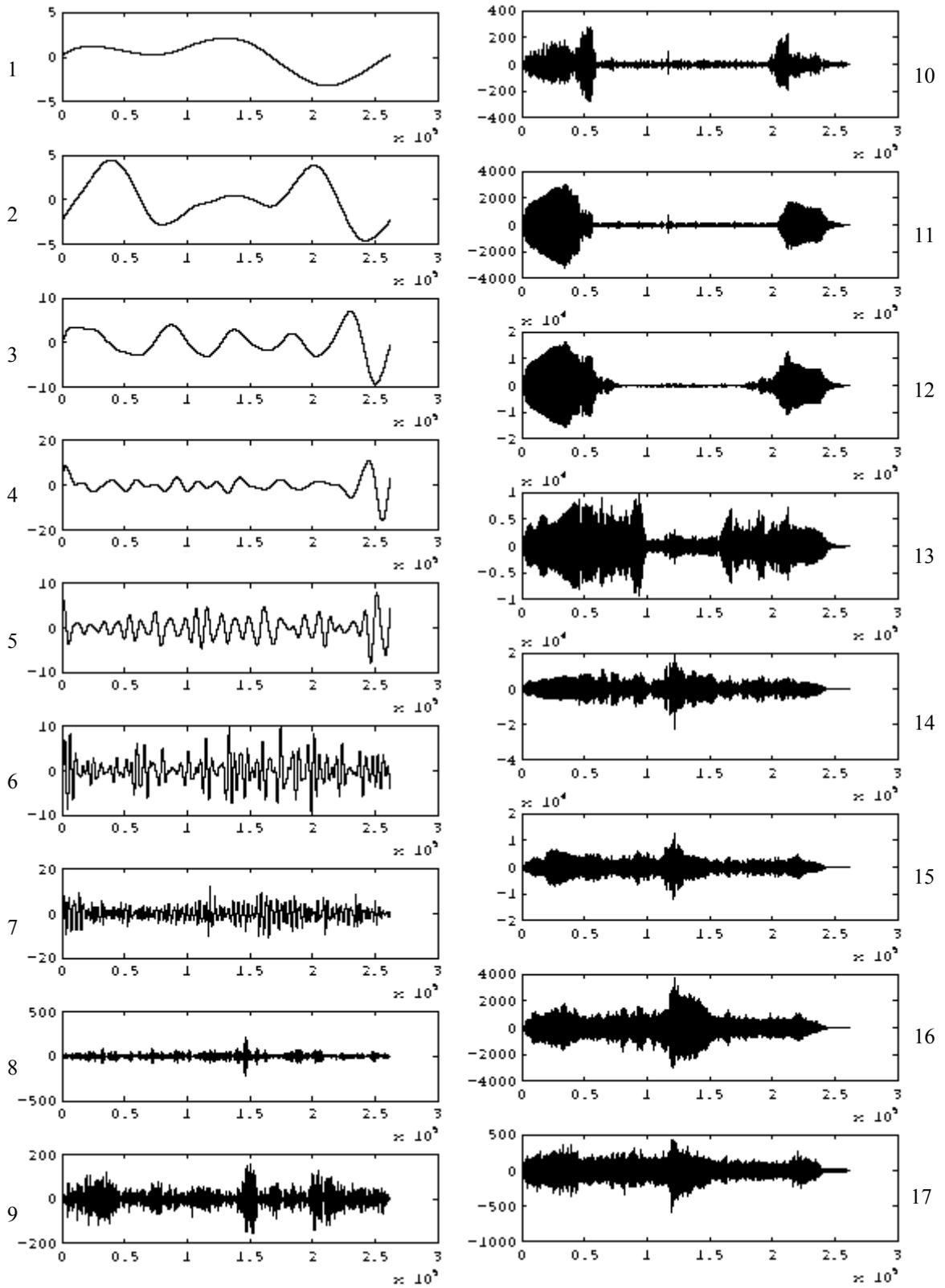
Mais de 99% da energia contida nesta sequência de ataques em *spiccatos* concentra-se nos níveis de 12 a 15, sendo o nível 12 o mais energético para estes mi5 (E5). O padrão de distribuição no segundo gráfico mostra-se irregular, observando-se uma baixa participação energética no nível 9 em comparação aos níveis 7 e 8.

Análises posteriores de trechos sem sinais musicais (somente pausas, na ausência de emissões acústicas por parte dos instrumentos) mostraram que o nível 9 (para sinais de extensão 2^{17}) revelava um ruído de fundo grave (baixa frequência) pronunciado para alguns

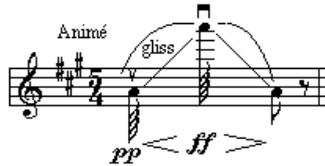
exemplos/gravações e bem atenuado em outros, e não correlacionado com os eventos musicais analisados. As séries harmônicas das notas deste exemplo não alcançam oitavas tão baixas quanto as cobertas pelos níveis 9 e inferiores. Somente componentes transientes e outros componentes de baixa frequência associados a gestos interpretativos poderiam ser visíveis nestes níveis, além de ruídos e componentes frequenciais espúrias que porventura fossem adquiridas durante o processo de gravação e preparo das amostras para análise, para os quais o nível 9 em particular mostrou-se bastante sensível.

O fato do nível 9 aqui revelar uma baixa energia é um indício de que o sinal original apresenta um menor nível de ruído em baixa frequência adquiridos no processo de gravação e preparo das amostras. De fato, a qualidade sonora do sinal original revela-se mais pura, isenta de ruídos e "*hiss*" (chiados).

Violino: Fraseado, Ligado, Normal, em Glissando - Análise com D16

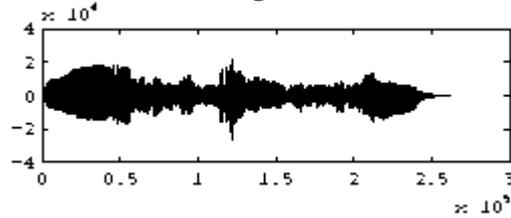


Evento: Interpretação de um fraseado ligado em glissando, extraído da peça *Daphnis et Chloé*, de Ravel. O trecho consiste numa interpretação ligada entre lá3 (A3) e lá5 (A5) transpondo-se duas oitavas em glissando, associado com uma indicação de dinâmica *fortissimo* na nota mais alta (veja partitura ilustrativa ao lado).



The musical score shows a treble clef with a key signature of two sharps (F# and C#) and a 4/4 time signature. The tempo marking is 'Animé'. The first measure contains a half note G4 (labeled 'pp'). The second measure contains a half note A4 (labeled 'ff') with a glissando line above it connecting it to a higher A5. The third measure contains a half note A5 (labeled 'ff').

Gráfico do sinal original:

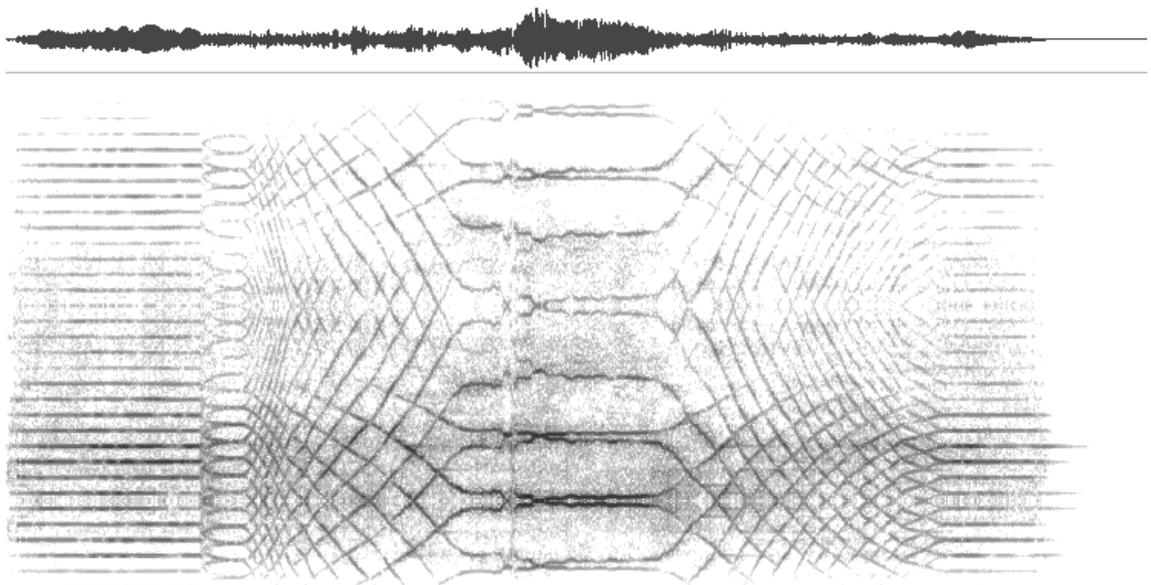


Extensão da amostra: 2^{18} pontos (262144). **Wavelet de análise:** D16

No nível de resolução mais fina (17) há claramente uma intensa "coloração de alta frequência", com uma densa população de frequências além de 10000 Hz. A percepção básica do som do nível é a de um ruído agudo, com pouca informação sobre os gestos musicais mas apresentando traços das séries harmônicas das notas (as tonalidades são perceptíveis). O ruído de fundo associado está bem abaixo do ruído de arco característico, e o nível mostra-se sensível a variações dinâmicas (intensidades é proporcional ao ruído de arco). A ausência deste nível numa reconstrução torna o som um pouco mais "granuloso", de uma textura sônica mais ríspida ou áspera, embora tal efeito seja sutil.

Ao nível 16 chama a atenção um efeito interessante onde algumas raias mais pronunciadas do espectro deslocam-se para cima, na direção das frequências mais altas, e outras para baixo durante a transição do glissando. A figura abaixo mostra um espectrograma do nível 16 ilustrando esse efeito.

Figura 3 - Espectrograma do nível 16 (fraseado, ligado, normal, glissando; Wavelet D16)

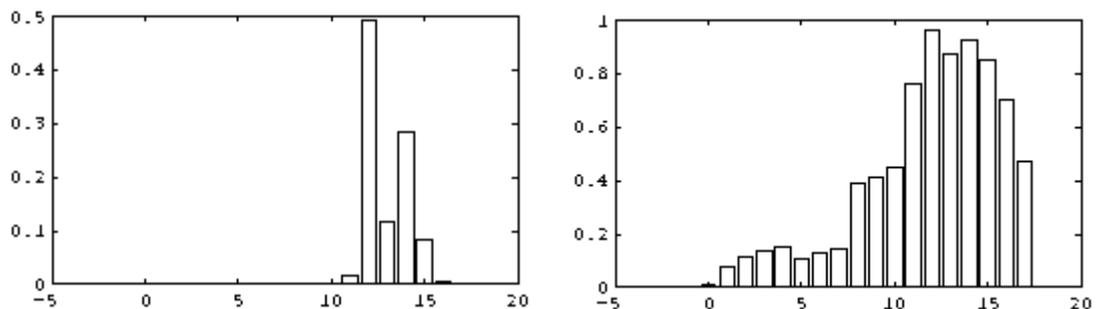


Note que nas regiões de sustentação das notas mais baixas em frequência (a primeira e a última) a separação entre raias é bem menor que para a nota central, onde as raias estão mais bem separadas. Verifica-se também claramente o deslocamento de algumas raias para cima e outras para baixo durante a transição dos glissandos. As menores amplitudes são localizadas nas regiões de transição. O espectrograma como ferramenta auxiliar na análise de um nível mostra-se bastante útil para localizar os instantes iniciais do glissando, bem como detalhes da evolução espectral das tonalidades.

Dos níveis 10 ao 13 (inclusive) nota-se nitidamente uma menor contribuição da nota central (lá5, A5), evidenciado pelas baixas amplitudes nas regiões centrais destes gráficos (veja gráficos da análise). Isso é devido aos esquemas de como as notas/tons se distribuem pelos níveis da análise, diferentes para as notas lá3 e lá5. A nota mais alta (lá5) possui uma série harmônica deslocada para as faixas de alta frequência, portanto distribuindo-se e concentrando a maior parte de sua energia nos níveis mais altos da análise. O nível 14 é o nível mais energético para a tonalidade lá5 (veja no gráfico de distribuição de energia em escala linear, logo abaixo). As tonalidades mais baixas (lá3, A3) apresentam um padrão de distribuição e concentração de energia pelos níveis da análise semelhante à tonalidade mais alta, mas deslocado dois níveis para baixo, uma vez que se situam a duas oitavas abaixo da nota mais alta, e cada nível da análise cobre uma oitava. A primeira e última notas portanto se distribuem e concentram maior parte de sua energia cerca de dois níveis abaixo, sendo particularmente o nível 12 o mais energético para esta tonalidade.

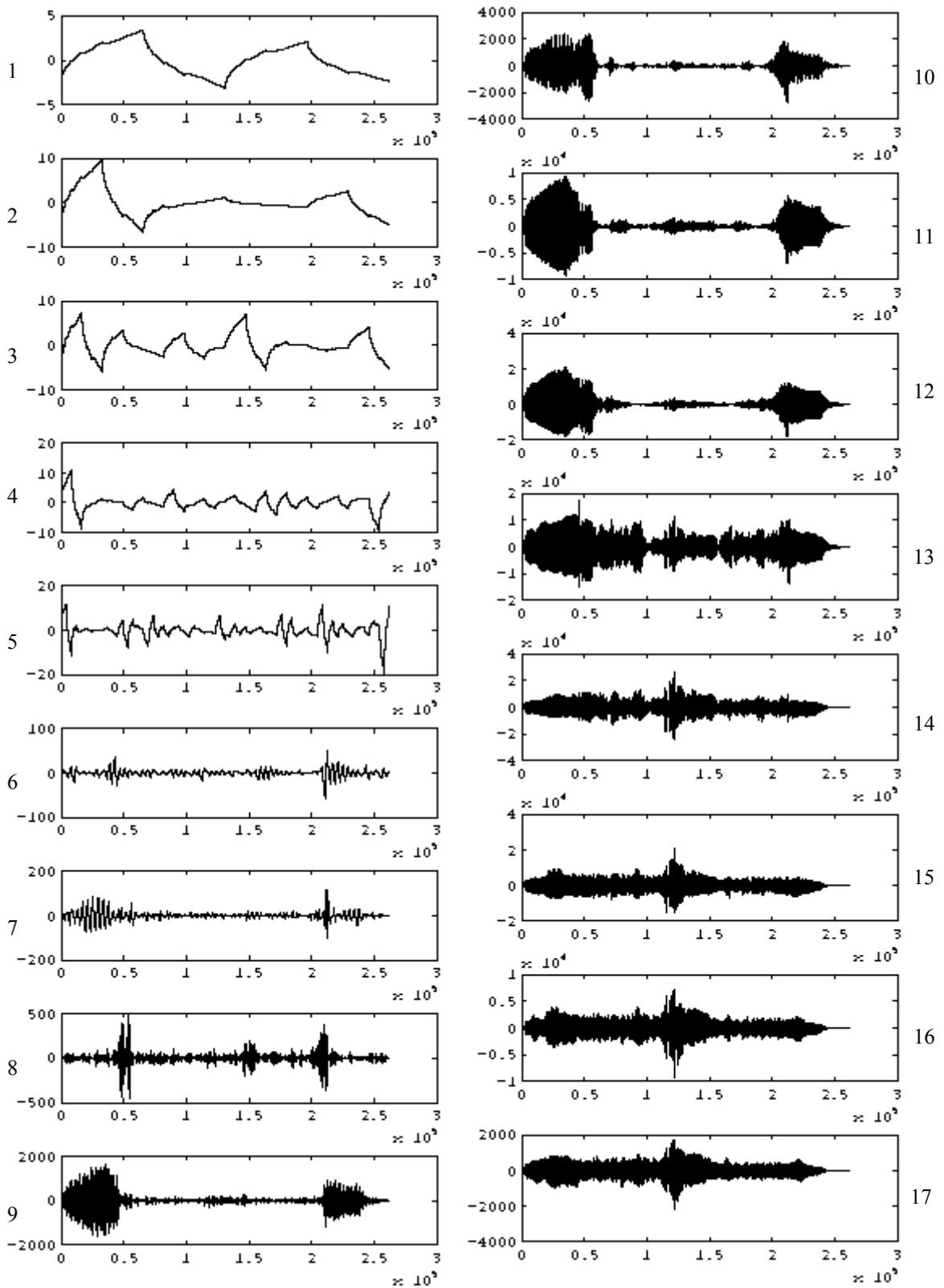
A audição dos níveis da análise mostra que os níveis mais energéticos concentram a maior contribuição da série harmônica dos sinais musicais, isto é, onde a maior parte da energia das vibrações musicais do instrumento se concentra. Além disso, nestes níveis ficam menos aparentes os gestos musicais que envolvem a produção de transientes, alterações bruscas e descontínuas.

Distribuição de energia:



Mais de 99% de toda a energia do exemplo concentra-se nos níveis 11 a 15. Os níveis 12 e 14 são os mais energéticos, intercalados por um nível 13 relativamente menos energético, exatamente por haver duas notas bem caracterizadas na interpretação deste exemplo: um lá3 (A3) e um lá5 (A5), distantes duas oitavas entre si, e portanto distribuindo-se pelos níveis da análise com padrões de distribuição de energia semelhantes porém deslocados por dois níveis. Como cada nível cobre uma oitava, é de se esperar que o nível mais energético para o lá3 esteja dois níveis abaixo do nível mais energético para o lá5. Os níveis mais baixos, especialmente os inferiores ao nível 7 (inclusive) não impactam a qualidade da reconstrução caso sejam desprezados, exatamente por não contribuírem significativamente em energia.

Violino: Fraseado, Ligado, Normal, em Glissando - Análise com D4



Evento: Mesmo fraseado ligado em glissando da análise anterior (fraseado ligado em glissando), agora analisado com a *Wavelet* D4. Trecho extraído da peça *Daphnis et Chloé*, de Ravel.

Gráfico do sinal original: (veja análise anterior)

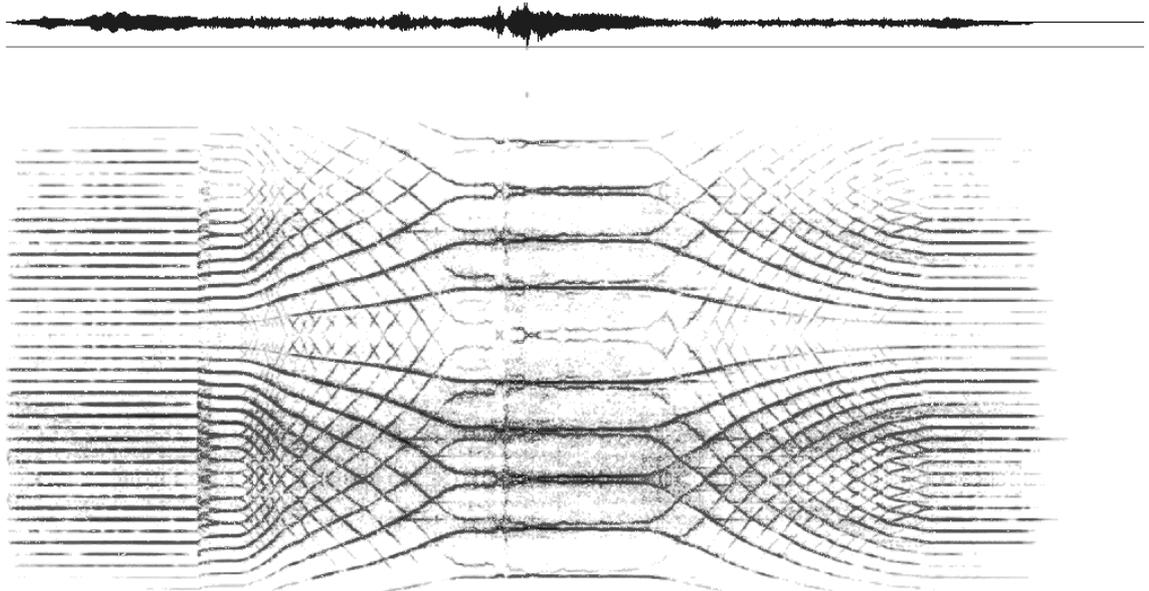
Extensão da amostra: 2^{18} pontos (262144). **Wavelet de análise:** D4

O nível 17, mais fino, apresenta uma alta coloração na região de alta frequência, com uma densa presença de frequências acima de 10000 Hz. A percepção auditiva do som do nível é a de um ruído agudo, com pouca informação musical sobre os gestos musicais associados no exemplo.

Ao nível 16, numa primeira audição percebe-se que existem alguns tons que sobem em frequência, enquanto outros descem. Um espectrograma do nível (veja Figura 4 abaixo) revela que algumas linhas mais intensas - normalmente percebidas como tons, cujos *pitchs* estão associados às frequências de pico do espectro - se deslocam para baixo durante o glissando (no sentido das baixas frequências), enquanto outras se movem no sentido inverso, para cima. Algumas destas raias sofrem um maior deslocamento em frequência devido ao glissando do que outras, que se deslocam de um intervalo menor. Nas regiões de sustentação das notas, onde os tons estão estáveis, o número de raias intensas para a nota mais alta (lá5) é menor e elas também estão mais bem separadas (em intervalos de aproximadamente 1810 a 1850 Hz), ao passo que para as notas mais baixas (lá3 do início e do fim) existe um maior número de raias, e a separação entre elas é bem menor, cerca de 4 vezes mais estreita (aproximadamente 430 Hz).

Na transição que caracteriza o glissando, o deslocamento das raias forma uma figura semelhante a uma interferência de frentes de onda. A figura a seguir mostra este espectrograma.

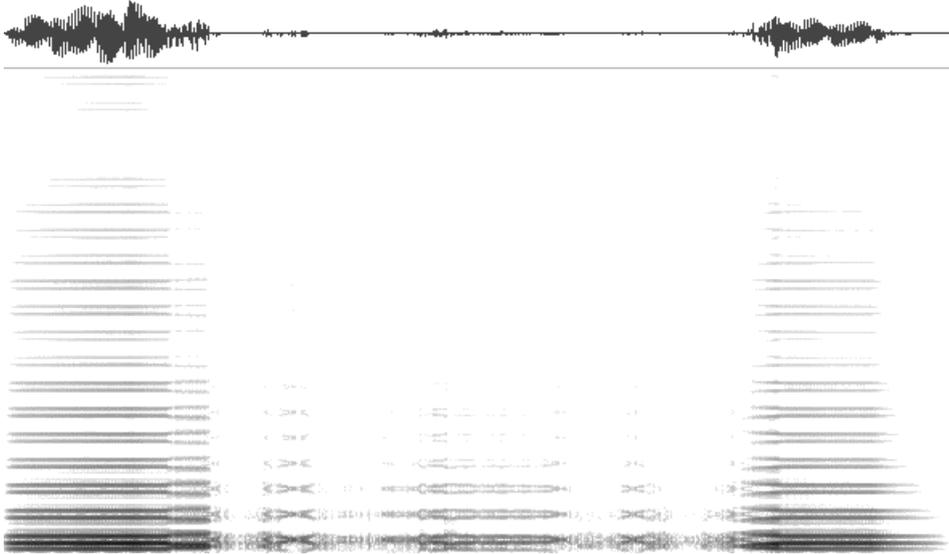
Figura 4 - Espectrograma do nível 16 (fraseado, ligado, normal, glissando; Wavelet D4)



Ao nível 11, as transições de glissando têm o mesmo efeito sobre a amplitude do sinal sonoro que teria a imposição de um envelope de amplitude sobre o sinal, isto é, pode-se considerar o glissando como um elemento modulador de amplitude neste nível. No glissando ascendente há uma diminuição da amplitude do sinal, e no descendente, o contrário: um aumento de amplitude no sinal. Isto é explicado pelo esquema de distribuição das notas/tons através de uma análise

Wavelet em multiresolução, onde as notas mais baixas concentram-se mais nos níveis mais baixos da análise. A figura abaixo mostra um espectrograma ilustrando o glissando visto no nível 11.

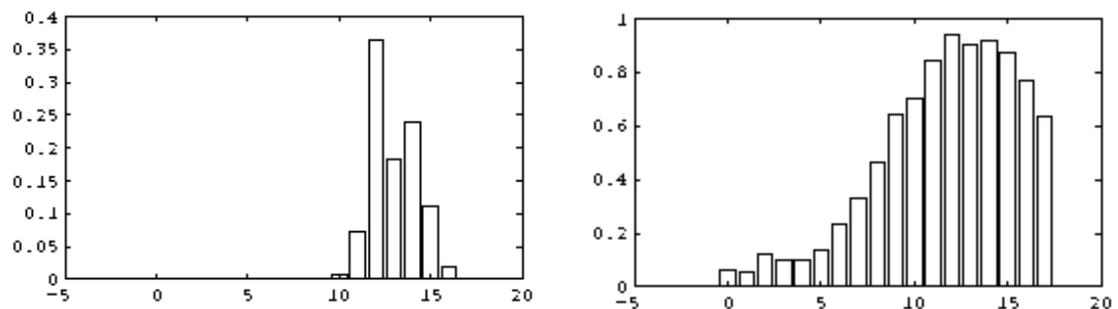
Figura 5 - Espectrograma do nível 11 (fraseado, ligado, normal, glissando; Wavelet D4)



O nível 10 assemelha-se em formato ao nível 11. A presença de certos "nódulos" no gráfico deste nível nos trechos de glissando evidencia os momentos quando as componentes de frequência cobertas por este nível tornam-se pronunciadas. O nível 9 já não permite mais extrair muitas informações sobre o efeito deste glissando. Somente as partes estáveis (sustentação) da primeira e segunda notas são audíveis, não havendo contribuição significativa deste nível ao conteúdo espectral da nota central (lá5).

O nível 8 destaca em suas três regiões de maiores amplitudes os momentos de início das transições, quando a intérprete opera sobre o dedilhado. Um exame mais cuidadoso mostra que os pontos de transição, que marcam o início ou o final de glissando, são localizados neste nível, sugerindo uma sensibilidade às variações de dedilhado e manipulação contínua do arco. Ao nível 6 já aparecem formatos de onda menos suaves, que se assemelham à *Wavelet* escolhida.

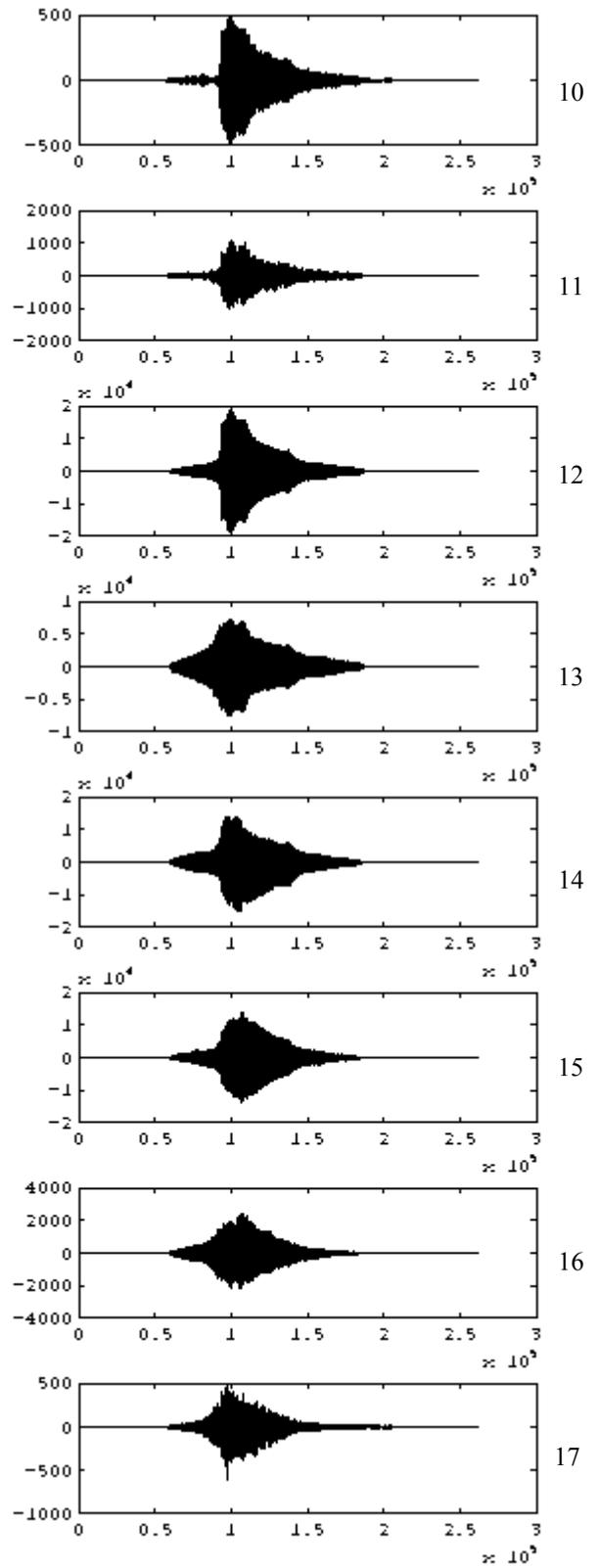
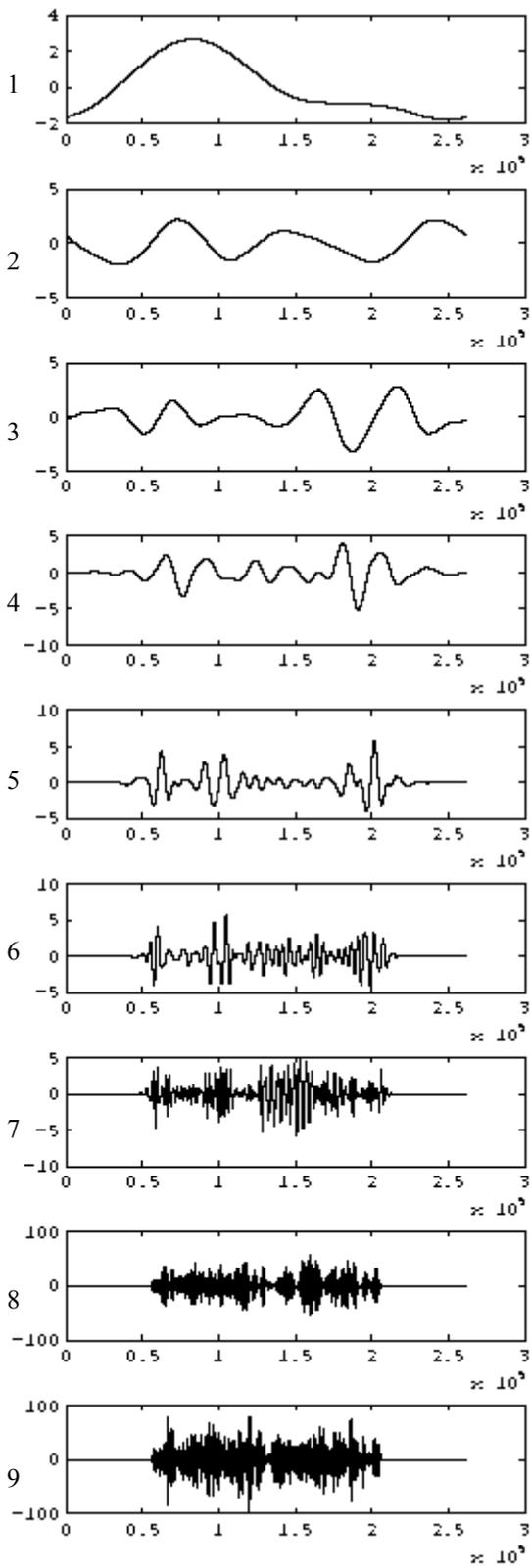
Distribuição de energia:



Mais de 99% de toda a energia se concentra nos níveis de 10 a 16, sendo os níveis 12 e 14 os mais energéticos. Quando a *Wavelet* de análise é menos regular (como a D4) a energia se distribui mais uniformemente pelos níveis, numa gradação mais suave entre a energia de um nível e do nível adjacente, como pode-se verificar comparando a distribuição (em escala

logarítmica) do segundo gráfico acima com seu correspondente na análise anterior (análise com D16).

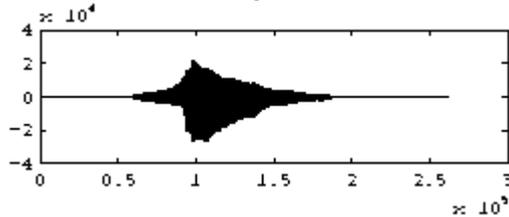
Violino: Fraseado, Ligado, Acentuado



Evento: fraseado, transição contínua (ligada) de uma nota normal (dó4 - C4) para uma nota acentuada (si3 - B3). O exemplo, notado em partitura ao lado, foi extraído de Piston (1955), à p.20 (ex.20). O interesse pelo trecho está na análise da transição para uma nota acentuada.



Gráfico do sinal original:



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

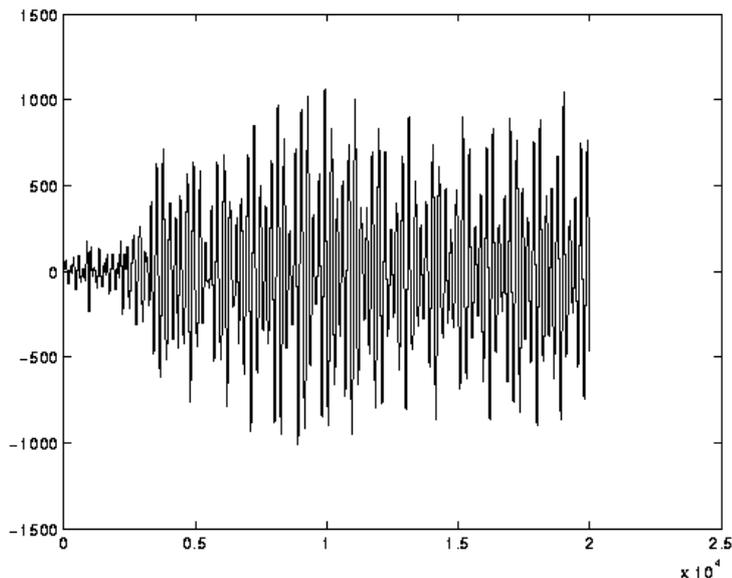
Níveis 17 e 16: Apresentam um envelope que acompanha a dinâmica variante do trecho. A intérprete, preparando-se para executar a nota acentuada (a segunda), produziu um aumento gradual na intensidade. Ambos os níveis mais finos, sensíveis ao ruído de arco, mostram que este é proporcional à dinâmica.

Nível 15: Transição audível destacada, na qual ouve-se um segundo *pitch* entrando imediatamente, sem um ataque típico. Note que este segundo *pitch*, um *pitch* importante na descrição do tom da segunda nota, aparece destacado neste nível, em bom som, porque essa nota, um si4 (B3), possui componentes intensas, energéticas, neste nível.

Nível12: Mais intenso, amplitudes médias sendo o dobro das do nível anterior

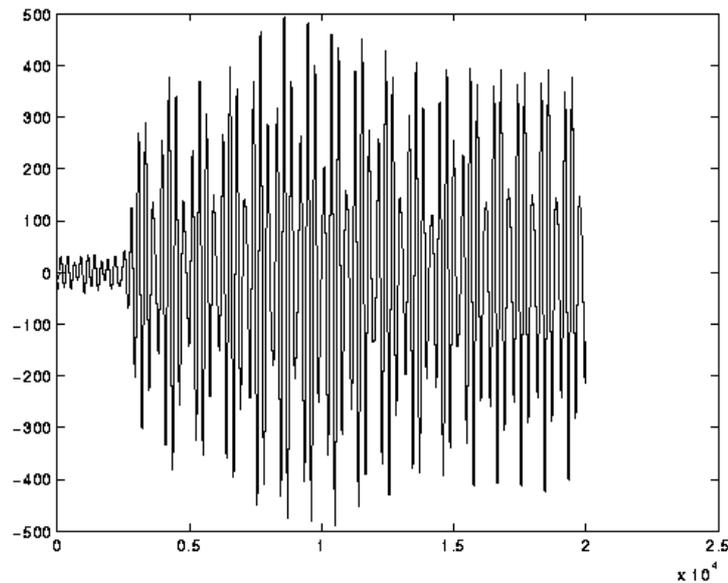
Nível11: Bem menos energético e mais grave que o nível anterior, próximo do limite da audição. No entanto, mostra mais destacado a presença de uma nota acentuada transitando em *legato* a partir de uma nota anterior não acentuada. O acento é expresso em termos da intensidade relativa no início da segunda nota, o que pode ser visto com mais proximidade na Figura 6 abaixo, mostrando o momento da transição e o pulso em cerca de 500 pontos de amplitude da primeira nota (do4, C4) para a segunda acentuada (B3).

Figura 6 - Detalhe magnificado da região de transição e o início da segunda nota (B3) (nível 11)



Nível10: Apesar da baixa energia comparado aos níveis mais finos e mais audíveis, a transição entre as notas neste nível fica melhor destacada, com um uma rampa de subida do final do C4 para o início de B3 bastante íngreme. A Figura 7 abaixo localiza com melhor resolução o momento de *onset* da segunda nota, e permite aferir visualmente a repentina elevação na taxa de variação de amplitude.

Figura 7 - Detalhe magnificado da região inicial do ataque da segunda nota (B3) (nível 10)



Um gráfico da derivada deste sinal (não determinístico) em função do tempo indicaria o ponto de transição, mas esta técnica de localização é mais necessária quando não se dispõe de uma ferramenta que permita localizar transientes facilmente no domínio do tempo, o que não se aplica à análise *Wavelet*, capaz de poder localizar transientes eficazmente na dimensão original (tempo).

Outra característica visualizada neste nível é a presença de uma modulação em amplitude na primeira nota, provavelmente em função de um leve vibrato.

Note também que a enorme presença de B3 no nível 10 se deve a ela ser uma nota mais grave que C4, e normalmente deve possuir componentes populando o nível 10.

Na maioria das vezes, as notas (tomadas as análises do período de sustentação, descontando-se os períodos transientes do ataque e colapsos) distribuem-se numa AWMR concentrando quase a totalidade de sua energia em 4 níveis consecutivos, possuindo pouca energia distribuída nos outros níveis da análise.

Dependendo do registro da nota, isto é, sua altura em frequência, que depende da sua posição nas oitavas da escala musical, elas deslocarão o seu vetor de distribuição pelos níveis da análise para cima (em direção a níveis mais agudos/finos) caso sejam mais altas, ou para baixo (em direção a níveis mais graves/grossos) caso sejam mais baixas.

Tomemos como exemplo a análise de uma nota mi4 (E4) tocada na corda lá do violino, com extensão da amostra de 2^{17} . Considere a análise *Wavelet* em multiresolução do período de sustentação da nota pura, quando já se estabeleceu um estado estacionário das vibrações do tom (a excitação/ressonância estão regulares, não havendo transientes relevantes, modulações ou efeitos especiais). A análise gerará 17 níveis, sendo o nível 0 o nível menos fino e o 16, o mais

fino. Constata-se, pela análise, que quase toda totalidade da energia da nota/tom se distribui por 4 níveis consecutivos, concentrando-se nos níveis 11, 12, 13 e 14, que também são os níveis mais claros e relevantes em termos de audibilidade, exatamente por reterem as frequências mais intensas e importantes na constituição espectral do tom.

Uma forma alternativa de se enumerar os níveis de uma AWMR, atendendo simplesmente a uma maior facilidade notacional, consistiria em tomar o nível mais fino como o primeiro nível (nível 1) e o menos fino ou menos resolvido como o nível X (se a amostra original possuir 2^x pontos). Esta nomenclatura alternativa poderia facilitar a identificação dos níveis mais concentradores de energia. No exemplo acima, os níveis mais concentradores seriam segundo esta nomenclatura os níveis 3, 4, 5 e 6.

Esta forma de enumeração é mais prática no sentido de ser independente do número de níveis ou da extensão da amostra sonora, uma vez que o nível mais fino sempre será o nível 1 e os outros subsequentes são numerados em ordem crescente, até o nível menos fino.

O fato mais importante no entanto é que a AWMR fornece meios para localizar o *onset* de uma nota acentuada dentro de um fraseado *legato*.

Nível 9: Primeiro nível que não faz referência ao envelope original. Também é um nível não relevante para termos de ressíntese, conforme pôde ser verificado reconstruindo-se o sinal original com a prévia eliminação da contribuição deste nível (zerando seus coeficientes) e comparando-se com o sinal original via audição. Também o ponto de enforcamento ou afunilamento observado no centro do sinal não corresponde à transição das notas neste exemplo, que ocorre antes.

Nível 8: Já muito grave, no limite da audição, sem relevância acústica.

Nível 7: Mostra estruturas periódicas possivelmente correlacionadas a gestos do arco, como um vibrato, e um aumento na pressão de contato no período de sustentação da segunda nota, a nota acentuada. O aumento da intensidade na segunda nota não está localizado a partir do seu *onset*, e sim após um curto espaço de tempo, já no período de sustentação. Isto revela pormenores expressivos da interpretação, uma vez que a intérprete deve ter esperado o ataque da segunda nota engatar para então acentuar a nota, ou então, permite levantar a hipótese de que este tipo de fraseado já possui um padrão gestual natural para a intérprete seguir.

Serão necessários estudos mais específicos para revelar e comprovar se gestos de arco podem ser caracterizados neste nível.

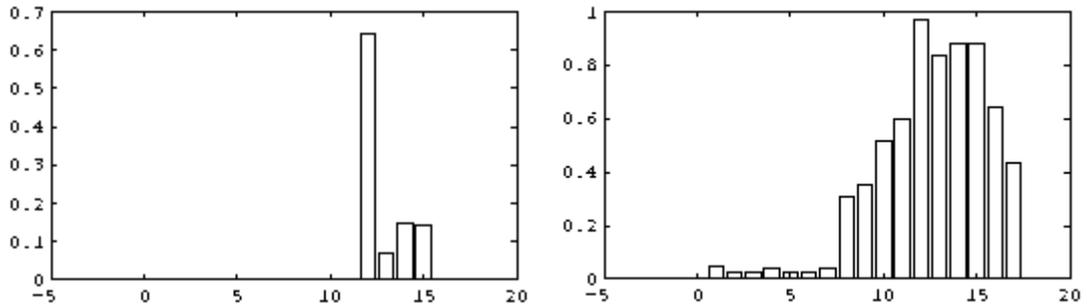
Nível 6: Revela picos nos ataques das duas notas. Revela os momentos de ataque e início da sustentação de cada uma das duas notas do exemplo, mesmo com precisão de localização já menor que nos níveis anteriores mais finos. Provavelmente a informação neste nível é sensível às mudanças de parâmetros físicos do arco sobre a corda (velocidade, direção e pressão), mas a confirmação desta hipótese exige a análise orientada exclusivamente a gestos de arco.

Nível 5: Também mostra as mesmas estruturas do nível anterior, mas (vantajosamente) descartando detalhes não importantes para caracterização e visualização dos picos. Nível inaudível.

Nível 3: Pode ser útil para revelar informações num nível musicológico, localizando períodos onde há temas repetidos ou progressões numa análise inteira de um movimento sinfônico. Induz-nos a sugerir que pode avaliar o número de eventos importantes no período analisado, ou

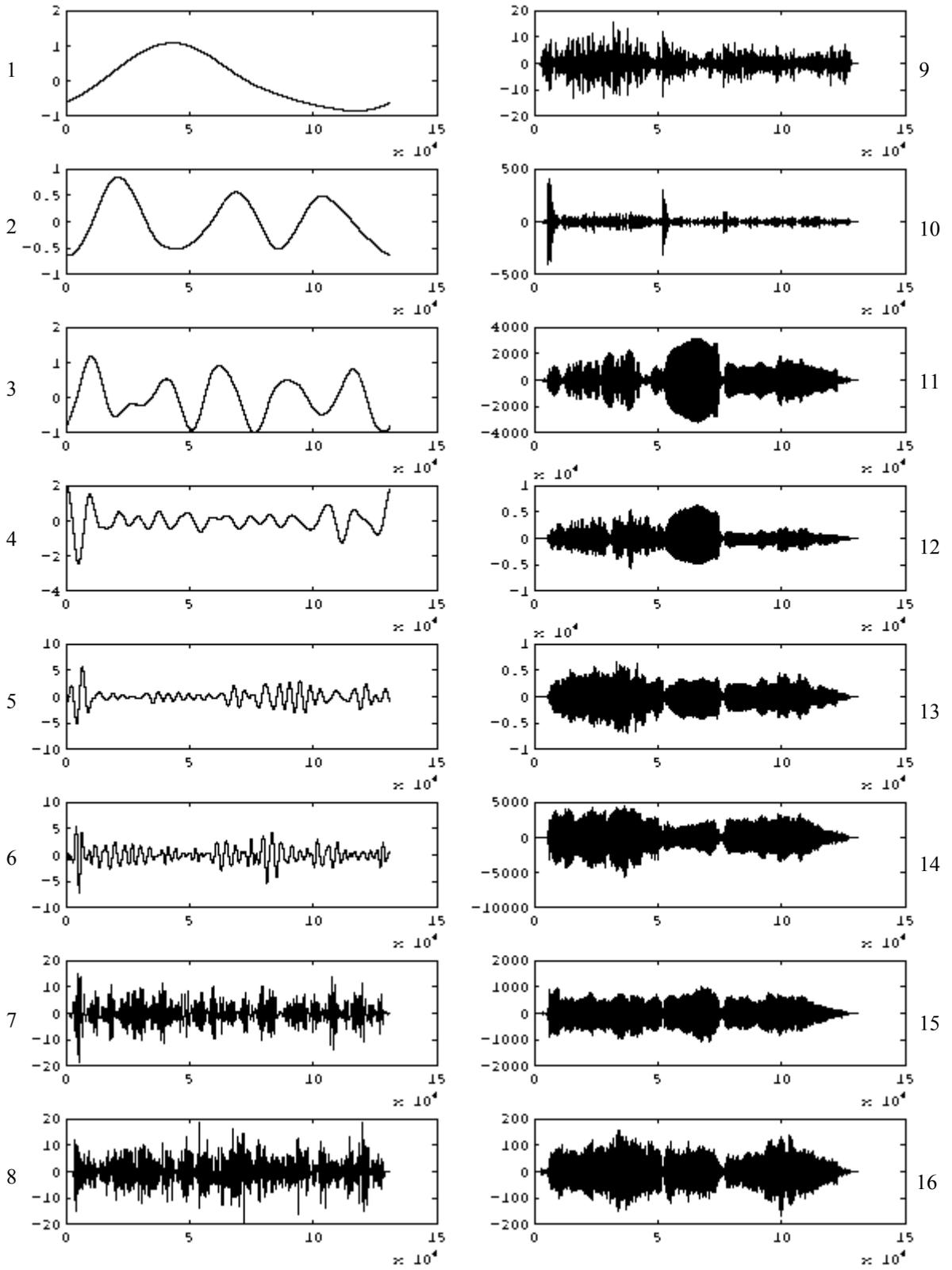
seja, neste caso, duas notas, uma vez que podemos discernir dois grupos ondulatórios mais destacados.

Distribuição de energia:

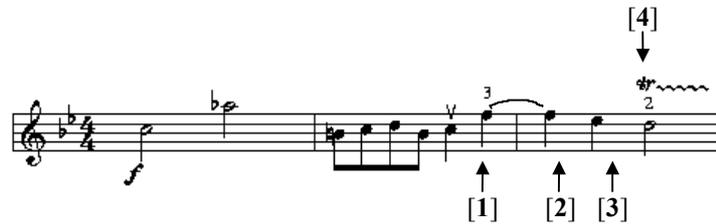


Quase toda a totalidade da energia deste exemplo concentra-se nos níveis de 12 a 15, com o nível 12 sendo o mais energético (veja gráfico com escala de intensidade linear à esquerda). O gráfico à direita, como nas análises anteriores, plota o logaritmo da energia de cada nível pelo logaritmo da energia total. Note que os níveis de 0 a 7 praticamente não contribuem energeticamente.

Violino: Fraseados Não ligados, Normais

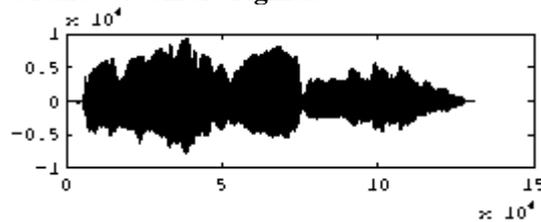


Evento: Os fraseados normais não ligados deste exemplo foram extraídos da interpretação dos 2º e 3º compassos, em *Allegro assai*, da peça "Sonata", de Tommaso Albinoni (Opus 2/IV). Veja abaixo a partitura correspondente ao trecho tocado (segundo e terceiro compassos somente):



Neste trecho há quatro eventos musicais de interesse (veja indicações na partitura acima): o ataque da primeira nota, que se dá ao final do 2o. compasso [1]; o fraseado não ligado da primeira nota do 3º compasso (fá4 - F4) para a segunda nota (mi bemol 4 - Eb4) [2]; o fraseado também em transição não ligada desta segunda nota para a última nota do compasso (ré4 - D4) [3]; e finalmente o trinado aplicado à esta última nota (D4) [4].

Gráfico do sinal original:



Os estrangulamentos visíveis no sinal acima são as fronteiras que delimitam as notas.

Normalmente, em fraseado não ligado, cada nota numa sequência requer uma mudança no sentido do arco, isto é, para cada nota alterna-se o sentido do movimento do arco, ora para cima, ora para baixo, e na fronteira é executado o gesto de chaveamento do sentido.

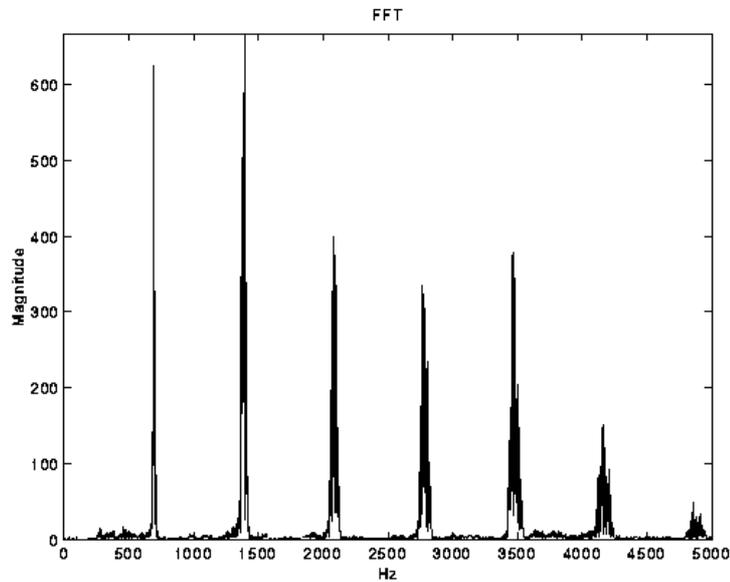
Neste exemplo específico foi solicitado à intérprete que evitasse vibratos na execução do trecho. É interessante notar que, ao omitir o efeito de vibrato (que é uma modulação automaticamente aplicada pelos violinistas) a intérprete também omitiu o trinado e observou com mais rigidez a grade de tempo (duração de cada nota).

A **primeira nota** do exemplo é um **Fá4** (F4), que segundo a tabela de tons/frequências de Itokawa e Kumagai (1952) corresponde a 698,4 Hz.² Uma **análise FFT sobre a extensão da primeira nota (F4)** mostra que os picos, em ordem decrescente de intensidade, estão em **1393,6 Hz (segundo harmônico)**, **694,4 Hz (fundamental)**, **2078,6 Hz (terceiro harmônico)**, próximo do terceiro harmônico teórico em 2083,2Hz), **2769 Hz (quarto harmônico)**, próximo do esperado em 2777,6 Hz), etc., seguindo a lógica esperada para instrumentos harmônicos.

A Figura 8 mostra a FFT desta primeira nota entre 0 e 5000 Hz, onde estão todos os seus harmônicos mais intensos e maior energia:

² Como a entonação no violino depende da posição do dedo sobre a corda, e também varia em função da afinação microtonal própria de cada intérprete, pode-se admitir uma faixa de variação em torno do *pitch* oficial

Figura 8 - FFT da primeira nota do exemplo (intervalo de 0-5000 Hz)



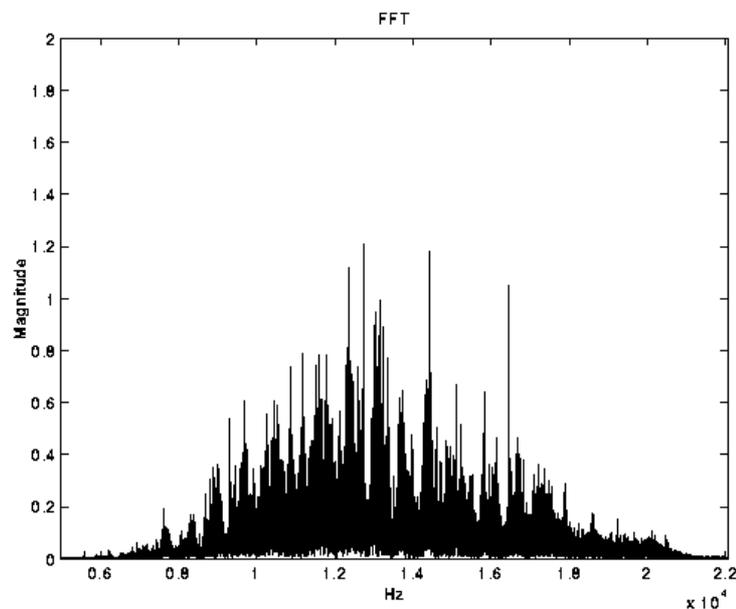
Os eventos de maior interesse de análise no momento são as transições não ligadas nos fraseados.

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

Análise dos níveis:

Nível 16: Compreende a evolução de uma densa população de frequências altas, acima de 5000 Hz. A Figura 9 mostra a FFT deste nível na faixa de 5000-22050 Hz.

Figura 9 - FFT do sinal ao nível 16 (mais fino)



Pela sua análise, verificamos que a FFT não ajuda a localizar características importantes musicalmente relacionadas ao exemplo, simplesmente nos fornecendo os valores relativos da contribuição de cada frequência presente, e nos informando que há uma densa população delas presentes no som do violino nesta execução.

Normalmente o nível mais fino retém o ruído característico do arraste do arco sobre a corda, e pode ser útil para localizar os momentos onde a velocidade do arco passa por zero, isto é, os

pontos onde ocorre a mudança de sentido do arco. Também permite inferir sobre a quantidade de atrito envolvida no arraste do arco em cada instante. Os pontos de estrangulamento visíveis no gráfico deste nível localizam os momentos em que o arco pára (velocidade nula) para alternar o sentido.

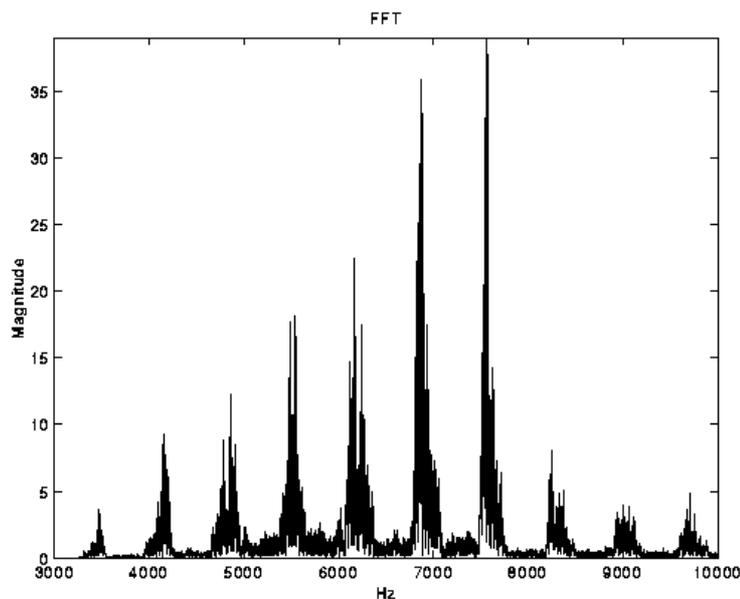
A identificação e diferenciação das tonalidades fica difícil por estarem fundidas ao som ruidoso. Levemente pode-se notar um vibrato via audição.

Nível 15: Envelope semelhante ao nível anterior; este nível já permite a percepção de tons diferentes para as notas. Níveis altos, como este, revelam fenômenos vibratórios associados com frequências altas, e que portanto colapsam velozmente na ausência da fonte de excitação. Estes níveis são úteis para localizar transientes, pontos de mudança bruscas no tom, e os delimitar os envelopes de cada nota, bem como caracterizar sua composição frequencial nesta escala/nível.

Passemos a uma análise da FFT da primeira nota do nível para avaliar sua estrutura espectral estática, que geralmente fornece informações sobre a série harmônica.

A figura abaixo mostra a FFT da primeira nota do nível 15, entre 3000 e 10000Hz.

Figura 10 - FFT da primeira nota ao nível 15 da análise (plotado entre 3000-10000 Hz)



Como a análise *Wavelet* isola faixas de frequências, e o nível 15 analisa frequências mais altas, não estão presentes os harmônicos mais baixos do tom.

As raias presentes sugerem a presença de uma série harmônica. Verifica-se, no entanto, um deslocamento em frequência das raias esperadas para a série harmônica desta nota.

O primeiro pico à esquerda se encontra em 3458,1 Hz, nas redondezas do **5º harmônico** da nota. A existência de um vibrato muito sutil e a influência dos transientes fora da região de sustentação da nota podem explicar o deslocamento desta raia do valor esperado (~14 Hz acima).

O segundo pico está em 4160 Hz (o 6º harmônico).

O terceiro pico está em 4854,4 Hz, deslocado de 2 Hz do **7º harmônico** da nota (4856,4 Hz).

O quarto pico está em 5534 Hz, nas redondezas do **8º harmônico**. É interessante notar que existem entre 5480-5544 Hz quatro picos com amplitudes próximas entre si, formando uma estrutura espelhada em torno de 5512 Hz.

O **quinto pico** está em **6170,6 Hz**, deslocado do **9º harmônico em 76,7 Hz**. Entretanto, um outro pico local mais adiante em 6247,3 Hz melhor corresponde ao 9º harmônico.

O **sexto pico** está em 6865,1 Hz, **deslocado do 10º harmônico em 70 Hz para baixo**. Há também um pico local em 6935 Hz, menor em intensidade, correspondendo ao 10º harmônico.

O **sétimo pico - o mais intenso desta FFT** - está em 7566,9 Hz, cerca de **58 Hz deslocado** do 11º harmônico, não sendo portanto representativo deste harmônico. Um pequeno nódulo um pouco mais à frente, com um pico local em torno de 7625 é mais representativo do 11º harmônico.

Verifica-se a presença da série harmônica na FFT da primeira nota do nível 15, mas também outras frequências são observadas, aparecendo a partir do 5º pico, e deslocadas da série harmônica. Estas componentes se concentram ao redor das raiais da série harmônica sem serem representativas desta.

Observando-se o padrão de distribuição destas componentes, e a forma como parecem se espelhar em torno de algumas frequências centrais, verificou-se que:

- Na raia associada ao 8º harmônico, há duas bandas laterais, aproximadamente de mesma intensidade, "espelhadas" ao redor de 5512,5 Hz, exatamente 1/8 da frequência de amostragem (44100 Hz).
- A linha espectral mais intensa da FFT - em 7566,9 Hz - é também uma banda lateral "espelhada" da banda associada ao 5º harmônico - em 34581 Hz, também ao redor de 5512,5 Hz.
- A linha em 6865,1 Hz é um "espelho" da linha em 4159,9 Hz, o 6º harmônico, espelhada em torno de 5512,5 Hz.
- A linha em 6170,6 Hz é um "espelho" da linha em 4854,4 Hz, o 7º harmônico, espelhada também em torno de 5512,5 Hz, 1/8 da frequência de amostragem.

Esta regra de formação, posteriormente verificada em vários níveis de todas as análises, é devida ao processamento do algoritmo piramidal que implementa a transformada *Wavelet* inversa (reconstrução). Particularmente, a replicação de bandas - ou **imageamento** - é descrita por Strang (1996), e causadas pelas etapas de superamostragem ($\uparrow 2$) no processo de reconstrução na árvore de filtros. O imageamento de bandas provoca o aparecimento de "bandas gêmeas" ou imagens dispostas espelhadas ao redor de divisores racionais da frequência de amostragem.

Nas discussões da dissertação, no capítulo 5, item 5.2 do texto principal, aborda-se este efeito com mais profundidade.

É possível reconhecer o tom original (F4) a partir da sua audição no nível 15. O timbre neste nível é ruidoso, desprovido de graves, mas mantém a integração da série harmônica, sendo possível ainda reconhecer o registro do tom original (a nota e oitava).

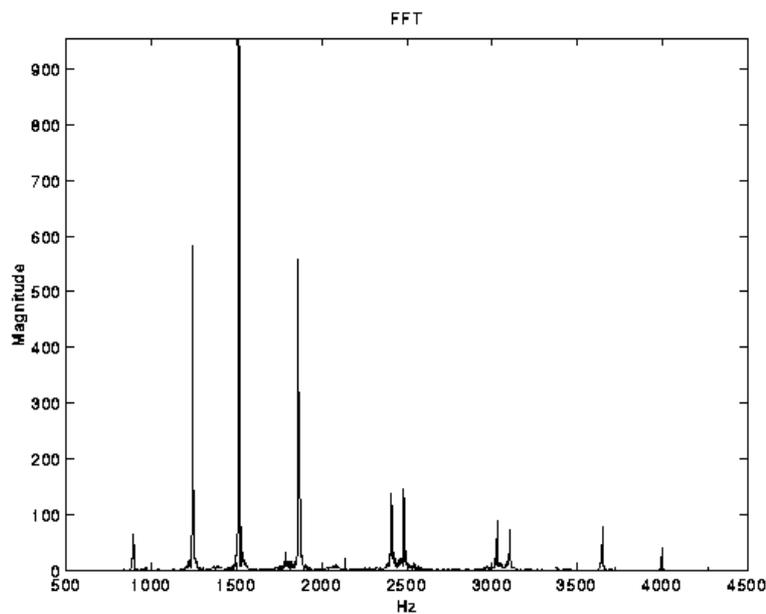
Isso mostra que o segundo nível mais fino de uma AWMR contém uma parcela da série harmônica original, e é capaz de manter sua integração, mesmo na existência de outras combinações de frequências. No caso específico que analisamos verificamos que há uma "série paralela" à série harmônica original também presente na nota.

Como no nível 16, é possível localizar os pontos de mudança de sentido do arco. Também é sensível a qualidade de um leve vibrato sobre a última nota.

Nível 13: Soa como um "rádio antigo", uma recepção AM de um rádio de baixa qualidade. O "rabicho" do ataque da primeira nota é praticamente desprezível neste nível. Os pontos de estrangulamento nesta faixa de frequências não possuem amplitudes tão pequenas como nos níveis anteriores. A primeira nota no nível 13 possui 4 harmônicos intensos da nota F4 original: o segundo (de maior amplitude), o terceiro, quarto e quinto (este último com o pico em 3433,9, cerca de 38 Hz abaixo do valor correto para o quinto parcial)

A figura abaixo mostra a FFT da segunda nota neste nível.

Figura 11 - FFT da segunda nota do nível 13



A nota original Eb4 possui *pitch* central em torno de 620,4 Hz (primeiro harmônico ou fundamental). O segundo harmônico, mais intenso, está em 1242,2 Hz, e o terceiro, em 1826 Hz.

A segunda nota no nível 13 (correspondente à Eb4 original) apresenta o seguinte espectro estático (veja figura acima): uma linha em 1242,2 Hz (segundo harmônico), pico mais intenso corresponde em 1514,1 Hz (3ª raia da esquerda para direita), 1826 Hz (terceiro harmônico). Também aqui verifica-se a ocorrência de raias espelhadas, desta vez em torno de 1/16 da frequência de amostragem, ou 2756,5 Hz: a linha em 2409 Hz é "gêmea" do 5º harmônico, em 3103,5 Hz; da mesma forma, a linha em 2481,7 Hz (o 4º harmônico para esta nota) é "gêmea" da linha que aparece em 3030,8 Hz.

A linha intensa em 1514,1 Hz não faz parte da série harmônica do tom, e consiste numa imagem do segundo harmônico (1242,2 Hz) espelhado ao redor de 1/32 da frequência de amostragem (1378,125 Hz).

Vejamos agora a última nota do exemplo: o D4 (ré4).

A FFT original da nota ré4 (D4) consta de uma série harmônica com picos em **594,2 Hz (fundamental, e mais intenso)**, em **2384,8 Hz (redondezas do quarto harmônico)**, juntamente com um pico muito próximo em 2379 Hz), em **1800 Hz (o terceiro harmônico)**, em **2962,2 Hz (redondezas do quinto harmônico)**, e em **1189 Hz, o segundo harmônico**.

Na FFT desta última nota **no nível 13** (figura não mostrada), o pico mais alto está em **2384,8 Hz** (quarto harmônico), em seguida vindo o terceiro harmônico, em **1800 Hz**

A presença de raias "gêmeas", espelhadas ao redor de frequências que são divisores racionais da frequência de amostragem (f_a), também é verificada:

- a linha em 956,2 Hz é imagem do terceiro harmônico, em 1800 Hz, ao redor de $1/32 f_a$
- a linha em 1567,2 Hz é imagem do segundo harmônico, em 1189 Hz, ao redor de $1/32 f_a$
- a linha em 2550,3 Hz é imagem do quinto harmônico, em 2962,2 Hz, ao redor de $1/16 f_a$
- a linha em 3127,7 Hz é imagem do quarto harmônico, em 2384,8 Hz, ao redor de $1/16 f_a$

Nível 12: A primeira nota aparece cheia de pulsos periódicos, envelope irregular. A segunda nota aparece com um envelope mais contínuo, sem tantas flutuações como a primeira nota, e apresenta as maiores amplitudes relativas às outras 2 notas. A terceira nota, originalmente a mais baixa das três em *pitch*, aparece com menores amplitudes relativas às outras duas.

Um novo aspecto merece ser abordado no que tange à percepção auditiva dos *pitchs* centrais de cada nota na sequência em que são tocados. No exemplo original o pequeno fraseado de 3 notas progride descendo a escala: a primeira nota é um fá4, a segunda um mi bemol 4, e a última um ré4, num claro **movimento descendente** em frequência.

No nível 12, entretanto, o que se percebe numa primeira audição é um **tema ascendente**: a sequência das três notas ao invés de descer a escala parece subir. Isso ocorre porque neste nível certas faixas de frequências em cada nota são destacadas de outras, evidenciando alguns modos harmônicos (e possíveis imagens), e alterando o *pitch* central percebido. Vejamos os *pitchs* predominantes em cada uma das notas nestes níveis (correspondendo às raias predominantes do espectro estático):

A **primeira nota** apresenta neste nível ***pitch* predominante** no **segundo harmônico** da nota original (F4): pico em **1388,2 Hz**, e uma segunda raia mais intensa na fundamental, em 694 Hz. A **segunda nota** apresenta neste nível ***pitch* predominante** em **1242,2 Hz** (**segundo harmônico** da nota Eb4), e um segundo *pitch* imediatamente menos intenso em 1514 Hz. Já a **terceira nota** apresenta ***pitch* predominante** em **783,9 Hz**, e um segundo *pitch* menos imediatamente menos intenso em 1189 Hz (segundo harmônico desta nota), e ainda um terceiro *pitch* em 1567 Hz.

Vejamos então se a sequência percebida deve ascender analisando-se os *pitchs* predominantes:

primeira nota: **1388 Hz** e 694 Hz.

segunda nota: **1242 Hz** e outro em 1514 Hz

terceira nota: **783,9 Hz**, depois 1189 Hz e 1567 Hz.

Pode-se perceber que a sequência é decrescente segundo os *pitchs* predominantes. Por outro lado, a sequência 1388-1514-1567 Hz conduz à percepção de uma série ascendente, e é possível que a sensação de ascendência seja clara para alguns e mascarada para outros, uma vez que podemos agrupar sequencialmente *streams* diferentes.³

Nível 11: O nível de flutuações/modulações é mais visível graficamente bem como audível. Flutuações mais evidentes aparecem especialmente na primeira nota. O nível é sensível a vibrações, e modulações.

³ A noção de "*streams*" vem da Psicoacústica e refere-se aos padrões e objetos que são agrupados pelo sistema auditivo numa única unidade percebida. Bregman (1990) aborda este conceito em profundidade.

As modulações na segunda nota apresentam uma frequência superior à das outras duas. As modulações podem ser devido menos ao vibrato, embora sutil, realizado pelos dedos, e mais às imagens espelhadas das bandas, que produzem um padrão de *modulação em anel*: um tipo de modulação de amplitude onde a portadora inexistente.

O nível revela capacidade de se localizar mudanças de nota, os pontos de transição, e caracterizar o colapso no final da última nota, com uma qualidade audível diferenciada para cada estrutura.

É um nível rico para se analisar auditivamente, uma vez que todos os objetos detêm propriedades sonoras diferenciadas, consequência direta da qualidade de nível energético (este é o segundo nível mais energético da amostra).

Apresenta semelhanças com o nível 12 no sentido de que há arranjos de tons diferentes em cada nota, e também diferentes percepções do agrupamento sequencial da sequência tocada.

Nível 10: É bastante relevante a capacidade deste nível de rastrear os gestos do arco nos momentos dos ataques de cada nota, dando inclusive uma indicação da intensidade de cada ataque. No tempo restante, onde não há transições bruscas na regularidade do movimento do arco, paradas ou mudança de sentido, o sinal neste nível se mantém numa intensidade vários graus de grandeza menor que nos pontos de ataques, como um ruído de baixa frequência, que parece derivar da ressonância na caixa do violino.

A quantidade de energia neste nível é cerca de 659 vezes menor que no nível 11, e este nível pode ser desprezado numa reconstrução, e o resultado sonoro ainda é praticamente o mesmo do som original.

Nível 9: O formato característico do envelope, visível nos níveis superiores ao 10, aqui é perdido completamente. O nível 9 não conserva este formato de envelope. Destaca as componentes vibratórias provenientes da caixa sob ressonância. A localização de estruturas, transitórios e *onsets* das notas também é deficiente neste nível.

No entanto, a amplitude das ondas neste nível apresenta certa dependência do nível de dinâmica impresso pelo violinista. Como a pressão do arco não influi significativamente na intensidade do som produzido, mas sim esta depende da velocidade e da posição da corda em que o arco é arrastado, seria interessante um estudo mais aprofundado para verificar a correlação da amplitude neste nível com parâmetros como velocidade e ponto de engate do arco na corda.

Nível 8: Semelhante ao nível 9. Ruído de baixa frequência, *pitch* predominante em 45 Hz, segundo raia mais destacada em 66 Hz, e uma presença menos intensa de várias frequências na faixa de 20-150 Hz..

Nível 7: Semelhante ao nível 8. Praticamente inaudível. Não oferece informação gráfica, nem audível sobre os gestos existentes. Sua ausência em reconstruções seletivas, como nos demais níveis abaixo mais grossos, não impacta a qualidade percebida do sinal em comparação ao som original..

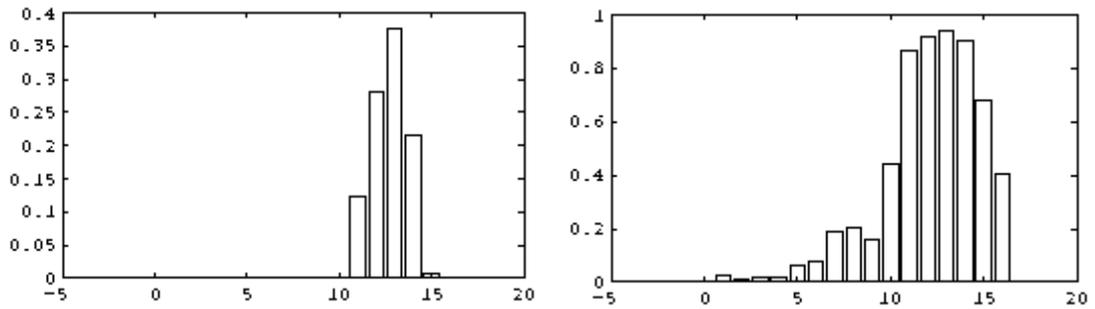
Nível 6: Picos na região do primeiro ataque. Inaudível. Revela periodicidade nas faixas de frequência onde atua este nível, mas a significância é desprezível em termos de energia.

Nível 5: Há modulação em amplitude, o primeiro ataque é forte, e a amplitude cresce no início da 3a. nota, que é mais baixa em frequência.. A localização entretanto é péssima nesse nível.

Nível 4: apresenta grande amplitude no início e no fim. A ondulação é praticamente constante ao longo de todo o exemplo, assim como a frequência de ondulação.

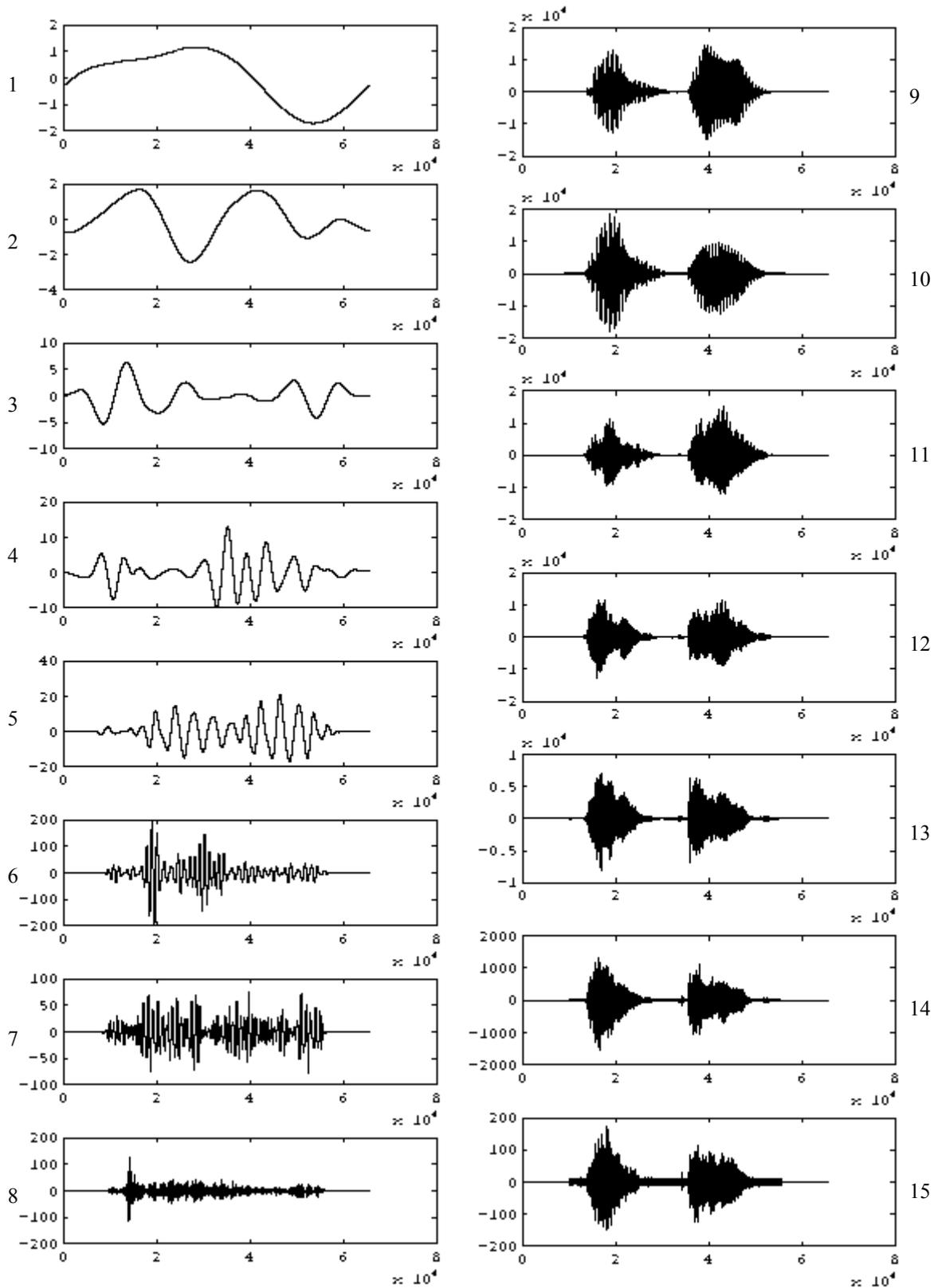
Os níveis 3, 2, e 1 já denotam as formas das *Wavelets* D16.

Distribuição de energia:

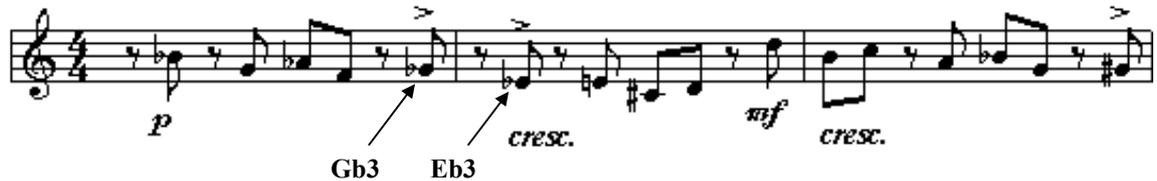


Mais de 99% de toda a energia do exemplo concentra-se nos níveis 11 a 15. O nível 13 é o mais energético. Os níveis mais baixos, especialmente os inferiores ao nível 9 (inclusive) não impactam a qualidade da reconstrução caso sejam desprezados nesta, exatamente por não contribuírem significativamente em energia.

Violino: Fraseado, Não Ligado, Acentuado

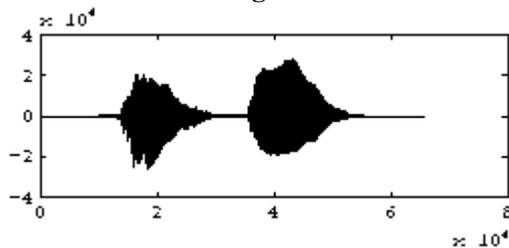


Evento: Fraseado não ligado acentuado. Interpretação exagerada de duas notas acentuadas: um sol3 bemol (Gb3) e um mi3 bemol (Eb3), indicados na partitura abaixo.



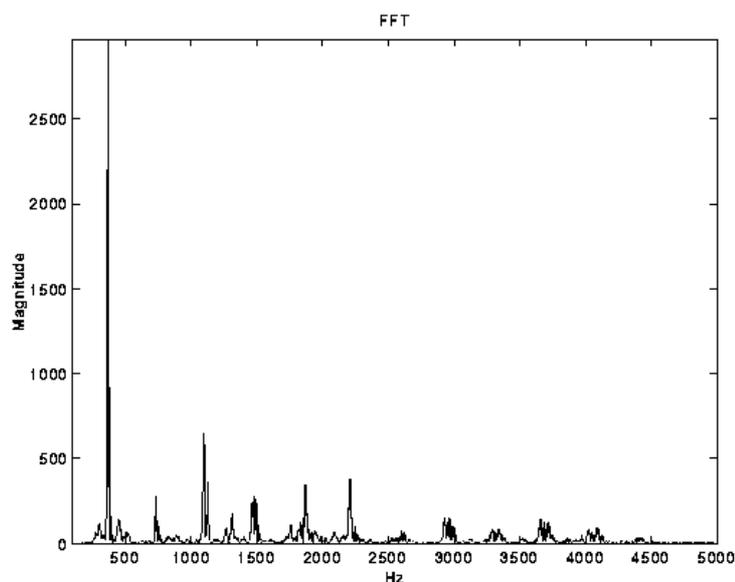
Neste exemplo⁴ o objetivo foi o de capturar o fraseado compreendendo duas notas acentuadas de importância neste tema melódico, e que foram executadas num trânsito não ligado (separadas ainda por pausa). Consistem pois num fraseado *non legato*, com acentos, tocados em dinâmica *piano*, crescendo suavemente.

Gráfico do sinal original:



A análise FFT da nota Gb3 revela pico máximo na sua **fundamental**, em **370,1 Hz**, segundo pico em **1100 Hz** (cerca de 10 Hz deslocado do esperado **terceiro harmônico** em 1110 Hz), logo depois vem o **sexto harmônico** (**2208 Hz**, deslocado de 12,6 Hz do harmônico esperado em 2220,6 Hz). A Figura 12 mostra a FFT desta nota, entre 100 e 5000 Hz:

Figura 12 - FFT da primeira nota do fraseado: sol3 bemol (Gb3)



A análise FFT da segunda nota, Eb3, revela pico máximo na **fundamental** em **311,6 Hz**; o segundo pico vem em **624,5 Hz** (o **segundo harmônico**), e em seguida o **quarto harmônico** em **1243 Hz** é o terceiro pico mais intenso.

⁴ extraído de Piston (1955), à p.22.

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

Análise dos níveis:

Nível 15: O sinal neste nível mostra a intensidade do atrito provocado pelo arco nos ataques acentuados. A forma da onda revela o envelope de amplitude da contribuição do arco arrastando sobre a corda tocada, variando sua pressão e velocidade para caracterizar o acento. A definição de um envelope característico de acentos não fica mascarada mesmo em dinâmicas *piano*, como é o caso deste exemplo. O acento na primeira nota é mais vigoroso provavelmente devido ao sentido do arco na execução desta nota (arco para baixo). A segunda nota revela uma pressão/velocidade de arco menor e mais suave que na primeira. Isto pode ser verificado pela audição das notas originais.

Também presente neste nível um ruído de fundo com amplitude praticamente constante do início ao fim. No sinal original, na transição entre as notas, verifica-se que a intérprete realizou um colapso forçado da primeira nota, a fim de respeitar a sua duração oficial e também executar a pausa (silêncio) prevista na partitura entre as notas acentuadas. A técnica usada para este fim foi forçar o colapso da nota, sem no entanto parar completamente o arco durante a pausa. No entanto esperar tal interpretação mecanicamente precisa de uma interpretação natural é querer também extrair a expressividade embutida na mesma, que dá naturalidade ao tema executado, aspectos que exatamente deseja-se compreender neste trabalho. Outro ponto a ser notado é que os momentos em que a amplitude do sinal neste nível aumenta consideravelmente (em virtude dos acentos do arco) não coincidem exatamente com o início das notas, isto é, o ponto em que a intérprete ataca a corda em cada nota, mas ocorre logo em seguida, já sobre a fase de sustentação.

Novamente observamos aqui comportamentos ligados à interpretação real nos violinos: os acentos em fraseados não ocorrem simultaneamente com o engajamento do tom correto, devendo esse ocorrer primeiro. Isto parece naturalmente óbvio, porque na mudança de uma posição a outra, a intérprete primeiro deve verificar a correção da posição e do tom que ela produz (o que normalmente um violinista experiente faz pelo tato e confirma pela audição) para então acentuá-lo, caso contrário poderia destacar pelo acento uma nota errada ou ligeiramente desafinada.

Outro aspecto deste nível é que ele não fornece quase informação auditiva alguma para se determinar o tom de cada nota, as duas notas expressam praticamente o mesmo registro tonal, e possuem qualidade de timbre idênticas a de um ruído áspero e agudo.

Nível 14: Este nível revela um envelope diferente do nível anterior, duas coisas notavelmente visíveis a primeira vista: (1) o ruído de fundo existente é cerca de 85 vezes menos intenso que no nível anterior (em média para o mesmo intervalo medido no nível 15), e os pontos que localizam melhor no tempo o *onset* dos acentos são mais definidos.

A energia média neste nível é também cerca de 90 vezes maior que no nível mais fino (15), uma medida de sua também maior importância na contribuição energética do sinal original.

Um pequeno objeto visível pouco antes da segunda nota sugere ser aquele o ponto onde o arco sofre uma parada e mudança de sentido, preparando-se para a próxima nota.

A percepção dos *pitchs* já é mais clara, e é possível inferir-se sobre a nota que está sendo tocada.

A partir deste nível em direção ao mais baixo percebemos sempre a capacidade da análise *Wavelet* em localizar notas não ligadas, e em alguns deles (como neste também) é possível verificar visual e auditivamente a presença de acentos.

A FFT deste nível não mostra nenhuma componente abaixo de 2000 Hz. A FFT da primeira nota neste nível apresenta pico em 6936 Hz, o segundo logo abaixo em 7303,8 Hz, frequências muito pouco salientes na análise FFT original desta nota (mas aparecem picos locais nesta faixa de 6000-7500 Hz na nota original). Na série harmônica esperada teríamos o 20º harmônico em 7402 Hz, cerca de 100Hz deslocado do segundo pico verificado acima.

Foi verificado que a análise *Wavelet* revela algumas "frequências explodidas" quando se omitem alguns níveis *wavelet*, isto é, na ausência das componentes pertencentes aos níveis eliminados algumas frequências, originalmente não relevantes, "explodem" em amplitude, muitas das quais não presentes no sinal original. Em parte isso se deve ao efeito de espelhamento provocado pela superamostragem ($\uparrow 2$) do algoritmo de reconstrução.

Algumas *Wavelets* podem realçar melhor características de umas faixas de frequência do que outras, e portanto achar a *Wavelet* que melhor mostraria as linhas espectrais da série harmônica da nota é um trabalho de pesquisa experimental (empírico).

Por exemplo, analisando FFT's da primeira nota no nível 14 em análises com *Wavelets* diferentes, observa-se que as frequências de pico máximo (e de uma forma geral a composição espectral mostrada na FFT) variam de *Wavelet* para *Wavelet*. Veja na tabela abaixo os diferentes picos máximos locais verificados, em ordem decrescente de amplitude, para FFT's da primeira nota no nível 14 para as *Wavelets* de análise abaixo discriminadas:

Análise FFT da 1ª nota (Gb3) no nível 14 para diversas *Wavelets* de análise: picos em ordem decrescente de amplitude *

<i>Wavelet</i>	1º pico	2º pico	3º pico	4º pico	figura **
D4	8816,5	7368,4	6927	8061,5	1-n4-d4.gif
D16	6934,4	7303,8	6610,7	6247,3	1-n14-d16.gif
Beylkin	6948,5	6216,4	6326,7	6999,6	1-n14-be.gif
Coiflet1	7368,4	8816,5	6936,4	8097,8	1-n14-c1.gif
Coiflet5	6944,5	7314,6	6216,4	7005 (6594)	1-n14-c5.gif

* primeira nota tomada entre os pontos amostrais 1-30000 do sinal; faixa de frequências analisada: 5000-15000 Hz

** figuras dos gráficos das FFT's, em formato GIF, disponíveis no CD-ROM

Comparações sônicas ao nível 14

A *Wavelet* D4 fornece um som mais agudo e mais áspero que D16 neste nível.

Com Beylkin, o nível apresenta-se ainda mais suave que D16, soando como um instrumento de percussão semelhante a um chocalho.

Com Coiflet 1, o nível soa quase igual a D4, e visualmente as formas de onda são quase idênticas. O nível obtido com Coiflet 5 soa parecido com D16, mas ligeiramente mais suave na textura, entre D16 e Beylkin.

Nível 13: Pela audição do nível 13, verifica-se que este "isola" uma faixa de frequências bastante importantes na formação do "corpo do timbre" do violino. Objetos musicais menores e detalhes gestuais próprios da execução do instrumento ficam mais claros e perceptíveis, como os gestos de arraste do arco, e a percepção mais fina do colapso das notas, sugerindo-nos a visualização do movimento executado pela violinista. Essa percepção fica reforçada pela

presença neste nível de fenômenos processados em frequências mais altas, que, pela natureza física dos tons produzidos por instrumentos acústicos, colapsam mais rapidamente que aqueles associados a frequências mais baixas.

Ainda pela análise auditiva pode-se perceber neste nível uma diferença tonal entre as duas notas, e pela composição harmônica já se torna possível inferir sobre os *pitchs* dos tons originais (mesmo sendo os *pitchs* deste nível diferentes daqueles das notas originais).

Graficamente pode-se também mensurar os pontos de máximo dos acentos, e rastrear os movimentos de arco como a parada e inversão de sentido (note o pequeno "nódulo" logo antes da segunda nota). Os envelopes das notas são característicos de notas acentuadas, cabendo as porções de maior amplitude aos momentos em que o acento é aplicado. Vale enfatizar que a amplitude do sinal neste nível corresponde à intensidade pontual de todas as frequências que contribuem neste nível.

Nível 12: Uma análise visual do seu gráfico sugere uma compartimentação das notas em 3 porções, com amplitudes relativas e padrões vibratórios diferentes.

A forma de onda da primeira nota sugere 3 porções com propriedades e formatos diferentes: a primeira, incluindo o ataque e as altas amplitudes do acento, uma segunda região mais central, ainda revelando o acento, e finalmente uma porção final, onde a amplitude decai devido ao colapso forçado. A segmentação destas regiões é feita aqui sem muito rigor, baseada na informação visual e auditiva que temos do sinal.⁵

Pode-se perceber que na primeira porção da nota as vibrações são mais intensas, e na passagem para a porção central um leve estrangulamento delimita as duas. A porção final inicia-se logo após a porção central, e vai até o colapso. O contorno de cada porção que forma o envelope da nota apresenta o formato típico de ondas moduladas em amplitude, onde a amplitude varia ciclicamente em função de uma de baixa frequência, gerando regiões de cristas e vales. Mas a existência de uma modulação em amplitude não é o único indício de compartimentação nas notas, haja visto a presença dos acentos sugerirem padrões vibrantes diferentes onde ele é mais intenso e onde é menos intenso.

O número de compartimentações que podem existir na verdade é maior que três, dependendo do ponto de vista e dos seus objetivos da análise, podendo estabelecer vários critérios diferentes para segmentar-se a nota em regiões.

A segunda nota, pelo mesmo critério adotado acima, pode ser compartimentada em 3 porções, uma de ataque, outra central onde o padrão de acentos se caracteriza, e uma final, onde a nota reduz progressivamente de intensidade até seu colapso total.

As circunstâncias de interpretação envolvidas neste exemplo levam-nos a supor que os gestos do arco estejam por detrás da compartimentação da nota, os fatores pressão, velocidade e forma de arraste do arco (ponto de apoio e contato arco-corda) traduzindo-se em porções de intensidade relativas diferentes, regiões mais ricas em frequências altas e regiões onde existam combinações de diferentes frequências. A própria forma de interpretação das notas no violino sugere que pelo menos 3 regiões com propriedades discerníveis existam na execução de fraseados não ligados e com acentos, como nestes exemplo.

⁵ o uso de métodos analíticos e/ou estatísticos para determinar os contornos de cada região será recomendável nos casos em que se deseje quantificar suas bordas para efeito de síntese musical.

Primeiro a violinista tem de preparar o arco para tocar a nota, isto é, ele ou ela deve parar o arco, inverter o sentido do movimento, e iniciá-lo do zero.⁶ Sabendo-se que a nota é acentuada, a intérprete acelera e aumenta a pressão do arco progressivamente, marcando com os dedos a posição que estima ser a correta. O tom neste momento é regular e normal, sendo o tempo suficiente para que a intérprete avalie a correção da afinação e aplique o acento, com um aumento de pressão e velocidade, traduzidos num repentino aumento de atrito na corda e na intensidade do som. Neste momento várias frequências mais altas aparecem no som, e a combinação destas produz o formato visual da região onde se vê (e ouve-se) o acento. Num momento posterior, o gesto inicial vigoroso do acento sofre uma perceptível redução, e um período mais regular, mas não por isso pouco intenso, dá lugar a um padrão mais regular de amplitudes e composição de frequências. Finalmente, em observância ao tempo de duração permitido em partitura para aquela nota, a intérprete dá início a uma nova combinação de gestos de punho e arco, alterando a forma de excitação do arco, a pressão e velocidade de arraste, e mesmo (em alguns casos, dependendo da intérprete) a posição em que o arco apoia-se no comprimento vibrante da corda. A intérprete prepara o colapso, e neste momento podem surgir outras combinações de frequências que podem associar-se entre si como numa modulação em amplitude, até o colapso final.

O aparecimento de combinações de frequências específicas em certas circunstâncias na interpretação de instrumentos de corda, bem como a influência da posição de apoio do arco sobre a corda e outros fatores físicos envolvidos no arraste (velocidade, pressão, ângulo de ataque do arco) são aspectos estudados por diversos pesquisadores e amplamente abordados em literatura específica sobre acústica de instrumentos musicais.⁷

Nível 11: Cerca de 1,34 vezes mais energético que o nível 12. Observa-se um envelope com amplitude relativa mais intensa da segunda nota em relação à primeira; a segunda nota é mais grave, é de se esperar que tenha maior contribuição energética nas frequências cobertas neste nível.

Artefatos transientes do arco não aparecem tão bem caracterizados como nos níveis anteriores. A compartimentação das notas em regiões de propriedades e amplitudes médias diferentes é mais visível na primeira nota. A segunda nota mostra um envelope não compartimentado, com a região central mais intensa. A contribuição do acento na segunda nota fica bem presente neste nível.

Ouve-se cada nota neste nível com um *pitch* diferente: 1100 Hz (3^o harmônico de Gb3) para a primeira nota, e 1243 Hz (4^o harmônico de Eb3).

⁶ No retorno à corda, esta pode estar ainda vibrando, e o re-engate irá alterar o seu padrão vibratório; a extensão da alteração dependerá do instante do re-engate e da pressão, velocidade, e apoio do arco.

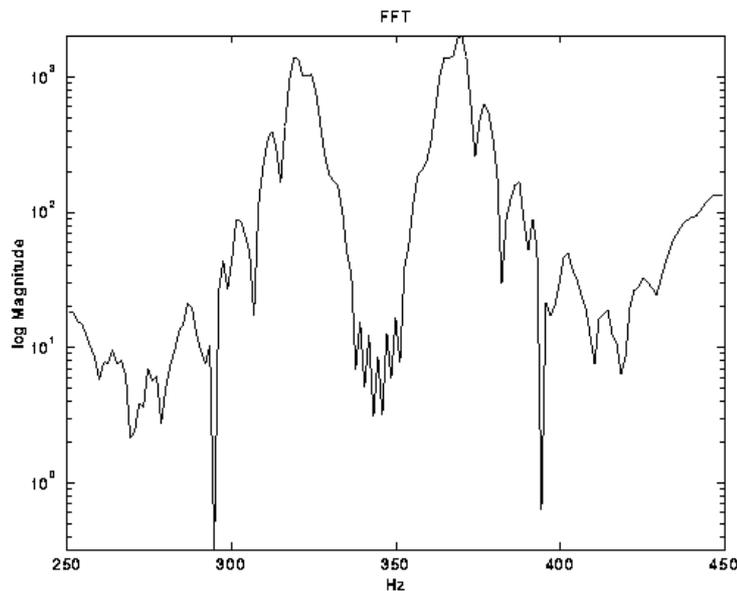
⁷ A mecânica da corda excitada pelo arco e o espectro associado a diferentes forma de arquear, bem como a resposta da caixa ressonante e de outras partes constitutivas do instrumento a essas excitações são tópicos cobertos por pesquisas em acústica musical, a fim de se compreender o funcionamento do violino e criar-se modelos paramétricos para o seu funcionamento, alguns destes baseados na teoria das linhas de transmissão elétricas (Schelleng, 1963), e nas importantes contribuições do físico alemão Hermann von Helmholtz (que elucidou os tipos de vibrações que distinguem uma corda arqueada de uma puxada, e explicou os fenômenos ressonantes em corpos volumétricos com duas aberturas, na sua famosa teoria das cavidades ressonantes). O leitor interessado pode referir-se a HUTCHINS (1975) como leitura específica para acústica de instrumentos da família dos violinos.

Nível 10: Cerca de 2,78 vezes mais energético que o nível 11. Este nível possui o envelope mais suave e de formato regular entre todos os níveis: as formas de onda para cada nota aderem perfeitamente ao modelo funcional adotado para as fases de uma nota: o ataque, a sobrelevação e decaimento (*release*), a sustentação e o decaimento (*decay*). Fica caracterizado nas duas notas o acento como a região de amplitudes maiores, especialmente na primeira nota, onde o acento foi mais energético. A figura mostra que as frequências contidas neste nível variam em amplitude proporcionalmente à intensidade do acento.

Outra estrutura que aparece mais regular por todas as duas notas é a de uma modulação em amplitude, que ocorre independentemente do acento.

Na Figura 13 abaixo mostra-se a FFT da primeira nota (Gb3) neste nível, traçada na faixa de 250-450 Hz onde se situam suas componentes mais importantes (o eixo vertical em escala logarítmica):

Figura 13 - FFT da primeira nota (Gb3) no nível 10 (traçada entre 250-450 Hz)



A componente mais intensa está em **370,1 Hz** (a **fundamental de Gb3**), havendo um pico local lateral à esquerda na mesma raia em torno de 364,7 Hz, cerca de 5,38 Hz da fundamental.

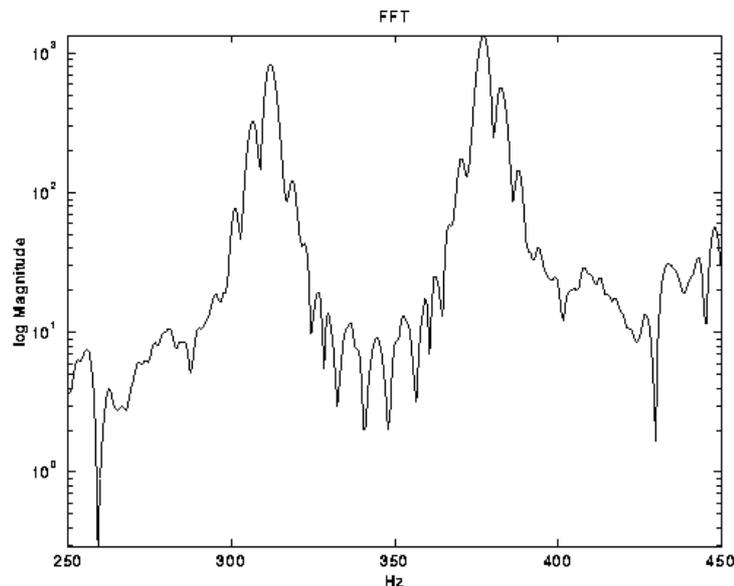
À esquerda localizamos outra raia, com pico em 318,96 Hz, distando 51,14 Hz da fundamental (esta é a frequência visível de oscilação no gráfico do nível 10 para a primeira nota). Um segundo pico à direita na mesma raia se encontra a também 5,38 Hz do primeiro, uma distância igual à já observada entre o primeiro e o segundo picos da segunda raia. Não menos coincidência, um terceiro pico local observado em cada uma das duas raias, se encontra a exatamente 6,73 Hz do pico principal. Este fato sugere que, a despeito das amplitudes relativas serem diferentes para cada raia, elas parecem se espelhar em torno de um eixo fictício em 344,53 Hz, exatamente $1/128 f_a$. O fenômeno equivale à ocorrência de uma espécie de modulação em amplitude onde a frequência portadora (em 344,53 Hz) está suprimida: a *modulação de anel*. A frequência de 344,53 Hz ("portadora") e 318,96 Hz ("banda lateral", o pico da primeira raia) não têm nenhum significado especial: não pertencem à série harmônica do tom Gb3, são respectivamente 100 e 82 vezes menores em amplitude que o pico máximo no espectro do sinal original, e nenhum corresponde à frequência de pico do filtro *Wavelet* deste

nível (apesar de estarem dentro da banda de passagem do mesmo), afastando definitivamente a hipótese de que o espelhamento das bandas seja uma modulação de envelope contida no sinal original. A causa do espelhamento acima verificado está no algoritmo de reconstrução *Wavelet*, particularmente nas fases de superamostragem, como já observado noutras análises.

É importante frisar que a escolha da *Wavelet* também influencia na determinação destas raiais espectrais e respectivas frequências de pico. Analisando-se a mesma nota no nível 10 gerado com a *Wavelet* de Beylkin, verifica-se que há uma certa semelhança no formato das bandas do espectro da nota, mas a primeira raia agora apresenta pico em torno de 368,75 Hz, cerca de 50 Hz deslocada da mesma raia/pico para a *Wavelet* D16 (em 318,96 Hz). O mesmo se repete analisando a segunda nota.

A Figura 14 abaixo mostra a FFT da segunda nota neste nível, traçada na faixa de 250-450 Hz onde estão as suas componentes mais intensas (o eixo vertical em escala logarítmica):

Figura 14 - FFT da segunda nota (Eb3) no nível 10 (traçada entre 250-450 Hz)



O primeiro pico à esquerda, em **312 Hz**, é a presença da **fundamental de Eb3**. O segundo pico, mais adiante em **377,5 Hz** é mais intenso que a fundamental, e dista cerca de 65 Hz da mesma. A audição do nível mostra claramente os dois *pitchs*, o segundo (377,5 Hz) sendo mais intenso e perceptível.

As duas raiais visíveis na figura espelham-se ao redor da frequência central de 344,86 Hz, com a diferença porém de que a segunda imagem apresenta maiores amplitudes relativas. Como na primeira nota, existência deste espelhamento equivale à ocorrência de uma modulação em amplitude, onde uma das frequências modulantes é a fundamental em 312 Hz, e a "portadora", em torno de 344,8 Hz, está suprimida.

A sugestão de uma frequência portadora em entretanto é discutível, porque esta frequência de 344,8 Hz não é representativa em nenhum aspecto: não pertence à série harmônica da nota Eb3, a sua amplitude no espectro da nota original é 334,6 vezes menor que a amplitude da fundamental, e além disso não é a frequência de pico no espectro da *Wavelet* Daubechies 16 no nível 10 (apesar de estar na faixa de cobertura desta *Wavelet*). O mesmo pode-se dizer da frequência mais intensa para a segunda nota neste nível (377,5 Hz): ela não pertence à série

harmônica de Eb3, apresenta amplitude cerca de 164 vezes menor que a amplitude da fundamental (isto é, sua participação no espectro da nota original é praticamente irrelevante), e também não é a frequência de pico do espectro da *Wavelet* do nível 10 (apesar de estar incluída na faixa de cobertura da *Wavelet*). Além disso a frequência de pico na análise *Wavelet* do nível vai variar segundo a *Wavelet* utilizada, mostrando essa escolha ser influente sobre o espectro do nível. Novamente, pois, atribui-se a causa deste imageamento de bandas ao algoritmo *Wavelet* de reconstrução.

O intervalo de frequência entre os dois picos é de 65,27 Hz, que é a frequência da ondulação visualizada na segunda nota do gráfico deste nível.

Nível 9: Tão energético quanto o nível anterior (somente 0,95 vezes menor). O formato do envelope neste nível é semelhante ao do nível 10 anterior: suave, identifica as regiões de ataque, decaimento, sustentação e colapso, e caracteriza as regiões onde há acento (a propósito, não localizado sobre o ataque mas pouco adiante).

Também como no nível 10, observamos "raias gêmeas" no espectro das notas, cujas distâncias em Hz entre os picos equivale às frequências de oscilação visíveis na figura do nível (assim como no nível anterior).

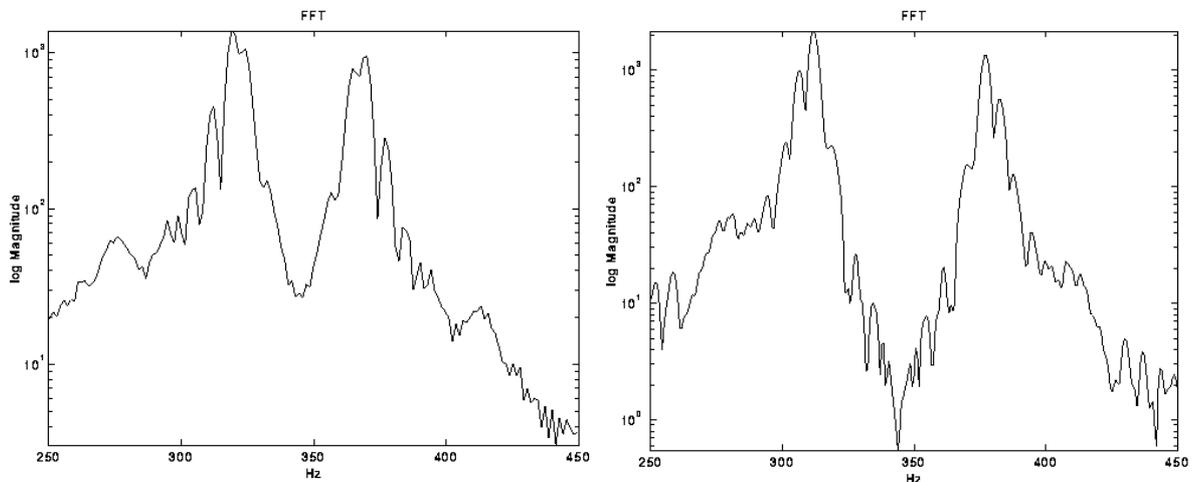
Note que na análise do espectro da primeira nota neste nível, ocorrem duas raias gêmeas, espelhadas em torno da mesma frequência observada no nível anterior. A primeira raia apresenta pico máximo em 318,96 Hz e a segunda em 370,1 Hz, os mesmos valores apresentados para o espectro da mesma nota no nível anterior 10. Entretanto a amplitude relativa da primeira raia (com pico em 318,96Hz) é mais intensa do que a segunda raia (com pico na fundamental), exatamente o contrário do que ocorria para o nível anterior, quando o pico maior do espectro estava na fundamental.

Analisando-se o espectro da segunda nota neste nível, e comparando-o com o mesmo espectro no nível anterior, também verifica-se a ocorrência de duas "bandas gêmeas" espelhadas em torno da mesma frequência e apresentando picos máximos e laterais nas mesmas frequências, a única diferença sendo que as amplitudes relativas estão invertidas: a primeira raia (com a fundamental em 312 Hz de Eb3) apresenta maiores amplitudes do que a segunda raia (com pico máximo local em 377,5 Hz).

O som da segunda nota no nível 9 fica dominado pelo *pitch* da fundamental (312 Hz), e apresenta, como era esperado, um brilho menor (menos frequências altas).

Conclui-se portanto que as frequências de pico máximo e locais para cada nota são as mesmas no nível 9 e 10, e que as raias dos espectros das notas estão espelhadas em torno das mesmas frequências, as bandas possuem formatos semelhantes, entretanto as amplitudes relativas estão invertidas de um nível a outro: a raia/banda mais intensa no nível 10 é a mesma menos intensa no nível 9. As FFT's da primeira e segunda notas neste nível (9) estão na figura abaixo:

Figura 15 - FFT's da 1ª (à esquerda) e 2ª (à direita) notas no nível 9 (entre 250-450 Hz)



Nível 8: Os níveis anteriores, que concentram a maior energia do sinal original, são aqueles que devem *a priori* ser os alvos de uma análise com *Wavelet packets*, para se determinar com mais resolução de frequência a composição frequencial e a localização de eventos musicais.

Este nível apresenta uma energia média cerca de 43630 vezes menor que o anterior.

O envelope do nível não guarda nenhuma semelhança com os envelopes observados anteriormente, não sendo possível localizar as etapas de cada nota, nem a pausa entre elas, nem os gestos de arco.

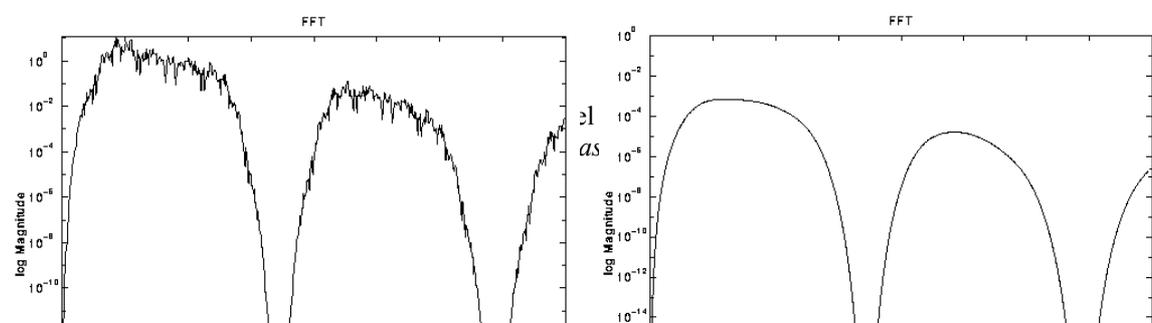
A amplitude no nível é razoavelmente regular por toda a sua extensão, incluindo aí os momento de pausa, que são preenchidos, um único destaque dado ao primeiro acento na primeira nota, que pode ser facilmente localizado neste nível, delimitando uma região onde a amplitude é mais intensa. O fato de que o acento da segunda nota não aparecer destacado se devem entre outros fatores não determinísticos, à diferença no conteúdo energético deste em relação ao acento da primeira nota (que foi realizado com um arco para baixo, mais vigoroso).

O som do nível lembra ruídos de motores de naves espaciais de filmes de ficção científica⁸, e seu espectro apresenta pico máximo de frequência em 100,93 Hz, um segundo pico em 71,93 Hz, e uma população de frequências mais densa entre 70-250 Hz, típica formação ruidosa.

Nível 7: O nível é 2,38 vezes mais energético que o anterior. Não é possível se localizar com precisão as notas presentes no nível, mas o formato de seu envelope traz um esboço dos envelopes das notas, podendo-se presumir, pelas regiões onde há estrangulamentos e onde as amplitudes são maiores, os pontos de ataques, centro da pausa, e as regiões de acentos. O colapso das notas é marcado com amplitudes relativas altas neste nível.

A influência do filtro *Wavelet* sobre a banda passante para este nível pode ser vista comparando-se a FFT do sinal no nível 7 com a FFT da *Wavelet* para o nível 7, isto é, a banda passante do nível *wavelet* ao nível 7. A figura abaixo mostra estas FFT's (plotadas para a faixa de 0-400 Hz, com a amplitude em escala logarítmica):

Figura 16 - FFT's do sinal no nível 7 (à esquerda) e da *Wavelet* D16 para o nível 7 (à direita)



Nível 6: Este nível é cerca de 2,7 vezes mais energético que o anterior. Praticamente inaudível. Seu envelope não permite localizar as notas e os eventos musicais associados com a resolução (temporal) dos níveis mais finos.

As informações isoladas neste nível retratam eventos de periodicidade entre 10 e 60 Hz. Pequenas modulações das variáveis físicas envolvidas na manipulação do arco podem estar descritas nesta faixa de frequências. Mostra-se sensível a certas componentes gestuais que ocorreram no centro da 1ª nota, e também sensível ao sinal de origem não instrumental (ruído de fundo ambiente), que se destaca na região de pausas (entre notas). As duas regiões do nível de maiores amplitudes relativas coincidem com o instante do acento na primeira nota, e a segunda região coincide com a região de colapso da primeira nota.

Aparentemente todos os dois gestos de arco contém ou são capazes de excitar frequências ressonantes na cavidade da caixa do violino em frequências na faixa de 10 e 60 Hz.

Nível 5: Excepcionalmente pouco energético, praticamente inaudível. Revela frequências que não constituem a série harmônica audível do som original, onde se concentra a esmagadora maioria da energia acústica.

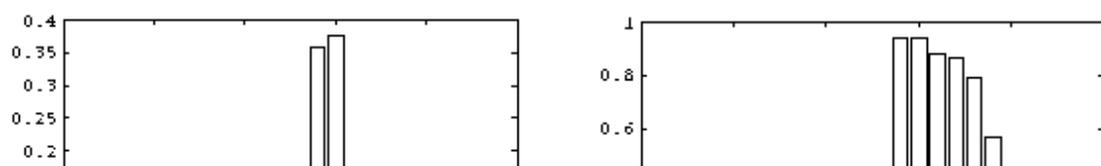
Os eventos registrados neste nível são provenientes de fenômenos subsônicos, na faixa de 5 a 30 Hz. A FFT do nível mostra um pico em 12,11 Hz, uma frequência subsônica, com um segundo pico local à esquerda aos 9,42 Hz.

O nível 5 pode funcionar (neste exemplo) como um contador de eventos acústicos importantes, como notas: existem duas estruturas visíveis no nível, isto é, dois períodos da sua frequência mais importante, coincidindo as regiões de amplitudes maiores (cristas) com as regiões onde existem as notas no sinal original (como pode-se facilmente observar superpondo-se os sinais original e deste nível).

Nível 4: Inaudível. Um "contador de eventos de baixíssima periodicidade", onde inexistem componentes da faixa acústica. Pode servir como um "contador de pausas", dada a coincidência entre os seu máximos e os momentos onde não existe a geração de vibrações sonoras acústicas audíveis.

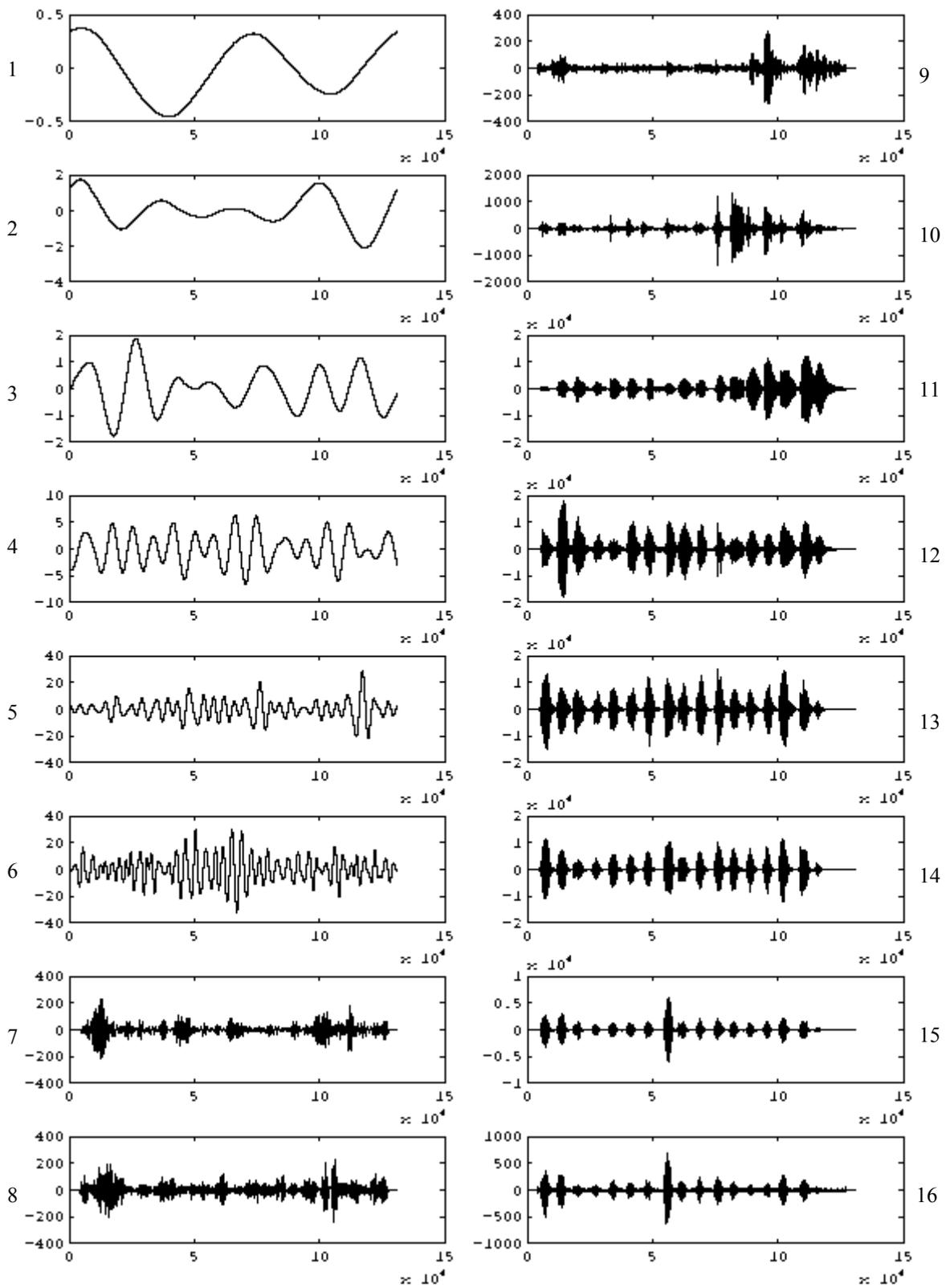
Nível 3: Seu formato não elucida nada sobre os eventos acústicos envolvidos. Pode entretanto ser útil para analisar partes orquestrais mais longas, localizando trechos de dinâmica variadas, ou trechos com progressões específicas onde haja a presença de certas classes de gestos musicais repetitivos, etc. É desprezível numa reconstrução, assim como o são os níveis anteriores 4,5 e 6.

Distribuição de energia:



Mais de 99% de toda a energia deste exemplo de fraseado acentuado não ligado se concentra dos níveis 9 ao 13, sendo o nível 10 o mais energético.

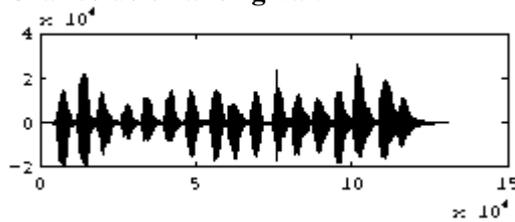
Violino: Fraseados, Não Ligados, com *Spiccato*



Evento: Sequência de *staccatos* do tipo *spiccato* num fraseado melódico não ligado. O trecho a ser analisado abaixo foi extraído da Sinfonia nº 1 de Beethoven, e na parte transcrita em Piston (1955) consta que a dinâmica inicial da frase é *piano*, e a velocidade *Andante con moto*. Abaixo a partitura notada do trecho em questão (adaptada da figura 10, à p.15 de Piston (1955)):



Gráfico do sinal original:



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

Análise dos níveis:

O nível 16, o mais fino da AWMR neste exemplo, encampa uma banda larga, e retém a maior parte da energia do ruído de arco. Tal ruído, fruto do atrito gerado no arraste do arco sobre as cordas metálicas do violino, aparece neste nível separado das frequências mais baixas desenvolvidas na cavidade ressonante do violino (e demais partes vibrantes). A despeito de quão vigorosa seja a vibração das ondas longitudinais nas cordas do violino, o som produzido por elas somente seria quase inaudível, dada a baixa irradiação de energia acústica envolvida. Isto pode ser compreendido quando se verifica que a superfície das cordas é muito pequena para comunicar um apreciável deslocamento de ar.⁹

O atrito gerado no arraste apresenta uma espectro portanto de baixa densidade energética, mas de banda larga, aproximando-se do que se pode designar pela nomenclatura genérica de "ruído".

As componentes presentes neste nível produzem um envelope cujo formato tende a ser mais retangular comparado ao envelope das notas no som original, com as rampas de subidas e descidas mais íngremes, e durando intervalos de tempo ligeiramente inferiores aos verificados nas notas originais.

Verifica-se a existência de um ruído de fundo de regularidade e amplitude constantes, presente durante toda a duração do exemplo. A análise do material de fundo presente na gravação, na ausência de qualquer evento musical ou ruído relevante captado pelo microfone, apresentou o mesmo ruído de fundo persistente, indicando que ele não faz parte dos sons gerados na interpretação do instrumento.

Essas características observadas neste nível mais fino levam à apreciação de 3 aspectos relevantes tratados a seguir:

⁹ Hutchins, C.M. *The physics of violins*, Nov.62, in: HUTCHINS (1975).

PRIMEIRO ASPECTO

Em alguma etapa entre a gravação e a produção das amostras adicionou ruído de fundo ao som. Ruídos podem ter sido produzidos devido a várias causas, entre as quais citamos as mais prováveis:

1. **ruído ambiente do estúdio.** O estúdio pode ter apresentado baixo isolamento acústico em alta e baixa frequências.
2. **sensibilidade do microfone.** O microfone utilizado foi um Neuman direcional, sensível a movimentos sutis e capaz de captar uma variedade de eventos acústicos e pequenos ruídos causados pela respiração, a própria manipulação do instrumento, manipulação de papéis, arraste de pés, etc.
3. **qualidade do DAT utilizado na gravação.** O DAT TEAC modelo DA-P20 é um DAT semi-profissional, e pode-se notar em várias gravações por ele feitas a presença de ruídos, seja devido à qualidade e precisão do seu conversor analógico-digital, do circuito de codificação PAM (*Pulse Amplitude Modulation*) ou de eventuais pré-filtros analógicos.
4. **Transferência de meio.** A material sonoro gravado originalmente em estéreo, em fita DAT, com taxa de amostragem de 48 KHz e largura de resolução de 16 bits/amostra, foi lido por outro equipamento (DAT Silicon Graphics) e transferido para disco rígido de computador, armazenado neste meio como arquivos de som estéreo, no formato proprietário da Apple (AIFF). A codificação em AIFF não altera o conteúdo original de informação, somente adiciona um cabeçalho (*header*) específico às palavras digitais tal que sejam reconhecidas pelo sistema operacional como arquivos de som (AIF).
5. **Conversão de taxa de amostragem.** A gravação original foi feita com taxa de amostragem de 48000 amostras/segundo (48 KHz), dois canais (estéreo). Previamente à realização das análises as amostras sofreram uma mudança de taxa de amostragem, sendo convertidas para 44100 amostras/segundo (44,1 KHz, padrão CD), monoaural (um canal). Não foi realizada, entretanto, nenhuma conversão de largura de resolução dinâmica, originalmente em 16 bits e mantida inalterada.
6. **normalização das amostras.** Os arquivos AIFF convertidos foram editados, e fragmentados em trechos menores, contendo somente os eventos analisáveis desejados. Seguindo a edição, foram normalizados em amplitude, num processo que requer a localização da maior amplitude no intervalo e a amplificação do trecho proporcionalmente à maior amplitude. Apesar de ser uma etapa necessária para se preparar as amostras para a análise, o processo pode realçar ruídos gerados nas fases anteriores, amplificando-os.

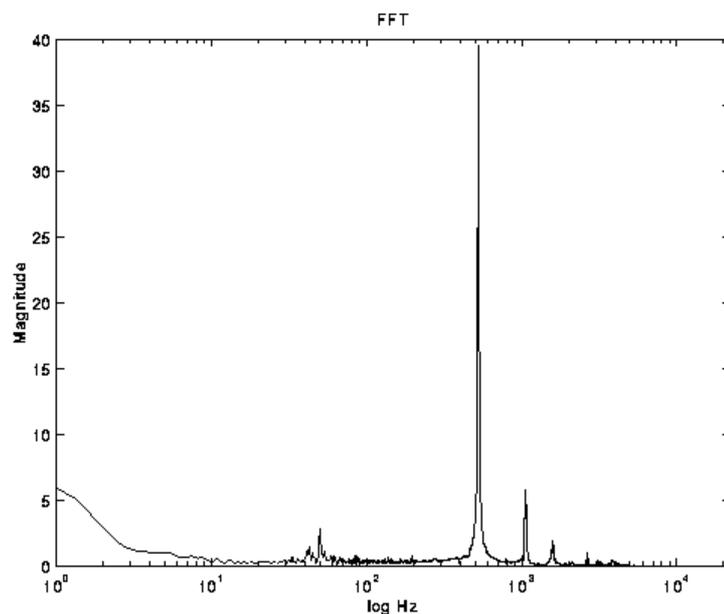
Dentre estas, a transferência de meio digital, conversão de taxa de amostragem e normalização das amostras são as etapas com maior potencial para geração de ruído. Também o próprio ruído de fundo acrescido dos sons oriundos de fontes que não os instrumentos musicais, mas que exibam intensidades próximas ao nível de amplitude dos sons musicais quando estes estão baixos (como a respiração, vozes, arraste dos pés, manipulação de papéis, etc.) constituem uma classe de ruídos que se pronunciam em diversas amostras gravadas.

A análise do nível *wavelet* mais fino (16) revela que há contribuições no final do colapso, mas que findo completamente as vibrações na corda ainda resta um ruído persistente de fundo com amplitude média constante.

Acompanhe a análise de um trecho ruidoso cobrindo uma região de pausa. Este trecho começa com o final do colapso da última nota executada com *spiccato*, e encampa um intervalo de pausa (silêncio) do instrumento.

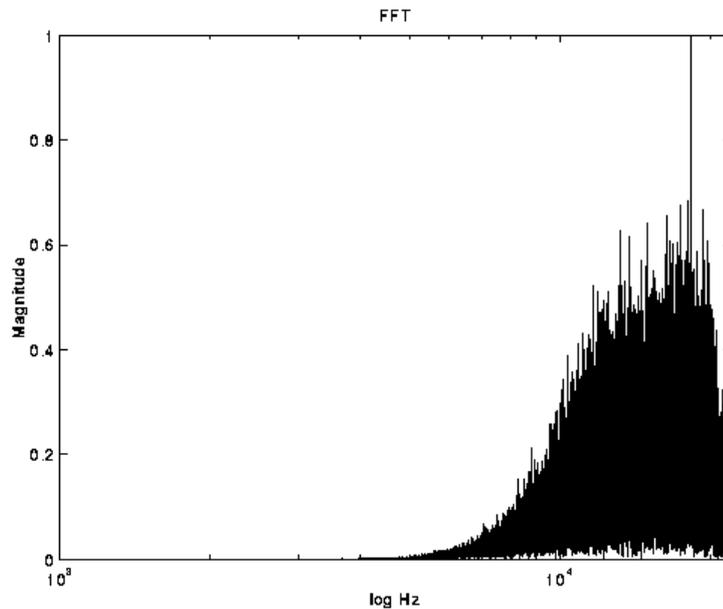
A análise FFT desta região revela a contribuição da sequência final do colapso da nota, mostrando sua série harmônica, com sua raia mais intensa em 525,54 Hz, como também a contribuição do ruído de fundo, menos intenso, e que se espalha pelas altas e baixas frequências, como ilustrado no gráfico abaixo (com o eixo horizontal das frequências em escala logarítmica):

Figura 17 - FFT da região de ruído extraída a partir do final da última nota do exemplo



A análise FFT do nível *wavelet* mais fino deste trecho ruidoso após a última nota do exemplo (analisado com Daubechies 16) revela um espectro denso nas altas frequências, caracterizando o caráter de ruído. A figura abaixo mostra esta FFT, com as frequências em escala logarítmica e a magnitude normalizada em 1 (isso porque as amplitudes absolutas dos componentes são baixas demais para serem visualizadas num gráfico com magnitude absoluta):

Figura 18 - FFT do nível *wavelet* mais fino para o trecho ruidoso após a sequência de *spiccato*s



A filtragem *Wavelet* extraiu a contribuição da sequência final do colapso do tom de violino, dando destaque às frequências mais altas presentes no sinal. A escolha da *Wavelet* influi significativamente no espectro do nível mais fino, uma vez que alterado o filtro *Wavelet* para este nível também fica alterada a banda passante, e as frequências realçadas serão outras.

Note que a faixa de frequências destacada na Figura 18 acima não está relevante na FFT da região ruidosa (figura anterior), e que o envelope frequencial sugerido pela figura segue o contorno da banda passante para a *Wavelet* D16 mais fina. Vê-se, portanto, pela curva ascendente no espectro em direção às altas frequências, que o filtro *Wavelet* realça as altas frequências neste nível mais fino.

Isto pode representar a primeira vista uma desvantagem de se usar filtros *Wavelet*, na medida que a informação espectral apresentada por um nível pode dar a ilusão de que certa faixa de frequências é muito significativa (ou não) para a composição do sinal. Todavia, o usuário do filtro *Wavelet* deve tomar conhecimento prévio destas propriedades dos filtros *Wavelet*, a fim de aproveitar o seu lado positivo, qual seja o de revelar os detalhes do sinal numa banda específica de frequências.

SEGUNDO ASPECTO

Durante o *spiccato* o tempo em que o arco se ausenta da corda é o suficiente para permitir que as vibrações na corda se extingam completamente havendo portanto curtos intervalos de silêncio entre elas (pelo menos é o que se passa para as componentes de alta frequência envolvidas no tom e na interpretação do gesto). Segundo indica a análise do nível mais fino 16, o arco realmente se ausenta da corda periodicamente durante o *spiccato*, com ciclo ativo de pausa (salto do arco) de aproximadamente 53% (para este exemplo), sugerindo que o gestual envolvido na movimentação do arco possui também ciclo ativo próximo de 50%.¹⁰

¹⁰ Valores de ciclo ativo baseados unicamente numa estimativa gráfica sobre a forma de onda.

TERCEIRO ASPECTO

O formato ligeiramente mais quadrado das notas no *spiccato* neste nível deve-se ao fato de que as frequências altas iniciam e morrem mais rápido na ausência da excitação (arraste do arco).

Outro fato digno de nota no gráfico deste nível é a intensidade diferenciada da 8ª nota da sequência em relação às outras notas. A observação do gráfico do som original não revela indícios visuais de que esta nota seja mais intensa ou destacada de alguma forma das outras. A audição do som original entretanto revela haver um destaque nesta nota, pontuando com expressividade o tema desta passagem de Beethoven. A intérprete na execução desta nota empregou um arco para baixo no retorno do salto entre a sétima e a oitava nota, conferindo-lhe mais vigor. Tal gesto transferiu uma maior energia à corda intensificando sua resposta nas altas frequências, o que fica caracterizado neste nível.

Portanto, ataques pontuais e energéticos do arco às cordas intensificam a energia dos parciais mais altos.

Nível 15: Também revela a intensidade relativa maior da 8ª nota nesta faixa de frequências em relação às demais. O ruído de fundo, persistente, se encontra presente com uma amplitude bem inferior ao nível anterior.

Ampliando-se o gráfico da figura do nível, para examinar com mais detalhes as regiões entre notas, verifica-se que o "intervalo morto" entre elas (descontando-se o ruído residual de fundo) é menor que no nível anterior, quando o ciclo ativo estimado para a pausa era de aproximadamente 53%. Nesta faixa de frequências as vibrações persistem por mais tempo sem a excitação, e alongam a sequência final do colapso das notas. O ciclo ativo da pausa medido para neste exemplo (utilizando a mesma metodologia de medida adotada no nível anterior) varia entre 15 e 20% do período total do gesto musical. A estimativa gráfica empregada todavia serve somente como guia, admitindo uma margem de erro que pode ser significativa quando não se extrai o ruído de fundo que contamina o sinal.

Já é possível pela audição deste nível inferir sobre os tons reais das notas originais.

Nível 14: Também apresenta um ciclo ativo menor que o nível 16, e o colapso das notas é mais suave, menos quadrado. A intensidade relativa das notas é também mais regular que nos níveis anteriores, mais sensíveis a sutis variações nos parâmetros físicos do arraste (velocidade, pressão, ponto e ângulo de contato arco-corda).

A audição fornece indícios de um ruído de fundo persistente, levando-nos a confirmar realmente sua presença nesta gravação específica.

Nível 12: Apresenta um som límpido e claro, de textura menos ruidosa. Dois fatores concorrem para esta percepção: o nível é o mais energético, destacando melhor a contribuição harmônica dos tons, e parte de sua cobertura em frequência coincide com a região de melhor audibilidade humana: entre 1000 e 3000 Hz.¹¹

A segunda nota é a mais intensa, e a oitava nota, mais intensa nos níveis anteriores, agora já não se destaca. As notas mais baixas em tom aparecem com menor amplitude relativa às outras.

¹¹ KOBAC apud HUNGRIA (1995). *Fisiologia da Audição* (cap. 30).

Nível 11: A amplitude relativa das notas aumenta da metade até o fim do trecho musical, estando as maiores magnitudes nas últimas notas.

Nível 10: Apresenta também as maiores amplitudes relativas na segunda metade do exemplo, sendo a 11^a e 12^a notas as mais intensas. A audição do nível revela a percepção de um batimento percussivo para cada nota. A localização e delimitação do envelope de cada nota neste nível não é possível: a maioria das notas se fundem, ficando a descrição do envelope neste nível para *spiccato*s mais difícil de ser definida ou descrita para efeito de síntese, exatamente pela sua irregularidade na forma.

Nível 9: Revela um ruído grave, disperso por todo o exemplo, que não guarda nenhum formato semelhante com os envelopes observados nos níveis anteriores, o que impede qualquer delimitação (localização) das notas. Sugere ainda que sua origem não esteja correlacionada com os eventos musicais ou sons produzidos pelo instrumento.

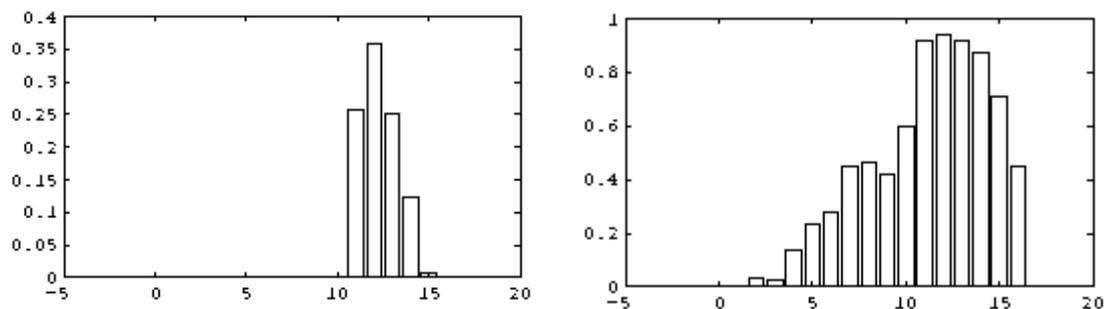
Amplitudes mais intensas aparecem nas notas da segunda metade, aproximadamente as mesmas notas intensas do nível anterior.

Nível 8: Sem localização de eventos. Pontos altos no início (sobre a segunda (lá4) e terceira (sol4) notas, principalmente) e no final, para a 15^a nota (ré4).

Nível 7: Primeiras notas mais intensas. Mais regular e bem comportadas as oscilações neste nível, com envelopes mais suaves. Presença de uma frequência modulante.

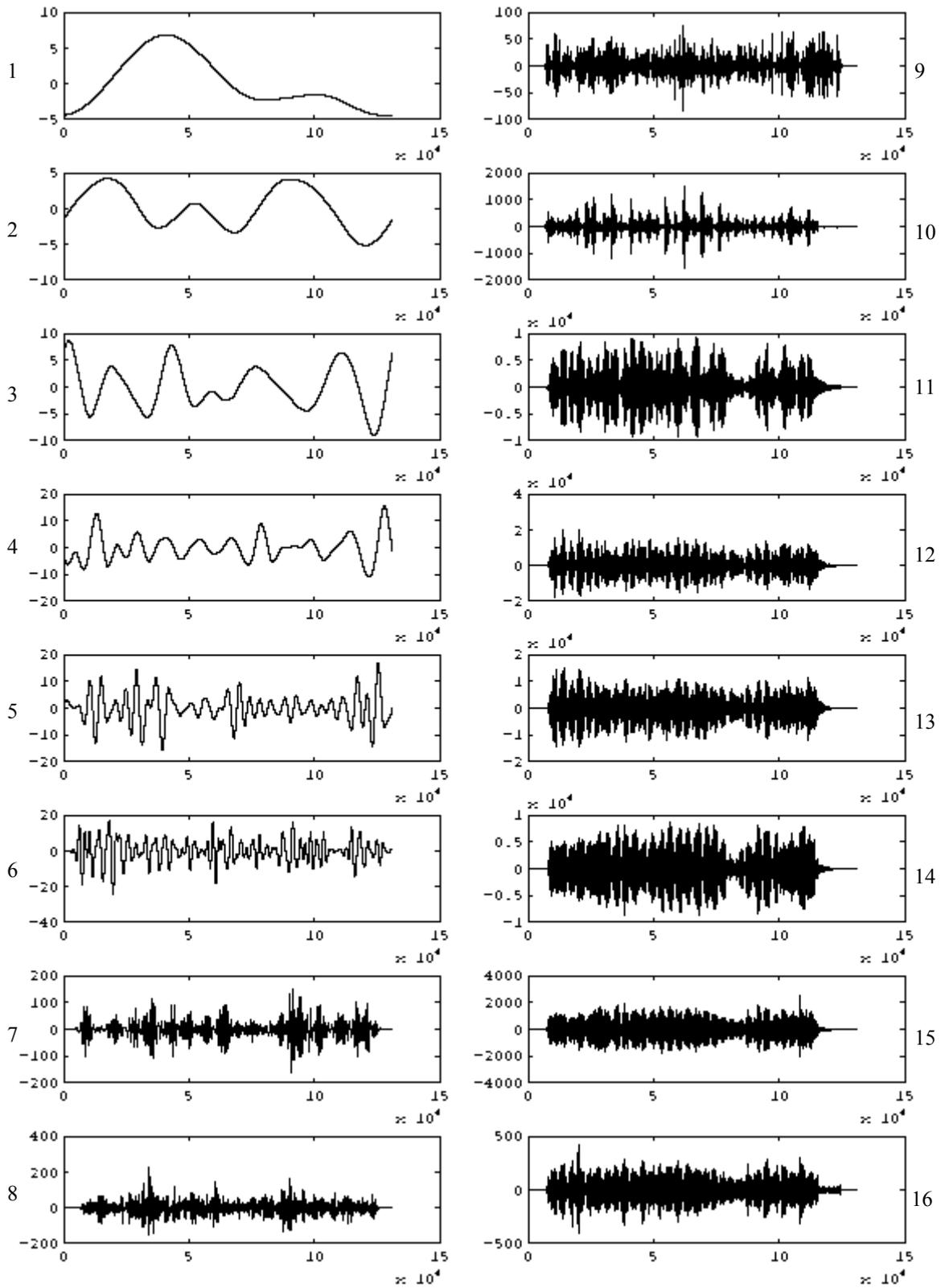
Nível 6: Envelope irregular. A FFT do nível fornece a frequência predominante de 11,77 Hz no nível, já subsônica. Nível inaudível.

Distribuição de energia:



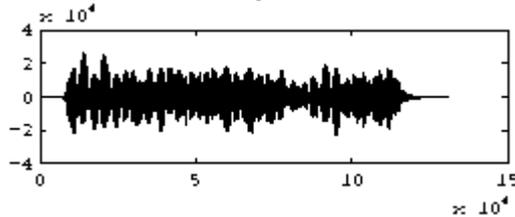
Os gráficos acima mostram a distribuição de energia pelos níveis da análise *Wavelet*, primeiro em escala linear de intensidade, e segundo o logaritmo da energia de cada nível pelo logaritmo da energia total. Mais de 99 % de toda a energia do exemplo concentra-se nos níveis de 11 a 15, sendo o nível 12 o mais energético.

Violino: Fraseado, Não ligado, Trêmulo de arco - nota mi4 (E4) na corda lá



Evento: trêmulo de arco, repetição acelerada da nota mi4 (E4) na corda lá. A interpretação deste efeito musical num fraseado consiste no emprego do arco em movimentos curtos e rapidamente alternando-se o sentido, de forma a repetir a nota tocada várias vezes num determinado espaço de tempo.

Gráfico do sinal original:



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

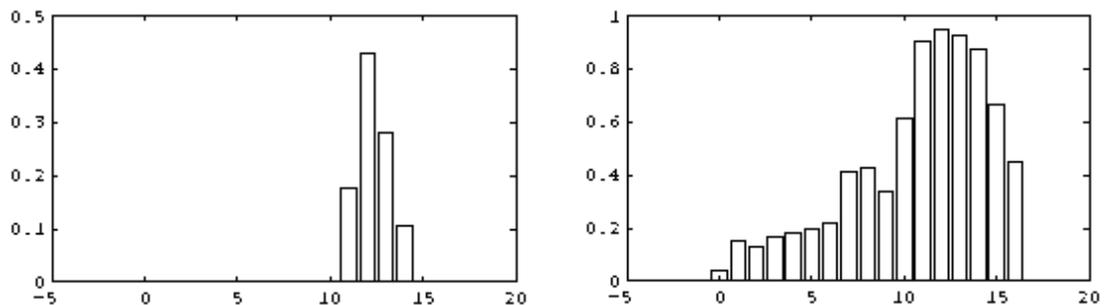
Análise dos níveis:

O nível 16, o mais fino, mostra o ruído característico do arco; apresenta uma distribuição de densidade de frequência significativa por uma larga banda.

O nível 10 permite a localização dos pontos de ataque à corda durante o chaveamento rápido e periódico de sentido do arco, indicando com transientes mais intensos os gestos mais incisivos de re-engate.

O nível 9 apresenta um ruído de fundo grave e contínuo, semelhante àquele observado (ouvido) no nível 10 da análise anterior (sustentação com trinado).

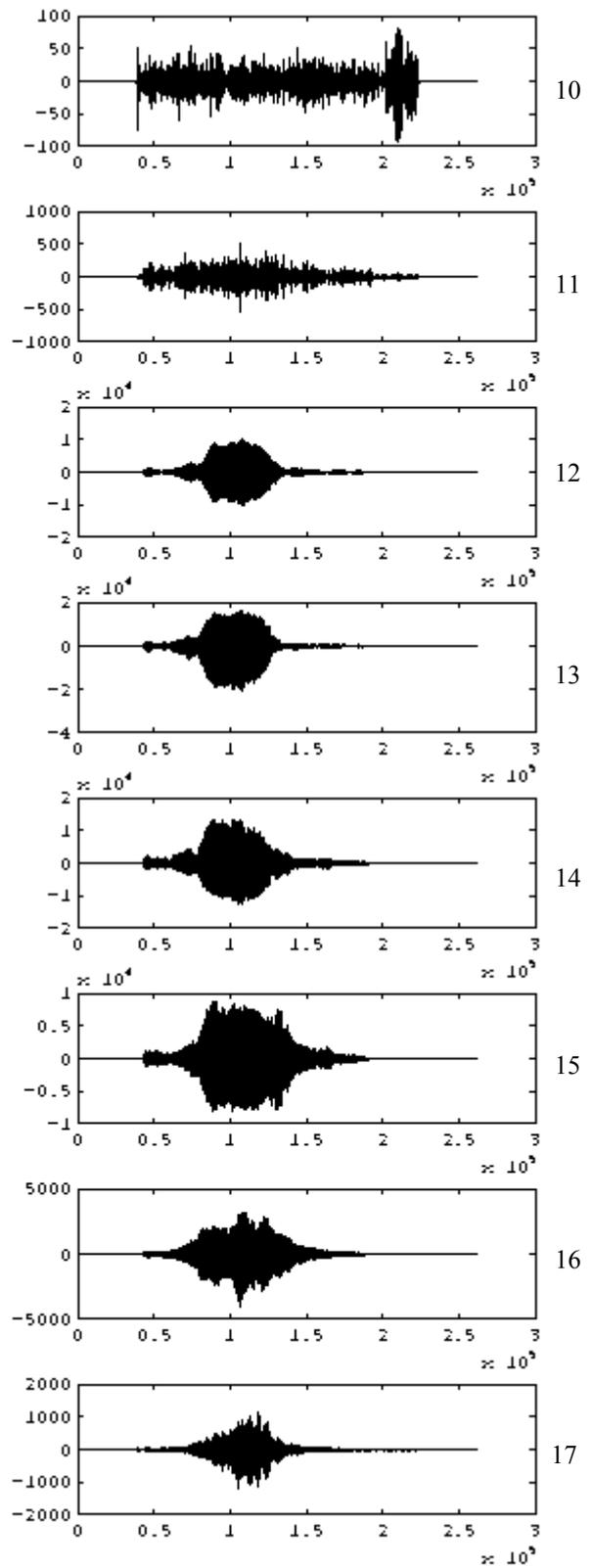
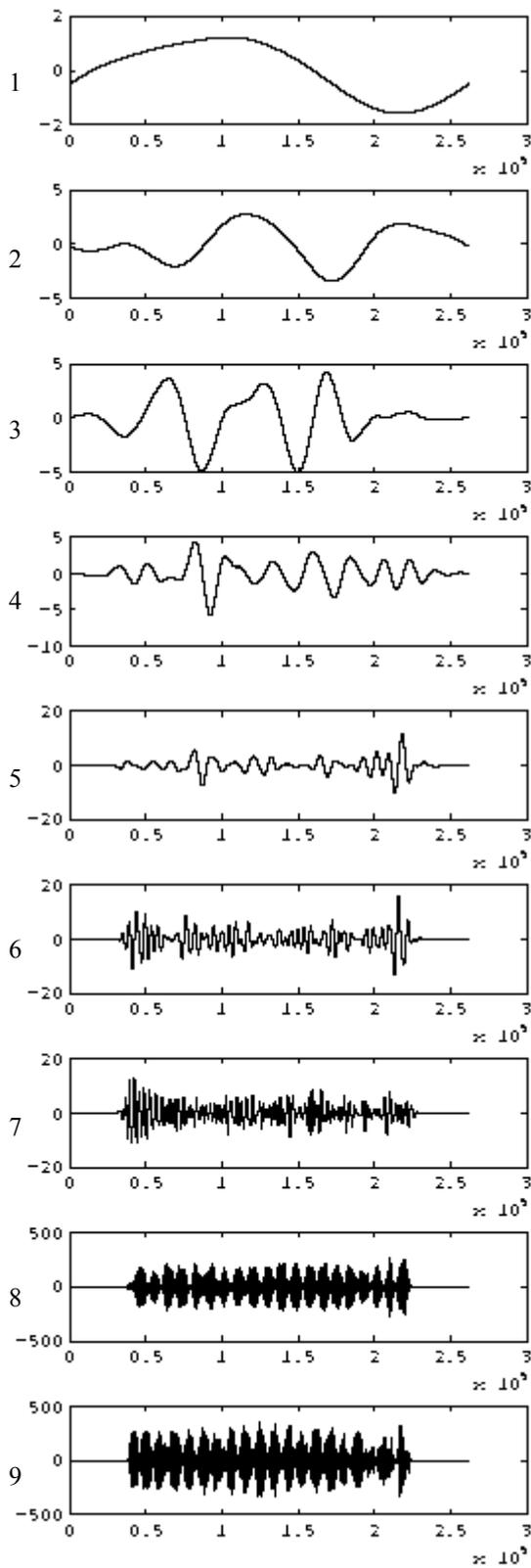
Distribuição de energia:



Mais de 99% de toda a energia do exemplo concentra-se nos níveis 11 a 14. O nível 12 é o mais energético. Os níveis mais baixos, especialmente os inferiores ao nível 6 (inclusive) não impactam a qualidade da reconstrução caso sejam desprezados nesta, exatamente por não contribuírem significativamente em energia.

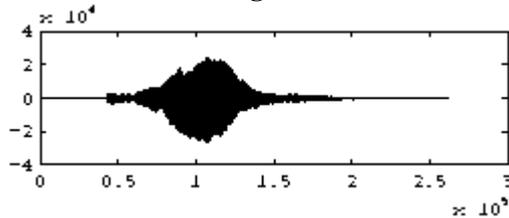
O três primeiros níveis mais finos, concentrando aproximadamente 40% de toda a energia, mostram que na interpretação do trêmulo os gestos musicais associados produzem intensa quantidade de ruído de arco e impõem constante variação às vibrações excitadas na corda. Toda essa manipulação reflete-se numa maior turbulência no sinal tocado e potencializam a qualidade e quantidade percebida de ruído.

Violino: Sustentação com Dinâmica



Evento: Sustentação da nota mi4 (E4) na corda lá, com dinâmica variando de *pianíssimo* a *fortíssimo* e retornando ao *pianíssimo*.¹²

Gráfico do sinal original:



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

Análise dos níveis:

Nível 17: Populado pelo ruído do arco, verifica-se que o nível de ruído de arco cresce proporcionalmente com a dinâmica. Este é um fenômeno esperado, uma vez que a dinâmica (ou intensidade) é determinada pela pressão e velocidade de arraste do arco sobre a corda, parâmetros que afetam a amplitude dos microdeslocamentos da corda (amplitude da vibração).

Nível 16: Semelhante ao formato do nível anterior mais fino. Diferenças em amplitude em certos pontos da sustentação evidenciam a irregularidade na excitação de algumas frequências mais altas quando a dinâmica é maior.

Nível 15: A conformação do envelope deste nível, onde se esboçam 3 porções com intensidades médias diferentes, sugere que em dinâmicas mais fortes (intensas) há uma maior contribuição de harmônicos mais altos, não tão relevantes em dinâmicas fracas. As 3 porções são: uma **central**, onde a dinâmica é mais forte, logo possui uma intensidade média maior (amplitudes na faixa de 5000 a 10000 unidades); uma **segunda** porção, na vizinhança de ambos os lados da porção central (à esquerda, na direção do início, e à direita, em direção ao final), cuja intensidade é menor (não atingindo 4000 unidades de amplitude); e finalmente uma **porção de menor amplitude** (inferior a 2000 unidades), nas vizinhanças do início e do fim. Em verdade, a segunda porção é mais indefinida, estando na transição entre a mais intensa e a mais fraca. Note também que o envelope do som original possui também estes formatos.

Nível 14: Também evidencia as três porções descritas acima.

Nível 13: Neste nível destaca-se a existência de 2 regiões de amplitude média diferentes: a central, onde a dinâmica é forte, e outra mais fraca na periferia de ambos os lados da região central. Na parte inicial, é possível notar o nódulo do ataque inicial, e uma zona de transição em dinâmica crescente, e em seguida a região mais intensa (central). Após a região central, na direção do colapso, há uma zona de transição mais curta, com uma rampa de descida mais acentuada na direção do colapso final.

Nível 12: Possui também regiões distintas de amplitude média, mas o desenho do envelope revela detalhes de pequenas modulações de amplitude nesta faixa de frequências.

A segmentação deste exemplo em regiões de amplitudes médias distintas pode ser verificada em vários níveis, denotando que harmônicos presentes em todas estas níveis variam de intensidade proporcionalmente à dinâmica da nota. O nível 12, todavia, parece destacar mais duas regiões com amplitudes distintas: uma **central**, com amplitudes de 5000-10000 unidades, e uma **lateral** com amplitudes variando de 0-2500, com valor mediano em torno de 79 entre os pontos

¹² Dinâmica, em Música, refere-se à intensidade ou volume sonoro da nota ou passagem musical.

140000-262000, e baixando para aproximadamente 36 entre os pontos 160000-262000, valores bem mais baixos que o valor médio calculado na região central, de 2715 unidades (entre os pontos 80000-135000, cerca de 34 vezes mais intenso que 79, atingindo 3313 unidades numa faixa mais central entre 90000-125000).

Nível 11: Este é o primeiro nível cujo envelope não guarda correlação com o formato do envelope original ou dos níveis anteriores mais finos.

Algumas reconstruções seletivas foram realizadas:

1. feita zerando-se os coeficientes do nível 11. O resultado sonoro ficou praticamente idêntico ao som original.
2. feita fazendo os coeficientes do nível 11 três vezes mais intensos. O resultado sonoro foi um reforço aos graves ressonantes na caixa do violino.

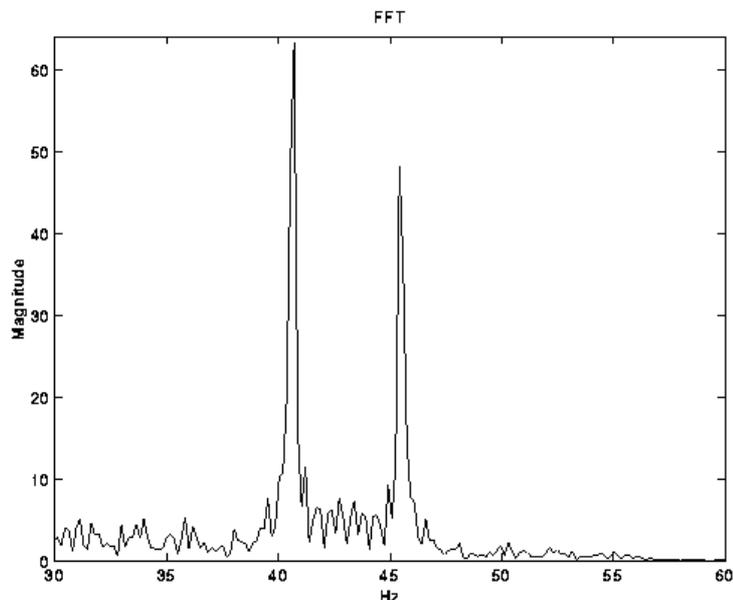
Nível 10: Também sem guardar qualquer semelhança com o envelope original ou dos níveis anteriores, este nível apresenta som muito grave, com um leve aumento de amplitude no final, quando o arco está variando a pressão e/ou velocidade de contato e abandona a corda. Este nível não revela indícios de que a dinâmica está variando: apresenta baixa correlação com a variação dinâmica. Não parece revelar portanto componentes frequenciais dependentes da pressão do arco (da pressão de contato). No entanto é sensível a alterações na velocidade do arco e ao seu abandono à corda.

Nível 9: Formato do envelope aproximadamente quadrilátero, imune a variações dinâmicas, mas apresentando uma estrutura periódica: uma modulação em amplitude visível graficamente e audível.

A moduladora apresenta uma frequência em torno de 5 Hz, o que pode ser verificado pela simples diferença ou distância (em frequência) entre as linhas (picos) espectrais de sua FFT (não mostrada).

O **nível 8** também possui uma forma de onda periódica em ~5Hz. A Figura 19 mostra o espectro deste nível, a diferença entre os picos sendo a modulação de 5 Hz.

Figura 19 - FFT do nível 8



Nível 7: Inaudível. Possui uma pronúncia no início, na parte do ataque. O espectro distribui-se entre 5 e 30Hz, com picos em frequências subsônicas. Pode ser utilizado para avaliar aspectos gestuais, que implicam em movimentos ou transientes de periodicidade subsônica.

Nível 6: Menores intensidades no centro, onde a dinâmica é mais forte. Há estruturas periódicas cujas amplitudes são maiores nas extremidades do sinal. O pico mais alto ocorre no final do som, semelhantemente ao observado no nível 10.

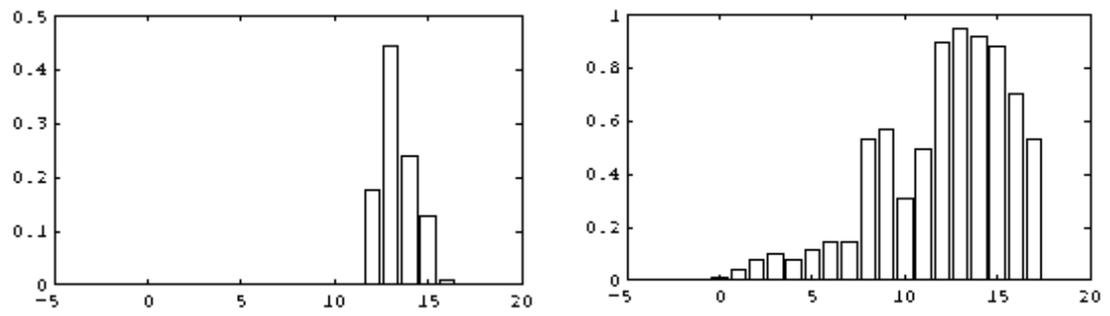
As menores amplitudes coincidem com as regiões onde a produção energética do tom (série harmônica) é maior. Isto mostra que a contribuição deste nível é atenuada quando a dinâmica do tom é maior, sugerindo que dele participam ruídos que se codificam nos bits mais baixos das amostras digitais, e são normalmente "encobertos" quando a dinâmica cresce¹³. A reconstrução desprezando este nível não provoca diferenças perceptíveis do som comparado ao original.

Reconstruções seletivas adicionais realizadas:

1. níveis 17 e de 0 a 6 (inclusive) zerados. Reconstrução quase idêntica ao som original, exceto por uma sutil diferença na granularidade das frequências mais altas. Omitindo-se todos estes níveis, em especial o 17, a reconstrução parece ser um pouco mais intensa na região central do que o som original.
2. níveis 17 e de 0 a 10 (inclusive) zerados. Reconstrução quase idêntica à original.
3. níveis de 0 a 5 zerados, 16 e 17 zerados, 12 zerado, o resto intacto. Aparecem duas vozes, uma próxima da nota do violino original, e outra, mais alta em frequência, numa faixa onde não percebia-se antes componentes tão altas. Aparentemente, eliminando o nível 12 provoca uma ruptura na fusão de alguns harmônicos mais altos, que passam a ser percebidos como uma outra voz muito aguda. Ouvindo-se a nota original em seguida, é possível perceber a existência da frequência muito aguda, mas ela fica destacada quando se reconstrói como acima.
4. idem anterior, mas com o nível 12 intacto. Aparentemente o nível 12 não provoca sozinho a desintegração perceptual que leva ao destaque da componente muito aguda, pois ela fica destacada aqui também, embora mais fraca.

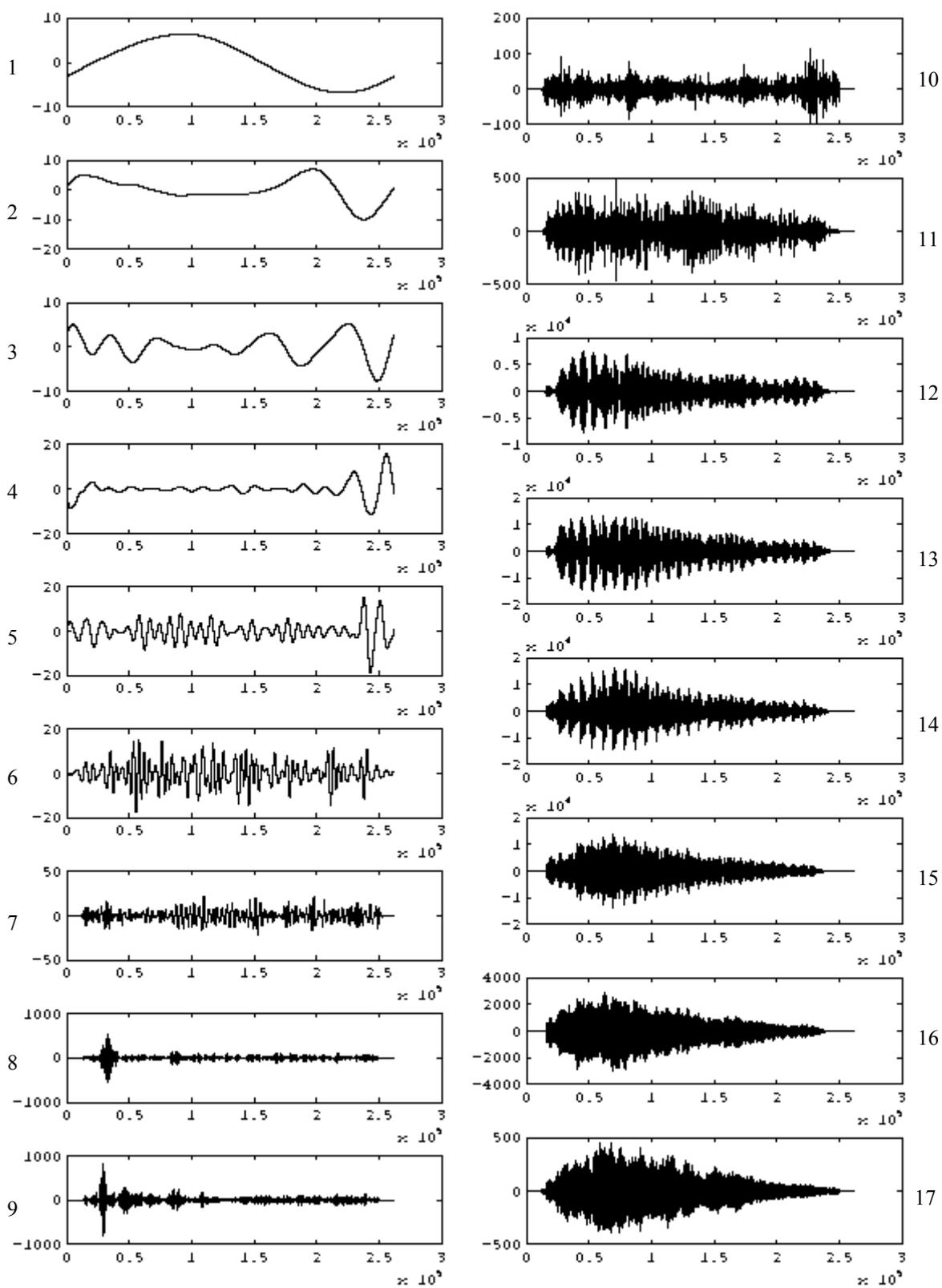
Um espectrograma (tipo *waterfall*) da 3ª reconstrução acima é mais "claro" e "despoluído" que o da figura original. Das duas frequências mais intensas na nota, a frequência mais baixa (em torno de 640 Hz) fica mais fraca na 3ª reconstrução comparada com a original. Isto porque níveis mais baixos foram eliminados. Também muito visível a eliminação dos níveis mais altos, o 16 e 17, o que faz o espectrograma não ir adiante dos 10000Hz: o gráfico da *waterfall* está desprovido de componentes nesta faixa. Quanto à frequência mais aguda, ela se destaca mais no som original do que na 3ª reconstrução: há mais componentes em torno de 7000 Hz no som original do que nesta reconstrução.

¹³ Essa constatação sugere que *Wavelets* poderiam ser utilizadas em sistemas redutores de ruídos de quantização e modeladores de ruído (sistemas de *dithering*).

Distribuição de energia:

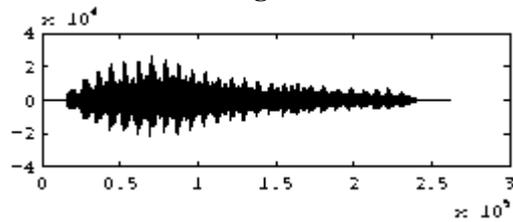
Cerca de 99% de toda a energia do sinal musical original concentra-se nos níveis de 12 a 16, sendo o nível 13 o mais energético.

Violino: Sustentação com Vibrato - nota mi4 (E4) na corda lá



Evento: Sustentação com vibrato da nota mi4 (E4) tocada na corda lá.

Gráfico do sinal original:



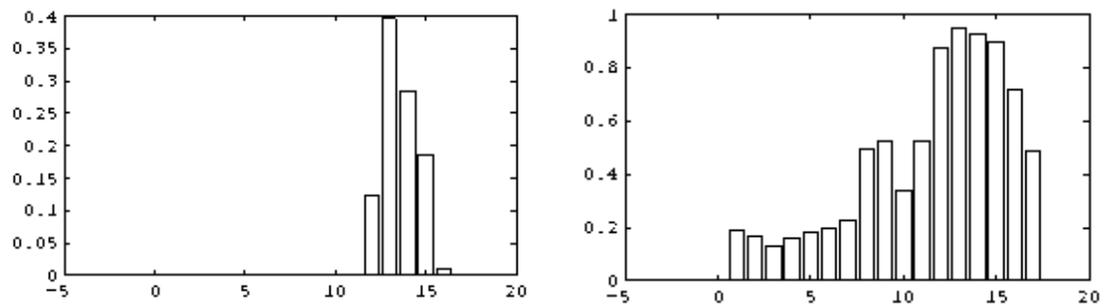
Extensão da amostra: 2^{18} pontos (262144). **Wavelet de análise:** D16

Análise dos níveis:

O nível 17 apresenta muito ruído, característico do arraste/atrito do arco. Ao nível 16 já é perceptível o vibrato.

Nos níveis de 12 a 15 observa-se a predominância da flutuação do vibrato se impondo às formas de onda. O nível 10 é ruidoso, de formato irregular, não guardando semelhança com o envelope da nota tocada. Nos níveis 8 e 9 o transiente associado ao ataque da nota é bem localizado.

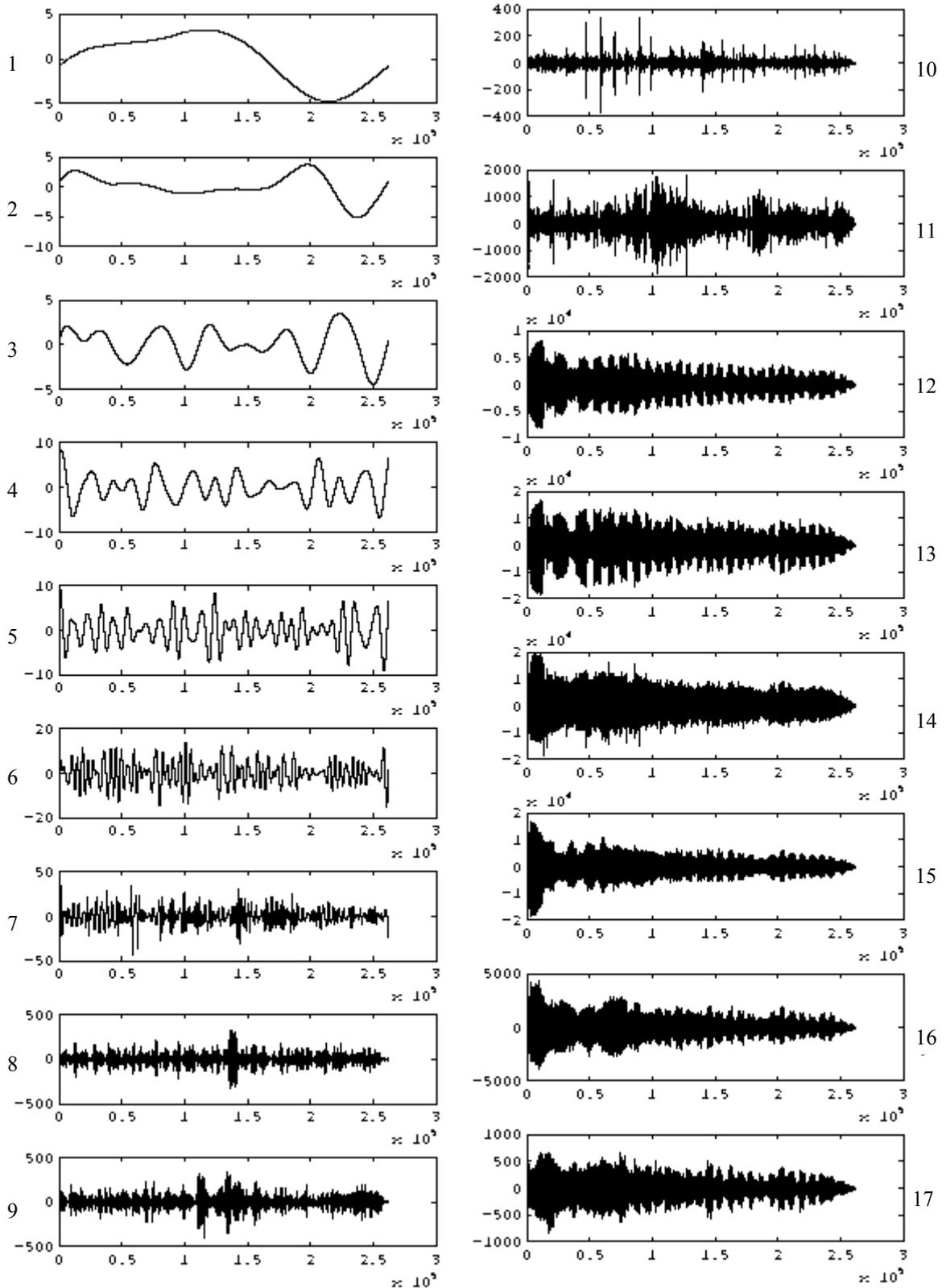
Distribuição de energia:



Mais de 99% de toda a energia do exemplo concentra-se nos níveis 12 a 16. O nível 13 é o mais energético. Os níveis mais baixos, especialmente os inferiores ao nível 7 (inclusive) não impactam a qualidade da reconstrução caso sejam desprezados nesta, exatamente por não contribuírem significativamente em energia.

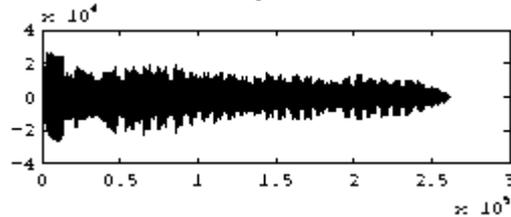
Os níveis mais energéticos concentram mais da qualidade vibratória do vibrato, como se pode observar pelas representações gráficas e pela audição dos níveis.

Violino: Sustentação com trinado - notas mi⁴ (E₄) e fá[#]₄ (F[#]₄)



Evento: Sustentação com trinado envolvendo as notas mi 4 (E4) e fá sustenido 4 (F#4)¹⁴, na corda lá. Amostra editada sem acolchoamento com zeros (*zero padding*), o trecho musical analisado apresentando extensão exata em potência de dois.

Gráfico do sinal original:



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

Análise dos níveis:

Níveis 17 e 16: O ruído de arco é o componente predominante. O efeito do trinado é perceptível.

Níveis 15 e 13: O trinado é bem percebido auditivamente, como dois tons se alternando

Nível 10: Ruído grave, contínuo, sobreposto em determinados pontos com transientes ou pulsos. Os pulsos ocorrem na transição do mi4 para o fá#4, quando o dedo tem de tocar a corda para diminuir seu comprimento vibrante. Espectros deste nível, na região de ruído, e na região dos pulsos, revelam relações de uma quinta perfeita (3:2) entre as frequências mais pronunciadas.

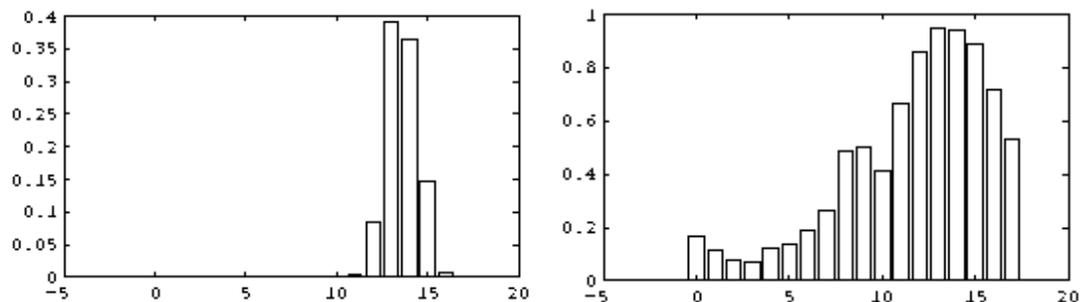
Nível 9: Muito grave em frequência, informações auditivas precárias; o gráfico não apresenta nenhuma informação adicional relevante. O nível 8 assemelha-se ao 9, mas apresenta um pulso central pronunciado, e é o último nível audível.

Nível 7: Apresenta uma forma de onda cheia, preenchendo todo o exemplo, sem no entanto distinguir porções especiais. O nível 6, contendo somente os detalhes que não são evidenciados no nível anterior, já mostra algumas estruturas distintas.

Nível 5: O seu espectro mostra que cobre somente as frequências de 0 a 10 Hz presentes no sinal analisado. O nível 4 revela uma forma de onda mais suave que a do nível anterior. Nos níveis mais grossos subsequentes esta característica é comum, uma vez que a forma de onda da *Wavelet*, uma D16 suave, vai se impondo.

Os níveis abaixo não fornecem informações relevantes sobre o evento em curso (trinado). Assinalam nestas escalas a região onde a nota se desenvolve, sem precisão de localização dos eventos.

Distribuição de energia:



¹⁴ O símbolo "#" em Música refere-se a um semitom acima. Assim, F#4 está um semitom acima de F4.

Mais de 99% da energia do exemplo concentra-se nos níveis de 11 a 16, sendo o nível 13 o mais energético.

Segundo gráfico (à direita): logaritmo da energia de cada nível da análise pelo logaritmo da energia total (escala logarítmica, normalizada).