

## ESTUDO COMPARATIVO SOBRE FILTRAGEM DE SINAIS INSTRUMENTAIS USANDO TRANSFORMADAS DE FOURIER E WAVELET

**Roberto Kawakami Harrop Galvão**

Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, São José dos Campos - SP

**Mário César Ugulino de Araújo\***, **Teresa Cristina Bezerra Saldanha** e **Valeria Visani**

Departamento de Química, CCEN, Universidade Federal da Paraíba, CP 5093, 58051-970 João Pessoa - PB

**Maria Fernanda Pimentel**

Departamento de Engenharia Química, CTG, Universidade Federal de Pernambuco, Recife - PE

Recebido em 11/10/00; aceito em 14/3/01

COMPARATIVE STUDY ON INSTRUMENTAL SIGNAL DENOISING USING FOURIER AND WAVELET TRANSFORMS. A comparative study of the Fourier (FT) and the wavelet transforms (WT) for instrumental signal denoising is presented. The basic principles of wavelet theory are described in a succinct and simplified manner. For illustration, FT and WT are used to filter UV-VIS and plasma emission spectra using MATLAB software for computation. Results show that FT and WT filters are comparable when the signal does not display sharp peaks (UV-VIS spectra), but the WT yields a better filtering when the filling factor of the signal is small (plasma spectra), since it causes low peak distortion.

Keywords: instrumental signal denoising; Fourier transform; wavelet transform.

### INTRODUÇÃO

No contexto de processamento de sinais, entende-se por filtragem a remoção de ruído visando melhorias na interpretação e utilização dos dados.

Entre os ruídos que normalmente contaminam respostas instrumentais, podem ser citados o Johnson, o Shot, o Flicker e o ambiental<sup>1</sup>. Os métodos mais comumente utilizados em Química para remoção de tais interferências são os filtros de média *boxcar*, média móvel e Savitsky-Golay<sup>1</sup>. Embora de fácil implementação, tais filtros tendem a distorcer picos estreitos<sup>2</sup>, o que pode comprometer a análise, por exemplo, de sinais espectroscópicos.

Nesse contexto, melhorias podem ser obtidas através do uso da Transformada de Fourier (TF)<sup>1</sup>, conforme descrito no recente tutorial de Cerqueira et al<sup>2</sup>. Uma outra alternativa consiste no uso da Transformada Wavelet (TW)<sup>3-9</sup>, ferramenta de processamento de sinais introduzida nos anos 80 para superar algumas limitações da TF. A literatura<sup>10</sup> relata que a filtragem via TW, em contraste com métodos de suavização polinomial, não desloca nem deforma pontos notáveis (máximos, mínimos e inflexões) do sinal a ser filtrado, apresentando ainda vantagens sobre a TF quando o sinal possui um baixo "fator de preenchimento" (*filling factor*)<sup>11</sup>, isto é, apresenta picos alternados com regiões de baixa intensidade.

Este trabalho tem por objetivo fazer um estudo comparativo do desempenho da TF e TW em problemas de filtragem de sinais instrumentais. Para tal, a teoria pertinente e os procedimentos de filtragem via TF e TW são brevemente descritos. A título de ilustração, é investigada a aplicação dessas técnicas a espectros de absorção molecular UV-VIS e de emissão em plasma.

### ASPECTOS TEÓRICOS E LIMITAÇÕES DA TRANSFORMADA DE FOURIER

Em sua versão discreta, a TF de um sinal  $f(x)$  com  $N$  pontos, é calculada como

$$F(\omega) = \sum_{x=0}^{N-1} f(x) \exp(-i\omega x) \quad (1)$$

onde  $x$  é uma variável cuja natureza depende da técnica instrumental utilizada na geração do sinal (por exemplo, comprimento de onda no caso de espectroscopia). Vale ressaltar que alguns autores empregam um fator de normalização de  $1/N$  na Equação 1, mas a expressão aqui apresentada é a empregada pelo software MATLAB. O parâmetro  $\omega$ , denominado frequência, assume os seguintes valores:

$$\omega = 2\pi \frac{k}{N}, \quad k = 0, 1, \dots, N-1 \quad (2)$$

onde  $k$  é chamado "índice de frequência".

Como resultado da aplicação da TF, obtém-se um conjunto de  $N$  coeficientes, indexados por  $k$ . Anulando os coeficientes em que a presença do ruído é preponderante (normalmente em frequências elevadas) e aplicando a Transformada de Fourier Inversa (TFI), obtém-se um sinal mais "limpo".

Uma das limitações da TF encontra-se no fato de que ela não permite analisar em separado diferentes trechos do sinal. Assim, caso um trecho seja extremamente ruidoso ou contenha pontos anômalos, o processamento de todo o sinal é comprometido. A fim de resolver esse problema, foi introduzida a Transformada de Fourier Janelada (TFJ), que consiste em dividir o sinal em regiões e aplicar a TF a cada uma delas. Matematicamente, a TFJ de um sinal  $f(x)$  é dada por<sup>12</sup>

$$F(p, \omega) = \sum_{x=0}^{N-1} f(x) w(x-p) \exp(-i\omega x) \quad (3)$$

onde  $w(x)$  é uma função de janelamento, responsável pela delimitação do trecho que está sendo considerado no sinal. A posição da janela dentro do sinal é dada pelo parâmetro  $p$ . A título de ilustração, a Figura 1 mostra um espectro dividido em 8 janelas retangulares.

No exemplo da Figura 1, a função  $w(x)$  realiza cortes abruptos, dividindo o sinal em janelas bem delimitadas. Contudo, esse procedimento causa um problema conhecido como

\* e-mail: laqa@quimica.ufpb.br

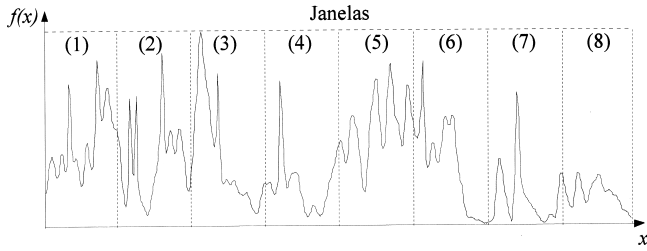


Figura 1. Espectro dividido em 8 regiões por janelas retangulares.

“efeito de borda” ou “fenômeno de Gibbs”<sup>13</sup>, que pode ser minimizado usando-se janelas com decaimento suave, tais como a gaussiana:

$$w(x) = e^{-x^2/2} \tag{4}$$

Todavia, a recuperação do sinal a partir de uma TFJ que empregue janelas desse tipo é uma tarefa complexa<sup>12</sup>, o que dificulta sua aplicação à filtragem de ruído.

Além do efeito de borda, outro problema relevante é a escolha da largura para a função de janelamento. Suponha, por exemplo, que sejam usadas janelas retangulares com  $M$  variáveis  $x$ . Então, cada janela dá origem a  $M$  coeficientes Fourier e, portanto, a preservação do número total de  $N$  variáveis requer o uso de  $N/M$  janelas. Note-se que, aumentando  $M$ , melhora-se a análise de cada janela (maior número de coeficientes Fourier), mas perde-se resolução espacial (menor número de janelas). Esse problema é conhecido na área de processamento de sinais como Princípio da Incerteza de Heisenberg<sup>2,12</sup>. Para ilustrar a dificuldade em se escolher uma janela de largura conveniente, pode-se notar na Figura 1 que a região 2 precisaria ser dividida em janelas mais estreitas que a região 8.

Para contornar os problemas de efeito de borda e escolha da largura da janela, foi introduzida a TW<sup>3-7</sup>. Como será visto a seguir, a TW utiliza janelas sem cortes abruptos, à semelhança da gaussiana, mas que, ao contrário desta, permitem uma fácil reconstrução do sinal a partir dos coeficientes da transformada. Além disso, a TW utiliza janelas de largura variável, que podem assim ser mais bem ajustadas às características de cada trecho do sinal.

**TRANSFORMADA WAVELET**

**Histórico**

Embora desde o começo do século a literatura registre tratamentos matemáticos semelhantes à TW (Alfred Haar, 1910, apud Daubechies<sup>4</sup> e Meyer<sup>7</sup>), foi no final da década de 70 que a mesma passou a ter uma identidade própria. Nessa ocasião, o francês Jean Morlet, trabalhando para a companhia de petróleo Elf Aquitaine, propôs uma modificação na TF, para melhor tratar sinais geofísicos. Por falta de embasamento matemático, Morlet inicialmente encontrou muitos opositores. Esses, como lembra Daubechies<sup>9</sup>, teciam críticas do seguinte tipo: “Se isso estivesse correto, já estaria nos livros de matemática. Como não está, provavelmente é inútil”. No entanto, alguns anos depois, as “wavelets” de Morlet atraíram a atenção de Yves Meyer, um matemático que ajudou a enriquecer e amadurecer a nova teoria, encontrando paralelos surpreendentes com diversos outros campos da matemática, antes estudados separadamente<sup>7</sup>. Logo em seguida, Stéphane Mallat, um estudante de processamento de imagens, desenvolveu um algoritmo para calcular a TW de forma computacionalmente eficiente, abrindo então as portas à comunidade de processamento de sinais<sup>5,14</sup>. Como se vê, desde seu início, a teoria de wavelets se mostrou interdisciplinar, o que em parte explica a

grande popularidade adquirida nos últimos anos e a opinião de alguns autores<sup>15</sup> que a reputam como o “evento matemático de maior relevância na década de 80”.

**Formulação Matemática**

A TW de um sinal  $f(x)$  é expressa como

$$Wf(a,b) = \int_{-\infty}^{+\infty} f(x)\psi_{a,b}(x)dx \tag{5}$$

que, aproximando a integral por um somatório, torna-se

$$Wf(a,b) = \sum_{x=0}^{N-1} f(x)\psi_{a,b}(x) \tag{6}$$

para um sinal discreto com  $N$  pontos. A função  $\psi_{a,b}(x)$ , chamada “wavelet”, é derivada de uma função  $\psi(x)$  através da seguinte transformação:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right) \tag{7}$$

Há uma ampla gama de possíveis escolhas para a função  $\psi(x)$ , denominada “wavelet-mãe”. Duas possíveis alternativas, que foram usadas neste trabalho, estão apresentadas na Figura 2.

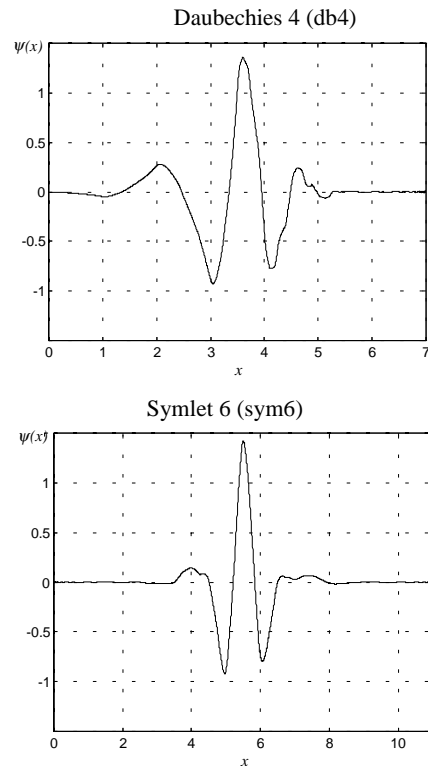


Figura 2. Exemplos de wavelets empregadas neste trabalho.

Comparando-se as Equações 3 e 6, vê-se que ambas realizam o produto do sinal por funções que apresentam oscilações dentro de uma janela. A diferença encontra-se na natureza dos parâmetros das transformadas. Na TFJ,  $p$  e  $w$  representam, respectivamente, posição e frequência, estando fixada a largura da janela. Na TW,  $b$  também representa posição (ou “translação” da wavelet), mas  $a$  (chamado parâmetro de “escala”) está associado à largura da janela, como se pode ver na Figura 3.

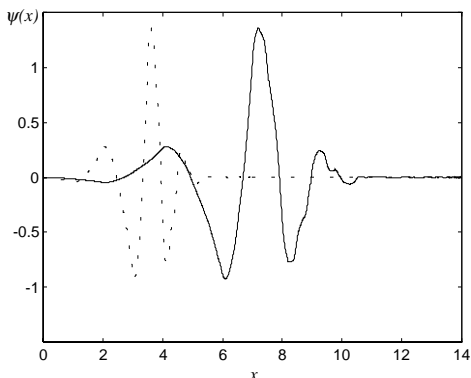


Figura 3. Efeito de modificações na escala para a wavelet db4. Em tracejado, encontra-se a wavelet-mãe e, em linha cheia, a wavelet com escala  $a = 2$ .

**Algoritmo de decomposição em árvore**

Como não é viável calcular a TW para todos os possíveis valores de  $a$  e  $b$  no conjunto dos números reais, é de praxe fazer a seguinte restrição:

$$a=2^m, \quad b=n2^m \tag{8}$$

onde  $m$  e  $n$  são números inteiros<sup>4</sup>. Esse procedimento conduz a uma estrutura de escalas e translações chamada “diádica”, que guarda semelhanças com a notação musical, em que potências de dois estão relacionadas com intervalos (oitavas) e durações das notas. Nesse caso, o resultado da aplicação da TW sobre um sinal é um conjunto de coeficientes wavelet indexados por  $m$  (nível de escala) e  $n$  (índice de translação). Pode-se mostrar<sup>4</sup> que dessa forma a informação do sinal é preservada e o número de coeficientes é igual ao número original de variáveis, como na TF. Além disso, torna-se possível obter os coeficientes de uma forma rápida e computacionalmente eficiente, através do algoritmo de decomposição em árvore proposto por Mallat<sup>4,5,14,16</sup> e representado esquematicamente na Figura 4.

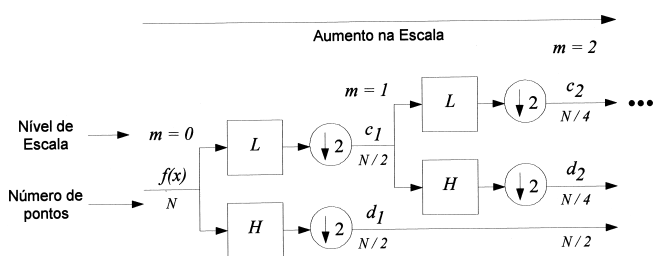


Figura 4. Representação do algoritmo de decomposição wavelet em árvore.

Nessa figura, os blocos  $L$  e  $H$  representam filtros digitais passa-baixas e passa-altas, respectivamente, que estão associados à wavelet adotada na análise. O símbolo  $(\downarrow 2)$  representa a operação de sub-amostragem, ou decimação, que consiste em eliminar todos os coeficientes de índice par (0, 2, ...) de uma seqüência. A sub-amostragem garante que o número total de pontos permaneça constante após a TW.

**Tabela 1.** Seqüências de filtragem associadas à wavelet-mãe db4.

$k$	0	1	2	3	4	5	6	7
$l(k)$	-0,011	0,033	0,031	-0,187	-0,028	0,631	0,715	0,230
$h(k)$	-0,230	0,715	-0,631	-0,028	0,187	0,031	-0,033	-0,011

Em cada nível de escala  $m$ , geram-se dois conjuntos: um conjunto  $d_m$  de coeficientes wavelet (também chamados, neste contexto, de “detalhes”) e um conjunto  $c_m$  de coeficientes ditos de “aproximação”. O conjunto  $c_m$  passa em seguida por uma nova divisão L/H, de maneira a gerar coeficientes  $c_{m+1}$  e  $d_{m+1}$  e assim sucessivamente. As operações de filtragem e decimação podem ser resumidas nas equações abaixo, em que as seqüências  $l(k)$  e  $h(k)$ , de  $K$  pontos cada uma, são responsáveis, respectivamente, pela filtragem passa-baixas e passa-altas.

$$c_{m+1}(n) = \sum_{k=0}^{K-1} l(k)c_m(2n-k+1)$$

$$d_{m+1}(n) = \sum_{k=0}^{K-1} h(k)c_m(2n-k+1) \tag{9}$$

A título de exemplo, as seqüências de filtragem associadas à wavelet-mãe db4 (Figura 2) possuem  $K = 8$  pontos, como mostra a Tabela 1.

Vale ressaltar que a expressão da wavelet-mãe  $\psi(x)$  não é usada no algoritmo de decomposição em árvore. Com efeito, wavelets como as apresentadas na Figura 2 não apresentam uma formulação matemática explícita, sendo obtidas de forma numérica<sup>4</sup> a partir das seqüências  $l(k)$  e  $h(k)$ .

A saída da árvore de decomposição depende do número de níveis de escala (também chamados “níveis de resolução”) empregados, como ilustrado na Figura 5.

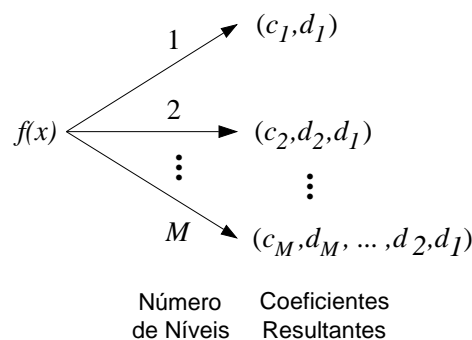


Figura 5. Resultado da decomposição de um sinal  $f(x)$  em coeficientes wavelet, em função do número de níveis de escala empregado.

Do ponto de vista químico, os detalhes estão associados a características bem localizadas do sinal, tais como picos estreitos, enquanto as aproximações correspondem a uma “linha média” do sinal, removidos os detalhes. Quanto mais níveis de detalhes são removidos, mais suavizada fica a aproximação.

Um segundo algoritmo em árvore pode ser usado para reconstruir o sinal a partir dos coeficientes (Transformada Wavelet Inversa - TWI) como mostra a Figura 6.

Nessa árvore, o símbolo  $(\uparrow 2)$  representa a operação de inserção de zeros entre os pontos de uma seqüência. A operação de reconstrução dos coeficientes  $c_{m-1}$  a partir de  $c_m$  e  $d_m$  pode ser resumida pela equação

$$c_{m-1}(n) = \sum_{k=0}^{N_m-1} [l(n-2k+K-2)c_m(k) + h(n-2k+K-2)d_m(k)] \tag{10}$$

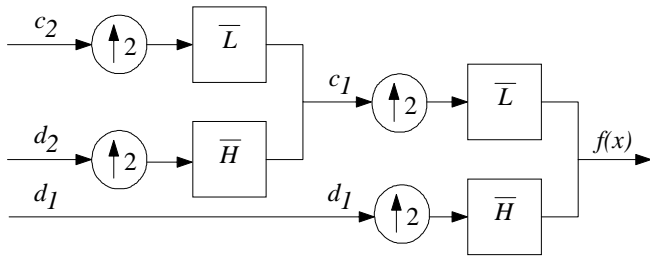


Figura 6. Árvore de reconstrução wavelet com dois níveis.

onde as seqüências  $d_m$  e  $c_m$  possuem  $N_m$  pontos cada uma. É possível mostrar que, sob certas condições<sup>4,14</sup>, as seqüências de filtragem usadas na TWI são iguais às seqüências  $l(k)$  e  $h(k)$  invertidas, isto é,

$$\bar{l}(k) = l(K - 1 - k) \text{ e } \bar{h}(k) = h(K - 1 - k) \quad (11)$$

Note-se ainda que, na Equação 10, considera-se  $\bar{l}(k) = \bar{h}(k) = 0$  para  $k < 0$  ou  $k \geq K$ .

Um interessante paralelo pode ser traçado entre o esquema de reconstrução wavelet e processos de visão, que eram o campo de pesquisa original de Mallat. Ao olharmos uma pessoa a distância, é possível distinguir apenas os contornos de seu corpo (aproximação). Contudo, ao nos aproximarmos, passamos a perceber mais características (detalhes) do indivíduo, tais como o formato de seu rosto, depois seus olhos e assim por diante. Ou seja, nesse processo, a formação da imagem passa pela incorporação sucessiva de detalhes em escalas cada vez menores.

Para mais detalhes sobre a TW, recomendam-se os tutoriais de Rioul e Vetterli<sup>6</sup> e também de Alsberg *et al*<sup>8</sup> e, com maior rigor matemático, a obra clássica de Daubechies<sup>4</sup>.

**Filtragem empregando TW**

O processo de filtragem baseado na TW (vide Figura 7) consiste em fazer iguais a zero os coeficientes de detalhe que, em módulo, são menores que um certo limiar de corte. Ou seja, em cada nível  $m$  da árvore de decomposição,

$$d_m^{filtrado}(n) = \begin{cases} d_m(n), & \text{se } |d_m(n)| > \text{limiar}_m \\ 0, & \text{se } |d_m(n)| \leq \text{limiar}_m \end{cases} \quad n = 0, 1, \dots, N_m - 1 \quad (12)$$

Note-se que podem ser usados limiares diferentes para cada nível. Após esse processo de corte, usa-se a TWI para obter o sinal filtrado.

**APLICAÇÃO**

Nesta seção, será inicialmente apresentado um problema de filtragem de um espectro simulado, visando ilustrar as potenciais vantagens da TW em relação à TF. Em seguida, será empregado um espectro de absorção molecular UV-VIS de uma mistura dos complexos de  $Co^{+2}$ ,  $Cu^{+2}$ ,  $Mn^{+2}$ ,  $Ni^{+2}$  e  $Zn^{+2}$  com 4-(2-piridilazo)resorcinol (PAR), obtido nas condições experimentais detalhadas por Saldanha *et al*<sup>17</sup>. Por fim, será considerado um espectro de emissão em plasma de uma amostra de aço-liga certificada (BCS-464) que contém Mn, Cr, Ni e Fe. Esse espectro foi obtido conforme descrito por Pimentel *et al*<sup>18</sup>. Como será visto, a absorção molecular e a emissão em plasma ilustram situações bastante distintas do ponto de vista de processamento de sinais.

No Apêndice, é apresentado o programa de filtragem via TW usado neste trabalho, escrito em linguagem MATLAB, que emprega o MATLAB Wavelet Toolbox<sup>5</sup>. A filtragem via TF foi também implementada em MATLAB, seguindo os mesmos passos descritos por Cerqueira *et al*<sup>2</sup>.

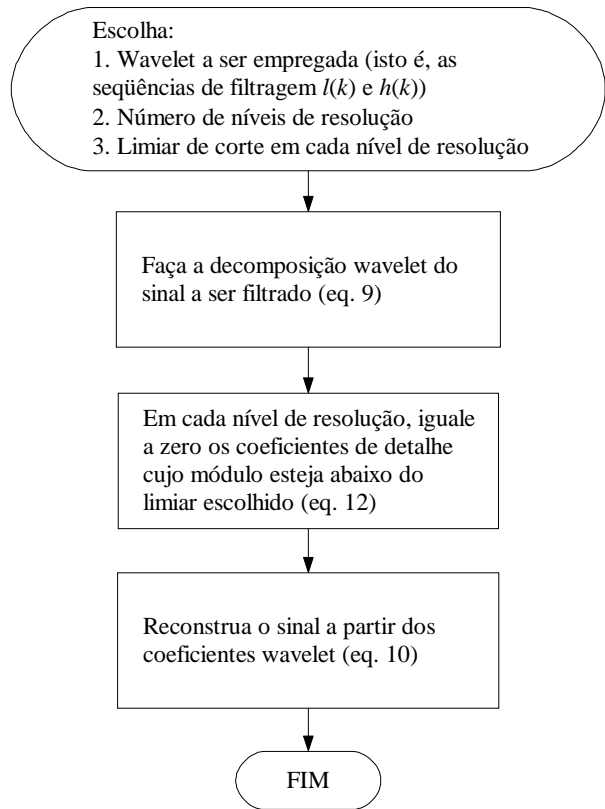


Figura 7. Fluxograma do algoritmo de filtragem wavelet.

**Filtragem de um espectro simulado**

Na Figura 8a, é apresentado um espectro formado por um pico triangular seguido de um período da função  $sen^2(x)$ . A Figura 8b traz o mesmo espectro, acrescido de um ruído branco com desvio-padrão de 0,02. Este primeiro exemplo tem por objetivo mostrar que um problema fundamental da filtragem, qual seja, a distorção de picos estreitos, pode ser amenizado pelo uso da TW.

As Figuras 9a e 9b apresentam a TF, em módulo, para os espectros original e ruidoso, respectivamente. Vale ressaltar que, devido à simetrização dos sinais, recomendada por Cerqueira *et al*<sup>2</sup>, a TF resulta em 300 coeficientes (o dobro do número inicial de comprimentos de onda). Contudo, as figuras apresentam apenas os 150 primeiros coeficientes, visto que os demais são iguais a estes, em módulo<sup>19</sup>.

Como se vê na Figura 9b, o efeito do ruído é mais evidente em altas frequências. Para filtrá-lo, deve-se definir uma frequência de corte adequada, de modo a não causar distorções no sinal após a TFI. Neste caso, a dificuldade consiste no fato de que o pico triangular introduz coeficientes de alta frequência ( $k = 70$  a  $110$  e  $k = 130$  a  $150$ , aproximadamente), como mostra a Figura 9a. Assim, qualquer tentativa de eliminar o ruído causará necessariamente uma distorção nesse pico. Esse problema está demonstrado na Figura 10, que mostra o valor do pico do sinal filtrado, em função da frequência de corte adotada. Nota-se que, quanto mais coeficientes de alta frequência são eliminados na filtragem, maior a distorção causada no pico.

As Figuras 11a e 11b apresentam o resultado da aplicação da TW aos espectros das Figuras 8a e 8b, empregando uma wavelet-mãe da família Symlet (sym3)<sup>4,5</sup> e dois níveis de resolução. Como se vê, o ruído se manifesta principalmente nos coeficientes de detalhe ( $d_1$  e  $d_2$ ). Como explicado anteriormente, o processo de filtragem consiste em fazer iguais a zeros todos os coeficientes de detalhe que, em cada nível, tiverem módulo inferior a um dado limiar.

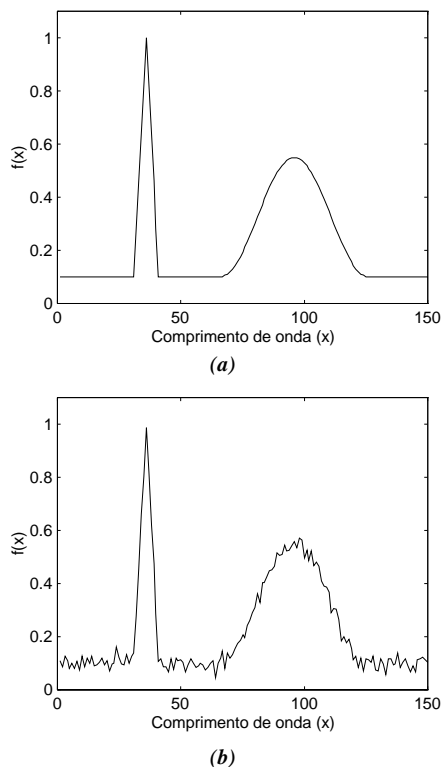


Figura 8. Espectro simulado antes (a) e após a introdução de ruído (b).

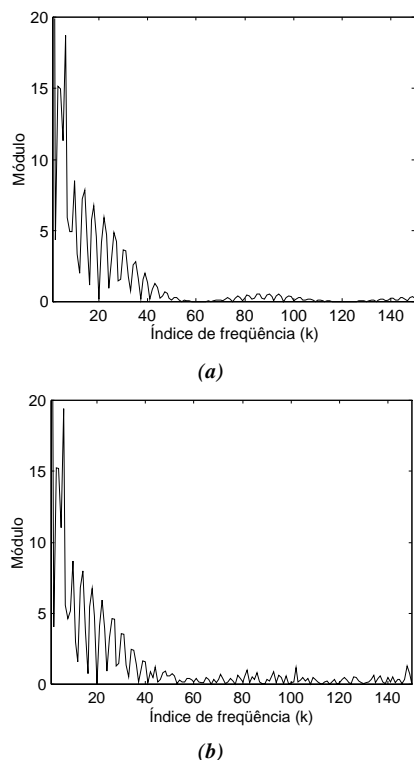


Figura 9. TF do espectro simulado antes (a) e após a introdução de ruído (b).

Contudo, assim como na TF, a eliminação de um número muito grande de coeficientes pode causar distorções no sinal. Esse fato está ilustrado na Figura 12, que mostra o valor do pico do sinal filtrado em função do limiar de corte, expresso como uma porcentagem do coeficiente de maior módulo em cada nível.

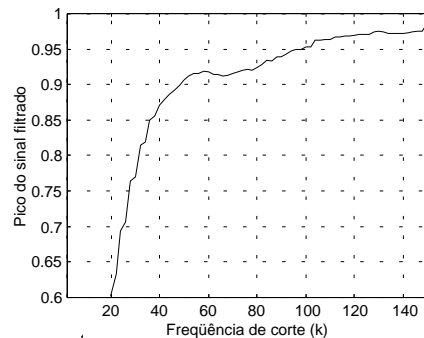


Figura 10. Valor do pico do sinal filtrado em função da frequência de corte da TF do espectro simulado.

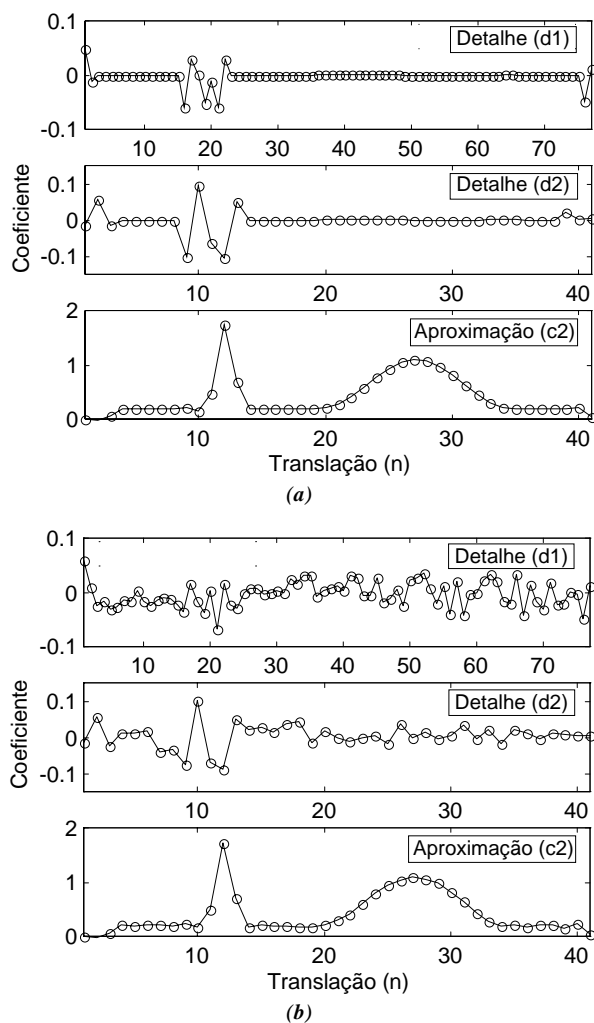


Figura 11. TW do espectro simulado antes (a) e após a introdução de ruído (b).

A fim de comparar a filtragem via TF e via TW, suponha que a maior distorção tolerável seja tal que o pico do espectro filtrado seja de 0,94. Nesse caso, a frequência de corte para a TF seria  $k = 88$  (vide Figura 10) e o limiar de corte para a TW seria de 80% (vide Figura 12). O resultado do corte pode ser visualizado nas Figuras 13a e 13b e os sinais reconstruídos através da TFI e da TWI, nas Figuras 14a e 14b. Como se pode notar, para um mesmo nível de distorção no pico do sinal, a filtragem via TW foi visivelmente melhor.

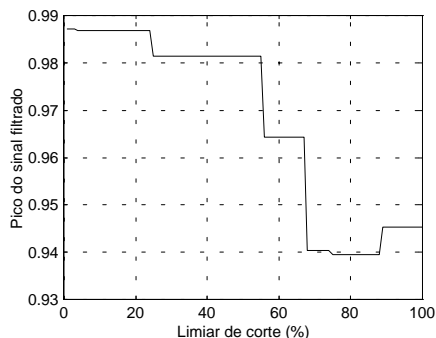
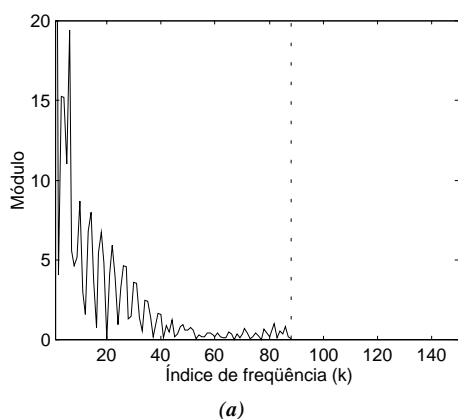
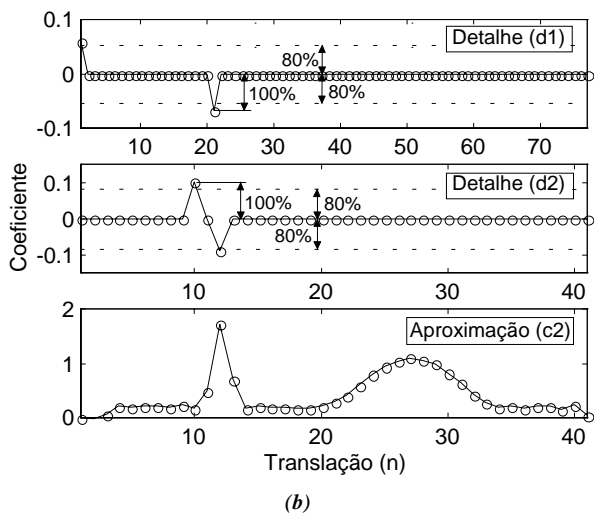


Figura 12. Valor do pico do sinal filtrado em função do limiar de corte da TW, expresso como uma porcentagem do coeficiente de maior módulo em cada nível de detalhe.



(a)

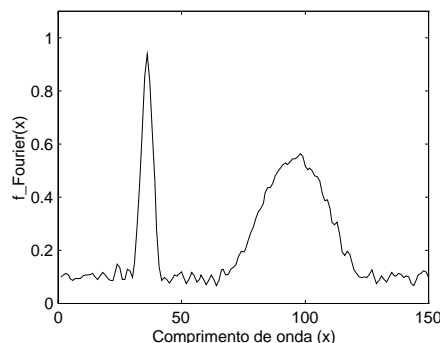


(b)

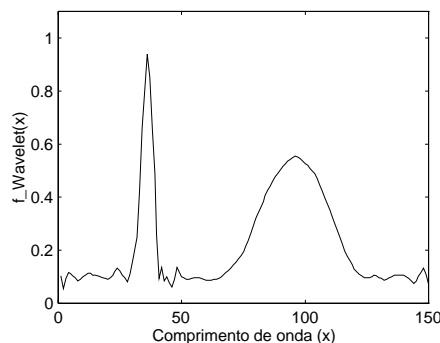
Figura 13. Resultado do corte sobre a TF (a) e TW (b) do espectro ruidoso. A frequência de corte está indicada por uma barra vertical tracejada em (a) e os limiares de corte, por barras horizontais tracejadas em (b).

**Filtragem de um espectro de absorção molecular UV-VIS**

Na Figura 15a é mostrado um espectro de absorção molecular UV-VIS de uma mistura contendo complexos de  $Co^{2+}$ ,  $Cu^{2+}$ ,  $Mn^{2+}$ ,  $Ni^{2+}$  e  $Zn^{2+}$  com PAR. A este sinal, foi artificialmente adicionado um ruído branco com desvio-padrão de 0,01 (Figura 15b). As Transformadas de Fourier em módulo para o espectro UV-VIS sem e com ruído são mostradas, respectivamente, nas Figuras 16a e 16b. A fim de investigar quantitativamente o efeito de variações na frequência de corte, propõe-se aqui o uso da raiz do erro quadrático médio (*root-mean-square error - RMSE*),



(a)



(b)

Figura 14. Espectros filtrados via TF (a) e via TW (b).

que mede a qualidade da filtragem comparando o espectro original  $f(x)$  com o espectro filtrado  $f_f(x)$ . O *RMSE* é expresso como

$$RMSE = \sqrt{\frac{1}{N} \sum_{x=0}^{N-1} [f(x) - f_f(x)]^2}$$

A Figura 16c mostra o valor do *RMSE* obtido em função da frequência de corte. Nota-se que, para frequências de corte baixas, o *RMSE* é elevado, indicando uma grande diferença entre o espectro original e o filtrado. Neste caso, a filtragem causa distorção no sinal. Para frequências de corte mais altas, o *RMSE* é elevado porque não se está removendo o ruído. A frequência de corte ótima é  $k = 16$  ( $RMSE = 4,8 \times 10^{-3}$ ), que está indicada na Figura 16b por uma barra vertical tracejada. Aplicando-se então a TFI, obtém-se o espectro filtrado (linha cheia na Figura 16d), que não apresenta grandes diferenças em relação ao espectro original (linha tracejada na Figura 16d).

Para fins de comparação, foi aplicada a Transformada Wavelet aos mesmos espectros apresentados nas Figuras 15a e 15b. A fim de selecionar a melhor função wavelet para o propósito de filtragem, foram testadas as famílias Daubechies (db1 a db8) e Symlet (sym2 a sym8)<sup>4,5</sup>. Para cada uma das wavelets dessas famílias, foram investigados os seguintes parâmetros: número de níveis de decomposição (até o valor máximo possível para cada wavelet, dado pela função WMAXLEV do Matlab Wavelet Toolbox) e limiar de corte (25%, 50% ou 75% do maior coeficiente wavelet em cada nível de detalhe).

O menor *RMSE* ( $4,8 \times 10^{-3}$ ) foi obtido para a wavelet db4 (Figura 2a) com 2 níveis de decomposição e limiar de corte de 50%. O resultado da transformada nessas condições para o sinal original e com ruído encontra-se, respectivamente, nas Figuras 17a e 17b. Os limiares para os detalhes  $d_2$  e  $d_1$  estão representados por barras horizontais tracejadas na Figura 17b e o resultado do corte é mostrado na Figura 17c. Aplicando-se a TWI, obtém-se o espectro filtrado (linha cheia na Figura 17d), que não apresenta grandes diferenças em relação ao espectro original (linha tracejada na Figura 17d).

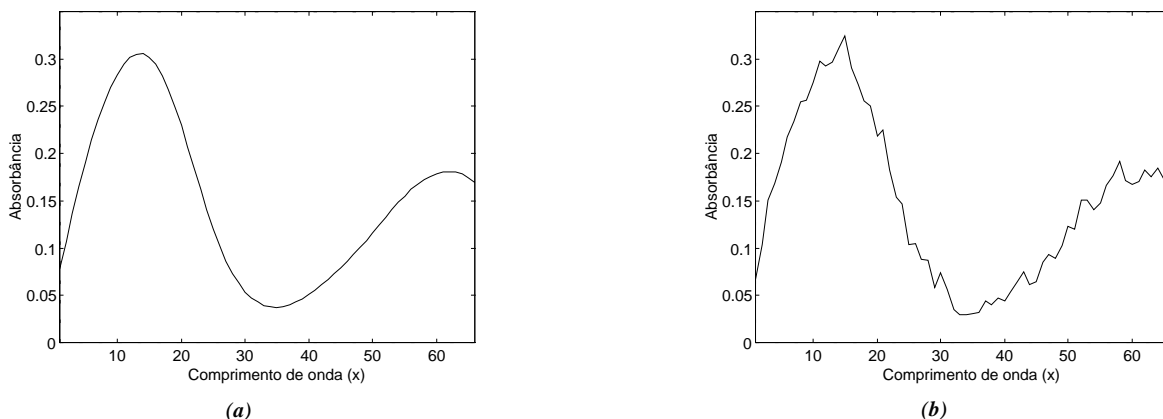


Figura 15. Espectro UV-VIS original (a) e após a introdução de ruído (b).

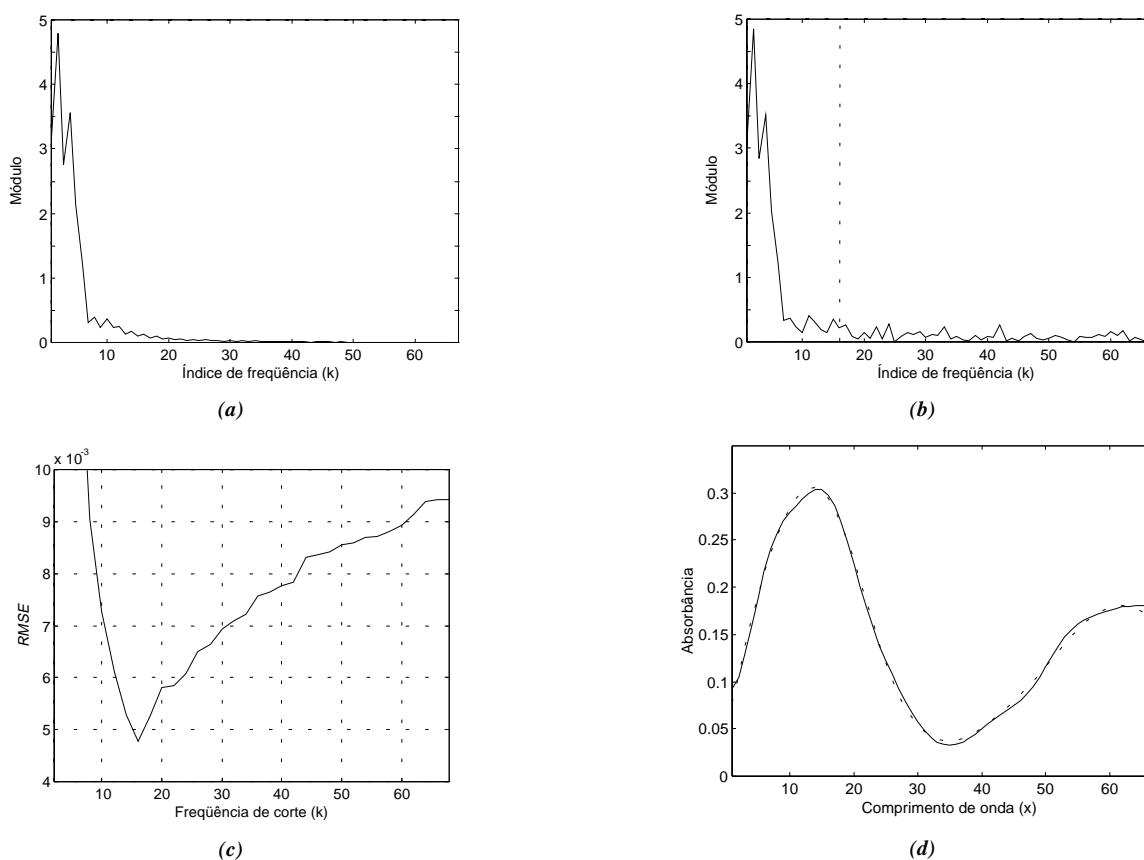


Figura 16. TF do espectro UV-VIS original (a) e após a introdução de ruído (b). RMSE em função da frequência de corte (c). Espectro filtrado (linha cheia) e espectro original (linha tracejada) (d).

Neste exemplo, a filtragem via TW não trouxe melhorias em relação à filtragem via TF, como se pode notar pelo *RMSE*, igual para os dois. A razão se encontra no fato de que o sinal estudado era suave, não apresentando muitos pontos notáveis (máximos, mínimos e inflexões).

#### Filtragem de um espectro de emissão em plasma

O processo de filtragem será agora ilustrado usando um espectro de natureza substancialmente diferente do espectro UV-VIS. Com efeito, o espectro de emissão em plasma possui picos de emissão pronunciados, que podem vir a ser deformados pelo processo de filtragem. Uma vez que o *RMSE* não é adequado para medir distorções nos picos, já que ele mede a qualidade global da filtragem, propõe-se o uso do erro absoluto máximo (*EAM*), expresso como

$$EAM = \max |f(x) - f_f(x)|, x = 0, 1, \dots, N-1 \quad (14)$$

visando quantificar a maior distorção causada no espectro pela filtragem.

As Figuras 18a e 18b apresentam, respectivamente, o espectro original e o espectro acrescido de um ruído branco com desvio padrão de 80. Observando a TF do espectro original (Figura 19a), nota-se que os coeficientes em alta frequência não se aproximam tanto de zero como na TF do espectro UV-VIS (Figura 16a). Isso se deve à presença de picos estreitos no espectro de emissão em plasma. Neste caso, a separação entre sinal e ruído não é tão clara, como se pode ver na Figura 19b e, portanto, a filtragem sempre acarretará uma distorção nos picos, com conseqüente perda da informação nos mesmos. Tomando-se por exemplo a frequência de

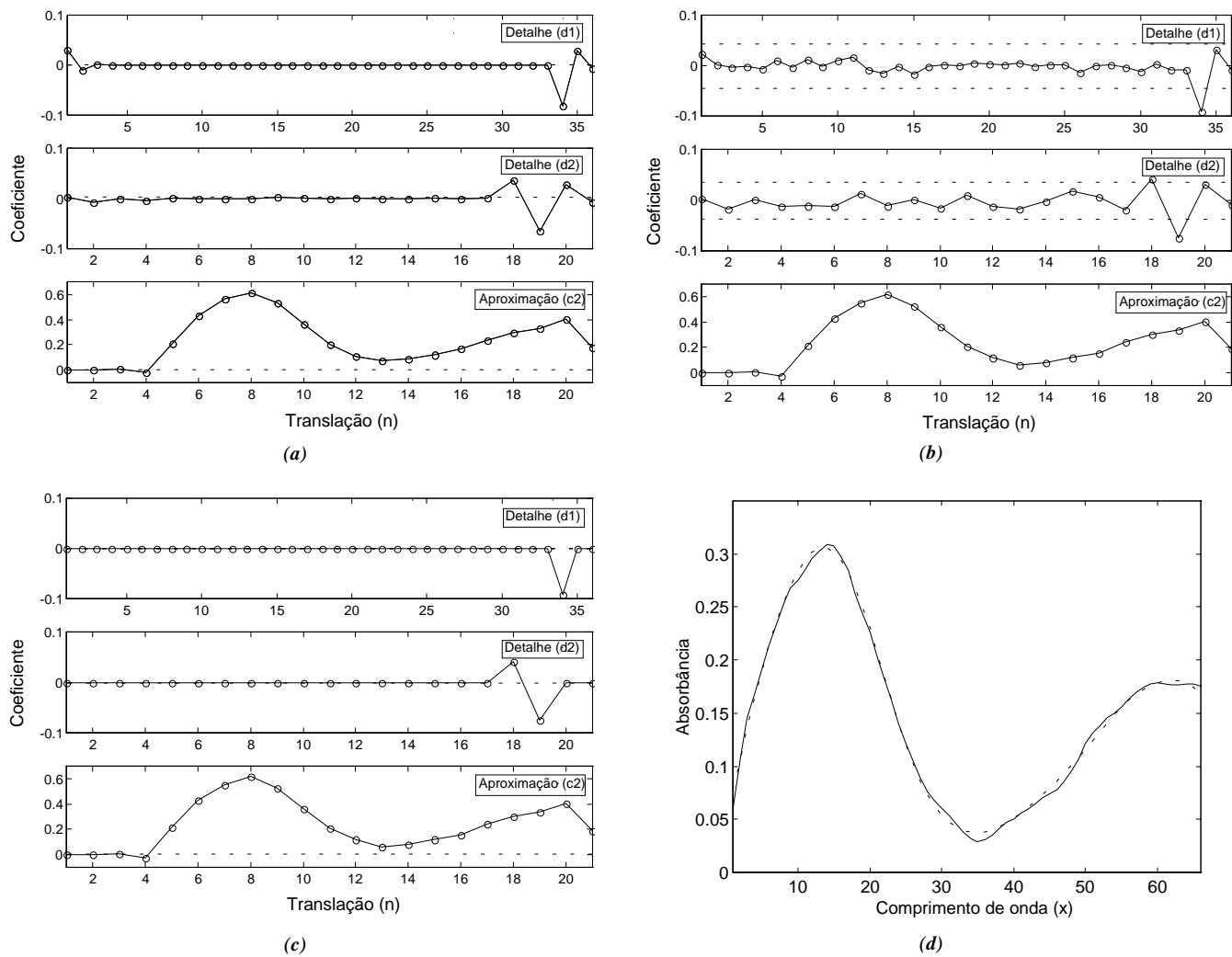


Figura 17. TW do espectro UV-VIS original (a), após a introdução de ruído (b) e após o corte (c). Espectro filtrado (linha cheia) e espectro original (linha tracejada) (d).

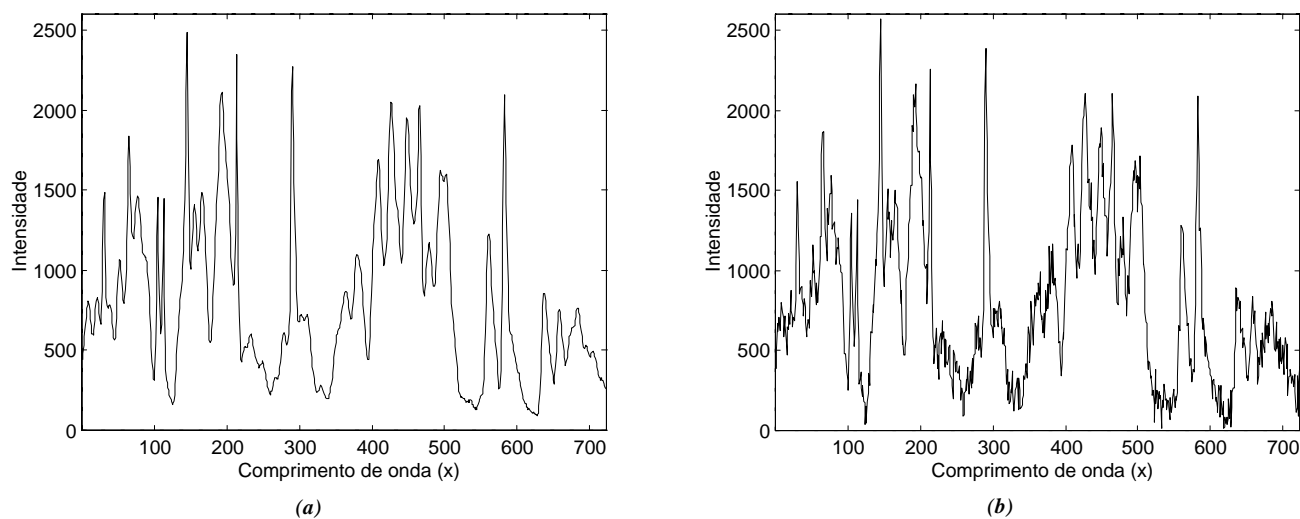


Figura 18. Espectro de emissão em plasma original (a) e após a introdução de ruído (b).



corte escolhida com base no mínimo do *RMSE* (Figura 19c) e representada por uma barra vertical na Figura 19b, obtém-se, após a TFI, o espectro mostrado na Figura 19d (*RMSE* = 61). Ao compararmos o espectro filtrado (linha cheia na Figura 19d) com o espectro original (linha tracejada na Figura 19d), verifica-se que a filtragem causou achatamento em alguns picos, destacados por elipses tracejadas na Figura 19d. Esse fenômeno pode ser mais facilmente compreendido analisando-se as curvas das Figuras 20 e 19c. Para a frequência de corte selecionada, o valor do *EAM* é de 447. À medida que a frequência de corte é reduzida, o *EAM* aumenta, indicando uma maior distorção nos picos. Por outro lado, se a frequência de corte é aumentada, de modo a reduzir o *EAM*, o ruído não é removido, acarretando um aumento no *RMSE* (Figura 19c). É mostrado a seguir que este problema pode ser minimizado quando a filtragem é feita via TW.

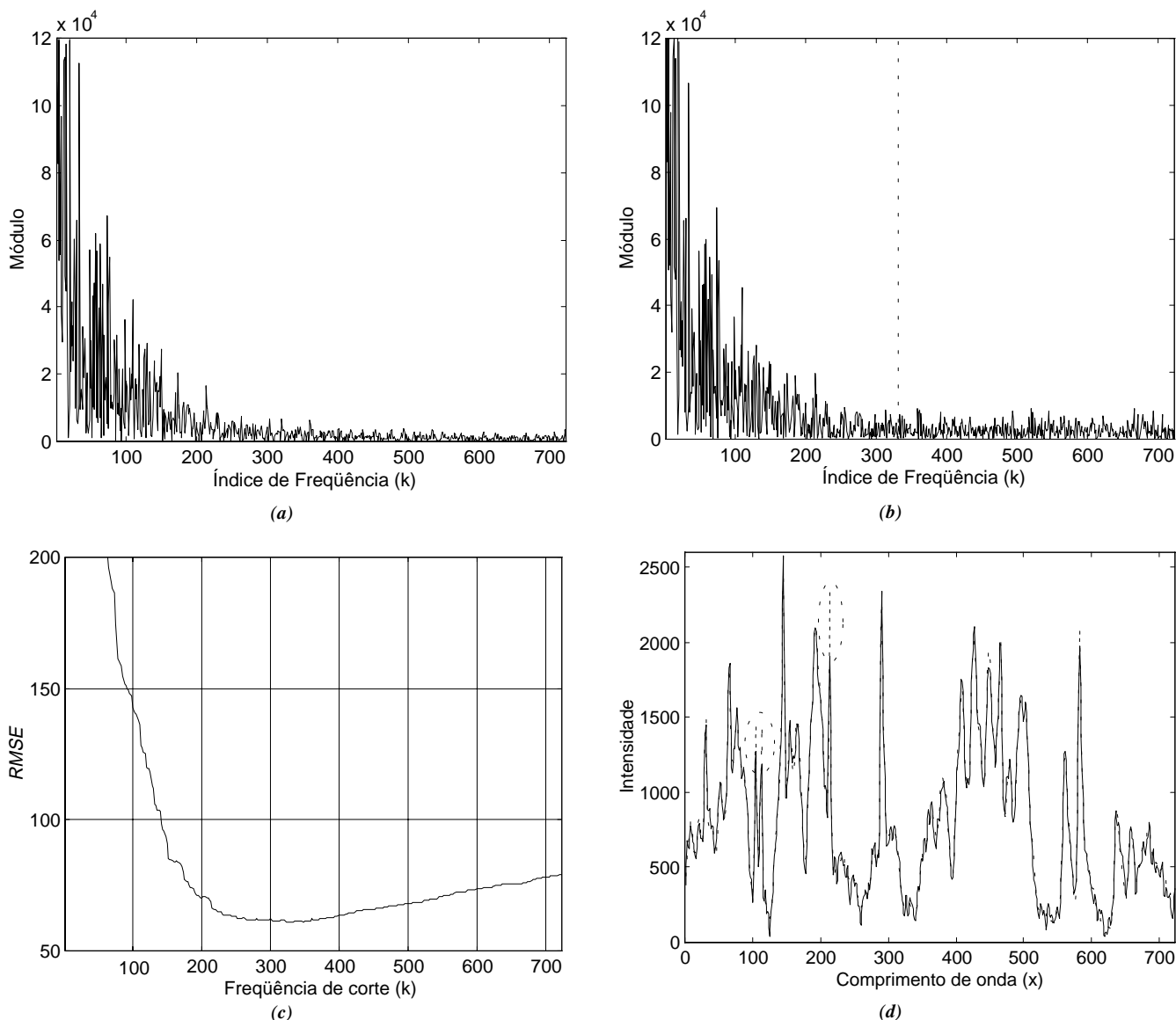
Seguindo-se um procedimento semelhante ao utilizado no espectro UV-VIS, concluiu-se que a melhor wavelet para filtragem do espectro de emissão em plasma é a Symlet 6, com

2 níveis de decomposição (Figura 2b) e limiar de corte de 50% (*RMSE* = 56). Possivelmente, a escolha recaiu sobre essa wavelet devido a suas características de simetria, também apresentadas por raias de emissão<sup>1</sup>.

O resultado da TW encontra-se nas Figuras 21a e 21b. Os limiares de corte para  $d_1$  e  $d_2$  estão representados por barras horizontais tracejadas na Figura 21b e o resultado do corte é mostrado na Figura 21c.

A Figura 21d apresenta o espectro filtrado após a TWI. Observa-se que, no global, a filtragem via TW (Figura 21d) foi melhor que a produzida pela TF (Figura 19d), o que se traduz em termos quantitativos por um *RMSE* de 56 contra 61. Além disso, o *EAM* é agora de 215, contra o valor de 447 obtido com a TF, como resultado de um menor achatamento dos picos.

Uma comparação final entre os resultados da filtragem via TF e TW pode ser feita observando-se a Figura 22. Como se vê, a TF provoca erros elevados nos picos de emissão mais estreitos, enquanto a TW conduz a um erro de filtragem mais “uniforme” ao longo do espectro.



**Figura 19.** TF do espectro de emissão em plasma original (a) e após a introdução de ruído (b). *RMSE* em função da frequência de corte (c). Espectro filtrado (linha cheia) e espectro original (linha tracejada) (d).

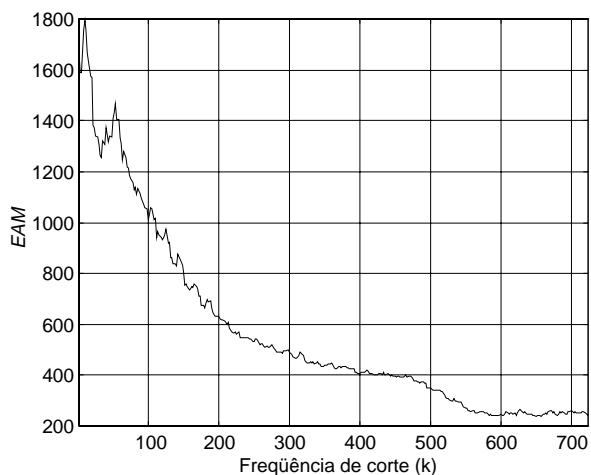


Figura 20. EAM em função da frequência de corte da TF do espectro de emissão em plasma.

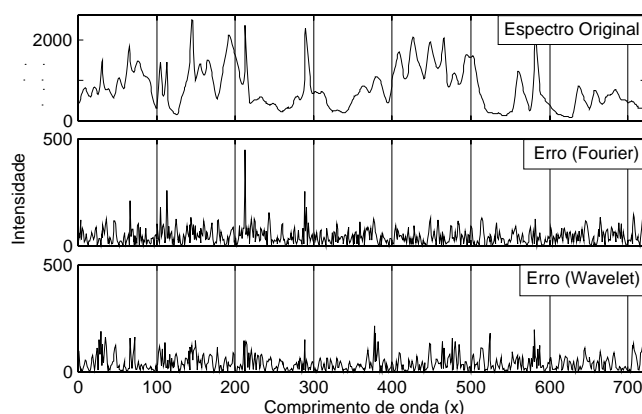
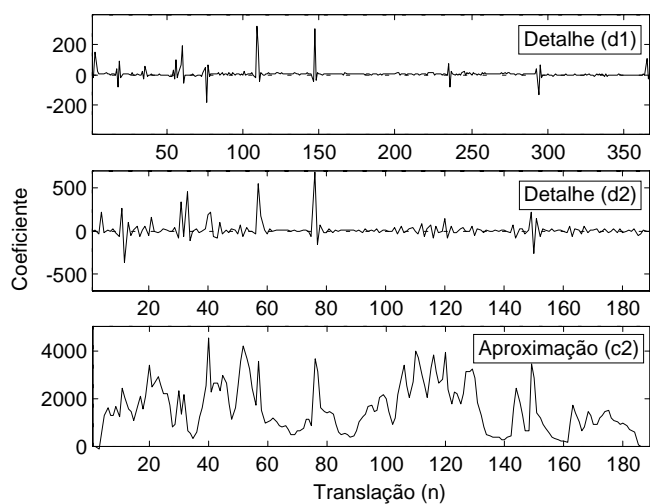
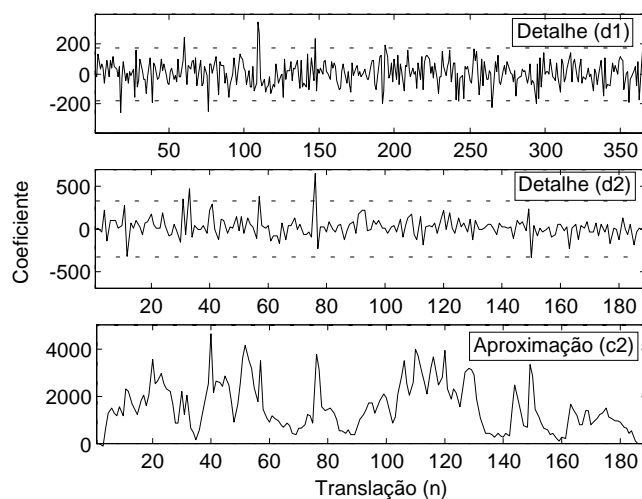


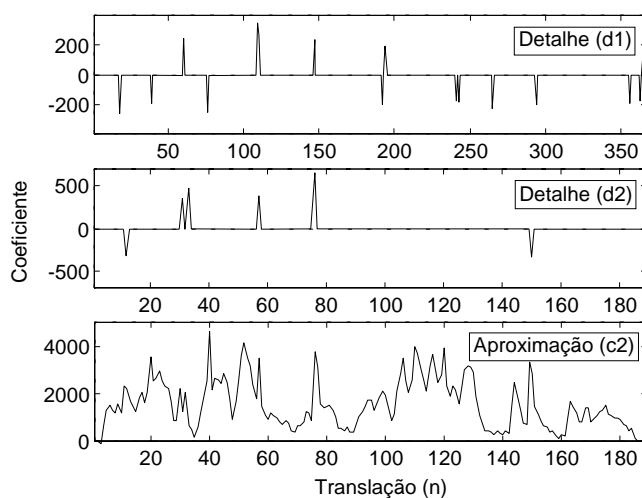
Figura 22. Comparação entre os erros de filtragem (em módulo) via TF e via TW. Em ambos os casos, os parâmetros de filtragem foram ajustados de modo a minimizar a raiz do erro quadrático médio (RMSE).



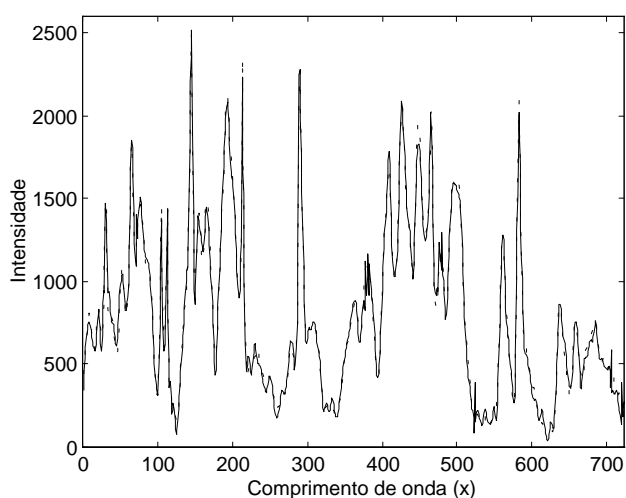
(a)



(b)



(c)



(d)

Figura 21. TW do espectro de emissão em plasma original (a), após a introdução de ruído (b) e após o corte (c). Espectro filtrado (linha cheia) e espectro original (linha tracejada) (d).

## CONCLUSÃO

Neste artigo, a filtragem de sinais instrumentais usando a Transformada de Fourier e a Transformada Wavelet é apresentada de forma simples e sucinta. Um estudo comparativo da aplicação destas técnicas de filtragem, mostra que, para sinais que apresentam poucos pontos notáveis, tais como espectros de absorção molecular UV-VIS, os resultados da filtragem via TF e TW são semelhantes. Contudo, para sinais com baixo “fator de preenchimento” (*filling factor*), como espectros de emissão em plasma, o uso da TW conduz a melhores filtrações, principalmente por não distorcer o formato de picos estreitos.

É importante notar que, neste trabalho, os parâmetros RMSE e EAM foram empregados com fins didáticos, isto é, para estabelecer uma base formal de comparação entre a TF e a TW. Na prática, contudo, como o sinal sem ruído não é conhecido, não é possível calcular um “erro de filtragem”. Assim, a escolha dos parâmetros (wavelet a ser empregada, número de níveis de resolução e limiar de corte) deve ser guiada por outros critérios. Infelizmente, embora diversas ferramentas possam ser usadas nesse sentido (gráficos de energia e histogramas de coeficientes, por exemplo), não se pode prescindir da sensibilidade do analista em julgar a qualidade do sinal filtrado. Vale ressaltar, contudo, que a escolha do ponto de corte na filtragem via TF também não é trivial, especialmente quando o ruído se espalha por uma faixa larga de frequências.

Um critério mais objetivo poderia ser utilizado se o sinal instrumental tiver por finalidade a determinação quantitativa de alguma grandeza, tal como a concentração de uma espécie química. Nesse caso, o erro na determinação dessa grandeza, em um conjunto de amostras de calibração, poderia ser usado como uma medida indireta da qualidade da filtragem. Essa possibilidade será investigada em trabalhos futuros do nosso grupo de pesquisa.

## AGRADECIMENTOS

Os autores expressam a sua gratidão pelos auxílios e bolsas da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES and CNPq (Projeto Nordeste de Pós-Graduação e Pesquisa - Processo 521172/97-7).

## REFERÊNCIAS

1. Skoog, D. A.; Holler, F. J.; Nieman, T. A.; *Principles of Instrumental Analysis*; 5.ed.; Harcourt Brace & Co.; Philadelphia, 1998.
2. Cerqueira, E. O.; Poppi, R. J.; Kubota, L.; Mello, C.; *Quim. Nova* **2000**, *23*, 690.
3. Depezynski, U., Jetter, K., Molt, K. e Niemoller, A.; *Chemo. Intell. Lab. Sys.* **1999**, *49*,151.
4. Daubechies, I.; *Ten Lectures on Wavelets*, SIAM; Philadelphia, 1992.
5. Misiti, M.; Misiti, Y.; Oppenheim, G.; Poggi, J. M.; *Wavelet Toolbox User's Guide*; The Mathworks: Natick, 1996.
6. Rioul, O., Vetterli, M.; *IEEE Signal Processing Magazine* **1991**, *8*, 14.
7. Meyer, Y.; *Wavelets-Algorithms and Applications*; SIAM; Philadelphia, 1993.
8. Alsberg, B. K.; Woodward, A. M.; Kell, D. B.; *Chem. Intell. Lab. Sys.* **1997**, *37*, 215.

9. Daubechies, I.; *Proc. IEEE* **1996**, *84*, 510.
10. Smrcok, L.; Durik, M.; Jorik, V.; *Powder Diffraction* **1999**, *14*, 300.
11. Pen, U. L.; *Phil. Trans. R. Soc. London A - Math. Phys. Eng. Sci.* **1999**, *357*, 2561.
12. Qian, S.; Chen, D.; *Joint Time-Frequency Analysis- Methods and Applications*; Prentice Hall PTR; Upper Saddle River, 1996.
13. Gottlieb, D.; Shu, C. W.; *Siam Review* **1997**, *39*, 644.
14. Strang, G.; Nguyen, T.; *Wavelets and Filter Banks*; Wellesley-Cambridge Press; Wellesley, 1996.
15. Abbate, A.; Koay, J.; Frankel, J.; Schroeder, S. C.; Das, P.; *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control* **1997**, *44*, 14.
16. Mallat, S.; *IEEE Trans. Acoust., Speech, and Signal Processing* **1989**, *37*, 2091.
17. Saldanha, T. C. B.; Araújo, M. C. U.; Neto, B. B.; Chame, H. C.; *Anal. Lett.* **2000**, *33*, 1187.
18. Pimentel, M. F.; Araújo, M. C. U.; Neto, B. B.; Pasquini, C.; *Spectrochim. Acta Part B* **1997**, *52*, 2151.
19. Oppenheim, A. V.; Schaffer, R. W.; *Discrete-Time Signal Processing*; Prentice-Hall; Englewood Cliffs, 1989.

## APÊNDICE

```
function ff = filtwav(f,lim,wname,niveis)
% Argumentos da função:
% f → Sinal a ser filtrado
% lim → Limiar de corte expresso como uma fração (0 a 1)
% do maior
% coeficiente em cada nível de detalhe
% wname → Nome da wavelet. Ex: 'db4'
% niveis → Número de níveis de resolução
% Obs: Requer o MATLAB Wavelet Toolbox

% Decomposição wavelet
[C,L] = wavedec(f,niveis,wname);
% Obtém os coeficientes de aproximação
A = appcoef(C,L,wname);
% Os coeficientes de aproximação são mantidos na filtragem
Cf = A;
for n = niveis:-1:1 % Para cada nível de detalhe
% Obtém os coeficientes de detalhe
D = detcoef(C,L,n);
% Calcula o módulo de cada coeficiente
MD = abs(D);
% Calcula o limiar de corte como uma fração
% do maior módulo
limiar = lim*max(MD);
% Encontra os coeficientes que, em módulo,
% estão abaixo do limiar
i = find(MD<limiar);
% Zera esses coeficientes
Df = D; Df(i) = zeros(1,length(i));
% Guarda o vetor de coeficientes resultante
Cf = [Cf Df];
end
% Obtém o sinal filtrado a partir dos coeficientes wavelet
ff = waverec(Cf,L,wname);
```