Bank of Canada Banque du Canada

Working Paper 2002-3 / Document de travail 2002-3

An Introduction to Wavelets for Economists

by

Christoph Schleicher

ISSN 1192-5434

Printed in Canada on recycled paper

Bank of Canada Working Paper 2002-3

January 2002

An Introduction to Wavelets for Economists

by

Christoph Schleicher

Monetary and Financial Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9

Contents

Acknowledgements

I would like to thank Paul Gilbert, Pierre St-Amant, and Greg Tkacz from the Bank of Canada for helpful comments.

Abstract

Wavelets are mathematical expansions that transform data from the time domain into different layers of frequency levels. Compared to standard Fourier analysis, they have the advantage of being localized both in time and in the frequency domain, and enable the researcher to observe and analyze data at different scales. While their theoretical foundations were completed by the late 1980s, the 1990s saw a rapid spread to a wide range of applied sciences. A number of successful applications indicate that wavelets are on the verge of entering mainstream econometrics. This paper gives an informal and non-technical introduction to wavelets, and describes their potential for the economic researcher.

JEL classification: C1
Bank classification: Econometric and statistical methods

Résumé

Les ondelettes sont des expansions mathématiques qui transforment les données du domaine temporel en différentes strates de fréquences. Elles présentent l'avantage, par rapport à l'analyse de Fourier standard, d'être localisées aussi bien dans le domaine temporel que dans celui des fréquences et de permettre au chercheur d'observer et d'analyser des données à différentes échelles. Les ondelettes, dont les fondements théoriques ont été mis au point à la fin des années 1980, se sont rapidement étendues à un large éventail de sciences appliquées au cours de la décennie 1990. À en juger par les résultats fructueux obtenus à l'égard d'un certain nombre d'applications, elles sont sur le point de s'ajouter à la panoplie d'outils couramment utilisés par l'économètre. Dans cette étude, l'auteur présente une introduction informelle et non technique aux ondelettes et expose les possibilités qu'elles ouvrent sur le plan de la recherche en économie.

Classification JEL : C1
Classification de la Banque : Méthodes économétriques et statistiques

1. Introduction

Wavelets can be compared to a wide-angle camera lens that allows one to take broad landscape portraits as well as zoom in on microscopic detail that is normally hidden to the human eye. In mathematical terms, wavelets are local orthonormal bases consisting of small waves that dissect a function into layers of different scale.

Wavelet theory has its roots in Fourier analysis, but there are important differences. The Fourier transformation uses a sum of sine and cosine functions at different wavelengths to represent a given function. Sine and cosine functions, however, are periodic functions that are inherently *non-local*; that is, they go on to plus and minus infinity on both ends of the real line. Therefore, any change at a particular point of the time domain has an effect that is felt over the entire real line. In praxis, this means that we assume the frequency content of the function to be stationary along the time axis. To overcome this restriction, researchers invented the windowed Fourier transform. The data are cut up into several intervals along the time axis and the Fourier transform is taken for each interval separately.

Wavelets, on the other hand, are defined over a finite domain. Unlike the Fourier transform, they are localized both in time and in scale (see Figure 1). They provide a convenient and efficient way of representing complex signals. More importantly, wavelets can cut data up into different frequency components for individual analysis. This scale decomposition opens a whole new way of processing data. As Graps (1995) states, wavelets enable us to see both the forest *and* the trees.

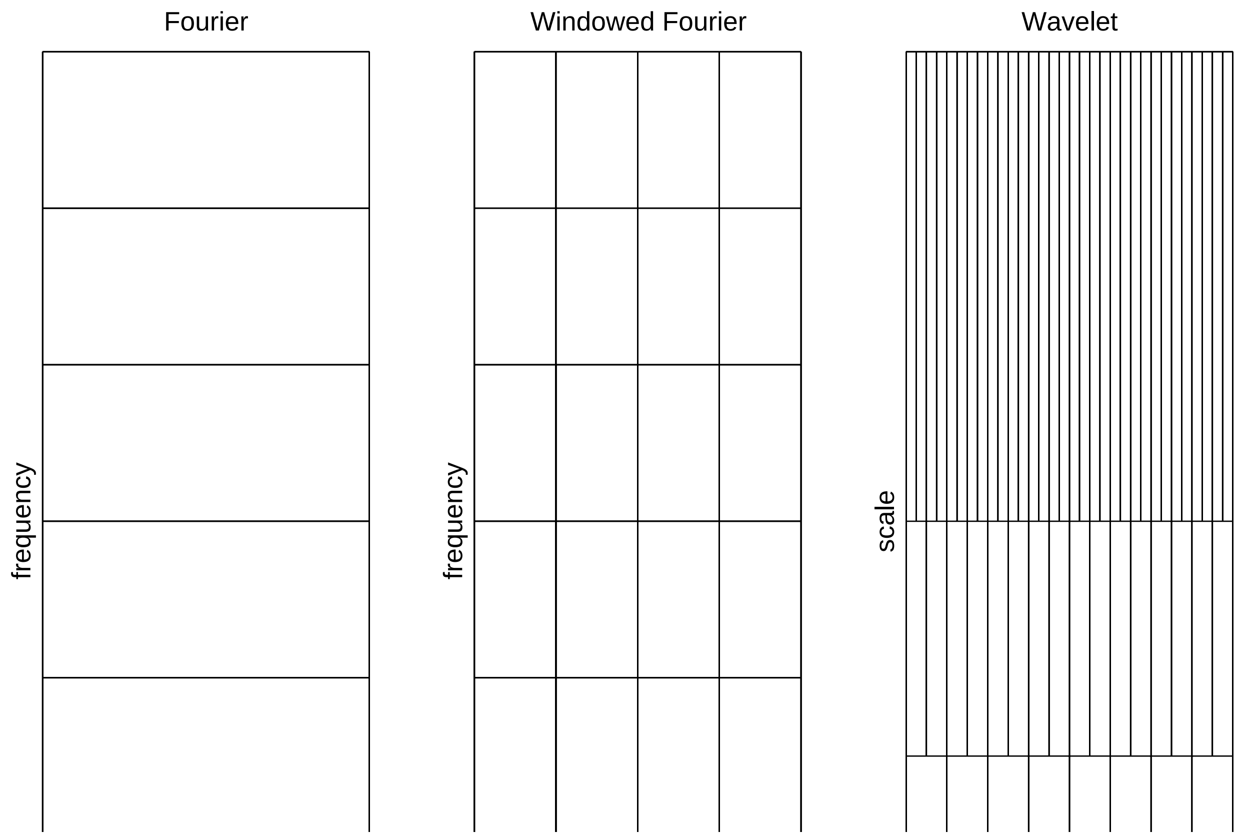
A wavelet basis consists of a father wavelet that represents the smooth baseline trend and a mother wavelet that is dilated and shifted to construct different levels of detail. This resembles the building plan of a natural organism that is based on self-similarity. At high scales, the wavelet has a small time support, enabling it to zoom in on details such as spikes and cusps, and on short-lived phenomena such as delta functions. At low scales, wavelets capture long-run phenomena. Their ability to adapt their scale and time support enables them to escape Heisenberg's curse; i.e., the law that says that one cannot be simultaneously precise in the time and the frequency domain.

After their theoretical foundations were completed by the late 1980s, wavelets began to enter the applied sciences. One of their first applications was in earthquake prediction. Wavelets provided a time dimension to non-stationary seismic signals that Fourier analysis lacked. Realizing their potential for compressing data, the FBI in 1992 reorganized its entire fingerprint database using wavelets.[1] Another early example, cited by Vidakovic and Mueller (1994), includes the de-noising

1. Most current image-compressing tools, like JPEG, are still based on Fourier analysis, but the new JPEG2000 standard will use both Fourier and wavelets. For sound compression, as in MP3, Fourier appears to be preferable, since sound signals naturally consist of sine and cosine signals.

of old recordings of Brahms playing his “First Hungarian Dance” on piano. Wavelets are now applied in a wide range of fields, from fractals and partial differential equations in mathematics to signal and image processing, speech recognition, software design, engineering, meteorology, and statistics, among others.

Figure 1: Time-Frequency Plane for Fourier and Wavelet Transform



Most wavelet-related research in economics has been done in the last few years and only a relatively small number of papers have been published. Judging from the large number of working papers on wavelets and their numerous statistical applications in physics, engineering, and biomedical research, the number of publications is expected to grow rapidly.

Their affinity to the Fourier transform makes wavelets an ideal candidate for frequency domain analysis in time-series econometrics. Conversely, their capability to simultaneously capture long-term movements and high-frequency details are very useful when dealing with non-stationary and complex functions. Wavelet estimators have also been used in connection with fractionally integrated processes that have long-memory properties. Decomposing a time series into different scales may reveal details that can be interpreted on theoretical grounds as well as be used to improve forecast accuracy. The theoretical point of interest is the observation that economic actions and decision-making take place at different scales. Forecasting seems to improve at the scale level, because forecasting models like autoregressive moving average (ARMA) or neural networks can extract information from the different scales that are hidden in the aggregate.

2. Wavelet Evolution

In 1807, the French mathematician Joseph Fourier asserted that any 2π-periodic function could be represented by a superposition of sines and cosines. The Fourier series and its coefficients are given by

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cdot \cos(kx) + b_k \cdot \sin(kx))$$

$$a_k = \int_0^{2\pi} f(x)\cos(kx)dx \text{ and } b_k = \int_0^{2\pi} f(x)\sin(kx)dx .$$

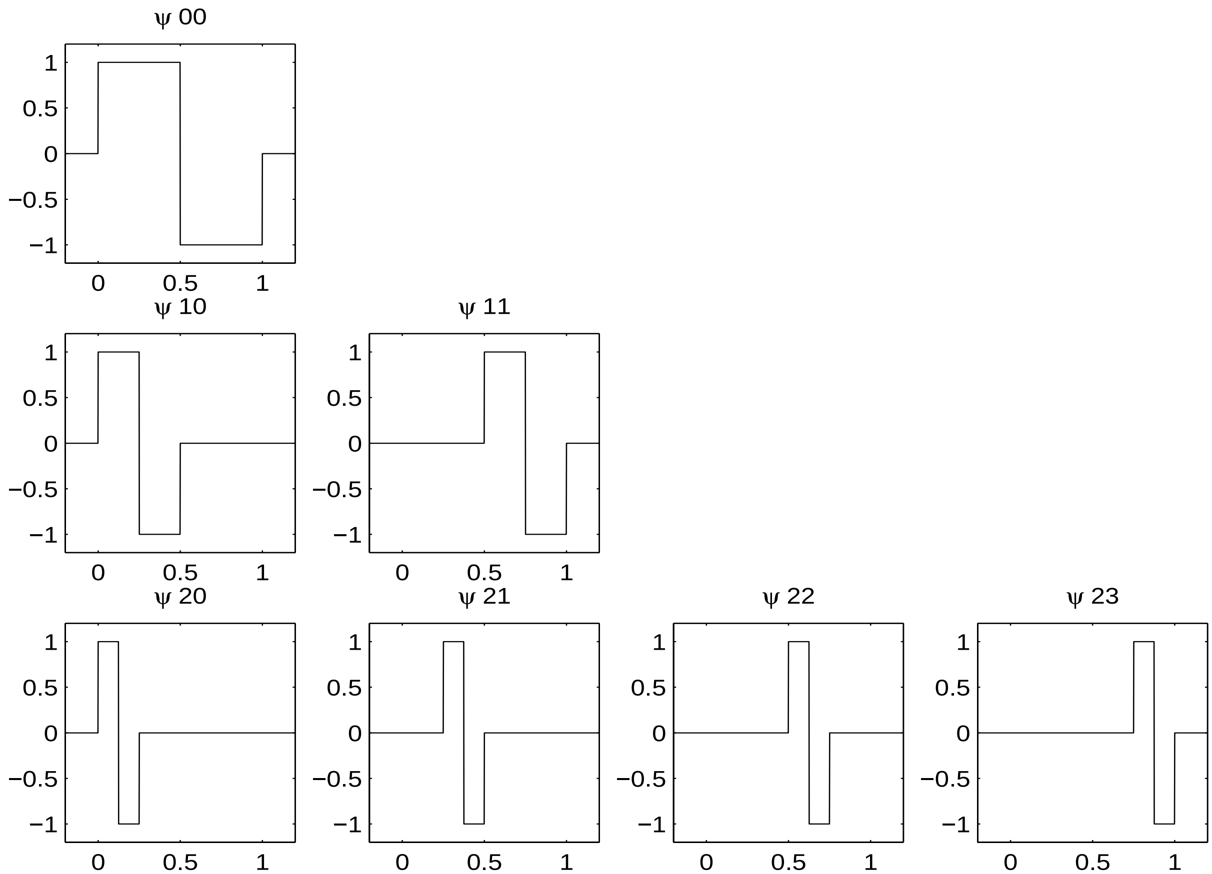
Using Euler's result that $e^{iy} = \cos(y) + i \cdot \sin(y)$, the Fourier transform $F(\omega) = \int_{-\infty}^{\infty} f(x)e^{-iwx}$ serves as a bridge between the time domain and the frequency domain. It gave rise to a new field of mathematics: frequency analysis.

Fourier transforms have played a dominant role in many areas of applied and pure science. Researchers, however, have kept looking for new mathematical structures that enabled them to localize functions in both the time and the frequency domain, a quest that led them from frequency analysis to *scale analysis*. The idea was to construct local basis functions that could be shifted and scaled and then reassembled to approximate a given signal.

In 1910, Alfred Haar constructed the first known wavelet basis by showing that any continuous function f(x) on [0,1] can be approximated by a series of step functions. The Haar basis is described in more detail in section 3. Some of its elements are shown in Figure 2, which illustrates the concept of dilation and translation: a basic *mother wavelet* ψ_{00} is dilated by narrowing its support and shifted (translated) along the unit interval. The individual elements are linearly independent (in fact, orthogonal) and can represent stepfunctions of increased fineness.

Haar's discovery played an important part in the work of the physicist Paul Levy, who in the 1930s used the Haar basis function to redefine Brownian motion. Since Brownian motion is non-differentiable by definition, the ability of the Haar basis function to describe small and complicated detail proved superior to the Fourier basis functions.

Figure 2: The Haar Wavelet at Three Dilation Levels



Further progress was made in the 1960s by the mathematicians Guido Weiss and Ronald R. Coifman, who were looking for common “assembly rules” to build all elements of a function space by its most primitive elements, called “atoms.” In 1980, Grossman, a mathematician, and Morlet, an engineer, gave the first formal definition of a wavelet in the context of quantum physics. The most commonly used definition, however, dates back to 1981, when Strömberg constructed orthonormal bases of the form $\psi_{jk}(x) = 2^{j/2}\,\psi(2^j x - k)$. Here, $\psi(x)$ is a mother wavelet that is dilated and translated with parameters j and k, respectively.

Wavelets entered mainstream science with Stephane Mallat’s work in digital signal processing. In 1985, he discovered the relationship between quadrature mirror filters (a pair of high- and low-pass filters) and orthonormal wavelet bases. Mallat’s multiresolution analysis builds on an iterative filter algorithm (called a pyramid algorithm) and it is the cornerstone of the fast wavelet transform (FWT), the wavelet pendant to the fast Fourier transform (FFT).[2] The last important

2. The FFT uses matrix factorization to decrease the number of operations.

step in the evolution of wavelet theory occurred in 1988, when Ingrid Daubechies constructed "consumer-ready" wavelets with a preassigned degree of smoothness.

3. A Bit of Wavelet Theory

Although they can have infinitely different shapes, all wavelets share the same basic construction plan. Given a mother wavelet $\psi(x)$, an orthonormal basis $\{\psi_{jk}(x)\}$ in $L^2(R)$, the space of square integrable functions, is defined by

$$\psi_{jk}(x) = 2^{j/2}\,\psi(2^j x - k).$$

The oldest and, for demonstration purposes, most useful wavelet is the Haar function, a step function defined as

$$\begin{aligned}\psi(z) &= 1\ (\,z \in [0, 0.5)\,)\\ &= -1\ (\,z \in [0.5, 1)\,).\end{aligned}$$

The parameters j and k dilate and translate the function, as shown in Figure 2 for $j = 0, 1, 2$, and $k = 0, .., 2^{j-1}$. Increasing j makes the Haar function finer, while k shifts it from the left to the right. ψ_{jk} is an orthonormal basis, because it

(i) is orthogonal $\int(\psi_{jk} \cdot \psi_{lm}) = 0, ((j \neq l) \vee (k \neq m))$, and

(ii) has an L^2 norm of unity $\int(2^{j/2}\psi(2^j x - k))^2 dx = 1$.

The scaling factor $2^{j/2}$ helps achieve the latter result.

Let y be a data vector with 2^n elements that can be represented by a piecewise constant function, $f(x)$ on [0,1].[3] The wavelet transformation of f(x) is then given by

$$f(x) = c_0\phi(x) + \sum_{j=0}^{n-1}\sum_{k=0}^{2^j-1} c_{jk}\psi_{jk}(x)\ .$$

Here, $\phi(x)$ is the father wavelet, also referred to as the scaling function that represents the coarsest components or the smooth baseline trend of the function. For the simplest wavelet, the Haar

3. Any vector y can be compressed into a function defined on the unit interval by applying the dilation formula $f(x) = \sum_{k=0}^{2^n-1} y_k \cdot 1(k2^{-n} \le x < (k+1)2^{-n})$.

wavelet, the father wavelet is just a horizontal line equal to one. For a more complicated wavelet, the Daubechies(4) wavelet (discussed in section 4), the father wavelet and the mother wavelet are shown in Figure 3. While the father wavelet integrates to one, the mother wavelet integrates to zero, reflecting the fact that it is used to represent differences in the data that average out to zero. Figure 4 gives an intuitive visualization of a Daubechies wavelet basis in the time-scale space. The father wavelet covers the whole time support at the lowest scale of resolution, while the mother wavelet is dilated and translated to capture different levels of fineness. The wavelet transform consists of the vector of all coefficients $w = [\ c_0, c_{00}, c_{10}, c_{11}, c_{20}, c_{21}, \ldots\ ]$.

Figure 3: The Daubechies(4) Father and Mother Wavelet

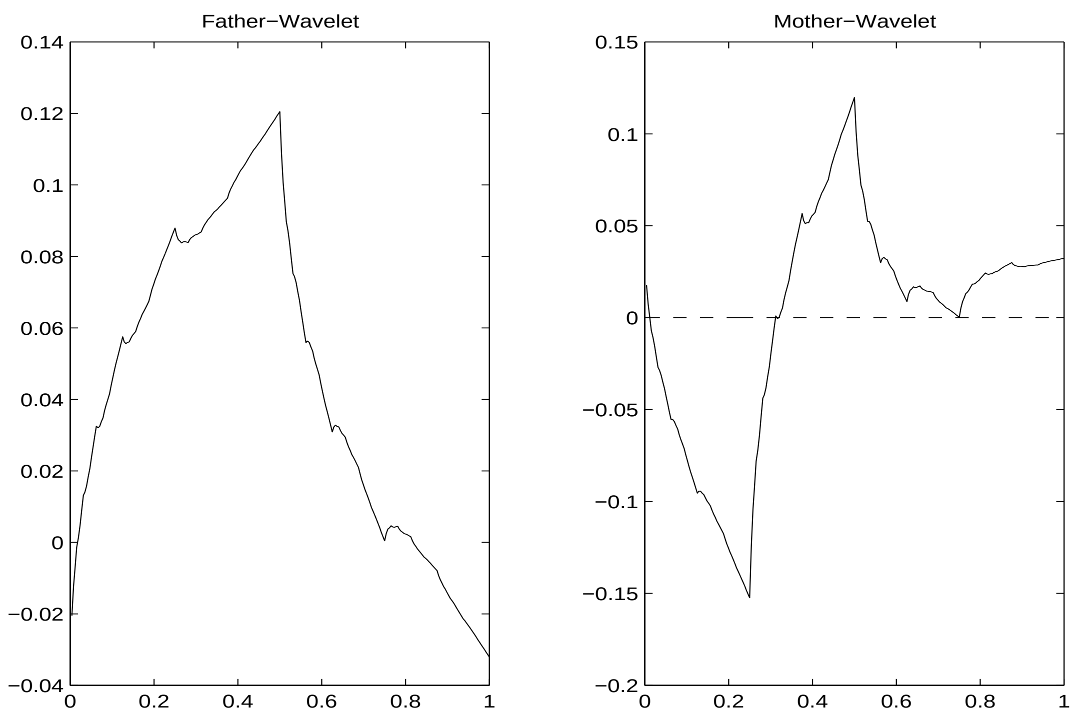
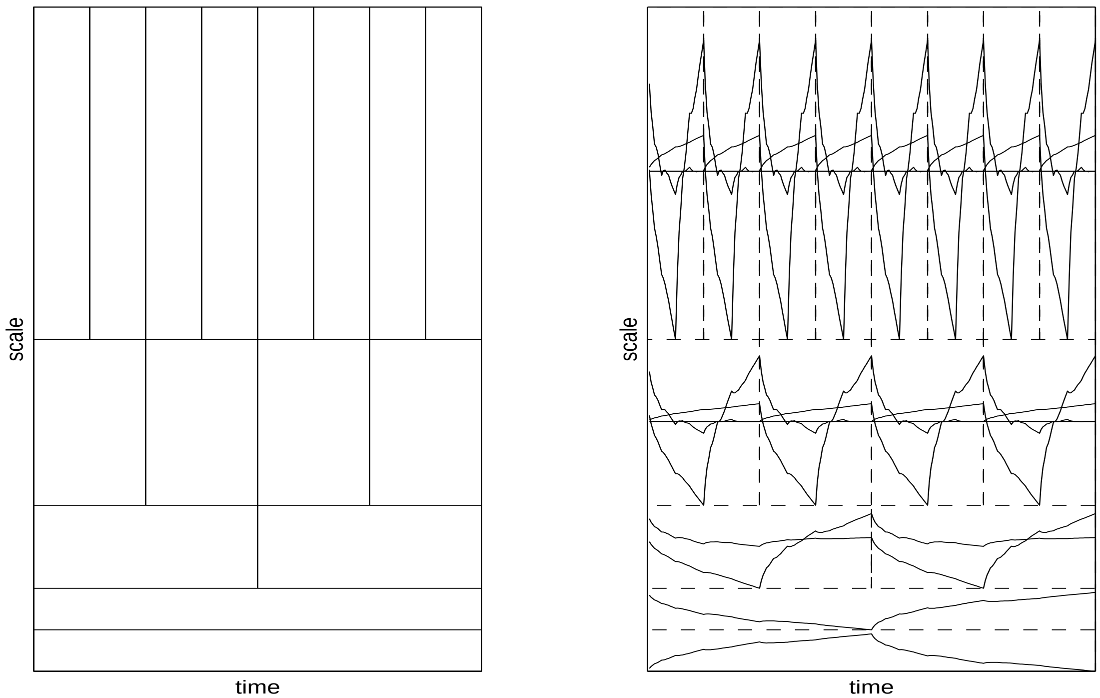


Figure 4: Daubechies(4) Wavelet Basis



Consider, for example, the four-dimensional vector $y = [\ 2, 5, 2, 7\ ]'$. Its wavelet representation is given by $y = W\,c$, where W contains the Haar basis vectors, $W = \begin{bmatrix} \phi_{00} & \psi_{00} & \psi_{10} & \psi_{11} \end{bmatrix}$:

$$\begin{bmatrix} 2 \\ 5 \\ 2 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & -1 & 0 & -\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_{00} \\ c_{10} \\ c_{11} \end{bmatrix}.$$

The matrix of basis vectors W can be inverted easily, since the inverse of any orthogonal matrix is equal to its transpose divided by 4.[4] The solution for the wavelet coefficients is then given by

4. In general, the inverse of a (real valued) orthogonal matrix A with dimension n is $\alpha A'$, with $\alpha = |A|^{-\frac{2}{n}}$.
For a complex orthogonal matrix $A^{-1} = modulus|A|^{-\frac{2}{n}} conjugate(A)'$.

$$\begin{bmatrix} c_0 \\ c_{00} \\ c_{10} \\ c_{11} \end{bmatrix} = \frac{1}{4} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 5 \\ 2 \\ 7 \end{bmatrix} = \begin{bmatrix} 4 \\ -1/2 \\ -\frac{3}{2\sqrt{2}} \\ -\frac{5}{2\sqrt{2}} \end{bmatrix}.$$

A similar representation could have been achieved by any orthogonal basis, such as the identity matrix I_4. However, the wavelet transform has the advantage of decomposing the data into different scales; that is, different levels of fineness. The vector of wavelet coefficients from our example consists of three levels (or scales): c_0, c_{00}, and $c_1 = (c_{10}c_{11})$. Setting the last level equal to zero and premultiplying with *W* sets the input vector *y* equal to [7/2 7/2 9/2 9/2]'; that is, the first two and the last two elements are averaged. In signal processing this is equivalent to applying a *low-pass filter*. Setting the last two levels equal to zero results in the transformed input vector [4 4 4 4]', the mean of *y*. Conversely, we could set all coefficients, except c_{00}, equal to zero and invert the transform by multiplying by *W*. The result would be the vector [-1/2 -1/2   1/2   1/2]', the difference between the mean and the second level of smoothness [7/2 7/2 9/2 9/2]'. Finally, setting all coefficients except c_{10} c_{11} equal to zero and reversing the transform gives the vector [-3/2   3/2 -5/2   5/2], the difference between the second level of smoothness and the original data. We can therefore use the wavelet decomposition to represent the vector *y* as the sum of its smooth component, S_2, and detail components D_2 and D_1:

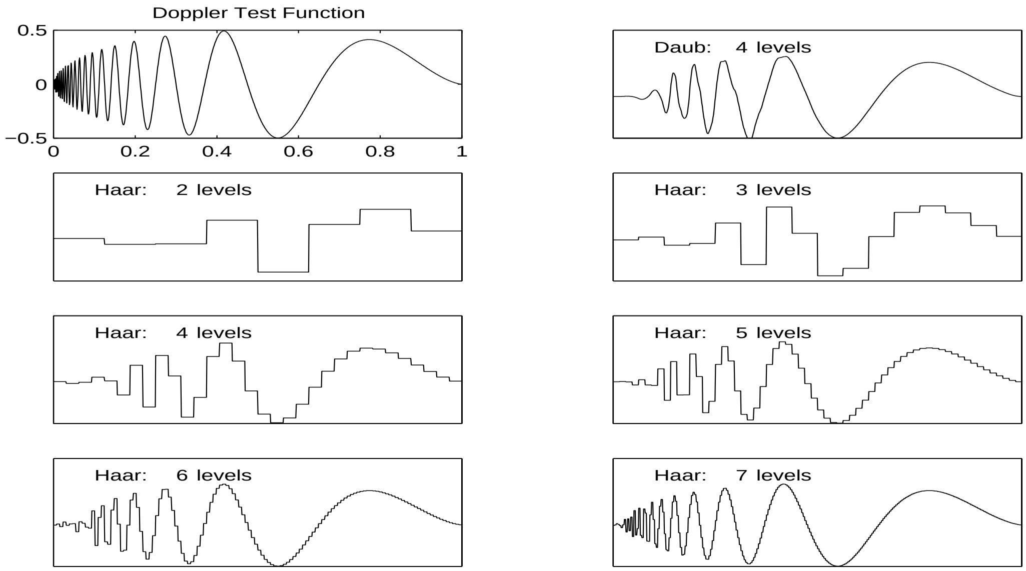
$$y = S_2 + D_2 + D_1 = \begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \end{bmatrix} + \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix} + \begin{bmatrix} -3/2 \\ 3/2 \\ -5/2 \\ 5/2 \end{bmatrix}.$$

Figure 5 shows the wavelet transformation for a more complicated function, the *Doppler* function (taken from Vidakovic 1999).[5] Each additional level doubles the resolution and adds more detail to the function. It is also clear that the Haar wavelet is not the optimal choice for continuous and smooth functions, because a Daubechies wavelet achieves a much better approximation at four levels of depth.

5. The Doppler function is defined as:

$$f(x) = x(1-x) \cdot \sin\frac{2.1\pi}{x+0.05}.$$

Figure 5: Approximation of the Doppler Function using Different Levels of Fineness[6]



For a comparison, we look at the discrete Fourier transform of the vector *y*,

$$f(x) = \sum_{k=0}^{3} a_k e^{ikx},$$

or in matrix form *f(x) = Fa*,

$$\begin{bmatrix} f(0\pi/2) \\ f(1\pi/2) \\ f(2\pi/2) \\ f(3\pi/2) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

As before, *F* constitutes an orthogonal basis and can be inverted by transposing its complex conjugate and dividing by 4:

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \frac{1}{4} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & (-i) & (-i)^2 & (-i)^3 \\ 1 & (-i)^2 & (-i)^4 & (-i)^6 \\ 1 & (-i)^3 & (-i)^6 & (-i)^9 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 5 \\ 2 \\ 7 \end{bmatrix} = \begin{bmatrix} 16 \\ 2i \\ -8 \\ -2i \end{bmatrix}.$$

6. Smooth wavelets like the Daubechies_4 wavelet (upper right corner) give better approximations to continuous functions than the Haar wavelet.

The Fourier coefficients represent the energy content of different frequencies. For a power spectogram, one uses the absolute value of the squared coefficients, in this case,

$$s = [\ 196, 4, 64, 4\ ]'.$$

3.1 Mallat's multiscale analysis[7]

Calculating wavelet expansions directly by matrix inversion is computationally intensive. A big breakthrough came in the mid-1980s when Mallat introduced methods from signal processing theory to wavelets. Using a technique called quadrature mirror filtering, he showed that any discrete wavelet transformation can be calculated rapidly using a cascade-like algorithm. This realization not only added to the general understanding of wavelets, it reduced the number of operations necessary for the transform to the order of n from the order of n log(n), thus making it faster than the FFT. Technically, a multiscale or multiresolution analysis projects a function on a set of closed subspaces:

$$\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \ldots$$

Furthermore, since the subspaces are nested, one can represent V_j as the direct sum of the coarsely approximated subspace V_{j-1} and its orthogonal complement, W_{j-1}:

$$V_j = V_{j-1} \oplus W_{j-1} .$$

The multiscale analysis then uses filters to split up a function $f_N \in V_N$ into different components that belong to subspaces V_{N-i} (i = 1,2,..) and their orthogonal complements:

$$f_N = f_{N-1} + g_{N-1} = \sum_{i=1}^{M} g_{N-M} + f_{N-M} .$$

Each represents a different scale of the function. One can think of the subspaces V_i as different levels of magnification, revealing more and more detail. They are self-similar, such that

$$f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}, j \in Z .$$

In the language of signal processing, we obtain the pair of functions f_{N-1} and g_{N-1} by applying a pair of *quadrature mirror filters* to the original function f_N. In a discrete setting, the quadrature mirror filters are a pair of sequences, $\{h(k)\}$ and $\{g(k)\}$, for $k \in Z$. $h(k)$ is a *low-pass* or *low-band*

7. In the following, the terms multiscale analysis and multiresolution analysis will be used interchangeably.

filter, while $g(k)$ is a *high-pass* or *high-band* filter. Intuitively, the low-pass filter makes the data smoother and coarser, while the high-pass filter retains the detailed information. The two filters are connected through the relation

$$g(n) = (-1)^n h(1-n) .$$

Each wavelet has a scaling function, or father wavelet, ϕ , that spans the space V_0 and can be represented by a linear combination of functions from the next subspace, V_1. Since the subspaces are self-similar and nested, there exists a relationship between scaling functions of any two neighbouring subspaces, V_j and V_{j+1}. This relationship is called the *scaling equation* or *dilation equation* and defines the filter coefficients:

$$\phi(x) = \sum_{k \in Z} h_k \sqrt{2}\phi(2x-k) .$$

In praxis, both filters are mappings from *l(Z)* to *l(2Z)*; that is, they transform a vector with n elements into two vectors with $n/2$ elements each, one of which contains the data smoothed by the low-pass filter, with the other containing the detail that was removed. Each wavelet can be characterized by a finite set of filter coefficients derived from the scaling equation that relates scaling functions of different subspaces, V_i, to each other. For the Haar wavelet, the filter coefficients are

$$h(1) = h(2) = \frac{1}{\sqrt{2}} \quad \text{and} \quad g(1) = \frac{1}{\sqrt{2}}, g(2) = \frac{-1}{\sqrt{2}} .$$

Using the vector f(2) = [2 5 2 7]’ from our earlier example, the filtered vectors are given by

$$f(1) = \begin{bmatrix} \frac{7}{\sqrt{2}} & \frac{9}{\sqrt{2}} \end{bmatrix} \text{ and } g(1) = \begin{bmatrix} \frac{-3}{\sqrt{2}} & \frac{-5}{\sqrt{2}} \end{bmatrix} .$$

$f(1)$ is a weighted average of the first two and the second two entries of $f(2)$, respectively, where the filter coefficients are used as weights. The same procedure is used for $g(1)$, except that one of the filter coefficients is negative, so that the weighted average is actually a difference, cutting out detail from $f(1)$. It is convenient to use operators H and G to denote the filter relations applied to a sequence $a = \{a_n\}$:

$$(Ha)_k = \sum_n h(n-k) a_n ,$$

$$(Ga)_k = \sum_n g(n-k)a_n .$$

Let the original signal be $c^{(n)}$, a vector with 2^n elements. Then, $c^{(n-1)} = H\, c^{(n)}$ and $d^{(n-1)} = G\, c^{(n)}$. Applying the low-pass filter twice yields $c^{(n-2)} = H^2\, c^{(n)}$ and $d^{(n-2)} = HG\, c^{(n)}$. Using the multiresolution analysis, the discrete wavelet transform of a sequence $y = c^{(n)}$ of length 2^n is another sequence of equal length, given by

$$w = [\ d^{(n-1)}, d^{(n-2)}, \ldots , d^{(1)}, d^{(0)}, c^{(0)}\ ] = [\ Gy, GHy, GH^2y, \ldots , GH^{n-1}y, GH^ny, H^ny\ ].$$

That is, the wavelet transform consists of all layers of detail, going from fine to coarse, stacked next to each other. Figure 6 shows the graphical interpretation of the multiresolution analysis for the Doppler function from Figure 2 for a Haar wavelet transformation.

To invert the wavelet transform, an inverse filtering procedure is applied. The operators $\hat{H}$ and $\hat{G}$ map a sequence from $l(2Z)$ into $l(Z)$, and each element is doubled and multiplied by the filter coefficients. Consider, for example, applying the inverse low-pass and the inverse high-pass filters to the vectors $f(1)$ and $g(1)$:

$$\hat{H} f(1) = \begin{bmatrix} \frac{7}{2} & \frac{7}{2} & \frac{9}{2} & \frac{9}{2} \end{bmatrix}, \hat{G} g(1) = \begin{bmatrix} \frac{-3}{2} & \frac{3}{2} & \frac{-5}{2} & \frac{5}{2} \end{bmatrix}.$$

Adding up the two expressions reproduces the initial vector, $f(2)$.

Figures 7 and 8 give a detailed example of the decomposition and the inverse transform of a 2^3 vector using the Haar wavelet.

The wavelet transformation in multiresolution analysis form is then given by $w = [\ c(0), d(0), d(1), d(2)\ ]$, or

$$w = \begin{bmatrix} \frac{13}{2\sqrt{2}} & \frac{3}{2\sqrt{2}} & -1 & \frac{5}{2} & \frac{5}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{-3}{\sqrt{2}} & \frac{-4}{\sqrt{2}} \end{bmatrix}.$$

To obtain the wavelet coefficients c_{jk}, one needs to multiply w by $2^{-N/2}$, such that for the equation $y = Wc$, c becomes

$$c' = \begin{bmatrix} \frac{13}{8} & \frac{3}{8} & -\frac{1}{2\sqrt{2}} & \frac{5}{4\sqrt{2}} & \frac{5}{4} & \frac{1}{4} & \frac{-3}{4} & -1 \end{bmatrix},$$

where W is the matrix representing the Haar wavelet basis for eight dimensions.

Figure 6: Multiresolution Analysis of the Doppler Function

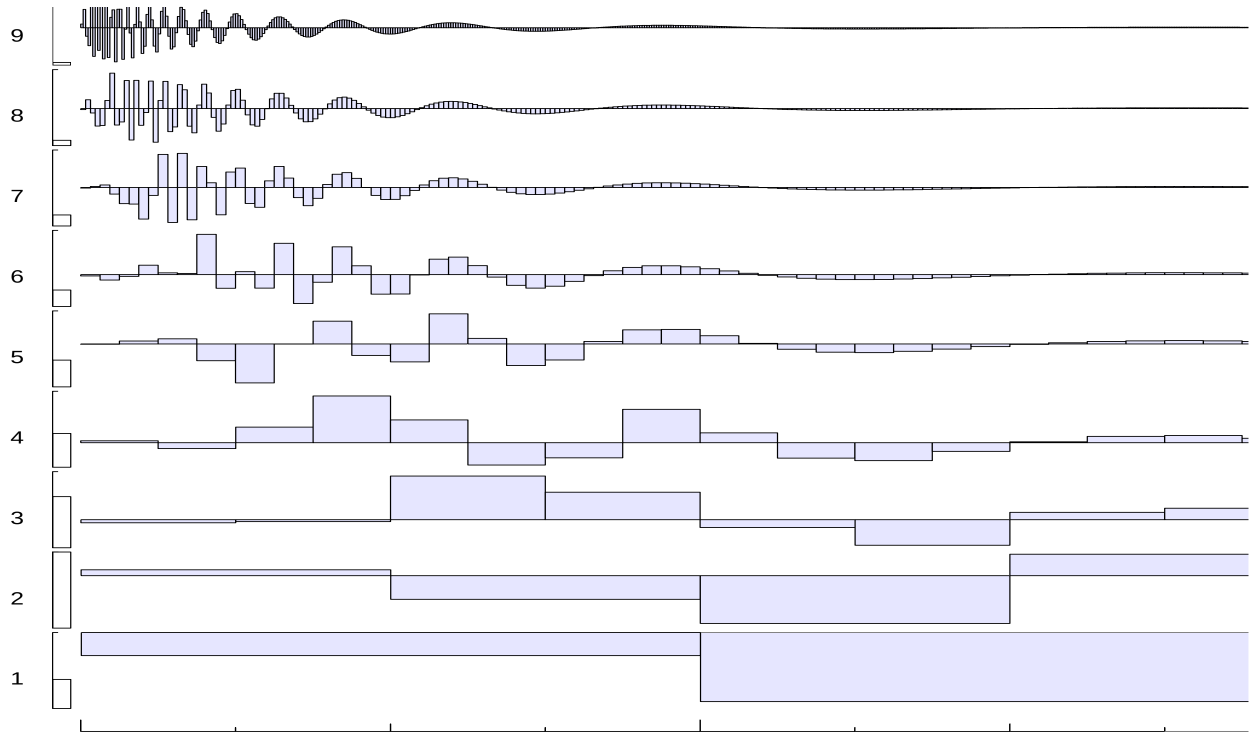


Figure 7: Multiresolution Analysis of the Vector y = [4 -1 3 2 1 4 -2 2], using the Haar Wavelet Filter Coefficients

pair of filters: $h = \left[\frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}}\right],\ g = \left[\frac{1}{\sqrt{2}}\ -\frac{1}{\sqrt{2}}\right].$

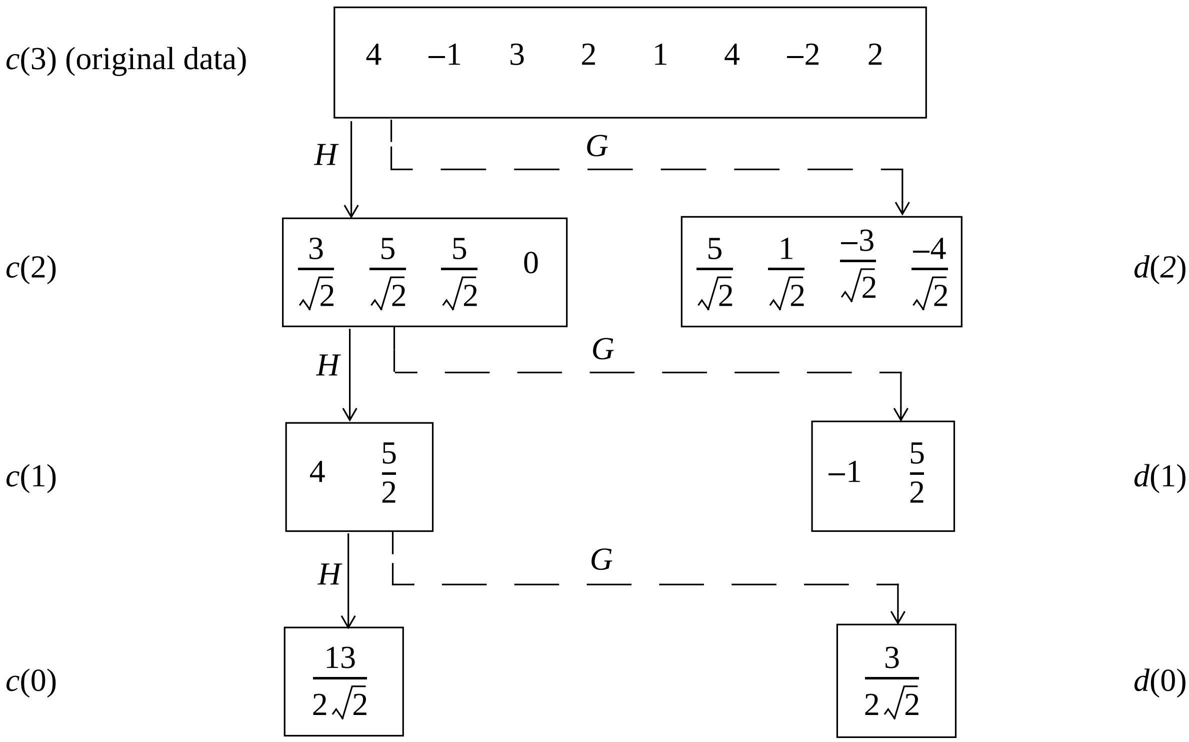


Figure 8: Reverse Wavelet Transform

$$c(0)\quad \boxed{\frac{13}{2\sqrt{2}}} \qquad \hat{H} \Rightarrow \boxed{\frac{13}{4} \quad \frac{13}{4}}$$

$$d(0)\quad \boxed{\frac{3}{2\sqrt{2}}} \qquad \hat{G} \Rightarrow \boxed{\frac{3}{4} \quad \frac{-3}{4}}$$

$$\Sigma: \boxed{4 \quad \frac{5}{2}}$$

$$c(1)\quad \boxed{4 \quad \frac{5}{2}} \qquad \hat{H} \Rightarrow \boxed{\frac{4}{\sqrt{2}} \quad \frac{4}{\sqrt{2}} \quad \frac{5}{2\sqrt{2}} \quad \frac{5}{2\sqrt{2}}}$$

$$d(1)\quad \boxed{-1 \quad \frac{5}{2}} \qquad \hat{G} \Rightarrow \boxed{\frac{-1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}} \quad \frac{5}{2\sqrt{2}} \quad \frac{-5}{2\sqrt{2}}}$$

$$\Sigma: \boxed{\frac{3}{\sqrt{2}} \quad \frac{5}{\sqrt{2}} \quad \frac{5}{\sqrt{2}} \quad 0}$$

$$c(2)\quad \boxed{\frac{3}{\sqrt{2}} \quad \frac{5}{\sqrt{2}} \quad \frac{5}{\sqrt{2}} \quad 0} \qquad \hat{H} \Rightarrow \boxed{\frac{3}{2} \quad \frac{3}{2} \quad \frac{5}{2} \quad \frac{5}{2} \quad \frac{5}{2} \quad \frac{5}{2} \quad 0 \quad 0}$$

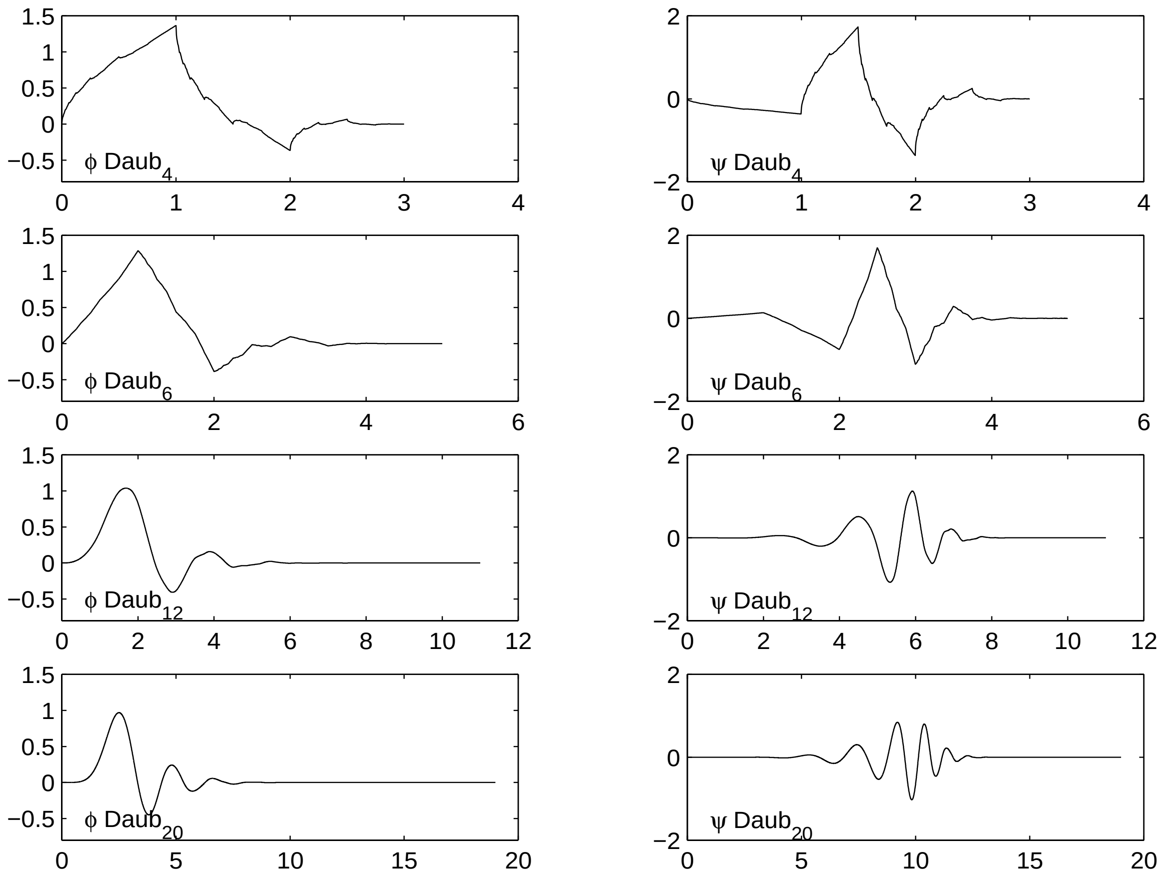
$$d(2)\quad \boxed{\frac{5}{\sqrt{2}} \quad \frac{1}{\sqrt{2}} \quad \frac{-3}{\sqrt{2}} \quad \frac{-4}{\sqrt{2}}} \qquad \hat{G} \Rightarrow \boxed{\frac{5}{2} \quad \frac{-5}{2} \quad \frac{1}{2} \quad \frac{-1}{2} \quad \frac{-3}{2} \quad \frac{3}{2} \quad -2 \quad 2}$$

$$\Sigma: \boxed{4 \quad -1 \quad 3 \quad 2 \quad 1 \quad 4 \quad -2 \quad 2}$$

4. Some Examples

We have already seen the Haar wavelet, but it has some serious limitations because of its discontinuity. In Figure 4 we saw that a Daubechies wavelet is much more convenient for approximating smooth functions. Daubechies wavelets were the first wavelet family with compact support and a preassigned degree of smoothness. They have an even number of filter elements, starting at 4. Wavelets within a family are usually denoted by the length of their filters. Figure 9 shows several specimens of the Daubechies family.

Figure 9: The Daubechies Family

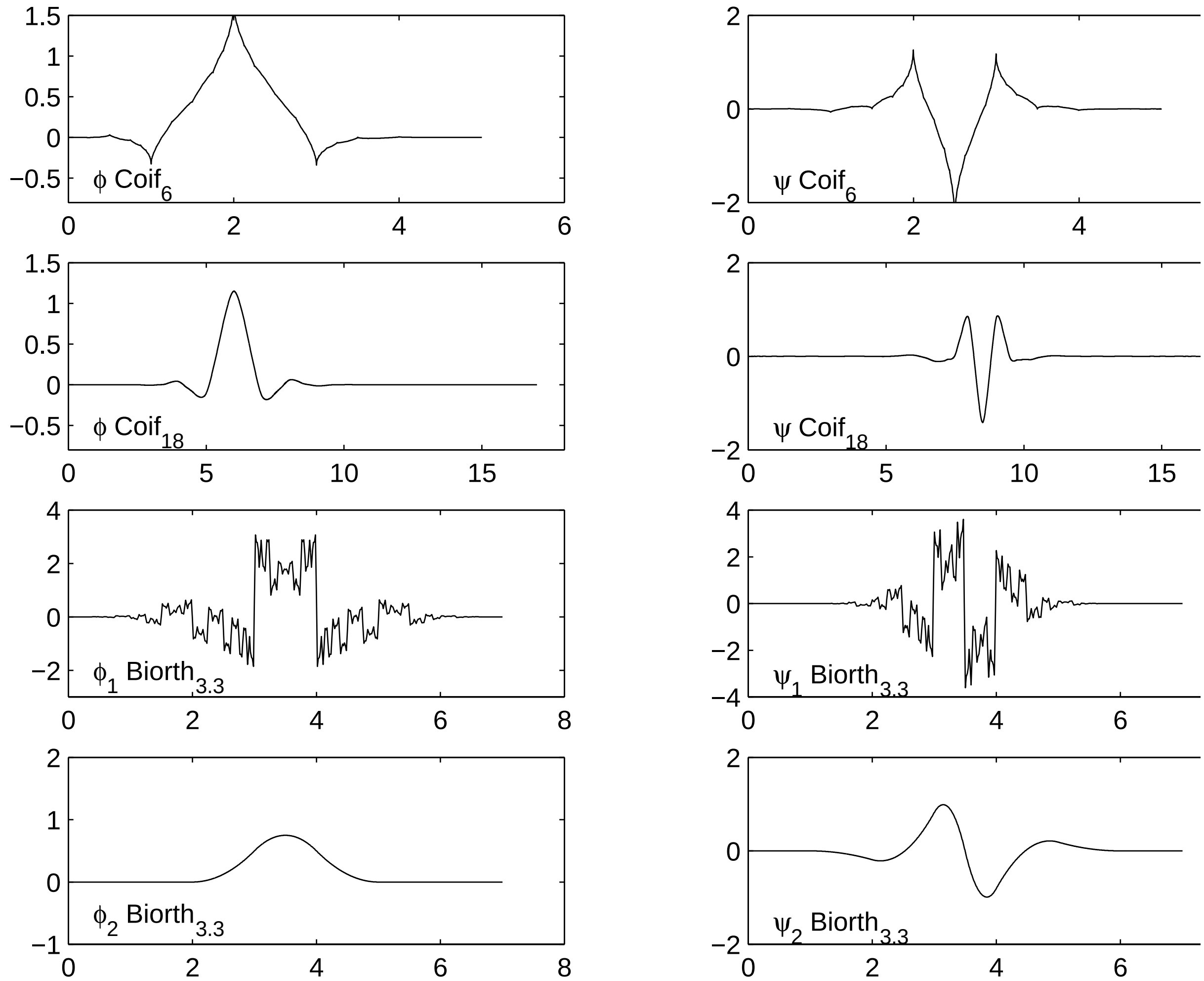


Increasing the number of filter elements increases the support of the wavelet and makes the wavelet smoother. The limiting case with two filter coefficients is the Haar wavelet. Daubechies wavelets are asymmetric, a necessary property for compactly supported wavelets.[8] It is, however,

8. In fact, the Haar wavelet is the only compactly supported orthonormal wavelet that is symmetric.

possible to construct wavelets that are closely related cousins to the Daubechies family and that are more symmetric. These wavelets are called least asymmetric wavelets, or symmlets. Another family of wavelets, coiflets (Figure 10), are even less asymmetric than symmlets. Named by Ingrid Daubechies after Ronald Coifman, these wavelets pay for their increased symmetry by a larger support. Compared to Daubechies wavelets, their support is $3L$-1 instead of $2L$-1, where L denotes the number of vanishing moments.[9]

Figure 10: Coiflets and Biorthogonal Symmetric Wavelets



Biorthogonal wavelets relax the assumption of a single orthogonal basis. Instead, they are defined as a pair of mutually orthogonal bases, neither of which is orthogonal. This relaxation allows for the construction of compactly supported symmetric wavelets. This property is especially desirable in image processing. Biorthogonal wavelets have primary and dual scaling and wavelet functions;

9. A wavelet ψ has N(>=2) vanishing moments if $\int x^n \psi(x)\,dx = 0$, $n = 0, 1, ..., N$-1.

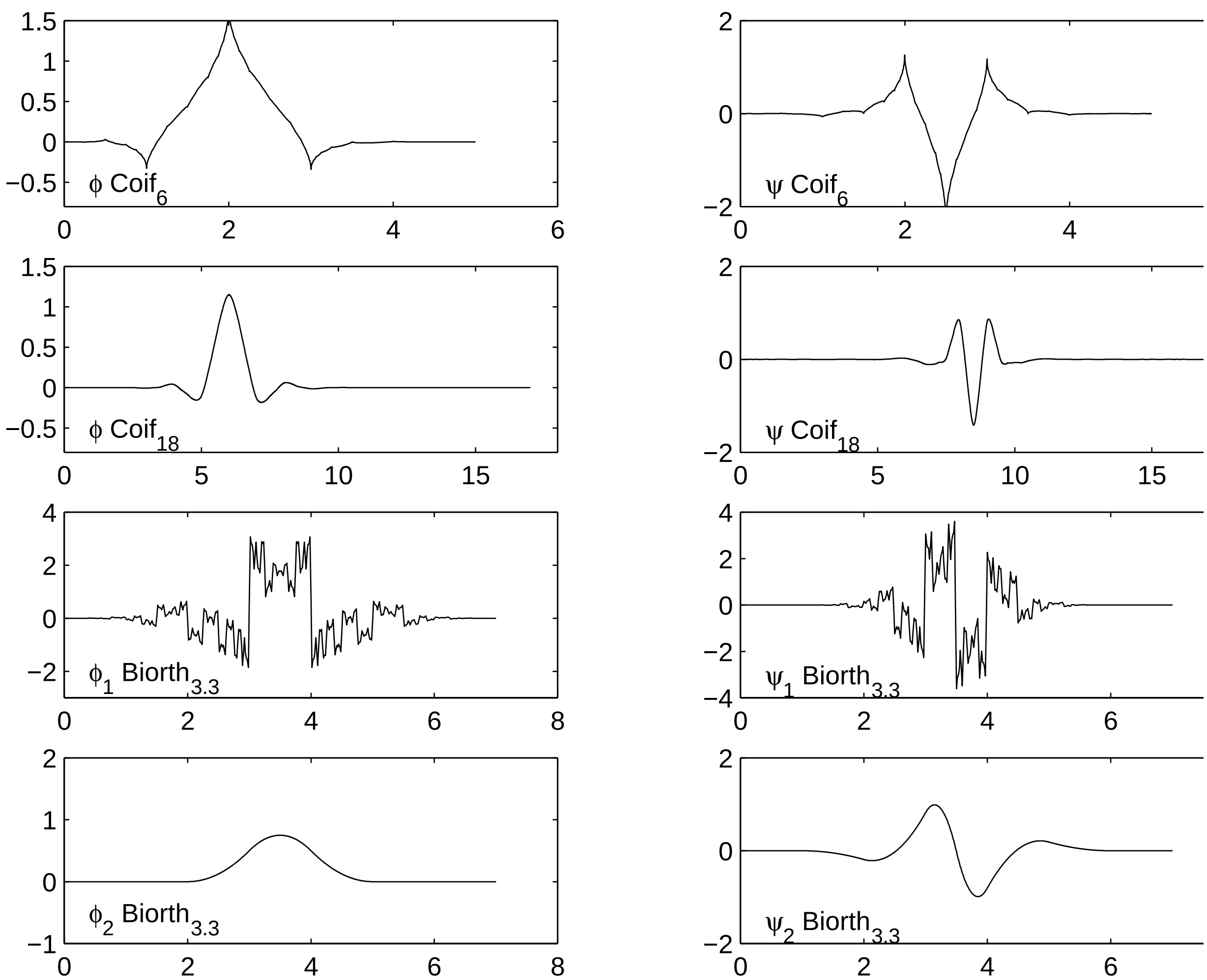
an example is shown in the last two rows in Figure 10 for BS_2.2. Figure 11 shows some coiflets and biorthogonal wavelets with different numbers of filter coefficients.

The following subsections describe some of the most useful ways in which wavelets can be applied in practice.

4.1 Filtering

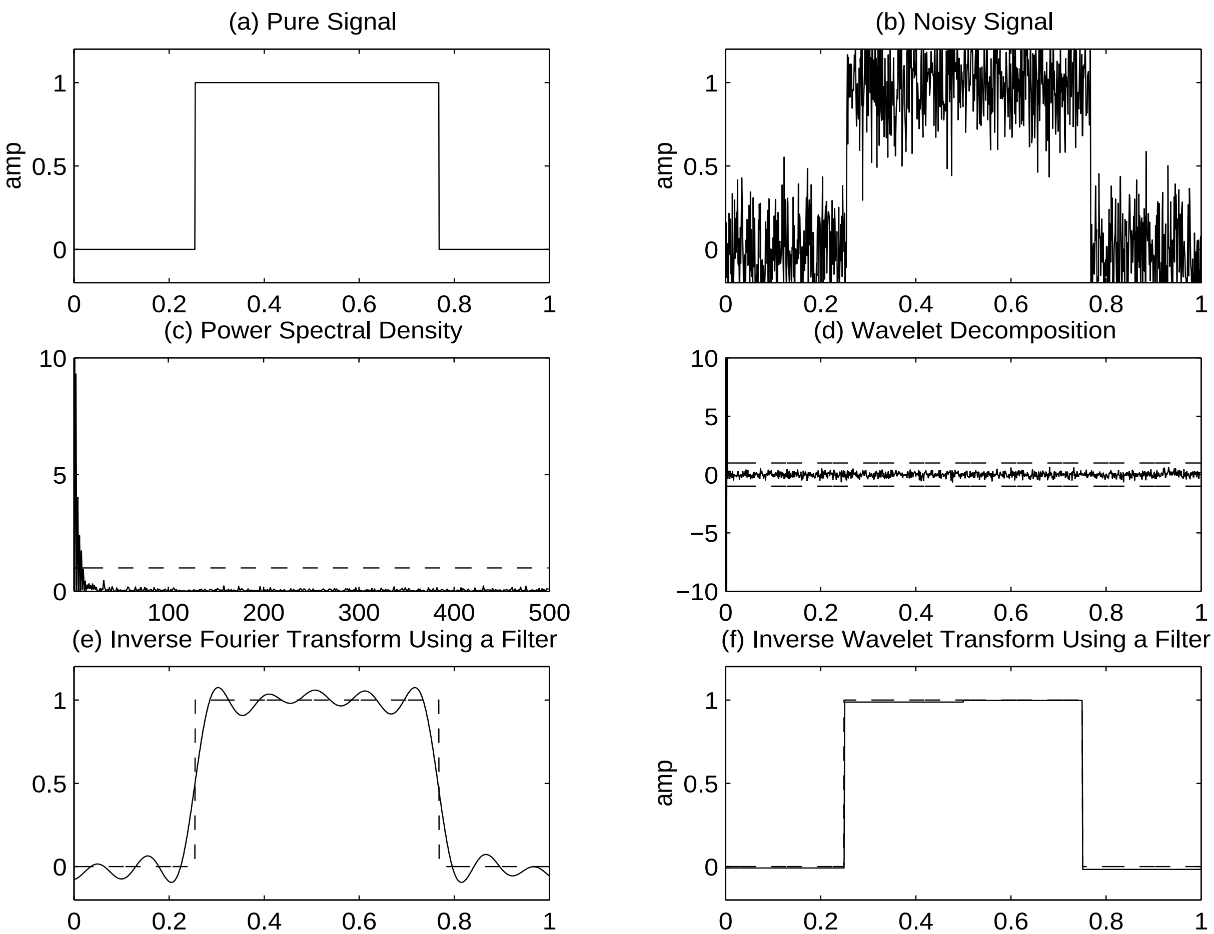
One of the main applications of frequency analysis concerns the filtering of noisy signals. Since white noise is uncorrelated at all lags and leads, it is distributed evenly over the frequency domain. Graph (a) in Figure 12 shows a digital impulse in the time domain that is covered by Gaussian white noise (Graph (b)). Graphs (c) and (d) show the power spectral density (normalized absolute value of the squared Fourier transform) and the wavelet transformation of the noisy signal.

Figure 11: Coiflets and Biorthogonal Symmetric Wavelets



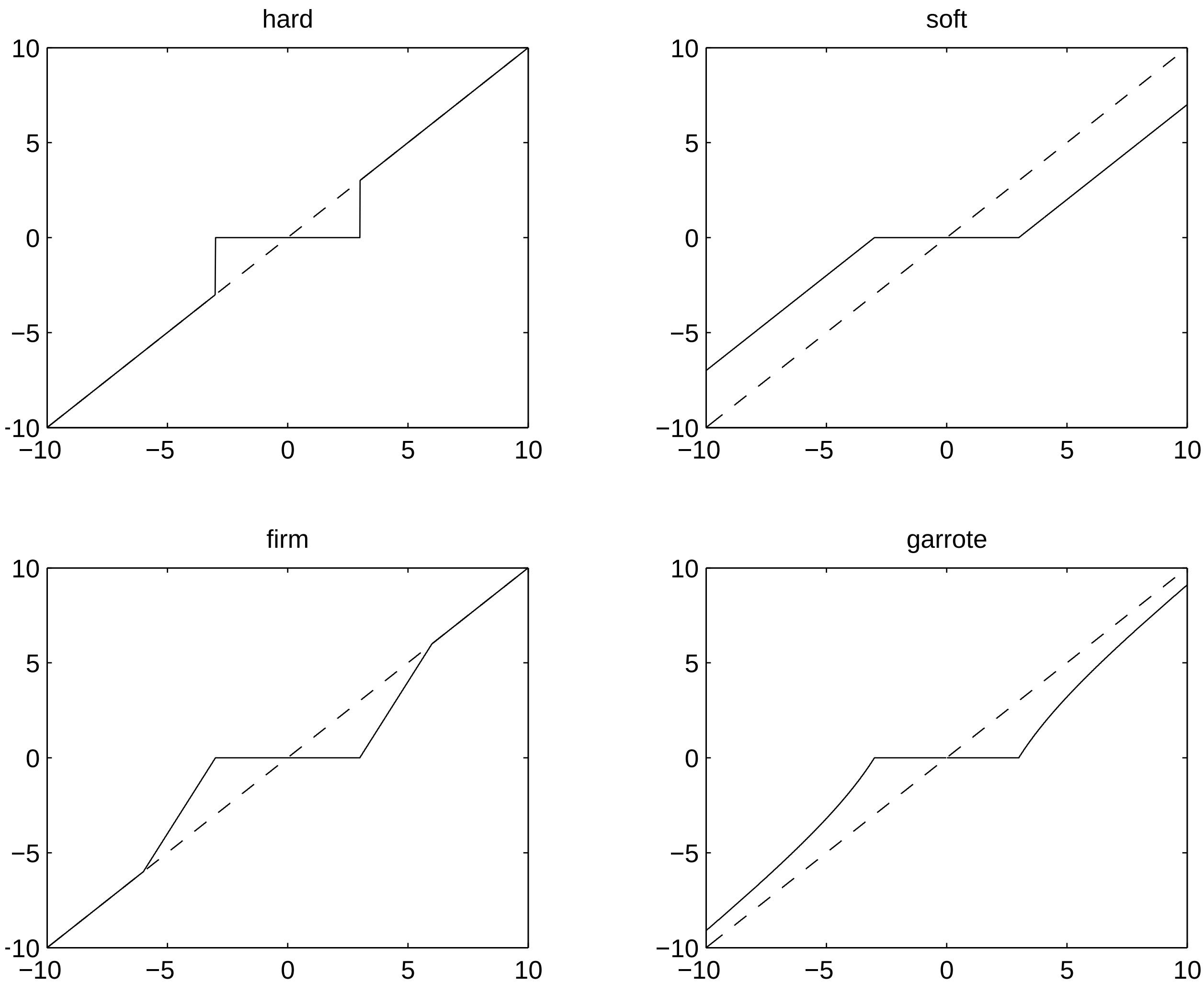
To filter out the white noise, all coefficients in the frequency domain below a certain bandwidth are set equal to zero. This process is called a “hard” thresholding rule. (Several other possibilities for thresholding are shown in Figure 13.) The filtered series are then transformed back into the time domain. A comparison between graphs (e) and (f) in Figure 12 shows the clearly superior performance of the inverse wavelet transform to restore the original signal. The performance is superior because the Fourier transform relies on a large number of layers to suppress the Gibbs effect (that is, the over- and undershooting at discontinuities), because its basis is non-local in the time domain and some of the higher layers are rubbed out in the filtering process. Another reason is that the Haar wavelet, which was used in this example, is particularly useful in approximating step functions like the rectangular pulse signal.

Figure 12: Filtering



The method of using thresholding rules to filter data was developed by Donoho and Johnstone (1994) and it is called wavelet shrinkage, or WaveShrink. Its main advantage is that the de-noising is carried out without smoothing out sharp structures such as spikes and cusps. One interesting application is in seismology, where researchers observe the water levels of wells to predict earthquakes. Wavelets can filter out the noise without removing the spikes that characterize changes in water levels prior to earthquakes.

Figure 13: Different Thresholding Rules

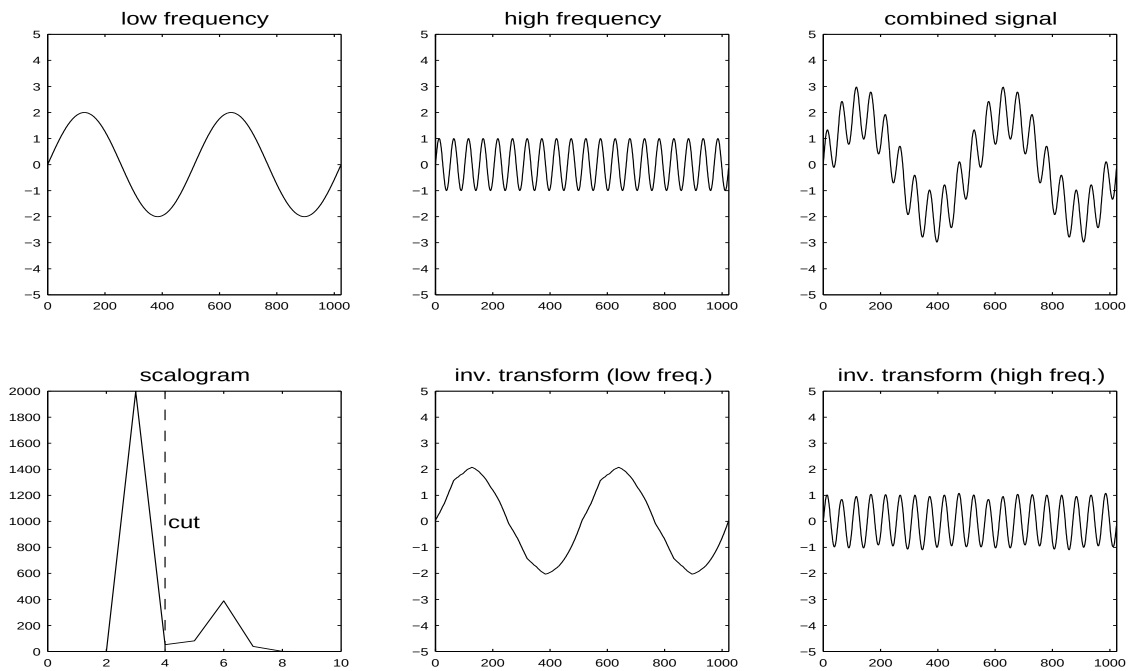


4.2 Separation of frequency levels

In another application, two sine functions with different frequencies are added up (see Figure 14, first row). To filter each component of the combined signal, we look at how much energy is contained in each scale of the wavelet transform. This can be done by adding up the squared coefficients in each scale to get a scalogram. The scalogram can be compared to the power spectral plot in Fourier analysis.

In the scalogram, we can observe two spikes: one at the third level and one at the sixth level. Since slow-moving, low-frequency components are represented by wavelets with larger support, we conjecture that level three represents the wavelet transform for function one, and that level six represents the wavelet transform for function two. To filter out function one, we keep only the first four levels and pad the rest of the wavelet transform with zeros. Then we take the inverse transform. Conversely, we keep levels five to nine for the second function and pad the first four levels with zeros. The last two graphs in Figure 14 show the filtered series given by the inverse transformations. In this particular case, a Fourier transform would be much more efficient, since the underlying functions are purely periodic. But wavelets are more useful when we are dealing with non-periodic data. Many economic data are likely generated as aggregates of different scales. Separating these scales and analyzing them individually provides interesting insights and can improve the forecasting accuracy of the aggregate series. Section 5 describes some applications.

Figure 14: Separation of Frequency Levels



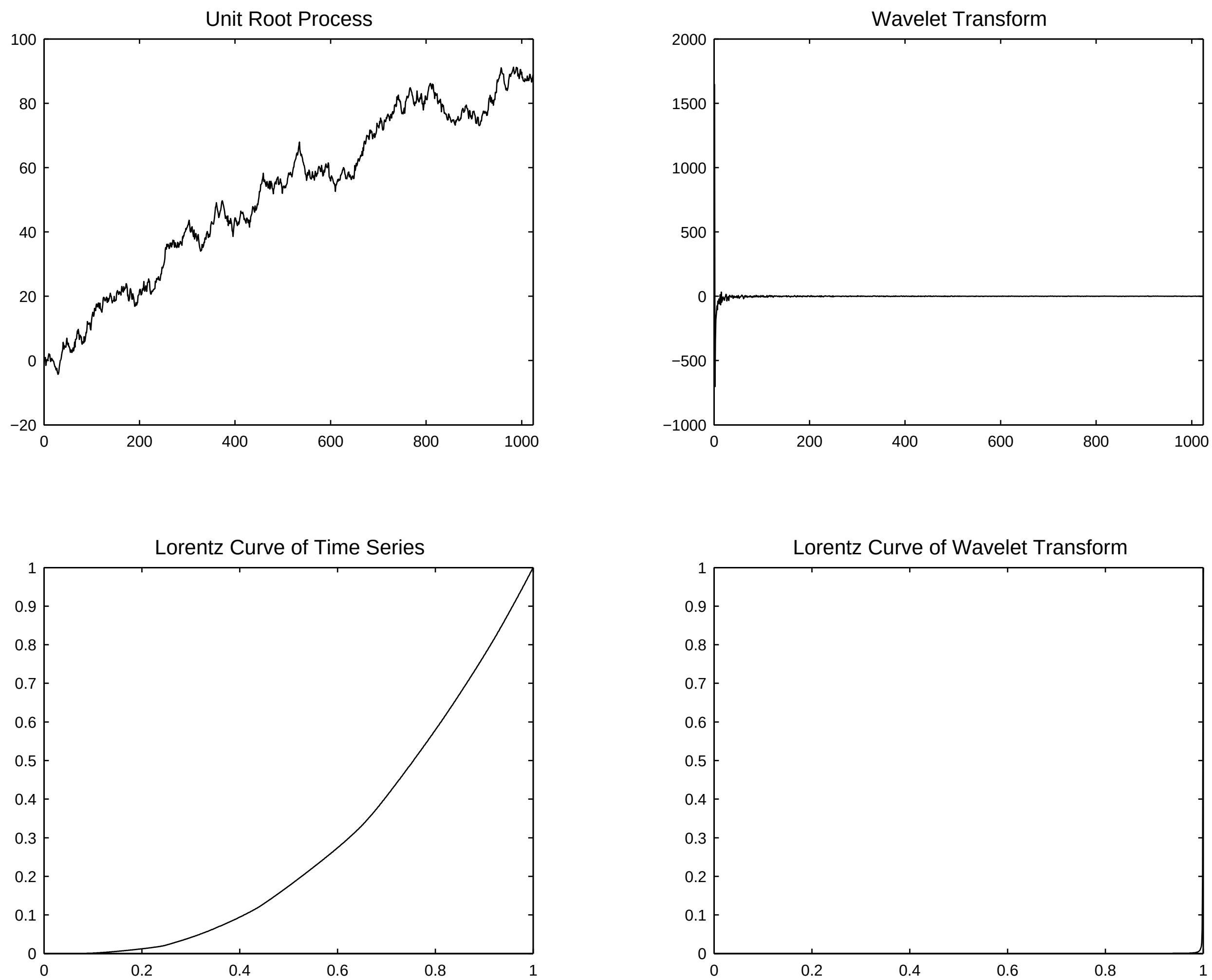
4.3 Disbalancing of energy

This is a relatively abstract concept, but it has very interesting implications. By referring to the energy of a signal or data vector, we mean its L^2-norm:

$$|y|^2 = \sum_i y^2{}_i .$$

Since the wavelet transform is an orthonormal operation, it preserves the total amount of energy. It distorts its distribution, however, making it more unequal. Figure 15 shows the time series and the wavelet transform of a unit root process. The lower scales hold most of the energy, a fact that makes intuitive sense, because the unit root process is characterized by lasting deviations from the current mean. The second row of Figure 15 shows the Lorentz curves for the two series.

Figure 15: Disbalancing of Energy



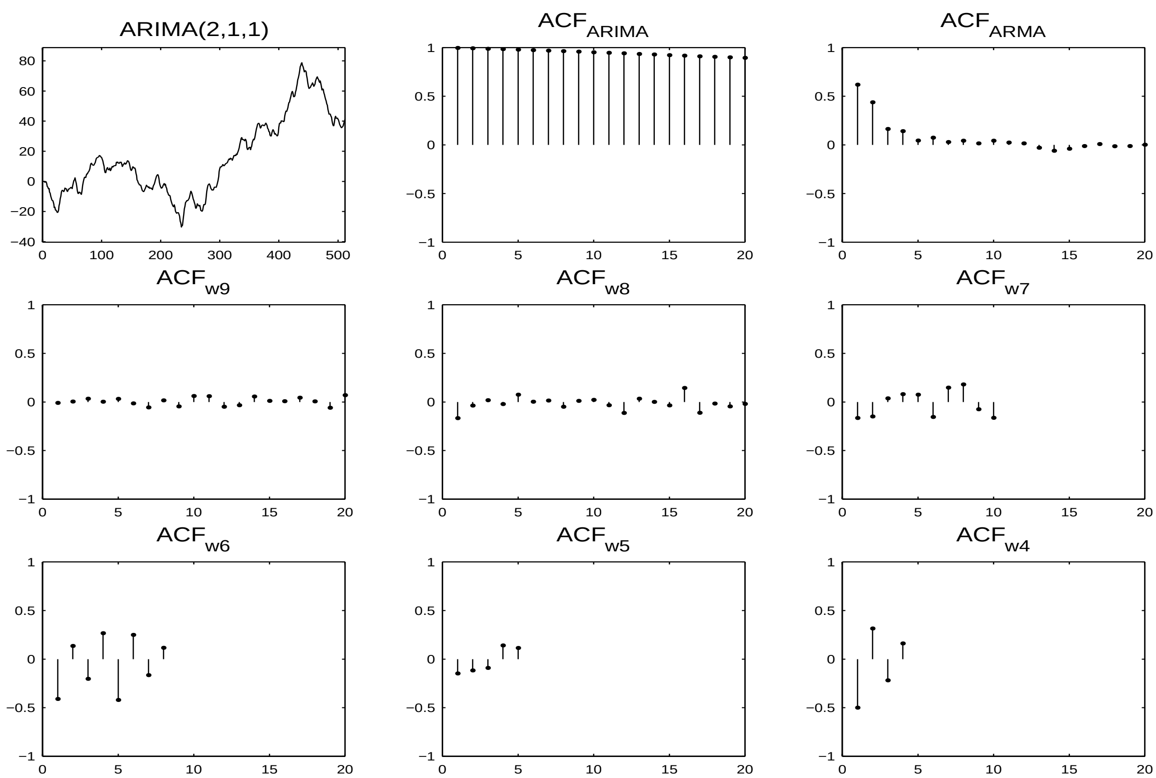
The Lorentz curve was initially developed by economists at the beginning of the twentieth century to study the distribution of income. It graphs the cumulative distribution against the quantiles. A 45-degree line would represent a completely homogeneous data set, whereas the curve shifts lower as inequality rises. The Lorentz curve of the wavelet transform is so close to the edge of the graph that it can hardly be seen.

Why do we want to disbalance the energy of a signal? There are several reasons. For data storage, this means that the signal can be well described by a fairly short sequence. In statistics disbalancing the energy can increase the variance between different distributions, thereby increasing the power of a test.

4.4 Whitening of correlated signals

The hierarchical construction of wavelets means that non-stationary components of time series are absorbed by the lower scales, while non-lasting disturbances are captured by the higher scales. This leads to a phenomenon that Vidakovic (1999) calls the whitening property. The example shown in Figure 16 uses an integrated autoregressive moving average (ARIMA) (2,1,1) model as a data-generating process, a realization of which is illustrated in the first graph. This kind of process contains a unit root; therefore, the autocorrelation function shows almost no decay over 20 periods. On the other hand, the autocorrelation of the differenced ARMA(2,1) model shows a drop after two periods, as predicted by the Box and Jenkins model selection criterion.

Figure 16: Whitening of Correlated Signals



The wavelet transform decomposes the time series into log2 (T) scales. In our case, the time series has 512 observations, so that we have nine scales. The last six graphs in Figure 16 show the autocorrelation functions of the six highest layers of the original ARIMA series (the bottom three layers have four or less elements, not enough information to compute the autocorrelation functions).

Note that the ninth scale looks almost like a white noise signal, while scales six and four, in particular, show clear signs of autocorrelation at all included lags. Moreover, these autocorrelation functions are oscillating, indicating a presence of a mean reverting autoregressive process. This mean reverting behaviour at the higher scales is a result of the absorbtion of the trend and other non-stationary components by the lower scales.

5. Applications for Economists

This section discusses some fields where wavelets have been applied to economic research. All applications are related to econometrics; however, it is quite possible that wavelets might also be applied in one way or another directly to theoretical research. Parts of this survey follow Ramsey (1996), who the reader is referred to for a more detailed treatment of some of the topics.

5.1 Frequency domain analysis

A general result, called the spectral representation theorem, states that any covariance-stationary process has a representation in both the time domain and the frequency domain. In terms of a time series that has a moving average representation, using the Fourier transformation, we can write:

$$x(t) = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \mu + \int_0^{\pi} [a(\omega)\cos(\omega t) + b(\omega)\sin(\omega t)]\, d\omega .$$

One of the first applications of wavelets to time-series analysis was the estimation of the spectrum density for stationary Gaussian processes (Gao 1993).[10] Neumann (1996) extends the analysis to non-Gaussian processes. Given the ability of wavelets to break down a time series on a scale-to-scale basis, each scale corresponding to a range of frequencies, they can be used in virtually all applications that were previously based on Fourier analysis. However, as Priestley (1996) points out, there is only an intuitive and very indirect connection between frequency and the scale.

The Fourier transform is based on periodic functions in the time domain, thus capturing weekly, monthly, or yearly cycles. However, many economic phenomena, such as business cycles, do not follow this strict periodicity, favouring the more flexible wavelet approach. Conway and Frame

10. The spectrum of a time series is given by:

$$S_x(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j} .$$

(2000), for example, use both Fourier and wavelet techniques to analyze the frequency content of output gaps generated by different methods ranging from structural VARs to mechanical filters.

5.2 Non-stationarity and complex functions

Nason and von Sachs (1999) and von Sachs and MacGibbon (2000) focus on signals with a possibly time-changing probability distribution; that is, locally stationary processes whose moments exhibit slow change. Quasi-stationarity can be defined as a situation in which more observations per unit period of time would lead to locally asymptotic convergence of the estimators. The authors suggest the use of a local minimum absolute deviation (MAD) estimator to estimate the variance of the wavelet coefficients within the quasi-stationary levels. An important consequence is the suitability of wavelet estimators for generalized autoregressive conditional heteroscedasticity (GARCH) processes.

Following the seminal paper of Donoho and Johnstone (1994), a whole set of estimators is based on the notion of wavelet shrinkage, filtering noise using thresholding rules. These estimators are discussed in detail in Vidakovic (1999, ch. 6). Gao (1997) applies this technique for the case of heteroscedasticity of unknown form. The model is given by

$$x = f(t) + \sigma(t)\varepsilon,$$

where ε is a vector of identically, independently distributed Gaussian disturbances. If $\sigma(t)$ is equal to the identity matrix, then the empirical wavelet coefficients, *w*, are also independent Gaussian random variables with the same variance. For the heteroscedastic case, the wavelet coefficients have the asymptotic distribution

$$w \sim N(Hf, HD^2H'),$$

where H is the wavelet transformation matrix, such that $w = Hx$.

Ramsey and Zhang (1997) pursue a similar idea, but along a different path. Instead of changing moments in the time domain, they are concerned with changes in the frequency domain over time. An analogy is given by human speech, each syllable of which involves a distinct set of frequencies that last for a short period of time only. Ramsey and Zhang use waveform dictionaries, a class of transforms that generalizes both windowed Fourier transforms and wavelets, and that are composed of individual time-frequency atoms. Applications to Standard & Poor’s (S&P) 500 stock price index and exchange rates with 16,384 observations each show evidence that frequencies wax and wane in strength over time, but that most of the power is concentrated in highly time-localized bursts of activity. These bursts, called chirps, are characterized by a rapid buildup of amplitude and rapid oscillation. The authors conclude that transmission of information is not a swift and effortless process, but that there are periods of adjustment represented by the chirps. Although the data can be represented by a relatively small number of atoms (about 100), there seems to be no way of forecasting the chirps.

5.3 Long-memory processes

The importance of differentiating between stationary I(0) and non-stationary I(1) processes has long been one of the focal points in theoretical and applied time-series analysis. In recent years, the attention of researchers has shifted towards fractionally integrated I(*d*) processes, that lie in the grey area between the two sharp-edged alternatives of I(0) and I(1). Specifically, when *d* lies between 0 and 0.5, the process still has a finite variance, but its autocovariance function decays at a much slower rate than that of a stationary ARMA process. Such processes are called long-memory processes. When *d* lies between 0.5 and 1, the variance becomes infinite, but the process still returns to its long-run equilibrium. The study of long-memory processes dates back to Hurst (1951) and has been applied to a number of economic time series in recent years. Jensen (1999) cites a wide range of applications, including real GDP, interest rates, stock market returns, option prices, and exchange rates.

A fractionally integrated process, I(*d*), can be defined by

$$(1-L)^d x(t) = \varepsilon(t),$$

where $\varepsilon(t)$ is white noise or follows an ARMA process.

Because long-memory processes have a very dense covariance matrix, direct maximum-likelihood estimation is not feasible for large data sets. Instead, the estimator most often used is based on a nonparametric approach, which regresses the log values of the periodogram on the log Fourier frequencies to estimate *d* (Geweke and Porter-Hudak (GPH) 1983).

McCoy and Walden (1996) find a log-linear relationship between the wavelet coefficients' variance and its scale and develop a maximum-likelihood estimator. Jensen (1999) develops a simpler, ordinary least-squares (OLS) estimator that is based on the observation that for a mean zero I(d) process, $|d| < 0.5$, the wavelet coefficients, d_{jk} (for scale (dilation) *j* and translation *k*), are asymptotically normally distributed with mean zero and variance $\sigma^2 2^{-2jd}$ as j goes to zero. Taking logs, we can estimate *d* using the linear relationship

$$\ln R(j) = \ln\sigma^2 - d\ln 2^{2j},$$

where R(j) is the sample estimate of the covariance in each scale. The wavelet estimators (OLS and maximum likelihood) have a higher small-sample bias than the GPH estimator, but Monte-Carlo experiments show that they have a mean-squared error that is about six times lower.

Mandelbrot and van Ness (1968) find self-similarities in fractional Brownian motions that have been observed in physical sciences as well as in financial time series. The ability of wavelets to dissect data into different scales makes it possible to detect these self-similar phenomena. Vidakovic (1999, p. 14, pp. 292) gives an overview of some of the related literature.

5.4 Time-scale decompositions: the relationship between money and income

It has long been recognized that economic decision-making is dependent on the time scale involved, and economists emphasize the importance of discerning between long-run and short-run behaviour. Wavelets offer the possibility of going beyond this simplifying dichotomy by decomposing a time series into several layers of orthogonal sequences of scales using Mallat's multiscale analysis. These scales can then be analyzed individually and compared across different series.

Ramsey and Lampart (1998a) use this method to examine the relationship between the money supply (M1 and M2) and output. The related literature has produced ambiguous and contradictory results regarding the Granger causality of the two variables, a fact that has been attributed to structural breaks and possible non-linearities. Ramsey and Lampart offer the alternative explanation that the contradictions may well be explained by the existence of overlaying timescale-structured relationships. To unmask these relationships, the authors use a wavelet transform to decompose the time series into a low-frequency base scale that captures the long-run trend and six higher-frequency levels. Here, two more interesting facets of wavelets are useful: (i) since the base scale includes any non-stationary components, the data need not be detrended or differenced; and (ii), the nonparametric nature of wavelets takes care of potential non-linear relationships without losing detail. Applying causality tests to the decomposed series, the authors find that, at the lowest timescales, income Granger causes money, but at business-cycle periods, money Granger causes income. At the highest scales, the Granger causality goes in both directions, suggesting some form of feedback mechanism. These results make intuitive sense and also explain why there are ambiguous causal relationships when timescales are aggregated.

A second important finding is that the causal relationship between different variables is non-stationary even along individual scales, since the two series are moving into and out of phase with each other. To explain these phase shifts and to differentiate them from structural breaks will be a further challenge for theoretical and applied researchers.

In a companion paper, Ramsey and Lampart (1998b) analyze the relationship between consumption and income at different timescales. As predicted by theory, they find that the slope coefficient relating consumption and income declines with scale.

5.5 Forecasting

Ariño (1998) and Ariño, Pedro, and Vidakovic (1995) describe a very simple approach for forecasting time series using wavelets. First, the time series is decomposed into different scales using the wavelet transform. Ariño shows that, by adding up the squared coefficients within each level, one can measure the energy content of each scale, similar to the power spectral density used in Fourier analysis. Using the properties of the multiscale analysis, the time series is then decomposed into two separate series. Each individual is then fitted using an ARIMA model and the aggregate forecast is obtained by adding up the individual forecasts. Ariño shows that his forecasts are preferable to a standard Box and Jenkins approach, but does not discuss the distributional properties of his forecast. A useful first step would be to test the wavelet estimator against other estimators using a Monte Carlo simulation.

A second field of application is the use of wavelets in connection with neural networks. Recently, wavelet networks have gained wide acceptance in physics, engineering, and biological research; however, their use for forecasting economic time series has been limited so far. Aussem and Murtagh (1997) and Aussem, Campbell, and Murtagh (1998) can find an improvement in the prediction of sunspots and the S&P 500 index. Similar to Ariño's approach, the time series is first decomposed into different scales. Each scale is then used to train a dynamic recurrent neural network and the individual forecasts are added up to obtain the combined forecast. Since neural networks need a lot of variation to extract information, only scales with a relatively high frequency can be used.

The main benefit of wavelets in forecasting appears to be their ability to reveal features in the individual scales that are dampened by the overlapping scales. It is therefore easier for ARIMA models or neural networks to extract periodic information in the individual scales.

6. Conclusions

Wavelets open a large, unexplored territory to applied economic researchers that can be roughly decomposed into three areas. The first area covers research that is related to Fourier and frequency analysis. While the Fourier transform maps from the time domain into the frequency domain, the wavelet transform decomposes a time series into a set of different scales, each of which can be loosely associated with a range of frequencies. The second area exploits several useful features of wavelets to improve statistical inference. These features are the ability to localize a function in both time and scale, to deal with non-linear and non-stationary environments, and to compress the energy content of a signal. The third area directly addresses the dissection of data into separate layers or scales. From a theoretical viewpoint, this is of special interest, since economic decisions and actions take place at different timescales that overlap. For forecasting purposes, there is evidence that the individual scales provide the forecasting mechanism (e.g., an ARIMA model or an artificial neural network) with more detailed information than the aggregate signal.

All three areas leave ample room for future research. For example, evidence for the improvement of forecasts by decomposing the time series is largely anecdotal and based on individual examples. A next step would be to calculate small sample properties and asymptotic distributions. The fact that wavelet transforms disbalance the energy of a signal could be used to construct more powerful tests; for example, for structural breaks or unit roots. A lot of statistical techniques, such as wavelet shrinkage estimators, that have been worked out and applied in biometrics and engineering could be applied to econometrics.

7. How to Get Started

There are a couple of easy and intuitive primers on wavelet theory; for example, Graps (1995) and Vidakovic and Mueller (1994). Vidakovic (1999) provides a more complete and technical, but accessible, treatment.

Most researchers use either MatLab or S-Plus to model wavelets. Both platforms offer commercial wavelet toolboxes as well as free add-ons. The examples and graphs used in this survey were made using Ojanen’s (1998) WaveKit toolbox for MatLab (*www.math.rutgers.edu/~ojanen/wavekit*). Another free MatLab toolbox is WaveLab, developed by Donoho et al. (1999) at Stanford (*www-stat.stanford.edu/~wavelab*). WaveLab has a very large set of commands and includes datasets and educational add-ons.

A good link to the newest developments and new publications in wavelet research is *www.wavelet.org*.

Bibliography

Ariño, M.A. 1998. “Forecasting Time Series via Discrete Wavelet Transform.” Unpublished manuscript.

Ariño, M.A., M. Pedro, and B. Vidakovic. 1995. “Wavelet Scalograms and Their Application in Economic Time Series.” Institute of Statistics and Decision Sciences, Duke University, Discussion Paper No. 94–13.

Aussem, A., J. Campbell, and F. Murtagh. 1998. “Wavelet-Based Feature Extraction and Decomposition Strategies for Financial Forecasting.” *Journal of Computational Intelligence in Finance* (March/April): 5–12.

Aussem, A. and F. Murtagh. 1997. “Combining Neural Network Forecasts on Wavelet-Transformed Time Series.” *Connection Science* 9(1): 113–21.

Conway, P. and D. Frame. 2000. “A Spectral Analysis of New Zealand Output Gaps Using Fourier and Wavelet Techniques.” Reserve Bank of New Zealand Discussion Paper No. 2000/06.

Daubechies, I. 1988. “Orthonormal Bases of Compactly Supported Wavelets.” *Communications on Pure and Applied Mathematics* 41: 909–96.

———. 1992. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics.

Davison, R., W.C. Labys, and J.-B. Lesourd. 1998. “Wavelet Analysis of Commodity Price Behavior.” *Journal of Computational Economics* 11: 103–28.

Donoho, D.L. and I.M. Johnstone. 1994. “Ideal Spatial Adaptation via Wavelet Shrinkage.” *Biometrica* 81: 425–55.

Donoho, D.L. et al. 1999. *WaveLab.* Stanford University, Department of Statistics. <URL: http://www-stat.stanford.edu/~wavelab>.

Gao, H.-Y. 1993. “Wavelet Estimation of Spectral Densities in Time Series Analysis.” PhD Dissertation, Department of Statistics, University of California, Berkeley.

———. 1997. “Wavelet Shrinkage Estimates for Heteroscedastic Regression Models.” Unpublished manuscript, MathSoft, Inc.

Geweke, J. and S. Porter-Hudak. 1983. “The Estimation and Application of Long Memory Time Series Models.” *Journal of Time Series Analysis* 4: 221–38.

Graps, A. 1995. “An Introduction to Wavelets.” *IEEE Computational Science and Engineering* 2(2): 50–61.

Hernandez, E. and G.L. Weiss. 1996. “A First Course on Wavelets.” CRC, Boca Raton.

Hong, Y. 1999. “One-Sided Testing for ARCH Effect Using Wavelets.” PhD thesis, Cornell University.

Hurst, H.E. 1951. “Long-Term Storage Capacity of Reservoirs.” *Transactions of the American Society of Civil Engineers* 116: 770–99.

Jensen, M.J. 1999. “Using Wavelets to Obtain a Consistent Ordinary Least Squares Estimator of the Long-Memory Parameter.” *Journal of Forecasting* 18: 17–32.

———. 2000. “An Alternative Maximum Likelihood Estimator of Long-Memory Processes using Compactly Supported Wavelets.” *Journal of Economic Dynamics & Control* 24: 361–87.

Lapenta, E.S., S.M. Abecasis, and C.A. Heras. 2000. “Discrete Wavelet Transforms for the Treatment of Financial Time Series.” Unpublished manuscript. <URL: http://www.sarabe1@infovia.com.ar>.

Mallat, S. 1989. “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11: 674–93.

Mallat, S. and Z. Zhang. 1993. “Matching Pursuits with Time-Frequency Dictionaries.” *IEEE Transactions on Signal Processing* (December): Vol. 41, No. 12.

Mandelbrot, B.B. and J.W. van Ness. 1968. “Fractional Brownian Motions, Fractional Noises and Applications.” *SIAM Review* 10(4): 422–37.

McCoy, E.J. and A.T. Walden. 1996. “Wavelet Analysis and Synthesis of Stationary Long-Memory Processes.” *Journal of Computational and Graphical Statistics* 5(1): 26–56.

Nason, G.P. and R. von Sachs. 1999. “Wavelets in Time Series Analysis.” *Philosophical Transactions of the Royal Society London, Series A: Mathematical, Physical and Engineering Sciences* 357: 2511–26.

Neumann, M.H. 1996. “Spectral Density Estimation via Nonlinear Wavelet Methods for Stationary Non-Gaussian Time Series.” *Journal of Time Series Analysis* 17: 601–33.

Ojanen, H. 1998. *WAVEKIT: A Wavelet Toolbox for Matlab.* Department of Mathematics, Rutgers University.

Percival, D.B. and A.T. Walden. 2000. *Wavelet Methods for Time Series Analysis*. New York: Cambridge University Press.

Priestley, M. 1996. “Wavelets and Time-Dependent Spectral Analysis.” *Journal of Time Series Analysis* 17: 85–103.

Ramsey, J.B. 1996. “The Contribution of Wavelets to the Analysis of Economic and Financial Data.” Unpublished manuscript.

Ramsey, J.B. and C. Lampart. 1998a. “Decomposition of Economic Relationships by Timescale Using Wavelets.” *Macroeconomic Dynamics* 2(1): 49–71.

———. 1998b. “The Decomposition of Economic Relationships by Time Scale using Wavelets: Expenditure and Income.” *Studies in Nonlinear Dynamics and Econometrics* 3(1): 23–42.

Ramsey, J.B., D. Usikov, and G.M. Zaslavskiy. 1995. “An Analysis of U.S. Stock Price Behaviour Using Wavelets.” *Fractals* 3(2): 377–89.

Ramsey, J.B. and Z. Zhang. 1997. “The Analysis of Foreign Exchange Data using Waveform Dictionaries.” *Journal of Empirical Finance* 4: 341–72.

Sachs, R. von and B. MacGibbon. 2000. "Non-Parametric Curve Estimation by Wavelet Thresholding with Locally Stationary Errors." *Scandinavian Journal of Statistics* 27: 475–99.

Sachs, R. von and M. Neumann. 2000. "A Wavelet-Based Test for Stationarity." *Journal of Time Series Analysis* 21: 597–613.

Spokoiny, V.G. 1996. "Adaptive Hypothesis Testing using Wavelets." *The Annals of Statistics* 24(6): 2477–98.

Strang, G. 1993. "Wavelet Transforms versus Fourier Transforms." *Bulletin (new series) of the American Mathematical Society* 28(2): 288–305.

Strichartz, R.S. 1993. "How to Make Wavelets." *American Mathematical Monthly* 100: 539–56.

Tkacz, G. 2000. "Estimating the Fractional Order of Integration of Interest Rates Using a Wavelet OLS Estimator." *Studies in Nonlinear Dynamics and Econometrics* 5: 19–32.

Vidakovic, B. 1999. *Statistical Modeling by Wavelets*. New York: John Wiley and Sons.

Vidakovic, B. and P. Mueller. 1994. "Wavelets for Kids, A Tutorial Introduction." Institute of Statistics and Decision Sciences, Duke University, Discussion Paper No. 95–21.

Whitcher, B.J. 1998. "Assessing Nonstationary Time Series Using Wavelets." PhD dissertation, Department of Statistics, University of Washington.

Bank of Canada Working Papers
Documents de travail de la Banque du Canada

Working papers are generally published in the language of the author, with an abstract in both official languages. ***Les documents de travail sont publiés généralement dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue*****.**

2002

2002-2	Asset Allocation Using Extreme Value Theory	Y. Bensalah
2002-1	Taylor Rules in the Quarterly Projection Model	J. Armour, B. Fung, and D. Maclean

2001

2001-27	The Monetary Transmission Mechanism at the Sectoral Level	J. Farès and G. Srour
2001-26	An Estimated Canadian DSGE Model with Nominal and Real Rigidities	A. Dib
2001-25	New Phillips Curve with Alternative Marginal Cost Measures for Canada, the United States, and the Euro Area	E. Gagnon and H. Khan
2001-24	Price-Level versus Inflation Targeting in a Small Open Economy	G. Srour
2001-23	Modelling Mortgage Rate Changes with a Smooth Transition Error-Correction Model	Y. Liu
2001-22	On Inflation and the Persistence of Shocks to Output	M. Kichian and R. Luger
2001-21	A Consistent Bootstrap Test for Conditional Density Functions with Time-Dependent Data	F. Li and G. Tkacz
2001-20	The Resolution of International Financial Crises: Private Finance and Public Funds	A. Haldane and M. Kruger
2001-19	Employment Effects of Restructuring in the Public Sector in North America	P. Fenton, I. Ip, and G. Wright
2001-18	Evaluating Factor Models: An Application to Forecasting Inflation in Canada	M.-A. Gosselin and G. Tkacz
2001-17	Why Do Central Banks Smooth Interest Rates?	G. Srour
2001-16	Implications of Uncertainty about Long-Run Inflation and the Price Level	G. Stuber
2001-15	Affine Term-Structure Models: Theory and Implementation	D.J. Bolder
2001-14	L'effet de la richesse sur la consommation aux États-Unis	Y. Desnoyers

Copies and a complete list of working papers are available from:
Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à :

Publications Distribution, Bank of Canada
234 Wellington Street, Ottawa, Ontario K1A 0G9
E-mail: publications@bankofcanada.ca
Web site: http://www.bankofcanada.ca

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9
Adresse électronique : publications@banqueducanada.ca
Site Web : http://www.banqueducanada.ca