A Wavelet Analysis for Stationary Processes

Chang Chiann and Pedro A. Morettin

University of São Paulo, São Paulo, Brazil

1 Introduction

Fourier analysis is now a powerful and well estabilished tool for analyzing time series. After the influential work of Blackman and Tukey (1959) and the introduction of the fast Fourier transform by Cooley and Tukey (1965), the field of spectral analysis of time series experimented a continuous growth. Recent references are Brillinger (1981) and Brockwell and Davis (1991) on the theoretical side, and Percival and Walden (1993) on the applied side. It is well known that Fourier analysis is well suited to the analysis of stationary series. For categorical time series an adequate tool is the Walsh-Fourier analysis. For details, see Morettin (1981) and Stoffer (1991).

Recently an enormous interest has emerged on the use of wavelets in several areas. For an interesting review of the history of wavelets we refer to Meyer (1993). A short review is given by Farge (1992). The basic fact about wavelets is that they are *localized* in time (or space), contrary to what happens to the trigonometric functions. This makes the wavelets to be ideal to handle non-stationarities and signals containing transients and fractal-type structures. Moreover, the wavelets allow us to analyze a series into both *time* and *scale*.

In Fourier analysis, the orthonormal system of the complex exponentials $\{w_n(x) = e^{inx}, n = 0, \pm 1, \cdots\}$ are obtained by dilation of the single function $w(x) = e^{ix}$ through $w_n(x) = w(nx)$. Thus, every periodic function, of period 2π, square integrable (belonging to $L^2(0, 2\pi)$), is generated by a superposition of the $w_n(x)$. We wish to extend this idea to $L^2(\Re)$, that is, generate this space from a single function, ψ say. This is done by translations and dilations of

ψ, namely considering the wavelets

$$\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k), \quad j,k \in Z,$$

where $Z = \{0, \pm 1, \cdots\}$.

Wavelets have been used mainly in the fields of signal processing, image coding and compression, and in certain areas of mathematics, as in solutions of partial differential equations or numerical analysis. For further references, see Mallat (1989), Grossman and Morlet (1984), Meyer (1990). The books of Chui (1992) and Daubechies (1992) are good mathematical references.

On the statistical point of view, wavelets have been used in the estimation of densities (Hall and Patil, 1993a,b, Donoho and Johnstone, 1990, Donoho, 1993), in nonparametric regression (Nason, 1994, Hall and Patil, 1993a,b) and in the estimation of spectral densities (Gao, 1993a,b, Moulin, 1993). Brillinger (1994a,b) uses wavelets to detect change of levels in hydrologic series and develops mean level function estimations based on wavelets. Nason and Silverman (1994) develop a software for the computation of the discrete wavelet transform.

In this paper we develop a wavelet spectral analysis for a stationary discrete process. In section 2 we give the basic ideas on wavelets. Section 3 presents the concept of wavelet spectrum and some examples. In section 4 we prove some asymptotic properties of the finite wavelet transform. The wavelet periodogram is discussed in section 5. Section 6 presents the method of scalegram and an harmonic time series is constructed to illustrate this method and in section 7 wavelet shrinkage is developed. Finally, in section 8 we present an application.

2 Wavelets

There is a "scale function" or "father wavelet" ϕ, which is a solution of the dilation equation

$$\phi(x) = \sum_k c_k \phi(2x - k), \tag{1}$$

where the constantes c_k satisfy:

. $\sum_k c_k = 2$;

. suppose that r is the order of the vanishing moments of ϕ, then for a maximal $r \geq 1$, $\sum_k (-1)^k k^j c_k = 0$, $0 \leq j \leq r-1$;

. $\sum_k k^{2r} c_k^2 < \infty$;

. $\sum_k c_k c_{k-2m} = 2\delta_{0,m}$ and $\delta_{0,m}$ is the Kronecker delta.

This function generates a family

$$\phi_{j,k}(x) = 2^{-j/2}\phi(2^{-j}x - k), \quad j,\ k \in Z. \tag{2}$$

From the multiresolution analysis, the sequence $\{\phi_{j,k}(x), j, k \in Z\}$ is an orthonormal family in $L^2(\Re)$.

See Meyer(1990), Daubechies(1992) and Strang(1989).

Properties:

1) $\int_{-\infty}^{\infty} \phi(x)dx = 1$;

2) The functions $\phi_{j,k}$, $j, k \in Z$, are orthonormal:

$$\int_{-\infty}^{\infty} \phi_{j1,k1}(x)\phi_{j2,k2}(x)dx = \delta_{j1,j2}.\delta_{k1,k2};$$

3) The function ϕ has bounded support with $r-1$ vanishing moments, for some $r \geq 1$:

$$\int x^j \phi(x)dx = \delta_{0,j},$$

for $j = 0, \cdots, r-1$ and

$$\int \mid x^r \phi(x) \mid dx < \infty.$$

The wavelet function ψ, some times termed "mother wavelet", may be obtained from the scale function ϕ by taking differences in the dilation equation (1):

$$\psi(x) = \sum_k (-1)^k c_{k+1} \phi(2x+k), \tag{3}$$

such that $\{\psi_{j,k}(x), j, k \in Z\}$ is an orthonormal basis of $L^2(\Re)$, where

$$\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k), \quad j,\ k \in Z. \tag{4}$$

Properties:

1) $\int_{-\infty}^{\infty} \psi(x)dx = 0$ (admissibility);

2) The functions $\psi_{j,k}$, $j, k \in Z$ are orthonormal:

$$\int_{-\infty}^{\infty} \psi_{j1,k1}(x)\psi_{j2,k2}(x)dx = \delta_{j1,j2}.\delta_{k1,k2};$$

3) The first $r-1$ moments of ψ vanish (cancellations):

$$\int x^j \psi(x)dx = 0,$$

for $j = 0, \cdots, r-1$ and

$$\int \mid x^r \psi(x) \mid dx < \infty;$$

4) ψ is orthogonal to all the translates of ϕ:

$$\int_{-\infty}^{\infty} \psi(x)\phi(x+m)dx = 0, \quad -\infty < m < \infty.$$

We suppose that ϕ and ψ are bounded and compactly supported because in practice, ϕ and ψ are typically compactly supported.

See Daubechies(1992).

Figure 2.1 shows some examples of function ψ.

The oldest and simplest example of a function ψ for which the $\psi_{j,k}$ defined by (4) constitute an orthonormal basis for $L^2(\Re)$ is the Haar function,

$$\psi(x) = \begin{cases} 1, & 0 \leq x < 1/2 \\ -1, & 1/2 \leq x < 1 \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

From (5), we have

$$\psi^{(H)}_{j,k}(x) = \begin{cases} 2^{-j/2}, & 2^j k \leq x < 2^j(k+\frac{1}{2}) \\ -2^{-j/2}, & 2^j(k+\frac{1}{2}) \leq x < 2^j(k+1) \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

3 A Wavelet Spectrum

In this section we define the wavelet spectrum of a stationary time series $\{X(t),\, t \in Z\}$, with respect to a wavelet ψ.

Assumption 1: We assume that $\{X(t), t = 0, \pm 1, \cdots\}$ is a discrete time, zero mean, second-order stationary time series. Let $\gamma(u) = E\{X(t)X(t+u)\}$, $u = 0, \pm 1, \cdots$, be the autocovariance function of $X(t)$.

In our considerations we will take $X(0)$, $X(1)$, $\cdots$, $X(T-1)$ with $T = 2^M$, as a sample of the series $\{X(t), t \in Z\}$, where M is a positive integer.

Assumption 2:

$$\sum_{u=-\infty}^{\infty} [1+ \mid u \mid] \mid \gamma(u) \mid < \infty. \tag{7}$$

Suppose Assumptions 1 and 2 hold, then the *wavelet spectrum of the series* $\{X(t),\, t = 0, \pm 1, \cdots\}$, at (j,k), *with respect to* ψ, is defined to be

$$\eta^{(\psi)}_{j,k} = \sum_{u=-\infty}^{\infty} \sum_{t=0}^{\infty} \gamma(u)\psi_{j,k}(t)\psi_{j,k}(t+ \mid u \mid), \quad j,k \in Z, \tag{8}$$

where $\psi_{j,k}(\cdot)$ is defined in (4).

Theorem 1: If Assumptions 1 and 2 hold, then $\eta_{j,k}^{(\psi)}$ is bounded and non-negative.

Proof.
As ψ is bounded and compactly supported, then there is $A \in \Re$, a finite bound, such that $\mid \psi \mid \leq A$ and $\mid support\psi \mid < \infty$. But we know that $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$, thus we have

$$\mid \psi_{j,k}(\cdot) \mid \leq 2^{-j/2}A$$

and

$$\mid support\psi_{j,k} \mid = 2^{j} \mid support\psi \mid .$$

We denote by V the variation of ψ.
Now, let

$$S_T^{(\psi)}(j,k) = \sum_{u=-T}^{T} \sum_{t=0}^{T} \gamma(u)\psi_{j,k}(t)\psi_{j,k}(t+ \mid u \mid), \quad j,k \in Z,$$

Next, with $u > 0$, (the negative case follows similarly),

$$\begin{aligned}
&\mid \sum_t \psi_{j,k}(t+u)\psi_{j,k}(t) - \sum_t \psi_{j,k}(t)\psi_{j,k}(t) \mid \\
&\quad \leq \sum_t \mid \psi_{j,k}(t) \mid\mid \psi_{j,k}(t+u) - \psi_{j,k}(t) \mid \\
&\quad \leq 2^{-j/2}A \sum_t \mid \psi_{j,k}(t+u) - \psi_{j,k}(t+u-1) + \psi_{j,k}(t+u-1) - \psi_{j,k}(t+u-2) + \\
&\quad\ + \psi_{j,k}(t+u-2) - \cdots + \psi_{j,k}(t+1) - \psi_{j,k}(t) \mid \\
&\quad \leq 2^{-j/2}A \sum_t \{\mid \psi_{j,k}(t+u) - \psi_{j,k}(t+u-1) \mid + \mid \psi_{j,k}(t+u-1) - \psi_{j,k}(t+u-2) \mid + \\
&\quad\ + \cdots + \mid \psi_{j,k}(t+1) - \psi_{j,k}(t) \mid\} \\
&\quad \leq 2^{-j/2}A2^{-j/2}V \mid u \mid \\
&\quad = 2^{-j}AV \mid u \mid .
\end{aligned}$$

Thus,

$$
\begin{aligned}
S_T^{(\psi)}(j,k) &= \sum_{u=-T}^{T} \gamma(u) \sum_{t=0}^{T} \psi_{j,k}(t)\psi_{j,k}(t+ \mid u \mid) \\
&= \sum_{u=-T}^{T} \gamma(u) \sum_{t=0}^{T} \psi_{j,k}(t)\psi_{j,k}(t) + \epsilon_T \\
&\leq \sum_{u=-T}^{T} \gamma(u) A^2 \mid support \psi \mid + \epsilon_T,
\end{aligned}
$$

because,

$$
\begin{aligned}
\sum_{t=0}^{T} \psi_{j,k}(t)\psi_{j,k}(t) &\leq (2^{-j/2}A)^2 \mid support \psi_{j,k} \mid \\
&\leq 2^{-j} A^2 2^j \mid support \psi \mid \\
&\leq A^2 \mid support \psi \mid .
\end{aligned}
$$

Also,

$$
\begin{aligned}
\mid \epsilon_T \mid &\leq \sum_{u=-T}^{T} \mid \gamma(u) \mid 2^{-j} AV \mid u \mid \\
&\leq 2^{-j} AV \sum_{u=-T}^{T} \mid u \mid\mid \gamma(u) \mid .
\end{aligned}
$$

Hence, by the Assumption 2, we have that

$$
\eta_{j,k}^{(\psi)} = \lim_{T \to \infty} S_T^{(\psi)}(j,k) < \infty.
$$

Now, let $I_{j,k}^{(\psi)} = [\sum_{t=0}^{T-1} X(t)\psi_{j,k}(t)]^2 \geq 0$. Then,

$$
\begin{aligned}
E[I_{j,k}^{(\psi)}] &= \sum_{t=0}^{T-1} \sum_{s=0}^{T-1} E[X(t)X(s)] \psi_{j,k}(t)\psi_{j,k}(s) \\
&= \sum_{t=0}^{T-1} \sum_{s=0}^{T-1} \gamma(t-s) \psi_{j,k}(t)\psi_{j,k}(s) \\
&= \sum_{u=-(T-1)}^{T-1} \gamma(u) \sum_{t=0}^{T-|u|} \psi_{j,k}(t)\psi_{j,k}(t+ \mid u \mid).
\end{aligned}
$$

Therefore, as $T \to \infty$, $E[I_{j,k}^{(\psi)}] \to \eta_{j,k}^{(\psi)}$ because Assumption 2 holds, hence $\eta_{j,k}^{(\psi)} \geq 0$.

Remark: $I_{j,k}^{(\psi)} = [\sum_{t=0}^{T-1} X(t)\psi_{j,k}(t)]^2$ is called the *second order wavelet periodogram* of the values $X(0), \cdots, X(T-1)$ with respect to ψ. It is an estimator of the spectrum $\eta_{j,k}^{(\psi)}$ and by Theorem 1, it is asymptotically unbiased.

We now give some examples of wavelet spectrum with respect to the Haar wavelet.

Example 3.1: If $\{\varepsilon(t),\ t = 0, 1, \cdots\}$, is a white noise sequence, with mean zero and variance σ^2, then,

$$\begin{aligned}
\eta_{j,k}^{(H)} &= \sum_{u=-\infty}^{\infty} \sum_{t=0}^{\infty} \gamma(u)\psi_{j,k}^{(H)}(t)\psi_{j,k}^{(H)}(t+\mid u\mid) \\
&= \sigma^2 \sum_{t=0}^{\infty} [\psi_{j,k}^{(H)}(t)]^2 \\
&= [2^{-j} I_{\{j<0, k\equiv 0 (mod 2^{-j})\}} + I_{\{j\geq 0\}}]\sigma^2 I_{\{k\geq 0\}}.
\end{aligned}$$

Hence, for $j \geq 0$, $k \geq 0$, $\eta_{j,k}^{(\psi)} = \sigma^2$, that is, constant. It suggests that if we find a flat spectrum for $j \geq 0$, $k \geq 0$, we may be in the presence of a white noise series.

Figure 3.1 shows $\eta_{j,k}^{(H)}$ of $\varepsilon(t)$, with $\sigma = 1$.

Example 3.2: Let us consider a first order moving average, MA(1),

$$X(t) = \varepsilon(t) - \beta\varepsilon(t-1),$$

with $\varepsilon_t \sim WN(0, \sigma^2)$. We obtain immediately

$$\gamma(u) = \begin{cases} (1+\beta^2)\sigma^2, & u = 0 \\ -\beta\sigma^2, & \mid u \mid = 1 \\ 0, & \mid u \mid \geq 2. \end{cases}$$

Then,

$$
\begin{aligned}
\eta_{j,k}^{(H)} &= \sum_{u=-\infty}^{\infty}\sum_{t=0}^{\infty}\gamma(u)\psi_{j,k}^{(H)}(t)\psi_{j,k}^{(H)}(t+\mid u\mid) \\
&= \gamma(0)\sum_{t=0}^{\infty}\psi_{j,k}^{(H)}(t)\psi_{j,k}^{(H)}(t) + 2\gamma(1)\sum_{t=0}^{\infty}\psi_{j,k}^{(H)}(t)\psi_{j,k}^{(H)}(t+1) \\
&= \{\gamma(0)[2^{-j}I_{\{j<0,k\equiv 0(mod 2^{-j})\}} + I_{\{j\geq 0\}}] + 2\gamma(1)[1-\frac{3}{2^j}]I_{\{j\geq 1\}}\}I_{\{k\geq 0\}}.
\end{aligned}
$$

In particular, with $\beta = 0.7$, we have $\eta_{j,k}^{(H)} = 0.09\sigma^2$, as $j \to \infty$ and $k \geq 0$.

Figure 3.2 shows $\gamma(u)$ and $\eta_{j,k}^{(H)}$, with $\beta = 0.7$ and $\sigma = 1$.

Example 3.3: Consider now a second order moving average, MA(2):

$$X(t) = \varepsilon(t) - \beta_1\varepsilon(t-1) - \beta_2\varepsilon(t-2),$$

where $\varepsilon_t \sim WN(0,\sigma^2)$. Then we have

$$
\gamma(u) = \begin{cases} (1+\beta_1^2+\beta_2^2)\sigma^2, & u = 0 \\ (-\beta_1(1-\beta_2))\sigma^2, & \mid u\mid = 1 \\ -\beta_2\sigma^2, & \mid u \mid = 2 \\ 0, & \mid u \mid \geq 3. \end{cases}
$$

Then,

$$
\begin{aligned}
\eta_{j,k}^{(H)} &= \sum_{u=-\infty}^{\infty}\sum_{t=0}^{\infty}\gamma(u)\psi_{j,k}^{(H)}(t)\psi_{j,k}^{(H)}(t+\mid u\mid) \\
&= \gamma(0)\sum_{t=0}^{\infty}\psi_{j,k}^{(H)}(t)\psi_{j,k}^{(H)}(t) + 2\gamma(1)\sum_{t=0}^{\infty}\psi_{j,k}^{(H)}(t)\psi_{j,k}^{(H)}(t+1) \\
&\quad + 2\gamma(2)\sum_{t=0}^{\infty}\psi_{j,k}^{(H)}(t)\psi_{j,k}^{(H)}(t+2) \\
&= \{\gamma(0)[2^{-j}I_{\{j<0,k\equiv 0(mod 2^{-j})\}} + I_{\{j\geq 0\}}] + 2\gamma(1)[1-\frac{3}{2^j}]I_{\{j\geq 1\}} \\
&\quad + 2\gamma(2)[1-\frac{6}{2^j}]I_{\{j\geq 2\}}\}I_{\{k\geq 0\}}.
\end{aligned}
$$

In particular, with $\beta_1 = 0.5$ and $\beta_2 = -0.3$, we have $\eta_{j,k}^{(H)} = 0.64\sigma^2$, as $j \to \infty$ and $k \geq 0$.

Figure 3.3 shows $\gamma(u)$ and $\eta_{j,k}^{(H)}$, with $\beta_1 = 0.5$, $\beta_2 = -0.3$ and $\sigma = 1$.

Example 3.4: Let us consider a first order autoregressive, AR(1):

$$X(t) = \alpha X(t-1) + \varepsilon(t), \quad |\, \alpha \,| < 1,$$

where $\varepsilon_t \sim WN(0, \sigma^2)$. Thus we obtain

$$\gamma(u) = \begin{cases} \frac{1}{1-\alpha^2}\sigma^2, & u = 0 \\ \alpha^u \gamma(0), & u \geq 1. \end{cases}$$

Then,

$$\begin{aligned} \eta_{j,k}^{(H)} &= \sum_{u=-\infty}^{\infty} \sum_{t=0}^{\infty} \gamma(u) \psi_{j,k}^{(H)}(t) \psi_{j,k}^{(H)}(t + |\, u \,|) \\ &= \gamma(0) \sum_{t=0} \psi_{j,k}^{(H)}(t) \psi_{j,k}^{(H)}(t) + 2\gamma(1) \sum_{t=0}^{\infty} \psi_{j,k}^{(H)}(t) \psi_{j,k}^{(H)}(t+1) \\ &\quad + 2\gamma(2) \sum_{t=0}^{\infty} \psi_{j,k}^{(H)}(t) \psi_{j,k}^{(H)}(t+2) + \cdots \\ &= \gamma(0)\{[2^{-j} I_{\{j<0, k \equiv 0 (mod 2^{-j})\}} + I_{\{j \geq 0\}}] + 2\alpha[1 - \frac{3}{2^j}] I_{\{j=1\}} \\ &\quad + \{2 \sum_{u=0}^{2(j-1)} \alpha^u [1 - \frac{3u}{2^j}] - 2 \sum_{u=2(j-1)+1}^{2^j - 1} \alpha^u [1 - \frac{u}{2^j}]\} I_{\{j>1\}}\} I_{\{k \geq 0\}} \end{aligned}$$

In particular, with $\alpha = 0.8$, we have $\eta_{j,k}^{(H)} = 25\sigma^2$ as $j \to \infty$ and $k \geq 0$.

Figure 3.4 shows $\gamma(u)$ and $\eta_{j,k}^{(H)}$, with $\alpha = 0.8$ and $\sigma = 1$.

4 The finite wavelet transform and its asymptotic properties

In this section we develop some stochastic properties of the finite wavelet transform of the sequence $X(t)$, $t = 0, 1, \cdots, T-1$. Here, $\{X(t),\, t \in Z\}$ is a zero mean, discrete stationary process and $T = 2^M$.

For $j = 1, 2, \cdots, M$ and $k = 0, 1, \cdots, (2^{(M-j)} - 1)$, we define the *finite wavelet transform* with respect to ψ as

$$d_{j,k}^{(\psi)} = \sum_{t=0}^{T-1} X(t)\psi_{j,k}(t). \tag{9}$$

Then,

$$E\{d_{j,k}^{(\psi)}\} = 0$$

and

$$\begin{aligned} Var\{d_{j,k}^{(\psi)}\} &= \sum_{t=0}^{T-1}\sum_{s=0}^{T-1} \gamma(t-s)\psi_{j,k}(t)\psi_{j,k}(s) \\ &= \sum_{u=-(T-1)}^{T-1} \gamma(u) \sum_{t=0}^{T-1-|u|} \psi_{j,k}(t)\psi_{j,k}(t+ \mid u \mid). \end{aligned}$$

If Assumption 2 holds, then $Var\{d_{j,k}^{(\psi)}\} \to \eta_{j,k}^{(\psi)}$, as $T \to \infty$.

For the Haar wavelet, the $d_{j,k}^{(H)}$ are given by

$$d_{j,k}^{(\psi)} = 2^{-j/2}[\sum_{t=k2^j}^{2^j(k+1/2)} X(t) - \sum_{t=2^j(k+1/2)}^{2^j(k+1)}],$$

which are local means of $X(t)$.

We now turn to the investigation of the covariance structure of the finite wavelet transform. The next result is immediate.

Theorem 2: Suppose Assumption 1 holds and

$$\sum_{t=0}^{\infty}\sum_{s=0}^{\infty} \mid \gamma(t-s) \mid < \infty. \tag{10}$$

Let

$$\eta_{(j,j'),(k,k')}^{(\psi)} = \sum_{u=-\infty}^{\infty}\sum_{t=0}^{\infty} \gamma(u)\psi_{j,k}(t+ \mid u \mid I_{\{u>0\}})\psi_{j',k'}(t+ \mid u \mid I_{\{u\leq 0\}}),$$

for $j, j', k, k' \in Z$. Then,

i) $E\{d^{(\psi)}_{j,k} d^{(\psi)}_{j',k'}\} \rightarrow \eta^{(\psi)}_{(j,j'),(k,k')}$, as $T \rightarrow \infty$;

ii) If $j = j'$, $k = k'$, then $\eta^{(\psi)}_{(j,j'),(k,k')} = \eta^{(\psi)}_{j,k}$.

Remark: $\eta^{(\psi)}_{(j,j'),(k,k')}$ is called the (*asymptotic*) *covariance of the wavelet transform* with respect to ψ.

Theorem 3: Suppose Assumptions 1 and 2 hold, then

$$E\{d^{(\psi)}_{j,k} d^{(\psi)}_{j',k'}\} = O(1), \quad as \quad T \rightarrow \infty,$$

for $T = 2^M$ and $(j,k) \neq (j',k')$.

Proof:

Let $d^{(\psi)}_{j,k} = \sum_{t=0}^{T-1} X(t)\psi_{j,k}(t)$,

$$\begin{aligned} E\{d^{(\psi)}_{j,k} d^{(\psi)}_{j',k'}\} &= \sum_{t,s=0}^{T-1} \gamma(t-s)\psi_{j,k}(t)\psi_{j',k'}(s) \\ &= \sum_{u=-(T-1)}^{T-1} \gamma(u) \sum_{t=0}^{T-1-|u|} \psi_{j,k}(t+ \mid u \mid I_{\{u>0\}})\psi_{j',k'}(t+ \mid u \mid I_{\{u\leq 0\}}). \end{aligned}$$

Next, with $u > 0$, (the negative case follows similarly),

$$\begin{aligned} &\mid \sum_t \psi_{j,k}(t+u)\psi_{j',k'}(t) - \sum_t \psi_{j,k}(t)\psi_{j',k'}(t) \mid \\ &\quad \leq \sum_t \mid \psi_{j',k'}(t) \mid\mid \psi_{j,k}(t+u) - \psi_{j,k}(t) \mid \\ &\quad \leq 2^{-j'/2} A \sum_t \mid \psi_{j,k}(t+u) - \psi_{j,k}(t+u-1) + \psi_{j,k}(t+u-1) - \psi_{j,k}(t+u-2) + \\ &\quad\ + \psi_{j,k}(t+u-2) - \cdots + \psi_{j,k}(t+1) - \psi_{j,k}(t) \mid \\ &\quad \leq 2^{-j'/2} A \sum_t \{\mid \psi_{j,k}(t+u) - \psi_{j,k}(t+u-1) \mid + \mid \psi_{j,k}(t+u-1) - \psi_{j,k}(t+u-2) \mid + \end{aligned}$$

$$
\begin{aligned}
&+\cdots+ \mid \psi_{j,k}(t+1)-\psi_{j,k}(t) \mid\} \\
&\leq 2^{-j^{'}/2}A2^{-j/2}V \mid u \mid \\
&= 2^{-(j^{'}+j)/2}AV \mid u \mid
\end{aligned}
$$

Thus,

$$
\begin{aligned}
E\{d_{j,k}^{(\psi)}d_{j',k'}^{(\psi)}\} &= \sum_{u=-(T-1)}^{T-1}\gamma(u)\sum_{t=0}^{T-1-|u|}\psi_{j,k}(t+\mid u\mid I_{\{u>0\}})\psi_{j',k'}(t+\mid u\mid I_{\{u\leq 0\}}) \\
&= \sum_{u=-(T-1)}^{T-1}\gamma(u)\sum_{t=0}^{T-1-|u|}\psi_{j,k}(t)\psi_{j',k'}(t)+\epsilon_T \\
&= O(1),
\end{aligned}
$$

as $T\to\infty$, because by Assumption 2, we have

$$
\begin{aligned}
&\mid \sum_u \gamma(u)\sum_t \psi_{j,k}(t)\psi_{j',k'}(t) \mid \\
&\quad\leq \sum_u \mid \gamma(u) \mid \sum_t \mid \psi_{j,k}(t)\psi_{j',k'}(t) \mid \\
&\quad\leq \sum_u \mid \gamma(u) \mid 2^{j/2}A2^{j^{'}/2}A \mid [support\psi_{j,k}]\cap[support\psi_{j',k'}] \mid \\
&\quad\leq \sum_u \mid \gamma(u) \mid 2^{(j+j^{'})/2}A^2 \mid [support\psi_{j,k}]\cap[support\psi_{j',k'}] \mid < \infty,
\end{aligned}
$$

and

$$
\begin{aligned}
\mid \epsilon_T \mid &\leq \sum_{u=-(T-1)}^{T-1} \mid \gamma(u) \mid 2^{-(j^{'}+j)/2}AV \mid u \mid \\
&\leq 2^{-(j^{'}+j)/2}AV\sum_{u=-(T-1)}^{T-1} \mid u \mid\mid \gamma(u) \mid < \infty.
\end{aligned}
$$

By Theorem 3 we see that the asymptotic covariance of the finite wavelet transform at two distinct pairs (j,k) and $(j^{'},k^{'})$ is not necessarily zero, which is in contrast to the trigonometric Fourier transform case, where the covariances are asymptotically zero.

Theorem 4: If $\psi(t)$ has support $[K_1, K_2]$, where $K_1 \geq 0$ and $K_2 \geq 0$ and $X(t)$ is a discrete time, zero mean, second order stationary time series with $\gamma(u) \neq 0$ for $\mid u \mid \leq U$, $U << T$, then, $E(d_{j,k_1}^{(\psi)} d_{j,k_2}^{(\psi)}) = 0$ for $\mid k_1 - k_2 \mid > K_2 - K_1 + \frac{U}{2^j}$, $j = 1, \cdots, M$ and $k = 0, 1, \cdots, (2^{M-j} - 1)$.

Proof:

Let

$$E(d_{j,k_1}^{(\psi)} d_{j,k_2}^{(\psi)}) = \sum_{u=-(T-1)}^{T-1} \gamma(u) \sum_{t=0}^{T-1-|u|} \psi_{j,k_1}(t+ \mid u \mid I_{\{u>0\}}) \psi_{j,k_2}(t+ \mid u \mid I_{\{u\leq 0\}}),$$

then,

$$[support \psi_{j,k_1}] = [2^j(K_1 + k_1) - \mid u \mid I_{\{u>0\}}; 2^j(K_2 + k_1) - \mid u \mid I_{\{u>0\}}]$$

and

$$[support \psi_{j,k_2}] = [2^j(K_1 + k_2) - \mid u \mid I_{\{u\leq 0\}}; 2^j(K_2 + k_2) - \mid u \mid I_{\{u\leq 0\}}].$$

It follows that

$$[support \psi_{j,k_1}] \cap [support \psi_{j,k_2}]$$

is empty if $\mid k_1 - k_2 \mid > K_2 - K_1 + \frac{U}{2^j}$.

Remark: By choosing a wavelet with large value for K_2 and small value for K_1, few wavelet coefficients satisfy the condition $\mid k_1 - k_2 \mid > K_2 - K_1 + \frac{U}{2^j}$, and then the $d_{j,k}^{(\psi)}$ are independent just over a smaller set.

We now come to a central limit theorem for the wavelet transform. Both the statement of the theorem and the proof follow Hannan(1973) very closely.

Theorem 5: Suppose $X(t)$ is strictly stationary and let $\mathcal{M}_t$ be the σ-field generated by $(X(j), j \leq t)$. Put

$$\alpha_j = [E\{E(X(t)/\mathcal{M}_{t-j}) - E(X(t)/\mathcal{M}_{t-j-1})\}^2]^{1/2}, \quad j \geq 0.$$

We assume that $\mathcal{M}_{-\infty}$ is trivial and that $\sum_{j=0}^{\infty} \alpha_j < \infty$.

(i) For $(j_1,k_1),\cdots,(j_p,k_p)$ and $m,n=1,\cdots,p$, suppose that

$$E\{d^{(\psi)}_{j_m,k_m} d^{(\psi)}_{j_n,k_n}\} \to \eta^{(\psi)}_{(j_m,j_n),(k_m,k_n)}, \quad as \quad T\to\infty.$$

Then,

$$\Phi \xrightarrow{\mathcal{D}} \mathcal{N}(0,\Delta),$$

where

$$\Delta = (\eta^{(\psi)}_{(j_m,j_n),(k_m,k_n)}),$$

and

$$\Phi = (d^{(\psi)}_{j_1,k_1},\cdots,d^{(\psi)}_{j_p,k_p})'.$$

(ii) In particular, we have

$$d^{(\psi)}_{j,k} \xrightarrow{\mathcal{D}} \mathcal{N}(0,\eta^{(\psi)}_{j,k}).$$

Proof:

(i) Let $Y^{(T)}(t)=\sqrt{T}\psi_{j,k}(t)$ and $d^2(T)=\sum_{t=0}^{T-1}\{Y^{(T)}(t)\}^2$. Then,

$$\begin{aligned} d^2(T) &= \sum_{t=0}^{T-1}\{Y^{(T)}(t)\}^2 \\ &= T\sum_{t=0}^{T-1}[\psi_{j,k}(t)]^2 \\ &= T+O(1), \end{aligned}$$

because

$$\sum_{t=0}^{T-1}[\psi_{j,k}(t)]^2 = 1+O(T^{-1}).$$

Now we have,

. $\lim_{T\to\infty} d^2(T)=\infty$;

. $\lim_{T\to\infty}\frac{|Y^{(T)}(T)|}{d(T)} = \lim_{T\to\infty}\frac{\sqrt{T}|\psi_{j,k}(T)|}{\sqrt{T}} = 0$, because of $|\ support\psi_{j,k}\ | < \infty$;

. $\lim_{T\to\infty}\frac{\{\sum_{m=1}^{T-t} Y^{(T)}(m)Y^{(T)}(m+t)\}}{d^2(T)} = \lim_{T\to\infty}\frac{T\{\sum_{s=0}^{T-1-t}\psi_{j,k}(s)\psi_{j,k}(s+t)\}}{T} = \rho(t).$

Thus we have,

$$\rho(t) = \lim_{T\to\infty}\{\sum_{s=0}^{T-1-t} \psi_{j,k}(s)\psi_{j,k}(s+t)\}.$$

Now,

$$\begin{aligned}
\eta_{j,k}^{(\psi)} &= \sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\gamma(t-s)\psi_{j,k}(t)\psi_{j,k}(s) \\
&= \lim_{T\to\infty}\sum_{t=0}^{T-1}\sum_{s=0}^{T-1}\gamma(t-s)\psi_{j,k}(s)\psi_{j,k}(t) \\
&= \lim_{T\to\infty}\sum_{u=-(T-1)}^{T-1}\gamma(u)\sum_{s=0}^{T-1-|u|}\psi_{j,k}(s)\psi_{j,k}(s+\mid u\mid) \\
&= \sum_{u=-\infty}^{\infty}\gamma(u)\{\lim_{T\to\infty}\sum_{s}\psi_{j,k}(s)\psi_{j,k}(s+\mid u\mid)\} \\
&= \sum_{u=-\infty}^{\infty}\gamma(u)\rho(\mid u\mid).
\end{aligned}$$

Thus, Theorem 4(i) can be proved in exactly the same way as Theorem 1(i) of Hannan (1973).

(ii) Immediate.

Remark: Theorem 5 is proved under a form of strong mixing condition, that is, asymptotic independence.

Another central limit theorem can be proved, under a different form of mixing assumption, based on cumulants. Define

$$C_x(u_1,\cdots,u_{k-1}) = Cum\{X(t+u_1),\cdots,X(t+u_{k-1}),X(t)\}$$

as the kth order cumulant of $X(t)$, $u_1,\cdots,u_{k-1} = 0,\pm1,\cdots$, $k = 2,3\cdots$ and $C_x(u_1) = \gamma(u_1)$.

Under the assumption that

$$\sum_{u_1}\cdots\sum_{u_{k-1}}\mid u_j\mid\mid C_x(u_1,\cdots,u_{k-1})\mid < \infty,$$

it can be proved that $d_{j,k}^{(\psi)}$ is asymptotically normal with mean zero and variance $\sum_u \gamma(u)$. Note that $\sum_u \gamma(u) = 2\pi f(0)$, where $f(\lambda)$ is the Fourier spectrum of $X(t)$. The proof follows along the same lines as in Brillinger (1994).

5 The wavelet periodogram

We have an observed sequence of values $X(0), \cdots, X(T-1)$ and based on these values we want to estimate $\eta_{j,k}^{(\psi)}$.

We first compute the finite wavelet transform with respect to ψ:

$$d_{j,k}^{(\psi)} = \sum_{t=0}^{T-1} X(t)\psi_{j,k}(t).$$

We have proved that $d_{j,k}^{(\psi)}$ is asymptotically normal $\mathcal{N}(0, \eta_{j,k}^{(\psi)})$ and this sugests that $\eta_{j,k}^{(\psi)}$ may be estimated by the statistic

$$I_{j,k}^{(\psi)} = (d_{j,k}^{(\psi)})^2 = [\sum_{t=0}^{T-1} X(t)\psi_{j,k}(t)]^2. \tag{11}$$

The statistic $I_{j,k}^{(\psi)}$ of (11) is called the *second order wavelet periodogram*, or more briefly *wavelet periodogram*, of the values $X(0), \cdots, X(T-1)$.

Remark: Let us define the " energy" in a time series as $\sum_{t=0}^{T-1} X^2(t)$, then, by using Parseval's theorem, we can decompose the energy in a time series with respect to scale j and time k by $I_{j,k}^{(\psi)}$ for $j = 1, \cdots, M$, $k = 0, 1, \cdots, 2^{(M-j)} - 1$ and $T = 2^M$.

In Theorem 1 we showed that $I_{j,k}^{(\psi)}$ is asymptotically unbiased. We also have,

Theorem 6: Let Assumption 2 be satisfied. Then,

$$E[I_{j,k}^{(\psi)}] = \eta_{j,k}^{(\psi)} + O(T^{-1}).$$

Proof:

We have,

$$I_{j,k}^{(\psi)} = [d_{j,k}^{(\psi)}]^2.$$

Then,

$$\begin{aligned}
E\{I_{j,k}^{(\psi)}\} &= \sum_{t=0}^{T-1}\sum_{s=0}^{T-1} \gamma(t-s)\psi_{j,k}(t)\psi_{j,k}(s) \\
&= \sum_{u=-(T-1)}^{T-1} \gamma(u) \sum_{t=0}^{T-1-|u|} \psi_{j,k}(t)\psi_{j,k}(t+\mid u \mid) \\
&= \sum_{u=-\infty}^{\infty} \gamma(u) \sum_{t=0}^{\infty} \psi_{j,k}(t)\psi_{j,k}(t+\mid u \mid) + \eta_T \\
&= \eta_{j,k}^{(\psi)} + \eta_T,
\end{aligned}$$

where

$$\begin{aligned}
\mid \eta_T \mid &= \mid \sum_{u=T}^{\infty} \gamma(u) \sum_{t=0}^{\infty} \psi_{j,k}(t)\psi_{j,k}(t+\mid u \mid) + \sum_{u=-\infty}^{-T} \gamma(u) \sum_{t=0}^{\infty} \psi_{j,k}(t)\psi_{j,k}(t+\mid u \mid) \\
&\quad + \sum_{u=-(T-1)}^{T-1} \gamma(u) \sum_{t=T}^{\infty} \psi_{j,k}(t-\mid u \mid)\psi_{j,k}(t) \mid \\
&\leq \mid \sum_{u=T}^{\infty} \gamma(u) \sum_{t=0}^{\infty} \psi_{j,k}(t)\psi_{j,k}(t+\mid u \mid) \mid + \mid \sum_{u=-\infty}^{-T} \gamma(u) \sum_{t=0}^{\infty} \psi_{j,k}(t)\psi_{j,k}(t+\mid u \mid) \mid \\
&\quad + \mid \sum_{u=-(T-1)}^{T-1} \gamma(u) \sum_{t=T}^{\infty} \psi_{j,k}(t-\mid u \mid)\psi_{j,k}(t) \mid \\
&\leq \mid S_1 \mid + \mid S_2 \mid + \mid S_3 \mid, say.
\end{aligned}$$

Now,

$$\begin{aligned}
\mid S_1 \mid &\leq \sum_{u=T}^{\infty} \frac{\mid u \mid}{T} \mid \gamma(u) \mid \sum_{t=0}^{\infty} \mid \psi_{j,k}(t)\psi_{j,k}(t+\mid u \mid) \mid \\
&\leq \frac{1}{T} \sum_{u=-\infty}^{\infty} \mid u \mid\mid \gamma(u) \mid \sum_{t=0}^{\infty} \mid \psi_{j,k}(t)\psi_{j,k-2^{-j}|u|}(t) \mid \\
&\leq \frac{1}{T} \sum_{-\infty}^{\infty} \mid u \mid\mid \gamma(u) \mid 2^{-j} A \mid [support \psi_{j,k}] \cap [support \psi_{j,k-2^{-j}|u|}] \mid .
\end{aligned}$$

Thus, $S_1 = O(T^{-1})$.

Similarly, $S_2 = O(T^{-1})$.

Now,

$$\begin{aligned}
\mid S_3 \mid &\leq \sum_{u=-(T-1)}^{T-1} \mid \gamma(u) \mid \sum_{t=T}^{\infty} \frac{\mid t \mid}{T} \mid \psi_{j,k}(t- \mid u \mid)\psi_{j,k}(t) \mid \\
&\leq \frac{1}{T} \sum_{u=-\infty}^{\infty} \mid \gamma(u) \mid \sum_{t=T}^{\infty} \mid t \mid\mid \psi_{j,k+2^{-j}|u|}(t)\psi_{j,k}(t) \mid \\
&\leq \frac{1}{T} \sum_{u=-\infty}^{\infty} \mid \gamma(u) \mid 2^{-j} A \mid [support \psi_{j,k+2^{-j}|u|}] \cap [support \psi_{j,k}] \mid M_{\psi},
\end{aligned}$$

where $-\infty < M_{\psi} = sup_{j,k}\{[support\psi_{j,k+2^{-j}|u|}] \cap [support\psi_{j,k}]\} < \infty$. Thus, $S_3 = O(T^{-1})$.

The covariance structure of the periodogram is given below.

Theorem 7: Suppose Assumptions 1 and 2 hold and

$$\sum_{l,m,n=-\infty}^{\infty} \mid C_x(l,m,n) \mid < \infty, \tag{12}$$

where $C_x(l,m,n) = Cum\{X(0), X(l), X(m), X(n)\}$ is the fourth order cumulant of $X(t)$. Then

$$Cov\{I_{j,k}^{(\psi)}, I_{j',k'}^{(\psi)}\} = 2\{\eta_{(j,j'),(k,k')}^{(\psi)}\}^2 + O(1), \quad as \quad T \to \infty.$$

Proof:

$$\begin{aligned}
I_{j,k}^{(\psi)} I_{j',k'}^{(\psi)} &= (d_{j,k}^{(\psi)})^2 (d_{j',k'}^{(\psi)})^2 \\
&= [\sum_{t_1=0}^{T-1} X(t_1)\psi_{j,k}(t_1)]^2 [\sum_{t_2=0}^{T-1} X(t_2)\psi_{j',k'}(t_2)]^2 \\
&= \sum_{t_1,s_1,t_2,s_2=0}^{T-1} X(t_1)X(s_1)X(t_2)X(s_2)\psi_{j,k}(t_1)\psi_{j,k}(s_1)\psi_{j',k'}(t_2)\psi_{j',k'}(s_2),
\end{aligned}$$

and hence,

$$E\{I_{j,k}^{(\psi)}I_{j',k'}^{(\psi)}\} = \sum_{t_1,s_1,t_2,s_2=0}^{T-1} E\{X(t_1)X(s_1)X(t_2)X(s_2)\}\psi_{j,k}(t_1)\psi_{j,k}(s_1)\psi_{j',k'}(t_2)\psi_{j',k'}(s_2).$$

From Hannan (1970, pag.23), we have,

$$\begin{aligned} E\{I_{j,k}^{(\psi)}I_{j',k'}^{(\psi)}\} &= \sum_{t_1,s_1,t_2,s_2=0}^{T-1}\{\gamma(t_1-s_1)\gamma(t_2-s_2)+\gamma(t_1-s_2)\gamma(s_1-t_2)+\gamma(s_1-s_2). \\ &\quad .\gamma(t_1-t_2)+C_x(s_1-t_1,t_2-t_1,s_2-t_1)\}\psi_{j,k}(t_1)\psi_{j,k}(s_1)\psi_{j',k'}(t_2)\psi_{j',k'}(s_2) \\ &= S_1+S_2+S_3+S_4, \quad say. \end{aligned}$$

$$\begin{aligned} S_1 &= \sum_{t_1,s_1,t_2,s_2=0}^{T-1}\gamma(t_1-s_1)\gamma(t_2-s_2)\psi_{j,k}(t_1)\psi_{j,k}(s_1)\psi_{j',k'}(t_2)\psi_{j',k'}(s_2) \\ &= \{\sum_{t_1,s_1=0}^{T-1}\gamma(t_1-s_1)\psi_{j,k}(t_1)\psi_{j,k}(s_1)\}\{\sum_{t_2,s_2=0}^{T-1}\gamma(t_2-s_2)\psi_{j',k'}(t_2)\psi_{j',k'}(s_2)\} \\ &= E(I_{j,k}^{(\psi)})E(I_{j',k'}^{(\psi)}). \end{aligned}$$

$$\begin{aligned} S_2 &= \sum_{t_1,s_1,t_2,s_2=0}^{T-1}\gamma(t_1-s_2)\gamma(s_1-t_2)\psi_{j,k}(t_1)\psi_{j,k}(s_1)\psi_{j',k'}(t_2)\psi_{j',k'}(s_2) \\ &= \{\sum_{t_1,s_2=0}^{T-1}\gamma(t_1-s_2)\psi_{j,k}(t_1)\psi_{j',k'}(s_2)\}\{\sum_{t_2,s_1=0}^{T-1}\gamma(s_1-t_2)\psi_{j,k}(s_1)\psi_{j',k'}(t_2)\} \\ &\rightarrow \{\eta_{(j,j'),(k,k')}^{(\psi)}\}^2, \quad as \quad T\rightarrow\infty. \end{aligned}$$

$$S_3 = \sum_{t_1,s_1,t_2,s_2=0}^{T-1}\gamma(s_1-s_2)\gamma(t_1-t_2)\psi_{j,k}(t_1)\psi_{j,k}(s_1)\psi_{j',k'}(t_2)\psi_{j',k'}(s_2)$$

$$
\begin{aligned}
&= \{\sum_{s_1,s_2=0}^{T-1} \gamma(s_1-s_2)\psi_{j,k}(s_1)\psi_{j',k'}(s_2)\}\{\sum_{t_1,t_2=0}^{T-1} \gamma(t_1-t_2)\psi_{j,k}(t_1)\psi_{j',k'}(t_2)\} \\
&\to \{\eta^{(\psi)}_{(j,j'),(k,k')}\}^2, \quad as \quad T \to \infty.
\end{aligned}
$$

Finally,

$$
S_4 = \sum_{t_1,s_1,t_2,s_2=0}^{T-1} C_x(s_1-t_1, t_2-t_1, s_2-t_1)\psi_{j,k}(t_1)\psi_{j,k}(s_1)\psi_{j',k'}(t_2)\psi_{j',k'}(s_2) < \infty,
$$

if (12) holds. Then $S_4 = O(1)$.

Thus,

$$
Cov(I^{(\psi)}_{j,k}, I^{(\psi)}_{j',k'}) = 2\{\eta^{(\psi)}_{(j,j'),(k,k')}\}^2 + O(1), \quad as \quad T \to \infty.
$$

In particular, if $j = j'$, $k = k'$, we obtain the variance

$$
Var[I^{(\psi)}_{j,k}] = 2\{\eta^{(\psi)}_{j,k}\}^2 + O(1).
$$

Remark: Theorem 7 shows that the wavelet periodogram is not a consistent estimate, since the variance does not $\to 0$ as $T \to \infty$.

The asymptotic distribution of the wavelet periodogram is obtained immediately from Theorem 5.

Theorem 8: Under the assumption of Theorem 5, $I^{(\psi)}_{j,k}$ is an asymptotically $\eta^{(\psi)}_{j,k}\chi^2_1$ variable.

Proof:
Since $d^{(\psi)}_{j,k}$ is asymptotically $\mathcal{N}(0, \eta^{(\psi)}_{j,k})$, then

$$
\frac{d^{(\psi)}_{j,k}}{[\eta^{(\psi)}_{j,k}]^{1/2}}
$$

is asymptotically $\mathcal{N}(0,1)$, hence its square is asymptotically χ_1^2.

We remember that the Fourier periodogram $I(\lambda) = \frac{1}{2\pi T} \mid \sum_{t=0}^{T-1} X(t)e^{-i\lambda t} \mid^2$ has, for each frequency $-\pi < \lambda < \pi$, $\lambda \neq 0$, an asymptotically $f(\lambda)\chi_2^2$ distribution, where $f(\lambda)$ is the (Fourier) spectrum of $X(t)$. Theorem 8 tells us that for each $j = 1, \cdots, M$ and each $k = 0, \cdots, (2^{(M-j)} - 1)$, $I_{j,k}^{(\psi)}$ has an asymptotically $\eta_{j,k}^{(\psi)}\chi_1^2$ variable. Hence, confidence intervals for $\eta_{j,k}^{(\psi)}$ can be constructed, in the same way they are for $f(\lambda)$.

6 The Scalegram

In (11) we have defined the wavelet periodogram with respect to ψ. Note that $I_{j,k}^{(\psi)}$ is defined for each (j,k), $j = 1, \cdots, M$ and $k = 0, \cdots, 2^{(M-j)} - 1$, $T = 2^M$.

We may define also the *scalegram at scale* j as being

$$S^{(\psi)}(j) = \sum_{k=0}^{2^{(M-j)}-1} [d_{j,k}^{(\psi)}]^2 = \sum_{k=0}^{2^{(M-j)}-1} I_{j,k}^{(\psi)}, \quad j = 1, \cdots, M. \tag{13}$$

We may think $S^{(\psi)}(j)$ as the energy of the wavelet coefficients at scale (or level) j. See Scargle (1993) and Arino and Vidakovic (1995).

Theorem 9: Under the assumption of Theorem 7, we have

$$E\{S^{(\psi)}(j)\} = \eta_{j,\cdot}^{(\psi)} + O(T^{-1}), \tag{14}$$

and

$$Cov\{S^{(\psi)}(j), S^{(\psi)}(j')\} = 2 \sum_{k=0}^{2^{(M-j)}-1} \sum_{k'=0}^{2^{(M-j')}-1} \{\eta_{(j,j'),(k,k')}^{(\psi)}\}^2 + O(1), \tag{15}$$

where

$$\eta_{j,\cdot}^{(\psi)} = \sum_{k=0}^{2^{(M-j)}-1} \eta_{j,k}^{(\psi)}. \tag{16}$$

Proof:

We have

$$E\{S^{(\psi)}(j)\} = \sum_k E\{I_{j,k}^{(\psi)}\} = \sum_k \eta_{j,k}^{(\psi)} + O(T^{-1}) = \eta_{j,\cdot}^{(\psi)} + O(T^{-1}),$$

and

$$\begin{aligned}
Cov\{S^{(\psi)}(j), S^{(\psi)}(j')\} &= E\{S^{(\psi)}(j)S^{(\psi)}(j')\} - E\{S^{(\psi)}(j)\}E\{S^{(\psi)}(j')\} \\
&= E\{\sum_k I_{j,k}^{(\psi)} \sum_{k'} I_{j',k'}^{(\psi)}\} - E\{\sum_k I_{j,k}^{(\psi)}\}E\{\sum_{k'} I_{j',k'}^{(\psi)}\} \\
&= \sum_k \sum_{k'} \{E(I_{j,k}^{(\psi)} I_{j',k'}^{(\psi)}) - E(I_{j,k}^{(\psi)})E(I_{j',k'}^{(\psi)})\} \\
&= \sum_k \sum_{k'} Cov(I_{j,k}^{(\psi)}, I_{j',k'}^{(\psi)}) \\
&= \sum_k \sum_{k'} 2\{\eta_{(j,j'),(k,k')}^{(\psi)}\}^2 + O(1).
\end{aligned}$$

$\eta_{j,\cdot}^{(\psi)}$ may be thought as the energy spectrum associated with scale j and $S^{(\psi)}(j)$ its estimator.

From Theorem 8, $I_{j,k}^{(\psi)}$ are distributed as a $\eta_{j,k}^{(\psi)}\chi_1^2$ and hence, using the properties of the chi-squared distribution, the distribution of $S^{(\psi)}(j)$ may be approximated by $a\chi_b^2$, where a and b are parameters to be determined.

Theorem 10: Under the assumptions of Theorem 7 and 8, $S^{(\psi)}(j)$ is asymptotically a $a\chi_b^2$ variable with

$$a \approx \frac{\{\sum_k \sum_{k'} \eta_{(j,j),(k,k')}^{(\psi)}\}^2}{\eta_{j,\cdot}^{(\psi)}}, \tag{17}$$

and

$$b \approx \frac{\{\eta_{j,\cdot}^{(\psi)}\}^2}{\sum_k \sum_{k'} \{\eta_{(j,j),(k,k')}^{(\psi)}\}^2}. \tag{18}$$

Proof:

The first two moments of the $a\chi_b^2$ variable are

$$E\{a\chi_b^2\} = ab,$$

and

$$Var\{a\chi_b^2\} = 2a^2b.$$

The distribution of $S^{(\psi)}(j)$ may be approximated by $a\chi_b^2$ because $I_{j,k}^{(\psi)}$ is asymptotically a chi-squared variable. Then the constants a and b may be computed from

$$b \approx \frac{2\{E(S^{(\psi)}(j))\}^2}{Var\{S^{(\psi)}(j)\}}, \tag{19}$$

and

$$a \approx \frac{E\{S^{(\psi)}(j)\}}{b}, \tag{20}$$

which gives the parameters of the approximating chi-squared distribution in terms of the first two moments of $S^{(\psi)}(j)$.

Now from (14) and (15) we obtain (17) and (18).

As an illustration, we consider an harmonic time series which is the sum of two perfect periodic series with periods 5 and 24,

$$X(t) = sin(\frac{\pi}{12}t) + 2sin(\frac{2\pi}{5}t),$$

for $t = 0, 1, \cdots, 2^{10} - 1$.

Part of the series $X(t)$ is presented in Figure 6.1 and the scalegram, as defined in (13), is shown in Figure 6.2. The presence of a peak in the scalegram of a time series at a low level of j means that a high frequency cycle (or small period) is present in the series. On the other hand, when there is a peak at a high level of j, it indicates that a low frequency cycle (or large period) is present in the series. Por example, in Figure 6.2, there is a peak at level 2, it corresponds to a period of 5 units and a peak at level 4 corresponds to a period of 24 units. The wavelet transform coefficients are shown in Figure 6.3 and the periodogram in Figure 6.4.

7 Wavelet shrinkage and smoothing

Shrinkage is basic in statistical work with wavelets. Using shrinkers to improve estimates has been around in statistics for many years, see Thompson (1968) and King (1972). Wavelet shrinkage acts as a smoothing operator. The idea behind thresholding is the removal of small wavelet coefficients considered to be noise. This kills small $d_{j,k}^{(\psi)}$ and keeps large $d_{j,k}^{(\psi)}$ that can be used to improve estimates.

The general procedure of wavelet shrinkage can be described into three steps:

1) take a wavelet transform of raw data (a time series) to get the empirical wavelet coefficients;

2) the empirical wavelet coefficients are shrunk by thresholding (all wavelet coefficients of magnitude less than the threshold are set to zero);

3) invert the wavelet transform, the processed empirical wavelet coefficients are returned to recover the smoothed, denoised signal.

The orthonormality and linearity of the wavelet transform tell us that the coefficients corresponding to the Gaussian white noise in the original signal remain the Gaussian white noise themselves. It does, however, get spread over all of the wavelet coefficients. This is why we shall shrink coefficients in addition to simple setting the small ones to zero.

There are many ways to threshold. Basically, the process of thresholding wavelet coefficients can be divided into two steps: the policy choice and the choice of a threshold.

The policy choice is the choice of the threshold function δ. Two standard choices are: **hard** and **soft** thresholding:

$$\delta_h(d_{j,k}^{(\psi)}, \lambda) = d_{j,k}^{(\psi)} I_{\{|d_{j,k}^{(\psi)}|>\lambda\}}, \qquad \delta_s(d_{j,k}^{(\psi)}, \lambda) = sgn(d_{j,k}^{(\psi)})(\mid d_{j,k}^{(\psi)} \mid -\lambda)_+,$$

respectively, where λ is the threshold.

The second step is the choice of a threshold. The choice of threshold, as well as the thresholding policy, is one of the most important issues in the wavelet shrinkage. Donoho and Johnstone (1992a,b,c, 1993) proposed a threshold λ, called universal threshold, $\lambda = \sigma\sqrt{2\log T}$, for recovering curves from noisy data, where the noise is assumed to have a Gaussian distribution and σ^2 is the noise level. Donoho, Johnstone, Kerkyacharian and Picard (1992) proposed a threshold $\lambda = \sqrt{j}$ for density estimation. Nason (1994) proposed a very interesting cross-validatory threshold selection procedure. The threshold is selected by minimizing a cross-validatory estimator of the integrated square error. Gao (1993b) proposed the use of thresholds $\lambda_{j,T}$ depending on sample size T, wavelet basis and resolution level j for estimating the log−spectrum of a stationary Gaussian time serie. The purpose of this thresholding level is to make the reconstructed log−spectrum as nearly noise free as possible. Donoho and Johnstone (1994) developed a technique, called SUREShrink, to select a threshold by minimizing the Stein's Unbiased Estimate of Risk (SURE) for recovering a function of unknown smoothness from noisy, sampled data.

In general, the wavelet shrinkage methods have a number of theoretical advantages, including near optimal mean squared error and near ideal spatial adaptation.

Other references in threshold selection and wavelet shrinkage applications are Donoho and Johnstone (1990), Donoho (1993) and Hall and Patil (1993a,b,c).

The estimators we consider in this paper are obtained by thresholding empirical coefficients:

$$d_{j,k}^{(\psi,\tau)} = \delta_h(d_{j,k}^{(\psi)}, \lambda). \tag{21}$$

Theorem 11: We suppose that $\lambda = \lambda_{j,k}$ and $\lambda_{j,k} = \sigma_{j,k}\sqrt{2\log T}$, where $\sigma_{j,k}^2 = Var\{d_{j,k}^{(\psi)}\}$. Then we have

$$E\{d_{j,k}^{(\psi,\tau)}\} = O((\log T)^{-1/2}),$$

and

$$Var\{d_{j,k}^{(\psi,\tau)}\} = O((\log T)^{-1/2}).$$

Proof:

Let

$$d_{j,k}^{(\psi,\tau)} = d_{j,k}^{(\psi)} I_{\{|d_{j,k}^{(\psi)}|>\lambda_{j,k}\}},$$

then,

$$E(d_{j,k}^{(\psi,\tau)}) = E(d_{j,k}^{(\psi)} I_{\{|d_{j,k}^{(\psi)}|>\lambda_{j,k}\}}),$$

and by Cauchy-Schwarz and Chebyshev inequalities, we have

$$\begin{aligned} \mid E(d_{j,k}^{(\psi,\tau)}) \mid &\leq [E((d_{j,k}^{(\psi)})^2)]^{1/2}[E(I_{\{|d_{j,k}^{(\psi)}|>\lambda_{j,k}\}})]^{1/2} \\ &\leq (P(\mid d_{j,k}^{(\psi)} \mid > \lambda_{j,k})^{1/2}(Var(d_{j,k}^{(\psi)}))^{1/2} \\ &\leq \frac{1}{\lambda_{j,k}}\sigma_{j,k}^2. \end{aligned}$$

Taking $\lambda_{j,k} = \sigma_{j,k}\sqrt{2logT}$, we obtain

$$E\{d_{j,k}^{(\psi,\tau)}\} = O((\log T)^{-1/2}).$$

Now,

$$\begin{aligned} \mid E[(d_{j,k}^{(\psi,\tau)})^2] \mid &\leq [E((d_{j,k}^{(\psi)})^4)]^{1/2}[E(I_{\{|d_{j,k}^{(\psi)}|>\lambda_{j,k}\}})]^{1/2} \\ &\leq (P(\mid d_{j,k}^{(\psi)} \mid > \lambda_{j,k})^{1/2}\{E[(d_{j,k}^{(\psi)})^4]\}^{1/2} \\ &\leq \frac{1}{\lambda_{j,k}}\{E[(d_{j,k}^{(\psi)})^2]Var(I_{j,k}^{(\psi)}) + (E[(d_{j,k}^{(\psi)})^2])^3\}^{1/2} \\ &\leq \{[\eta_{j,k}^{(\psi)} + O(T^{-1})][2\{\eta_{j,k}^{(\psi)}\}^2 + O(1)] + [\eta_{j,k}^{(\psi)} + O(T^{-1})]^3\}^{1/2}O((\log T)^{-1/2}). \end{aligned}$$

Hence,

$$E\{[d_{j,k}^{(\psi,\tau)}]^2\} = O((\log T)^{-1/2}),$$

from which it follows that

$$Var\{d_{j,k}^{(\psi,\tau)}\} = O((\log T)^{-1/2}).$$

8 Application

In this section the wavelet transform with respect to Haar wavelet is used to investigate the spectral components of electroencephalographic (EEG) sleep state patterns of infants whose mothers abstained from alcohol during pregnancy and infants whose mothers used moderate amounts of alcohol throughout pregnancy. Data were collected for a study of the effects of moderate maternal alcohol consumption on neonatal EEG sleep patterns. A detailed description of the study design, the methods used for measuring alcohol use, and the scoring of neonatal EEG sleep records can be found in Stoffer et al (1988).

Figure 8.1 shows the per minute EEG sleep record (for 120 minutes) of a normal infant approximately 24-36 hours after birth. The corresponding EEG sleep record of a infant whose mother used alcohol throughout pregnancy is shown in Figure 8.2. Here, sleep state is categorized (per minute) into one of six possible states using the labels: state 1: quiet sleep - trace alternant; state 2: quiet sleep - high voltage; state 3: indeterminate sleep; state 4: active sleep - low voltage; state 5: active sleep - mixed; state 6: awake.

We denote by UNEXPOSED the data shown in Figure 8.1 and EXPOSED the data shown in Figure 8.2. The wavelet transforms of UNEXPOSED and EXPOSED are presented in Figure 8.3 and 8.4 respectively. Consequently, Figure 8.5 and 8.6 show their wavelet periodograms. Here, in these figures, Resolution Level corresponds to level of $(M-j)$; in this case, $M=7$. These periodograms tell us how the energy in a time series is decomposed on a scale by scale basis at different times. Figure 8.5 and 8.6 tell us, for example, that the periodogram of scale 1 contributes more to energy in our time series than any of the other scales. On the other hand, in comparing the two periodograms, note that there are larger values of the periodogram of scale 1 in Figure 8.6 than in Figure 8.5.

Now let us consider what happens to reconstructed sleep state when we apply thresholding to the wavelet coefficients illustrated in Figures 8.3 and 8.4, in order to remove the noise presented in our time series. Here, we use the universal policy. Their thresholding wavelet periodograms are shown in Figures 8.7 and 8.8, and Figure 8.9 shows the reconstructed sleep state of

UNEXPOSED and EXPOSED.

Their scalegrams are presented in Figure 8.10. We note that there is a peak at a high level ($j = 5$) in these two scalegrams, indicating that a large period is present in both series. The Walsh-Fourier analysis of this data is given in Stoffer et al (1988).

References

Arino, M.A. and Vidakovic, B. (1995). On wavelet scalograms and their applications in economic time series. Preprint.

Blackman, R.B. and Tukey, J.W. (1958). The measurement of power spectra. Dover, NY.

Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory.* 2nd. Edition. Holden-Day, San Francisco.

Brillinger, D.R. (1994a). Some river wavelets. *Environmetrics.* Preprint.

Brillinger, D.R. (1994b). Some asymptotics of wavelet fits in the stationary error case. TR 415, Dept. Statistics, University of California, Berkeley.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods.* Second Edition. Springer-Verlag.

Chui, C.K. (1992). *An Introduction to Wavelets.* San Diego, Academic Press.

Cooley, J.W. and Tukey, J.W. (1965). An algorithm for the machine calculation of complex Fourier series. *Math. Comp., 19*, 297-301.

Daubechies, I. (1992) *Ten Lectures on Wavelets. CBMS Lecture Notes Series.* SIAM, Philadelphia.

Donoho, D.L. and Johnstone, I.M. (1990). Wavelets and optimal nonlinear function estimates. TR 281, Statistics Department, University of California, Berkeley.

Donoho, D.L. and Johnstone, I.M. (1992a). Minimax estimation via wavelet shrinkage. TR 402, Dept. Statistics, Stanford University.

Donoho, D.L. and Johnstone, I.M. (1992b). Adapting to unknown smoothness via wavelet shrinkage. TR, Dept. Statistics, Stanford University.

Donoho, D.L. and Johnstone, I.M. (1992c). Ideal spatial adaptation by wavelet shrinkage. TR 400, Dept. Statistics, Stanford University.

Donoho, D.L., Johnstone, I.M., Kerkyacharian, G. and Picard, D. (1992). Wavelet shrinkage: asymptopia ? TR 419, Dept. Statistics, Stanford University.

Donoho, D.L. (1993). Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data. Proceedings of Symposia in Applied Mathematics, Vol. 47, 173-205, AMS.

Farge, M. (1992). Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech., 24*, 395-457.

Gao, H.-Y. (1993a). Wavelet estimation of spectral densities in time series analysis. Ph.D. thesis, Univ.of California, Berkeley.

Gao, H.-Y. (1993b). Choice of thresholds for wavelet estimation of the log spectrum. TR 438, Dept. Statistics, Stanford University.

Grossmann, A. and Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal., 15*, 723-736.

Hall, P. and Patil, P. (1993a). On wavelet methods for estimating smooth functions. TR 12-93, Center for Mathematics and its Applications, Australian National University.

Hall, P. and Patil, P. (1993b). Formulas for mean integrated squared error of nonlinear wavelet based density estimation. TR 15-93, Center for Mathematics and its Applications, Australian National University.

Hall, P. and Patil, P. (1993c). On the choice of smoothing parameter, threshold and truncation in nonparametric regression by nonlinear wavelet methods. TR 21-93, Center for Mathematics and its Applications, Australian National University.

Hannan, E.J. (1970). *Multiple Time Series.* Weley, New York.

Hannan, E.J. (1973). Central limit theorems for time series regression. *Zeitshrift für Wahrscheinlichkeits Theorie and Gebiet, 26*, 157-170.

King, N. (1972). An alternative for the linear regression equation when the predictor variable is uncontrolled and the sample size is small. *Journal of the American Statistical Association, 67*, 217-219.

Mallat, S. (1989). A theory for multiresolution signal decompositions: the wavelet representation. *IEEE Trans. on Pattern Anal. Machine Intell., 11*, 674-693.

Meyer, Y. (1990). *Ondelettes et Opérateurs.* Tome I. Actualités Mathématiques. Paris: Hermann.

Meyer, Y. (1993). *Wavelets: Algorithms and Applications.* SIAM, Philadelphia.

Morettin, P.A. (1981). Walsh spectral analysis. *SIAM Review, 23*, 279-291.

Moulin, P. (1993). Wavelet thresholding techniques for power spectrum estimation. Preprint.

Nason, G.P. and Silverman, B.W. (1994). The discrete wavelet transform in S. *Journal of Computational and Graphical Statistics, 3*, 163-191.

Nason, G.P. (1994). Wavelet regression by cross-validation. TR 447, Dept. Statistics, Stanford University.

Percival, D.B. and Walden, A.T. (1993). *Spectral Analysis for Physical Application: Multitaper and Conventional Umvariate Techniques.* Cambridge: Cambridge University Press.

Scargle, J.D. (1993). Wavelet methods in astronomical time series analysis. Preprint.

Stoffer, D.S., Scher, M.S. Richardson, G.A., Day, N.L. and Coble, P.A. (1988). A Walsh Fourier analysis of the effects of moderate maternal alcohol consumption on neonatal sleep state cycling. *Journal of the American Statistical Association, 83*, 954-963.

Stoffer, D.S. (1991). Walsh-Fourier analysis and its statistical applications. *Journal of the American Statistical Association, 86*, 461-485.

Strang, G. (1989). Wavelets and dilation equations. *SIAM Review, 31*, 614-627.

Thompson, J.R. (1968). Some shrinkage techniques for estimating the mean. *Journal of the American Statistical Association, 63*, 113-122.

Captions:

Figure 2.1: Some examples of function ψ.

Figure 3.1: The autocovariance function $(\gamma(u))$ and the wavelet spectrum $(\eta_{j,k}^{(H)})$ with respect to the Haar wavelet of a white noise sequence, with mean zero and variance $\sigma^2 = 1$.

Figure 3.2: The autocovariance function $(\gamma(u))$ and the wavelet spectrum $(\eta_{j,k}^{(H)})$ with respect to the Haar wavelet of a MA(1) process, with $\beta = 0.7$ and variance $\sigma^2 = 1$.

Figure 3.3: The autocovariance function $(\gamma(u))$ and the wavelet spectrum $(\eta_{j,k}^{(H)})$ with respect to the Haar wavelet of a MA(2) process, with $\beta_1 = 0.5$, $\beta_2 = -0.3$ and variance $\sigma^2 = 1$.

Figure 3.4: The autocovariance function $(\gamma(u))$ and the wavelet spectrum $(\eta_{j,k}^{(H)})$ with respect to the Haar wavelet of a AR(1) process, with $\alpha = 0.8$ and variance $\sigma^2 = 1$.

Figure 6.1: An harmonic time series $X(t) = sin(\frac{\pi}{12}t) + 2sin(\frac{2\pi}{5}t)$.

Figure 6.2: The scalegram of $X(t)$ presented in Figure 6.1 with $T = 2^{10}$.