Two Reports

a) Some Examples of the Use of the Fourier Transform in the Analysis of Scientific Data

b) A Note on River Wavelets

By

David R. Brillinger

Statistics Department
University of California
Berkeley, CA

Technical Report No. 407
December 1993

*Research supported by National Science Foundation Grants DMS-9208683 and DMS-9300002

Department of Statistics
University of California
Berkeley, California 94720
Some Examples of the Use of the Fourier Transform in the Analysis of Scientific Data

David R. Brillinger

Abstract. The Fourier transform is of interest to mathematicians, to scientists and to statisticians both generally and particularly. It is important to the first group because generalizations seem inexhaustible and there are continually surprises. It is of interest to statisticians because it proves inordinately useful in the analysis of data. This paper presents some examples of the use of the discrete Fourier transform in the analysis of data from biophysics, seismology, chemistry and biology.

Key words and phrases: Fourier transform, electron microscopy, microtubule movement, nuclear magnetic resonance spectroscopy, seismic surface waves, shrinkage estimator, spectrum analysis, wavelet

"L’ étude approfondie de la nature est la source la plus féconde des découvertes mathématiques."

J. B. Fourier (1822), Théorie Analytique de la Chaleur

1. Introduction

The Fourier transform has a long and glorious history. Statisticians too have made some noteworthy contributions to that history, e.g. Slutsky (1934), Cramér (1942), Good (1958), Yaglom (1961), Tukey (1963), Hannan (1966), Diaconis (1988, 1989). As a concept and as a tool, the

David Brillinger is Professor of Statistics, University of California, Berkeley, CA 94720.
Fourier transform is pervasive in computing, mathematics, probability and statistics.

Classic works on the topic include: Wiener (1933), Bochner (1959,1960) and Zygmund (1968). Works on generalizations to abstract harmonic analysis include: Loomis (1953), Rudin (1962), Hewitt and Ross (1963), Katznelson (1976). More recent books are Terras (1988) and Körner (1989), the former particularly addressing the nonabelian case, the latter presenting a variety of historical examples.

In naive form, Fourier analysis is based on complex exponentials (characters),

$$\exp\{i\lambda t\}$$  \hspace{1cm} (1.1)

with frequency $\lambda$ in $G = \{-\pi < \lambda \leq \pi\}$ and time $t$ in $Z = \{0, \pm1, \pm2, \cdots\}$. $G$ and $Z$ are abelian groups under addition, dual to each other. The exponentials (1.1) may be used to approximate a $2\pi$-periodic function $g(\lambda)$, specifically

$$g(\lambda) = \sum_{t=-T}^{T} c_t e^{i\lambda t}$$  \hspace{1cm} (1.2)

with $T$ an integer and the coefficients given by

$$c_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda)e^{-it\lambda}d\lambda$$  \hspace{1cm} (1.3)

$t$ in $Z$. The coefficients (1.3) provide the best $L_2$ approximation of the form (1.2) to $g(\lambda)$. Other norms lead to other approximations. A broad class of approximations takes the form

$$\sum_{t} w_{t,T} c_t e^{i\lambda t}$$  \hspace{1cm} (1.4)

the $w_{t,T}$ being the so-called convergence factors of which Féjer’s, $w_{t,T} = 1 - 1t \mid /T$, is perhaps the most renowned. It can be insightful to
write (1.4) as
\[ \int W^T(\mu)g(\lambda-\mu)\,d\mu \] (1.5)
where the window (kernel) is given by
\[ W^T(\lambda) = \frac{1}{2\pi} \sum_{t} w_{t,T} e^{i\lambda t} \] (1.6)
for then one sees that (1.4) is a weighted average of the desired \( g(.) \). The expression (1.5) often allows direct study of the degree of approximation. Timan (1963), Butzer and Nessel (1971) are books specifically concerned with approximations based on Fourier expressions.

Some of the continuing elements of Fourier analysis are: translation, convolution, decomposition, approximation, orthogonality, duality. In particular the operations of translation and convolution
\[ g(\lambda-\mu), \quad \int W(\mu)g(\lambda-\mu)\,d\mu \]
only occur in physical situations. They behave simply under Fourier transform, as does the operation of scaling.

The Fourier transform unifies diverse procedures and problems. It is highly flexible. It leads to entities with physical interpretations. One can point to success stories including the following. Michaelson (1891a,b) measured visibility curves, essentially the modulus of a Fourier transform, and after an inversion thereby inferred that the red hydrogen line was a doublet. This inference of splitting ultimately led to important developments in quantum mechanics. Tidal components caused by the sun, moon and planets have been isolated by Fourier analysis, see Cartwright (1982), Bath (1974), Bracewell (1989). Katz and Miledi (1971) inferred the mechanism of acetylcholine release via a Fourier analysis. Bolt et al. (1982) saw a fault rupturing in an earthquake by an \( f-k \) spectral analysis. Finally it may be noted that R.R. Ernst received the 1991 Nobel Prize in
Chemistry for developing the technique of Fourier transform spectroscopy, see Amato (1991).

There are special computational, mathematical and statistical properties and surprises associated with the Fourier transform. These include: the central limit theorems for the stationary case with approximate independence at particular frequencies, the existence of fast Fourier transforms, Good (1958), Tukey (1963), Cooley and Tukey (1965), the need for convergence factors, the ideas of aliasing.

There are extensions to nonlinear operators and spectra corresponding with the physical occurrence of higher harmonics. These lead to extensions to non-Gaussian processes, and associated higher-order analyses.

The focus of the present work will be on the use of a Fourier transform in addressing four particular scientific problems. The examples highlight: approximation, shrinkage estimation, the method of stationary phase, central limit theorems and uncertainty estimation. The final example involves both wavelet and Fourier analysis. There is a section on open problems and an Appendix on uncertainty estimation in the wavelet case.

2. Physical Bases

Cycles, periods and resonances have long been noted in scientific fields such as astronomy, vibrations, tides, sound, light and crystallography. In technology they occur as phenomena in work with telephones, radio, TV and lasers. Natural operations that are to a good approximation linear and time invariant, (convolutions), occur commonly and these are the eigenoperations of Fourier analysis.

Fourier analysis is sometimes tied specifically to the physics of a problem - one observes the Fourier transform or its modulus directly. For
example Bazin et al. (1986) physically demonstrate the operations/concepts of translation, linearity, similarity, convolution and Parseval’s theorem for the Fourier transform via diffraction experiments with laser light. The Fourier transform here is formed via a lens, for from physical optics, one knows that the diffraction pattern that is formed in an objective lens focal plane represents the Fourier transform of the transmitted wave, see Goodman (1968), Glaeser (1985).

An important example arises in radio astronomy. Suppose one has an array of receivers. If $Y(x,y,t)$ denotes the radio field measurement made on the telescope located at position $(x,y)$ at time $t$, then

$$E \{ Y(x+u,y+v,t) \overline{Y(x,y,t)} \} = \int \int f(\alpha,\beta) e^{i(u\alpha+v\beta)} d\alpha d\beta \quad (2.1)$$

where $(\alpha,\beta)$ are the coordinates of the source of interest in the sky and $f(\alpha,\beta)$ is its brightness distribution, as a function of $(\alpha,\beta)$. In other words, the Fourier transform is just what one wants. The result (2.1) is known as the vanCittert-Zernike Theorem, see Born and Wolf (1964). One has in mind a small incoherent source, in the far-field, producing a plane travelling wave.

Nowadays in science there is much concern with nonlinear operations and phenomena. Impressively the classic trigonometric identity

$$[\cos \lambda t]^2 = \frac{1}{2} \cos 2\lambda t + \frac{1}{2} \quad (2.2)$$

is "demonstrated" in Yariv (1975) via a color plate showing red laser light becoming blue on passing through a crystal. The crystal involved, squares the signal as in (2.2). A wavelength of $6940\AA$ (red) becomes one of $3970\AA$ (blue). Bloembergen (1982), Moloney and Newell (1989) discuss such nonlinear aspects of light. The appearance of harmonics leads to a consideration of higher-order spectra in particular.
The Fourier transform is continually employed in the solution of equations of motion associated with physical phenomena and mathematicians have focussed on consequent cycles and harmonics. For example, Hirsch (1984) has remarked that "Dynamicists have always been fascinated (not to say obsessed) by periodicity." In that connection Ruelle (1989) makes effective use of the Fourier transform in the study of dynamic systems, specifically addressing aspects of chaos, periods and scaling.

A discussion of a variety of physical examples may be found in Lanczos (1966), Bath (1974), Bracewell (1989).

3. Stochastics and Statistics

3.1 Some roles

Statistics is about using numbers wisely. The Fourier transform is continually employed in this endeavor. In particular it is of use in: developing estimates and efficient estimates (Whittle (1952), Dzhaparidze (1986), Feuerverger (1990)), assessing goodness of fit (Feigin and Heathcote (1976)), deconvolving random measurements (Fan (1992)) and combining data (Brillinger et al. (1989, 1990)).

In the case of $Y(t), t = \pm 1, \pm 2, \cdots$ a second-order stationary process, following Cramér (1942), one has the Fourier representation

$$Y(t) = \int_{-\pi}^{\pi} e^{i\lambda t} dZ(\lambda)$$

with $Z(.)$ a random function having uncorrelated increments. The Cramér representation has the advantage of taking one directly to the frequency domain and thereby often making operations clearer, such as filtering. The series $Y(.)$ may be vector-valued, then so too will be $Z(.)$. Second- and higher-order spectra may be defined directly via $Z(.)$, eg. the power
spectrum $f(\lambda)$ at frequency $\lambda$ is given by

$$\text{cov}\{dZ(\lambda),dZ(\mu)\} = \delta(\lambda-\mu)f(\lambda)d\lambda d\mu$$  \hspace{1cm} (3.1)

$-\pi < \lambda \leq \pi$. If $Y(.)$ is Gaussian, so too is $Z(.)$.

Blackman and Tukey (1959), Bath (1974), Brillinger (1975) and Bloomfield (1976) are books focussing on the empirical Fourier analysis of time series.

3.2 Central limit theorems

In classic forms the central limit theorem is concerned with the distributions of sums of independent random variables

$$S_T = Y_0 + Y_1 + \cdots + Y_{T-1}$$

being approximately normal with variance $T\sigma^2$ for large $T$. It is usual to assume that the $Y$'s are identically distributed.

At some point engineers began promulgating a folk theorem to the effect that narrow-band noise is approximately Gaussian, see Leonov and Shiryaev (1960), Picinbono (1960), Rosenblatt (1961). One fashion to formulate this remark is as a statement that

$$S_T(\lambda) = Y_0 + e^{-i\lambda}Y_1 + \cdots + e^{-i\lambda(T-1)}Y_{T-1}$$  \hspace{1cm} (3.2)

$-\pi < \lambda \leq \pi$, is approximately (complex) normal. Under stationarity and mixing assumptions for the series $Y_t$, the variance of (3.2) is approximately

$$2\pi Tf(\lambda)$$  \hspace{1cm} (3.3)

with $f(\lambda)$ the power spectrum of (3.1). Surprisingly, the values of $S_T(\lambda)$ at distinct frequencies of the form $\lambda = 2\pi j/T$, are approximately independent. Problems involving stationary mixing processes may thus be converted into others involving (approximately) independent normal random variables. Empirical Fourier transforms such as (3.2) have many uses and
several are indicated in this paper.

Early work on the asymptotic properties of finite Fourier transforms includes that of Slutsky (1934), Leonov and Shiryaev (1960), Rosenblatt (1961), Good (1963), Hannan (1969), Brillinger (1970), Hannan and Thomson (1971), Hannan (1972).

There has been some consideration of the cases of long range dependence and stable distributions. References include: Rosenblatt (1981), Freedman and Lane (1981), Fox and Taqqu (1986), Yajima (1989), Shao and Nikias (1993). The case of random generalized functions, which includes for example point processes and random measures, is considered in Brillinger (1982).

3.3 Shrinking

Among surprises, in working with Fourier transforms, is the importance of convergence factors. These are the $w_{i,T}$ of (1.4). In (1.4) they shrink the coefficients of the $\exp(i\lambda t)$ towards 0 as $t$ increases. Such multipliers are also important in the stochastic case, see: Tukey (1959), Brillinger (1975), Bloomfield (1976), Dahlhaus (1984,1989).

A related concept is shrinking. In a regression context Tukey (1979) distinguishes three types of shrinking. Crudely: "first shrinkage" corresponds to pretesting and selection of regressor variables, "second shrinkage" corresponds to a type of Wiener filtering and "third shrinkage" corresponds to borrowing strength from other coefficients to improve the collection. Here the multipliers are not meant for attenuating high frequencies, rather they are meant for attenuating uncertain terms. A common characteristic is that the estimates become biased; however biased estimates have long been dominant in time series analysis.
Second shrinkage plays an important role in two of the examples that follow. A particular second shrinkage estimate, introduced in Tukey (1979), may be motivated as follows. Consider a classic regression model

\[ y = \beta x + \varepsilon \]

with \( b \) an estimate of \( \beta \) and \( s \) an estimate of its standard error. Consider seeking a multiplier \( m \) such that \( mbx \) is an improved estimate of \( \beta x \). The mean-squared error of the new estimate is

\[ x^2 E \{ (\beta - mb)^2 \} \]

which may be estimated by

\[ x^2 \{ (1-m)^2[b^2-s^2] + m^2s^2 \} \]

This is minimized by the choice \( m = 1 - s^2/b^2 \) and one would probably prefer to take \( m \) to be the positive part

\[ (1 - s^2/b^2)_+ \]

This multiplier has the reasonable property of being 0 for \( b \) less than its standard error.


4. Examples

In this section three examples are presented for which empirical Fourier analysis is basic.

4.1 Electron microscopy
Electron microscopy is a tool for studying the placements of atoms within molecules. It is mainly carried out with crystalline (periodic) material. One problem is to obtain improved images and that is the concern of this example. Glaeser (1985), Henderson et al. (1986), Hovmöller (1990) are references describing electron microscopy.

In the planar case, the principal theoretical concept is the projected (Coulomb) density distribution

$$V(x,y) = \sum_{h,k} F_{h,k} e^{2\pi i (hx + ky)/\Delta}$$  \hspace{1cm} (4.1)

$h,k = 0, \pm 1, \pm 2, \ldots$ with $(x,y)$ planar coordinates and with $\Delta$ the period of the crystal. The function $V(.)$ is real-valued and has various symmetries. The $h, k$ in (4.1) are referred to as the Miller indices, while the $F_{h,k}$ are referred to as structure factors. One wishes to estimate $V(x,y)$ over $0 \leq x, y < \Delta$.

The datum is an image, $Y(x,y)$, with $0 \leq x < X, 0 \leq y < Y$. The image may be written as

$$Y(x,y) = V(x,y) + \text{noise}$$

The empirical Fourier transform is

$$\hat{F}_{h,k} = \iint_{00} \frac{Y(x,y)e^{-2\pi i (hx + ky)/\Delta}}{dx dy}$$  \hspace{1cm} (4.2)

which may be written

$$\iint_{00} \sum_{m,n} Y(x + m \Delta, y + n \Delta)e^{-2\pi i (hx + ky)/\Delta} dx dy$$  \hspace{1cm} (4.3)

The expression (4.3) displays averaging taking place and a consequent sensible estimate. The synthesis corresponding to the analysis (4.2) is

$$\sum_{h,k} \hat{F}_{h,k} e^{2\pi i (hx + ky)/\Delta}$$  \hspace{1cm} (4.4)

$0 \leq x < \Delta, 0 \leq y < \Delta$. 
There are actually two methods to collect pertinent biophysical data—imaging as above and diffraction. In the example to be presented, diffraction experiments provide better estimates of the amplitudes $|F_{h,k}|$, see Henderson et al. (1986). These values, $R_{h,k}$, may be combined with naive phases, $\hat{\phi}_{h,k} = \text{arg} \hat{F}_{h,k}$, obtained in imaging experiments. A specific estimate is then

$$\sum_{h,k} R_{h,k} e^{i\hat{\phi}_{h,k}} e^{2\pi i (hx + ky) / \Delta}$$  \hspace{1cm} (4.5)

However there has been a concern to form an improved image. In this connection Blow and Crick (1959), Hayward and Stroud (1981) introduced "multipliers", $w(.)$, into expressions like (4.4), (4.5), forming

$$\hat{V}(x,y) = \sum_{h,k} w(|\hat{F}_{h,k}|/|\hat{\phi}_{h,k}|)R_{h,k} e^{i\hat{\phi}_{h,k}} e^{2\pi i (hx + ky) / \Delta}$$  \hspace{1cm} (4.6)

where the $\hat{\phi}_{h,k}$ are estimates of the standard errors of the $\hat{F}_{h,k}$. This is a second shrinkage estimate. Consideration of the mean-squared error, as at (3.4), leads to the multiplier

$$w(|\hat{F}|/\hat{\phi}) = \left[ 1 - \frac{\hat{\phi}^2}{|\hat{F}|^2} \right]_+$$  \hspace{1cm} (4.7)

which by analogy with Wiener filtering will be called the Wiener multiplier. By Bayesian arguments Blow and Crick (1959) and Hayward and Stroud (1981) were lead to the multiplier

$$w(\gamma) = \frac{\sqrt{\pi}}{2} \gamma \left[ I_0(\frac{\gamma^2}{2}) + I_1(\frac{\gamma^2}{2}) \right] e^{-\gamma^2 / 2}$$  \hspace{1cm} (4.8)

with $\gamma = |\hat{F}|/\hat{\phi}$, and $I_0, I_1$ modified Bessel functions, see Brillinger et al. (1989, 1990). It and (4.7) are graphed in Figure 1.

Estimates employing (4.7) and (4.8) are illustrated for images of the protein bacteriorhodopsin. This substance occurs naturally as a two-dimensional crystalline array within the cell membrane of
*Halobacterium halobium*. Together with accompanying lipid molecules, it is known as 'purple membrane'. This crystal is based on a hexagonal lattice. The estimates are presented in Figure 2 as contour plots. In the figure only the positive contours are shown. (Negative density features signify the absence of atoms and thus have no direct usefulness when the density map is interpreted.) The first panel of the figure shows the elementary estimate (4.5). The top right shows (4.6) with $w(.)$ of (4.8). The third, lower left, shows (4.6) with $w(.)$ of (4.7). The final panel provides an estimate based on combining 42 individual images, see Brillinger et al. (1989, 1990). This last image may be viewed as what the earlier images ascribe to be.

Through the inclusion of the multipliers, the peaks have become more substantial and better separated. Also, the estimates show better approximations to three-fold symmetry. Details of the data collection and further details of the analysis may be found in Brillinger et al. (1989,1990).

There are extensions to the 3D case, see Downing et al. (1990), Henderson et al. (1990), Wenk et al. (1992).

The Fourier transform is useful in this example firstly because of the lattice periodicities and secondly for the central limit theorem result suggesting specific estimates of the $\sigma_{hk}$ of (4.6).

4.2 Seismic surface waves

Various sound waves are transmitted through the Earth following a seismic disturbance, in particular surface (or Rayleigh) waves. These are vibrations whose energy is trapped and propagated just under the surface. The waves have sinusoidal form and are prominent in the later part of the seismogram, see the low velocity part of the vertical trace of Figure 3.
These waves have the interesting aspect of having been discovered mathematically. For basic details see Aki and Richards (1980) and Bullen and Bolt (1985).

Consider modelling the part of the seismogram where the Rayleigh waves occur. Let \( Y(x,t) \) denote the vibrations recorded at distance \( x \) from the earthquake source, as a function of time \( t \). With a layered crust model the theoretical seismogram is a solution of a system of differential equations with associated boundary conditions and may be represented as

\[
\int e^{-i(\lambda t - k(\lambda)x)} R(\lambda) d\lambda
\]  

(4.9)

Here, when \( x = 0 \),

\[
\int e^{-i\lambda t} R(\lambda) d\lambda
\]

represents the vibrations at the earthquake source. The solution (4.9) comes from substituting a particular solution \( \exp\{-i(\lambda t - kx)\} \) into the differential equations and matching boundary conditions. see Aki and Richards (1980). One writes \( k(\lambda) = \lambda/c(\lambda) \) with \( c(\lambda) \) the (phase) velocity with which the wave of frequency \( \lambda \) travels. The functions \( k(\cdot) \) and \( c(\cdot) \) depend on the transmission medium.

In the case that \( x \) is large one can use the saddle-point-method / method-of-stationary phase, see Cox and Barndorff-Nielsen (1989), on (4.9) to see the occurrence of the referred to sinusoidal form of the waves. Specifically for large \( x \), (4.9) is approximately

\[
R(\lambda_t) \exp\{-i(\lambda_t t - k(\lambda_t)x)\}
\]

(4.10)

with \( \lambda_t \) the solution of

\[
\frac{d}{d\lambda} [\lambda t - k(\lambda)x ] = 0
\]

that is \( k'(\lambda_t) = t/x = 1/U(\lambda_t) \). Here \( U(\lambda) \) is the group velocity, the velocity with which the energy travels, at frequency \( \lambda \).
Given an earth model, θ, that is a collection of layer depth, velocity and density parameters, one can compute the group velocity $U(\lambda|\theta)$, see Bolt and Butcher (1960), Aki and Richards (1980). For frequency $\lambda$ and parameter $\theta$ there may be several possible dispersion curves $U_n(\lambda|\theta)$, $n = 0, 1, 2, \ldots$ called modes. The phenomenon of waves with different frequencies travelling with different velocities is called dispersion. It is apparent in Figure 3 below. The concern of the example of this section is to estimate $\theta$.

The event studied originated in Siberia, 20 April 1989, and the trace was recorded at Uppsala, Sweden. Figure 3 provides a grey scale display of energy as a function of velocity and frequency. It is computed as

$$\left| \sum_{s=-S}^{S} h(s/S)Y(t-s)e^{-is\lambda} \right|^2$$

with $t = x_0/v$, $v$ velocity and $h(.)$ a convergence factor. One sees waves of about .07 cycles/second arriving first. The figure also shows the dispersion curves $U_n(\lambda|\hat{\theta})$ for one fitted earth model.

The velocity-frequency curves, $U_n(\lambda|\theta)$, may be inverted to frequency-time curves $\lambda = \lambda(t|\theta)$. To estimate $\theta$ one can then consider choosing $\theta, \alpha$ to minimize

$$\sum_{t} \left| Y(t) - \int e^{-i(\lambda t - k(\lambda|\theta)x_0)} R(\lambda|\alpha) d\lambda \right|^2$$

where $\alpha$ is some parametrization of the source motion. One approach is to approximating the integral (4.9), is to take $R(.)$ piecewise constant, linear in $\alpha$. Figure 4 provides the results of such an analysis. Graphed are the series, the fit and the residuals. In the research reported the standard errors were computed as in Richards (1961), focussing on the nonlinear parameters $\theta$ and acting as if the noise series was white. Details of this
and some other estimation procedures are presented in Bolt and Brillinger (1993), Brillinger (1993).

An improved procedure is needed, for the residual series of Figure 4 suggests the presence of signal-generated noise.

Even though this particular situation is clearly nonstationary, Fourier analysis has been basic to addressing it. This is a consequence of the presence of dispersion. It is also of some interest since one has a Fourier transform of 2 variables whose support lies on several curves, see Figure 3. This type of plot allows inference of the presence of higher modes and assessment of the fit as well.

4.3 NMR spectroscopy

Nuclear magnetic resonance is a quantum mechanical phenomenon employed to study the structure of various molecules. In an experiment, one creates a fluctuating magnetic field, \( X(t) \), encompassing a substance and then observes an induced voltage, \( Y(t) \). Hennel and Klinowski (1993) is one general reference.

If \( S(t) \) is a vector describing the state of the system at time \( t \), then the fluctuations are described by the Bloch equations

\[
\frac{dS(t)}{dt} = a + AS(t) + BS(t)X(t) \tag{4.12}
\]

and the measurements by

\[
Y(t) = e^cS(t) + \text{noise} \tag{4.13}
\]

with \( c \) depending on the geometry of the experiment. The principal parameters are frequencies of oscillation and decay rates. The parameters of interest sit in the matrices \( A \) and \( B \), see Brillinger and Kaiser (1992). The entries of \( A \) and \( B \) have physical interpretations, eg. the diagonal entries of \( A \) represent certain occupancy probabilities.
The equations (4.12) are interesting for being bilinear. They can be solved, symbolically, by successive substitutions, obtaining

\[ S(t) = C + \int_{t-s}^t e^{A(t-s)}CX(s)ds + \int_s^{t-s} e^{A(s-r)}Be^{A(s-r)}CX(r)X(s)drds + \cdots \]

with \( C = -A^{-1}a \). If \( A \) is written \( Ue^{\Lambda}U^{-1} \) with \( \Lambda \) diagonal, then the pulse response, \( S(t) \), is seen to be a sum of complex exponentials and various of their powers and products. The real parts of the entries of \( \Lambda \) will lead to the decay of these components, the imaginary represent resonance frequencies.

The problem is to estimate the parameters of (4.12) and thereby, for example, to characterize the substance. Some of the parameters may be estimated by cross-spectral analysis and others by a likelihood analysis.

Brillinger and Kaiser (1992) present results from a study of 2,3-dibromothiophene. The matrices \( A \) and \( B \) are 4 by 4 with complex-valued entries. The parameters include a coupling constant, \( J \) and frequencies \( \omega_A \) and \( \omega_B \). In the experiment the input employed was a sequence of pulses

\[ X(t) = \sum_j M_j \delta(t - j \Delta) \]

with \( \Delta = 1/150 \) sec., \( t \) in sec. and \( M_j \) the m-sequence given by \( M_j = M_{j-1}M_{j-4}M_{j-8}M_{j-12} \) starting at \( M_j = -1 \) for \( j = 1, \ldots, 12 \).

Figure 5 presents corresponding stretches of input and output together with the results of a cross-spectral analysis. Specifically the first-order transfer function estimate

\[ \hat{A}(\lambda) = \frac{\text{smooth}\{ \sum_{\mu=\lambda}^{\mu} t e^{-i\mu t} \} \{ \sum_{\mu=\lambda}^{\mu} X(t) e^{-i\mu t} \} \}}{\text{smooth}\{ \sum_{\mu=\lambda}^{\mu} t X(t) e^{-i\mu t} |^2 \} } \]
was computed. Theoretically its peaks are located at the frequencies

$$(\omega_A + \omega_B)/2 \pm J \pm \sqrt{J^2 + (\omega_A - \omega_B)^2}/2$$

and the widths of the peaks relate to a damping constant $T_2$.

In a more detailed analysis the parameters of the model, including initial state values, were estimated by least squares seeking

$$\min_{\theta} \sum_t |Y(t) - c^sS(t|\theta)|^2$$

$\theta$ referring to the unknown parameters. The state vector, $S(t|\theta)$ was evaluated recursively. Figure 6 shows the amplitude of the Fourier transform of the data and of the corresponding fit. There is an intriguing small peak just above 60 cycles/second which recurs when the data is broken down into contiguous segments. Details may be found in Brillinger and Kaiser (1992).

There are extensions of the cross-spectral approach to the 2, 3, 4, ... and higher dimensional cases, see Blümich (1985).

In this example the Fourier transform is useful for examining resonance, for assessing goodness of fit and for understanding the nonlinearity involved.

5. Wavelets

5.1. Basics.

Wavelet analysis is a technique enjoying a surge of contemporary discussion and is seen by some as a competitor of Fourier analysis. At the outset it may be remarked that in any case wavelets are intimately connected with Fourier analysis, see eg, Strichartz (1993), Benedetto and Frazier (1994). Wavelets are of practical importance because they can sometimes provide more parsimonious descriptions than Fourier ones.
Wavelets focus on local versus global behavior and in particular can pick up transient behavior. Basic is a (mother) wavelet $\psi(.)$ nonzero only on the unit interval $[0,1)$. Given a square-integrable function $g(x)$, one has an expansion

$$g(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \beta_{jk} \psi_{jk}(x)$$  \hspace{1cm} (5.1)$$

with

$$\psi_{jk}(x) = 2^{j/2} \psi(2^j x - k)$$  \hspace{1cm} (5.2)$$

and

$$\beta_{jk} = \int \psi_{jk}(x) g(x) dx$$

The family $\{\psi_{jk}(.)\}$ is orthonormal and complete, see eg. Daubechies (1992), Walter (1992,1993), Strichartz (1993), Benedetto and Frazier (1994).

The expansion (5.1) represents $g(.)$ in terms of functions with support individually on dyadic intervals $[k/2^j, (k+1)/2^j)$ for $j, k$ integers. Expression (5.1) suggests the approximation

$$g_n(x) = \sum_{j=-\infty}^{n-1} \sum_{k=-\infty}^{\infty} \beta_{jk} \psi_{jk}(x)$$  \hspace{1cm} (5.3)$$

This may be written in terms of a (father) wavelet, $\phi(.)$, as

$$g_n(x) = \int W_n(x,y) g(y) dy$$  \hspace{1cm} (5.4)$$

the kernel being

$$W_n(x,y) = \sum_{k=-\infty}^{\infty} \phi_{nk}(x) \phi_{nk}(y)$$  \hspace{1cm} (5.5)$$

with

$$\phi_{nk}(x) = 2^{n/2} \phi(2^n x - k)$$

This kernel tends to a delta function as $n \to \infty$, see Walter (1992). Expression (5.4) can be used to study the degree of approximation
directly. Expressions (5.4) and (5.5) are wavelet analogs of (1.5) and (1.6).

Just as there are fast Fourier transforms, there are fast wavelet transforms, Strang (1993). Also one could write $p 2^j$ for $2^j$ in (5.2) above, with no real change. The dynamic spectrum analysis of Example 4.2 is one type of wavelet analysis.

There are empirical versions of (5.3) for use when data $Y(t)$, $t = 0, \ldots, T-1$ are available. One computes, for example, the

$$\hat{\beta}_{jk} = \frac{1}{T} \sum_{t=0}^{T-1} \psi_{jk}(t/T)Y(t)$$  \hspace{1cm} (5.6)

In work to obtain improved estimates Donoho and Johnstone (1990), Hall and Patil (1993) create shrinkage estimates involving multipliers, which are referred to as "thresholders". These estimates take the form

$$\hat{g}_n(x) + \sum_{j=n}^{\infty} \sum_{k=-\infty}^{\infty} w(|\hat{\beta}_{jk}|/s_{jk}) \hat{\beta}_{jk} \psi_{jk}(x)$$  \hspace{1cm} (5.7)

with $0 \leq w(.) \leq 1$.

5.2 Microtubule Movement.

As an illustration of wavelet analysis, consider searching for jumps in records of microtubule movement. Microtubules are linear polymers basic to cell motility. A concern is whether movement is smooth, or rather via a series of jumps, see Malik et al. (1993).

If $Y(t)$ denotes the distance a microtubule has travelled at time $t$, the model considered is

$$Y(t) = \alpha t + g(t/T) + \text{noise}$$  \hspace{1cm} (5.8)

t = 0, \ldots, T-1 \text{ with } \alpha \text{ a parameter related to diffusion motion and } g(.) \text{ a step function.}$$
In the case of such a discontinuous function, a particular wavelet analysis is especially suitable, Haar wavelet analysis. This analysis is based on the functions

\[ \phi(x) = \begin{cases} 1 & \text{for } 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \]

and

\[ \psi(x) = \begin{cases} 1 & \text{for } 0 \leq x < 1/2 \\ -1 & \text{for } 1/2 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \]

In the Haar case the kernel is simply

\[ W_n(x, y) = 2^n \phi(2^n y - [2^n x]) \] (5.9)

and \( g_n(x) \) of (5.4) a local mean, \( g_n(x) = |I|^{-1} \int_I g(y) dy \), \( x \) being in the particular interval \( I = [m/2^n, (m+1)/2^n) \), see Fine (1949), Brillinger (1991), Walter (1992). The model (5.8) will be approximated by

\[ Y(t) = \alpha t + \sum_k \gamma_{nk} \phi_{nk}(t/T) + \text{noise} \] (5.10)

for some \( n \). Because of the presence of the term \( \alpha t \) in (5.10) the analysis in the present case is not so immediate, but still all that one needs are local means. The least squares estimates are obtained by regression of \( Y \) on the \( \phi_{nk}(t/T) \) and on \( t \) made orthogonal to the \( \phi_{nk} \).

In the experiments of concern samples were taken from the bovine brain, specifics may be found in Malik et al. (1993). The top panel of Figure 7 provides a data trace. Next is an estimate \( \hat{\phi}_{nk}(t/T) \) with \( w(u) \), of (5.7), \( \alpha = 1 \). Then an improved estimate with \( w(u) = (1 - 1/u^2)_+ \). The value of \( n = 3 \) was chosen having in mind a search for isolated jumps. Also indicated are approximate \( \pm 2 \) standard error limits around the fitted straight line. There is not much evidence for the presence of isolated jumps. The construction of the standard error estimate is described in the
Appendix.

The Fourier transform was used here to develop uncertainty estimates, following on an assumption that the noise was is stationary.

6. Discussion and Summary

"One can Fourier transform anything, often meaningfully."

J. W. Tukey

The principal interest of the examples of the paper has been in problem formulation and in addressing particular scientific questions. In each of the examples, an empirical Fourier transform has played a central role. With its broad collection of understood properties this transform has assisted the analyses greatly. The usefulness of second shrinkage, analogous to the use of convergence factors in Fourier approximation, is also noteworthy.

The particular groups of the examples have been abelian. General group theoretic ideas and empirical Fourier analysis have been discussed for other groups. For the case of the symmetric group see Diaconis (1988,1989) and Kim and Chapman (1993). For the locally compact abelian case see Brillinger (1982). For p-adics see Brillinger (1992). The use of p-adics in signal processing is discussed in Gorgui-Naguib (1990). For other cases see Hannan (1969). Key distinctions that arise are abelian versus nonabelian groups, compact versus locally compact groups, and whether $t$ is in a group or $Y$ is in a group.

There are other transforms that are useful in practice. These include: the Laplace, Hilbert, Stieltjes, Mellin, with some work having been done for abstract groups, see Loomis (1953).
The case of lacunary trigonometric series is somewhat like the case of point processes. Here the Fourier transform has a different form, eg. for point process data \( \{\tau_1 < \tau_2 < \ldots < \tau_N\} \) it is given by

\[
\sum_{j=1}^{N} \exp(-i\lambda \tau_j)
\]

\(-\infty < \lambda < \infty\). Such a transform is used in Rosenberg et al. (1989) for example.

Unemphasized, but important, topics include: the Poisson summation formula useful in understanding aliasing and the sampling theorem (Hannan (1965)), fast algorithms (Rockmore (1990)), spherical functions, uncertainty principles (Smith (1990)).

7. Some Open Problems

This section briefly lists a number of topics, motivated by the examples of the paper, that appear fruitful for more development.

Foremost among the topics calling out for further research is the theoretical and practical development of shrinkage estimates. The ideas are basic. The effects are substantial, see eg. Figure 2. One wonders about "optimal" choice of the multipliers/shrinkage factors. Perhaps optimal rates of convergence may be determined and then it be checked which multipliers lead to those. This paper has focussed on second shrinkage. Berger and Wolpert (1983) develop third shrinkage estimates in random function cases. Lillestol (1977) studies time series in one case.

In both the surface wave and nuclear magnetic resonance examples, examination of residuals suggests the presence of signal-generated noise. Better estimates are needed. Either because the ones used are inefficient or because the signal-generated noise is basic. In the latter case an appropriate likelihood function needs to be developed. If the noise is
indeed nonstationary and autocorrelated, then a novel of uncertainty estimation technique will be needed. In the case of the "improved" wavelet estimate, the uncertainty was estimated as if the shrinkage factors were constant, see Appendix. Perhaps a useful bootstrap procedure could be developed, based on an assumption of stationary innovations being present.

Quite a different type of problem is the following: develop the aliasing structure for higher-order spectra in the case of a process observed on a lattice. This will be particularly complicated in the case of lattices in $R^p$ with $p > 1$. Another problem is, in the case of image estimates, how to visualize the associated uncertainty.

The Fourier transforms studied have all been scalar-valued. There are central limit theorems for processes taking on values in a group. It would be of interest to obtain corresponding results for group-valued Fourier transforms, eg. in the p-adic case.

ACKNOWLEDGMENTS

The author thanks his collaborators in the examples of Section 4: Ken Downing, Bob Glaeser, Guy Perkins, Bruce Bolt and Reinhold Kaiser, respectively. The microtubule data studied in Section 5 was collected by Fady Malik and Ron Vail. This paper is based on talks presented at the opening ceremonies of the Fields Institute for Research in Mathematical Sciences at the University of Waterloo, Ontario, Canada and at the 1993 Winter Meeting of the Canadian Mathematical Society at Ottawa. The work was partially supported by the NSF Grants DMS-9208683 and DMS-9300002.
REFERENCES


Preprint, University of California, San Francisco.


MICHAELSON, A.A. (1891b). On the application of interference methods to spectroscopic methods - II. Phil. Mag. 34, 280-299.


APPENDIX

This section describes the construction of the error limits employed in Figure 7 the wavelet analysis of Section 5.2. What is wanted is an estimate of the standard error of the step function estimate \( \hat{g}_n(t/T) \). The use of the Haar wavelet leads to important simplifications.

The asymptotics developed have in mind \( T = VL \) with \( T = 2^N, V = 2^n, L = 2^{N-n} \) and \( V, L \to \infty \) as \( T \to \infty \).

Consider first the case without shrinkage. The situation is then that of linear regression (with stationary errors). The estimate presented in the middle panel of Figure 7 is ordinary least squares. (In many time series situations such estimates are asymptotically efficient.)
The model (5.10) is linear in $\alpha$ and the $\gamma_{nk}$. It may be written

$$y = X\gamma + Z\alpha + \varepsilon$$

with $Z = [t-t]$ and $X = [X_{tk}]$, with $X_{tk} = 1$ for $k/2^n \leq t/T < (k+1)/T$ and 0 otherwise. The least squares estimates may be written

$$\hat{\alpha} = (Z'PZ)^{-1}Z'Py$$  \hspace{1cm} (A.1)

$$\hat{\gamma} = (X'X)^{-1}X'(y - Z\hat{\alpha})$$  \hspace{1cm} (A.2)

with $P = I - X(X'X)^{-1}X'$. What is desired are the variances of the fitted values $X\hat{\gamma}$. Were the $\alpha$ term absent, the problem would be easy since the $X\hat{\gamma}$ are simply local means of $V = 2^n$ $Y$-values and, following (3.3), the large sample variance is approximately $2\pi f(0)/V$ where $f(.)$ is the power spectrum of the noise.

The fitted values have the form

$$X\hat{\gamma} = Hy - HZ\hat{\alpha}$$  \hspace{1cm} (A.3)

with $H = X(X'X)^{-1}X'$. The first term on the right in (A.3) is the local mean above. Its variance is approximately $2\pi f(0)/V$. Following the arguments of Theorem 5.11.1 in Brillinger (1975) the variance of the second term of (A.3) may be seen to be $O(V^{-1}L^{-1})$ and that term may be neglected as it has been assumed that $L \rightarrow \infty$. The needed power spectrum may be estimated from the residuals of the fit.

To obtain an estimate of the standard error in the case of a shrinkage estimate, where multipliers are inserted before the $\hat{\beta}_{nk}$, as is usual for example when obtaining standard error estimates for robust regression coefficients, it is assumed that the multiplier values are constant.
Legends

Figure 1. Graph of the multipliers (4.7) and (4.8) as a functions of the amplitude of the estimate divided by its estimated standard error.

Figure 2. Estimates of the basic cell of bacteriorhodopsin. The upper left panel is the naive estimate (4.5). The upper right panel is the estimate (4.6) with the multiplier (4.8). The bottom left panel is the estimate (4.6) with the multiplier (4.7). The last panel is (4.6), with (4.8), obtained by combining 42 individual images.

Figure 3. The Siberia-Upsalla dynamic spectrum as a function of frequency and velocity as computed from (4.11). The vertical trace is the seismogram as a function of velocity.

Figure 4. The top trace is the seismogram as a function of time. The middle is the fit based on (4.9). The bottom is the difference of these two.

Figure 5. Results of a nuclear magnetic resonance study of 2,3-dbt. The top right is a segment of the input and below is the corresponding output. The right column provides the estimated amplitude and phase of the (linear) transfer function.

Figure 6. The modulus of the Fourier transform of the output and of the corresponding fit derived from (4.11) and (4.12).

Figure 7. The top trace is the estimated movement of a microtubule as a function of time. The middle provides the fit of (5.10). The bottom panel employs the improved fit (5.7). The dashed lines provide approximate ±2 standard error limits.
Naive image

Blow-Crick image

Wiener image

Final image, \( n = 42 \)
Shrinkage factor

amplitude / standard error
Siberian Event at Uppsala April 20, 1989

Fit

Residuals
Microtubule movement

Local fit

Improved local fit
A Note on River Wavelets

David R. Brillinger

ABSTRACT

Wavelet analysis is described, and a Haar wavelet analysis is carried out, for the time series data of the flow rate of the Nile River at Aswan and also of the stages of the Amazon River at Manaus. A goal is looking for jumps in the average level. The errors around the average level are assumed stationary.

KEY WORDS: Amazon River stages, autocorrelation, change-point, Nile River flow, wavelets

1. INTRODUCTION

Problems of change are fundamental to studies of the environment and a variety of pertinent statistical techniques exist, see eg. MacNeill et al. (1991). In this work a method, based on wavelets, is applied to elicit change-points on data concerning the Nile and Amazon Rivers.

One goal is to illustrate this emerging technique of wavelet analysis. A second goal is to show how to obtain uncertainty measures of the results of such an analysis in a case with stationary serial dependence present. There is also mention of possible extensions and of improved procedures.

Wavelets are a contemporary tool for function approximation. They are competitors/collaborators with traditional Fourier analysis. In

1David Brillinger is Professor of Statistics, University of California, Berkeley, CA 94720.
particular they are useful for handling localized behavior, discontinuities, scale and shift transformations.

The idea of wavelets is illustrated with analyses of two river series: Nile River flow at Aswan from 1871 to 1970 and Amazon River stages at Manaus from 1901 to 1992.

Section 2 describes wavelet analysis generally. Section 3 indicates some statistical aspects of the technique. Section 4 presents worked analyses based on Haar wavelets for the Nile and Amazon series. The final section discusses the results and indicates some possible extensions.

2. WAVELETS

Wavelets focus on local versus global behavior. Basic is a function $\psi(.)$ nonzero only on the unit interval [0,1). Given a square-integrable function $g(x)$, one has the expansion

$$g(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \beta_{jk} \psi_{jk}(x)$$  \hspace{1cm} (2.1)

with

$$\psi_{jk}(x) = 2^{j/2} \psi(2^j x - k)$$  \hspace{1cm} (2.2)

and

$$\beta_{jk} = \int \psi_{jk}(x) g(x) dx$$

since the family $\{\psi_{jk}(.)\}$ is orthonormal and complete, see eg. Daubechies (1992), Walter (1992,1993), Strichartz (1993), Benedetto and Frazier (1994). The expansion (2.1) represents $g(.)$ in terms of functions with support individually on dyadic intervals $[k/2^j, (k+1)/2^j)$ with $j, k$ integers. Expression (2.1) may be written

$$g(x) = g_n(x) + \sum_{j=n}^{\infty} \sum_{k=-\infty}^{\infty} \beta_{jk} \psi_{jk}(x)$$  \hspace{1cm} (2.3)

where
\[ g_n(x) = \sum_{k=-\infty}^{\infty} \alpha_k \phi_k(x) \]  \hspace{1cm} (2.4)

with

\[ \phi_k(x) = 2^{n/2} \phi(2^n x - k) \]

for appropriate \( \phi(.) \) and

\[ \alpha_k = \int \phi_k(x) g(x) dx \]

The approximant \( g_n(x) \) has the representation

\[ g_n(x) = \int_{-\infty}^{\infty} W_n(x,y) g(y) dy \]  \hspace{1cm} (2.5)

with the kernel

\[ W_n(x,y) = \sum_{k=-\infty}^{\infty} \phi_k(x) \phi_k(y) \]

tending to a delta function as \( n \to \infty \), see Walter (1992). The expression (2.5) can be used to study the degree of approximation directly.

In the case of a function with discontinuities, a particular wavelet analysis is especially suitable, namely Haar wavelet analysis. This analysis is based on the functions

\[ \phi(x) = 1 \quad \text{for } 0 \leq x < 1 \]
\[ 0 \quad \text{otherwise} \]

and

\[ \psi(x) = 1 \quad \text{for } 0 \leq x < 1/2 \]
\[ -1 \quad \text{for } 1/2 \leq x < 1 \]
\[ 0 \quad \text{otherwise} \]

In the Haar case the kernel is simply

\[ W_n(x,y) = 2^n \phi(2^n y - [2^n x]) \]  \hspace{1cm} (2.6)

and \( g_n(x) \) is a local mean, \( g_n(x) = |I|^{-1} \int_I g(y) dy \) where \( x \) is in the particular interval \( I = [m/2^n, (m+1)/2^n] \), see Brillinger (1991).

Haar wavelet analysis will be performed in the examples below. In
other circumstances it is desirable that the functions $\phi(\cdot), \psi(\cdot)$ be smooth and other wavelets are appropriate, see eg. Benedetto and Frazier (1994). Also one could write $p 2^j$ for $2^j$ in (2.2) above, with no real change.

3. A STOCHASTIC MODEL AND SOME CONSEQUENCES

Suppose that data, $Y(t), t = 0, ..., T-1$ are available. Consider the model

$$ Y(t) = S(t) + E(t) $$

with $E(\cdot)$ stationary noise and $S(\cdot)$ a deterministic signal. In terms of the previous section $S(t) = g(t/T)$.

Given $n$, an elementary estimate of $S(t)$ suggested by (2.2), (2.3), is

$$ \hat{S}_n(t) = \sum_{k=-\infty}^{\infty} \hat{\alpha}_k \phi_k(t/T) + \sum_{j=n}^{\infty} \sum_{k=-\infty}^{\infty} \hat{\beta}_{jk} \psi_{jk}(t/T) $$

with

$$ \hat{\alpha}_k = \frac{1}{T} \sum_{t} \phi_k(t/T)Y(t) $$

$$ \hat{\beta}_{jk} = \frac{1}{T} \sum_{t} \psi_{jk}(t/T)Y(t) $$

see eg. Donoho and Johnstone (1990) or Hall and Patil (1993). Despite its appearance, expression (3.2) actually involves only a finite number of $k$ values. Following the discussion of the previous section, one might take simply the first term on the right of (3.2) as the estimate.

In the case of the Haar wavelet things simplify. In particular one can consider

$$ \hat{S}_n(t) = \sum_{k=-\infty}^{\infty} \hat{\alpha}_k \phi_k(t/T) = \sum_{l_t} Y(s) / \sum_{l_t} 1 $$

where $l_t$ is the dyadic interval of length $2^{-n}$ containing $t/T$. Computing such local means, in a search for change-points, seems intuitively plausible.
The statistics (3.3), (3.4), (3.5) are each linear in the $Y$'s, hence some sampling properties are directly available, eg. their large sample variances and distributions. In the case of (3.4) for example
\[ \text{var } \hat{\beta} = \int |\Psi(\lambda)|^2 f_{EE}(\lambda) d\lambda \] (3.6)
with $f_{EE}(\cdot)$ the power spectrum of $E(\cdot)$ at frequency $\lambda$ and
\[ \Psi(\lambda) = \frac{1}{T} \sum_{t=0}^{T-1} e^{-i\lambda t} \psi(t/T) \]
In the case of (3.5)
\[ \text{var } \hat{S}(t) = \frac{2\pi}{V} f_{EE}(0) \] (3.7)
where $V = T/2^n$, the number of data points in $I_t$, is assumed large. One can estimate $f_{EE}(0)$ in various ways. See for example Brillinger (1975) for estimates and the result (3.7). A particular estimate will be employed in the next section.

Donoho and Johnstone (1990) and Hall and Patil (1993) are concerned with estimates improving on (3.2). They modify the coefficients $\hat{\beta}$ in a nonlinear way. In an estimate employed in the next section, the coefficients, $\hat{\beta}$, will be shrunk to
\[ w(1|\hat{\beta}|/s) \hat{\beta} \]
where $s$ is an estimate of the standard error of $\hat{\beta}$ and $w(u) = 1$ for large $|u|$ and $= 0$ for small. This has the effect of leaving terms with $\hat{\beta}$ significantly large virtually unchanged, while reducing others and setting to 0 those less than their standard error.

Inserting the multipliers complicates the computation of standard errors, but it is hoped that this effect is secondary and in our computations the multipliers will be treated as constants.

In the Haar case things simplify. Suppose one employs the term in (3.2) with \( j = n \) as an improvement, then

\[
\hat{S}_{n+1}(t) = \hat{S}_n(x) + \sum_k \hat{\beta}_{nk} \psi_{nk}\left(t/T\right)
\]

and one can consider a shrunken estimate of the form

\[
\hat{S}_n(t) + w(\frac{\hat{\beta}_{nk}}{s_{nk}}) [\hat{S}_{n+1}(x) - \hat{S}_n(t)]
\]  

(3.8)

where \( k \) is such that \( 0 \leq 2^n x - k < 1 \) and \( s_{nk}^2 \) is an estimate of the variance of \( \hat{\beta}_{nk} \). The form of \( \hat{\beta}_{nk} \) is

\[
\frac{2^{n/2}}{T} [\Sigma Y(t) - \Sigma'' Y(t)]
\]  

(3.9)

where \( \Sigma' \) is over \( 0 \leq 2^n t/T - k < 1/2 \) and \( \Sigma'' \) is over \( 1/2 \leq 2^n t/T - k < 1 \). Expression (3.9) suggests taking \( 2\pi f_{EE}(0)/T \) for an estimate of its variance.

Turning to an approximate variance for (3.8), writing \( w = w(\frac{\hat{\beta}_{nk}}{s_{nk}}) \), (3.8) has the form

\[
[(1 \pm w)\Sigma' Y(t) + (1 \mp w)\Sigma'' Y(t)]/V
\]

suggesting the estimate

\[
(1 + w^2)2\pi f_{EE}(0)/V
\]  

(3.10)

This is the estimate employed in the next section.

4. EXAMPLES

The series of annual discharges of the Nile River at Aswan has been a common testbed for change-point techniques. A historical review of the data and of various analyses has been given in MacNeill et al. (1991). The data themselves are listed in Cobb (1978). The data are annual July-
June flows from 1871 to 1970 and are graphed in Figure 1. The estimate (3.5) is graphed in the bottom panel of the Figure with \( p = 25, \ T = 100, \ V = 25, \ L = T/V = 4, \ n = 2 \). The standard error of \( \hat{S}(t) \) is estimated, following (3.7), by \( \sqrt{2\pi \hat{f}_{\text{EE}}(0)/V} \) where

\[
\hat{f}_{\text{EE}}(0) = \frac{1}{L} \sum_{l=1}^{L} I^V(l, l) \frac{2\pi}{V}
\]

with \( I^V(\lambda, l) \) the periodogram of the \( l \)-th segment of the data. The \( \pm 2 \) standard error limits about the overall mean are shown by the dashed lines in the Figure. One sees that the first segment of the fit is well outside the limits. This is consistent with previous studies, MacNeill et al. (1991). These authors note that a dam was built at Aswan in the period 1899 to 1902.

A second example is based on the stages of the Amazon River at Manaus. Daily stages have been recorded since 1901. These data are described in Sternberg (1987) and in Brillinger (1989), where they were analysed for the presence of a monotonic trend. In the present case monthly values are employed, having reduced the seasonal effect by removing overall monthly means. \( T = 1024 \) values employed. The seasonally adjusted series is graphed in the top panel of Figure 2. The bottom panel provides the estimated autocorrelation function and \( \pm 2 \) standard error limits (as dashed lines). This display provides definite evidence for the presence of serial correlation that needs to be taken account of in uncertainty computations.

The top panel of Figure 3 gives (3.5) with \( T = 1024, \ V = 128, \ L = 8 \). The bottom panel gives the shrunken estimate (3.8), with \( w(u) = (1 - 1/u^{-2})_+ \). The dashed lines again give \( \pm 2 \) standard error limits, now computed using (3.10). There is evidence for an increased aver-
age in the years around 1970.

5. DISCUSSION AND EXTENSIONS

The analyses presented have involved both wavelet and Fourier analyses, the former to handle level shifts and the latter to handle autocorrelation in the errors.

The work shows that when Haar wavelet analysis is employed it takes a simple form, including the uncertainty estimation.

A problem is the choice of n. In Brillinger (1991) the AIC was employed, but that was based on white noise errors.

The confidence bounds in the Figures are marginal. Simultaneous bounds could be constructed in a manner extending Bjerve et al. (1985).

ACKNOWLEDGMENTS

The work was partially supported by the NSF Grant DMS-9300002.

REFERENCES


Figure legends

Figure 1. The top panel graphs the annual discharge of the Nile at Aswan for the period 1871 to 1970. The bottom panel graphs the fitted average level function. The dashed lines are marginal ±2 standard error bounds about the overall mean level.

Figure 2. The top panel is obtained by taking the monthly averages of daily values to obtain a monthly series, then removing the overall monthly means to get the residuals from seasonally adjusting. The bottom panel is the estimated autocorrelation function. The dashed horizontal lines give marginal ±2 standard error limits for a white noise series.

Figure 3. The top panel provided the estimate (3.5). The horizontal dashed lines are ±2 standard error marginal limits. The bottom panel is the "improved" estimate (3.8). The horizontal dashed lines are marginal ±2 standard error limits.
Nile discharge at Aswan

Local fit
Amazon monthly stages with seasonal removed

Auto correlation function
Local fit of Amazon stages

Improved local fit