J. Appl. Probab. **38A,** 701–717 (2001)
Printed in England

INTEGRATED PROCESSES AND THE DISCRETE COSINE TRANSFORM

ROBERT B. DAVIES,[1] *Statistics Research Associates Limited, Wellington*

Abstract

A time-series consisting of white noise plus Brownian motion sampled at equal intervals of time is exactly orthogonalised by a discrete cosine transform (DCT-II). This paper explores the properties of a version of spectral analysis based on the discrete cosine transform and its use in distinguishing between a stationary time-series and an integrated (unit root) time-series.

Keywords: Beta-optimal test; Brownian motion; DCT-II; discrete cosine transform; integrated process; random walk; spectrum; time-series; unit root

AMS 2000 Subject Classification: Primary 62M10
Secondary 62M15; 91B84

1. Introduction

A time-series $X_1, X_2, \dots$ is said to be an *integrated process* if it is non-stationary but the differences $X_2 - X_1, X_3 - X_2, \dots$ are stationary. A simple example is the sum of a white noise process and a Brownian motion sampled at equal intervals of time. This paper addresses the problem of distinguishing between an integrated process and a stationary process. There are two situations.

In the first, the hypothesis is that we observe a stationary sequence of normal random variables and the alternative is that it is the sum of a stationary sequence of normal random variables and a Brownian motion.

The second is that the hypothesis is that it is an integrated process and the alternative is that it is stationary. An example is the unit root testing problem (e.g. Dickey *et al.* 1986).

There is a vast literature on these problems, particularly the second. This paper provides a new approach through the use of the DCT-II version of the discrete cosine transform. This transform exactly orthogonalises the translation invariant part of a Brownian motion plus white noise, and so provides a new tool for both the theory and practice of the analysis of integrated processes.

The cosine transform is defined in Section 2 which also presents properties of a variant of the spectrum based on this transform and properties of the transform when applied to a first order autoregressive process. Section 3 investigates tests for the two

Received December 2000
[1] Postal address: Statistics Research Associates Limited, P.O. Box 12 649, Thorndon, Wellington, New Zealand. Email: `robert@statsresearch.co.nz`

situations described earlier. The tests are based on local optimality and beta-optimality (Davies, 1969).

In the Appendix we derive the likelihood ratios for the translation- and scale-independent transforms of the data, the statistical properties of the discrete cosine transform and the formulae for the scale-independent locally- and beta-optimal tests.

Notation: A' denotes the transform of a matrix A; 1_n denotes an n-dimensional column vector of ones and I_n denotes the $n \times n$ identity matrix. The Frobenius norm $(\sum\sum A_{j,k}^2)^{1/2}$ of a matrix A is denoted by $\|A\|_F$.

2. Cosine transform

Suppose $\{W_i\}$ is a standard Brownian motion, that is the differences $W_i - W_{i-1}$ are independent standard normal random variables with $W_0 = 0$. Let

$$X_i = \mu + \xi W_i + \epsilon_i, \tag{1}$$

where the ϵ_i are independent normal random variables with mean 0 and variance σ^2. Then the X_i, for $i = 1, \ldots, n$, have variance–covariance matrix $\xi^2\Sigma + \sigma^2 I_n$ where Σ is as defined in (A.7). It follows from Theorems A.3 and A.4 in the appendix that

$$F_j = (2/n)^{1/2} \sum_{k=1}^{n} X_k \cos \frac{\pi j(k - \frac{1}{2})}{n} \qquad (j = 1, \ldots, n-1), \tag{2}$$

are independent with zero means and variances

$$\sigma^2 + \frac{\xi^2}{4\sin^2(\frac{1}{2}\pi j/n)}\,. \tag{3}$$

Note that we are calculating $n - 1$ *translation invariant* terms and the orthogonal decomposition applies only to the *translation invariant* part of the process. The transformation (2) is known as the *discrete cosine transformation-II* (DCT-II). See Rao and Yip (1990) and Van Loan (1992) for further details and methods of computation. The DCT-II is well-known in the communications and information processing literature but has received only occasional attention in the statistical literature. It has properties similar to those of the usual discrete Fourier transform

$$C_j = (2/n)^{1/2} \sum_{k=1}^{n} X_k \cos(2\pi jk/n); \quad S_j = (2/n)^{1/2} \sum_{k=1}^{n} X_k \sin(2\pi jk/n). \tag{4}$$

Unlike the usual Fourier transform it gives an exact orthogonalisation of the process (1). Thus the DCT-II (2) may be a better starting point for carrying out a frequency analysis of a process with a Brownian motion component than the Fourier transform (4).

2.1. Discrete frequencies

Does the DCT-II separate out discrete frequency components? Suppose

$$X_k = \cos\{\pi\theta(k - \tfrac{1}{2})/n + \phi\},$$

i.e. X_k is a discrete frequency with $\frac{1}{2}\theta$ cycles over the period of observation. Then F_j is equal to

$$\begin{cases} \dfrac{1}{\sqrt{2n}}\left[\dfrac{\sin\{\frac{1}{2}\pi(\theta+j)\}\cos\{\frac{1}{2}\pi(\theta+j)+\phi\}}{\sin\{\frac{1}{2}\pi(\theta+j)/n\}} + \dfrac{\sin\{\frac{1}{2}\pi(\theta-j)\}\cos\{\frac{1}{2}\pi(\theta-j)+\phi\}}{\sin\{\frac{1}{2}\pi(\theta-j)/n)\}}\right] & \theta \neq j, \\ \sqrt{(n/2)}\cos\phi & \theta = j. \end{cases}$$

If $|\theta - j|$ is small compared with n, $\theta \neq j$, and $\sin\{\frac{1}{2}\pi(\theta+j)/n\}$ is not small, then

$$F_j = (2n)^{1/2}\left[\frac{\sin\{\frac{1}{2}\pi(\theta-j)\}\cos\{\frac{1}{2}\pi(\theta-j)+\phi\}}{\pi(\theta-j)} + O\Big(\frac{1}{n}\Big)\right].$$

For comparison, consider the standard Fourier transform (4). If $|\theta - 2j|$ is small compared with n, $\theta \neq 2j$, and $\sin\{\frac{1}{2}\pi(\theta+2j)/n\}$ is not small, then

$$(C_j^2 + S_j^2)^{1/2} \approx (2n)^{1/2}\sin\{\tfrac{1}{2}\pi(\theta-2j)\}/\{\pi(\theta-2j)\}.$$

If $\theta = j$ then $(C_j^2 + S_j^2)^{1/2} = (n/2)^{1/2}$. If $\theta = 2j$ only the jth pair of terms of (4) are affected by the discrete frequency. The corresponding result for (2) is that for $\theta = j$, when $\phi = 0$ only the jth term is affected, but if $\phi \neq 0$, then other terms are affected. This is a less satisfactory result than for the regular Fourier transform since one is unlikely to be able to arrange the sampling period so that known cycles affect the value of the transform for just one value of j.

When $\theta \neq 2j$ the range of points affected is similar for both transforms: see Table 1. The values of the transform values have been divided by $(n/2)^{1/2}$. The first column in the table shows $\theta - j$ for the cosine transform and $\theta - 2j$ for the usual Fourier transform.

Table 1: Response to discrete frequency component.

$\theta - (j \text{ or } 2j)$	F_j		$(C_j^2 + S_j^2)^{1/2}$
	$\phi = 0$	$\phi = \frac{1}{2}\pi$	
0.0	1.00	0.00	1.00
0.5	0.64	−0.64	0.90
1.0	0.00	−0.64	0.64
1.5	−0.21	−0.21	0.30
2.0	0.00	0.00	0.00
2.5	0.13	−0.13	−0.18
3.0	0.00	−0.21	−0.21
3.5	−0.09	−0.09	−0.13
4.0	0.00	0.00	0.00
4.5	0.07	−0.07	0.10
5.0	0.00	−0.13	0.13

2.2. Continuous spectrum

Let $\{X_k : k = 1, \ldots, n\}$ be a sample from a stationary time-series with auto-covariances $c_0, c_1, \ldots$. Define the spectrum

$$\lambda_j = c_0 + 2\sum_{k=1}^{\infty} c_k \cos\frac{\pi jk}{n} \tag{5}$$

for $j = 1, \ldots, n-1$. Then it is shown in the appendix that under certain conditions the F_j are approximately independently normally distributed with zero mean and variance given by (5). Thus F_j^2/λ_j are approximately independently chi-squared distributed with one degree of freedom.

More specifically, suppose

$$\sum_{j=1}^{\infty} |c_j| < \infty, \tag{6}$$

Λ is an $(n-1) \times (n-1)$ diagonal matrix with $\Lambda_{j,j} = \lambda_j$ and $\mathcal{L}$ denotes the actual variance–covariance matrix of the F_j. Then it follows from Theorem A.5 that

$$\lim_{n\to\infty} n^{-1/2} \|\mathcal{L} - \Lambda\|_F = 0 \tag{7}$$

and

$$\lim_{n\to\infty} \max_{j,k} |(\mathcal{L} - \Lambda)_{j,k}| = 0. \tag{8}$$

Equation (8) gives the asymptotic joint distribution of a small group of the F_j and is what is required to show that a suitably smoothed version of the F_j provides an estimate of the spectrum (5). Davies (1973, 1983) derives asymptotic results for fitting parametric models to stationary normal time-series. One can base the estimates on either the auto-covariances and the raw data or on the spectrum and Fourier transform of the data. Formula (7) is one of the formulae we need to show that we can also base the estimates on the F_j and λ_j. See Davies (1973, lemma 3.1 (iv)).

Note that Theorem A.5 applies only to symmetric series of a_j. Hence the DCT-II is less effective for analysing multivariate series as the cross-covariances are not likely to form a symmetric series. In particular, in co-integration studies using the DCT-II, it may be necessary to fit short term dependencies between the series as a parametric model separately from the spectral analysis.

2.3. Autoregressive process

Suppose X_k is generated by a first order autoregressive process:

$$X_{k+1} = \alpha X_k + \epsilon_k,$$

where the ϵ_k are independently and normally distributed with zero mean and unit variance. Then $c_j = \alpha^{|j|}/(1-\alpha^2)$ and

$$\lambda_j = 1/\{(1-\alpha)^2 + 4\alpha \sin^2(\tfrac{1}{2}\pi j/n)\}.$$

We are interested in the performance of the approximations given in the previous section, and in particular, when α is close to one. Let $\mathcal{L}$ and Λ denote the exact and approximate variance–covariance matrices of the F_j as in the previous section. Then Theorem A.6 shows that $(\mathcal{L} - \Lambda)_{j,k} = 0$ if $j + k$ is odd, and otherwise, i.e. for even $j + k$,

$$\frac{(\mathcal{L} - \Lambda)_{j,k}}{\sqrt{\lambda_j \lambda_k}} = -\frac{4\alpha(1-\alpha)\{1-\alpha^n(-1)^j\} \cos(\frac{1}{2}\pi j/n) \cos(\frac{1}{2}\pi k/n) \sqrt{\lambda_j \lambda_k}}{n(1+\alpha)}. \tag{9}$$

As expected, (9) tends to zero as $\alpha \to 1$ since the DCT-II provides an exact orthogonalisation of the Brownian motion process. For fixed α formula (9) is of order $1/n$. Unfortunately the approximation is not uniform (i.e. for any sample size n, there will be a value of α for which the approximation is poor). Let $\alpha = 1 - \gamma/n$ and consider the limit for fixed j,

$$\lim_{n\to\infty} \frac{\mathcal{L}_{j,j}}{\lambda_j} - 1 = -\frac{2\gamma\{1-(-1)^j e^{-\gamma}\}}{\gamma^2 + \pi^2 j^2}. \tag{10}$$

Values of this function for $j = 1, 2, 3, 4$ and various values of γ are given in Table 2.

Table 2: Error in cosine and sine transforms

	Fractional error in variance						
	Cosine transform				Sine transform		
$j =$	1	2	3	4	2	4	6
γ							
0.00	0.00	0.00	0.00	0.00	2.00	2.00	2.00
0.50	−0.16	−0.01	−0.02	−0.00	1.56	1.57	1.57
1.00	−0.25	−0.03	−0.03	−0.01	1.23	1.26	1.26
1.50	−0.30	−0.06	−0.04	−0.01	0.98	1.02	1.03
2.00	−0.33	−0.08	−0.05	−0.02	0.79	0.84	0.86
3.00	−0.33	−0.12	−0.06	−0.03	0.52	0.60	0.62
4.00	−0.31	−0.14	−0.08	−0.05	0.35	0.45	0.47
6.00	−0.26	−0.16	−0.10	−0.06	0.17	0.27	0.30
8.00	−0.22	−0.15	−0.10	−0.07	0.10	0.18	0.21
10.00	−0.18	−0.14	−0.11	−0.08	0.06	0.12	0.16
15.00	−0.13	−0.11	−0.10	−0.08	0.02	0.05	0.08
20.00	−0.10	−0.09	−0.08	−0.07	0.01	0.03	0.05

One can also calculate the same limit when the *cosines* in equation (2) are replaced by *sines*. This is of interest when comparing the performance of the cosine transform with the usual Fourier transform, which contains both sine and cosine terms. The formula corresponding to (10) is

$$\frac{2\pi^2 j^2\{1-(-1)^j e^{-\gamma}\}}{\gamma(\gamma^2+\pi^2 j^2)}.$$

Values of this function for $j = 2, 4, 6$ are also listed in Table 2. Odd values of j are not relevant to the sine transform as the standard Fourier transform is used only for even values of j.

For the cosine transform the difference between $\mathcal{L}_{j,j}$ and λ_j is serious only when $j \leq 3$. Even then, the accuracy is probably good enough for exploratory analysis, but for an accurate analysis a correction might be required. On the other hand with the sine transform the difference is large enough to be a problem even for exploratory analysis and does not die out for larger values of j. In the testing situation, we are probably not very concerned about the error for $0 < \gamma < 10$ since the power of any test for distinguishing between $\gamma = 0$ and any value with $0 < \gamma < 10$ is not large. See Table 4. In this table $\gamma = \xi n = 10$ gives a power of around 50%.

3. Testing for an integrated process

In this section I consider the problem of distinguishing between an integrated process and a stationary normal process. There are two different approaches to the problem.

The first uses the model of Harvey (1989, p.19). Suppose $\{W_i\}$ form a standard Brownian motion, $\{Z_i\}$ is a stationary process and

$$X_j = \xi W_j + Z_j. \tag{11}$$

The objective is to test the hypothesis $\xi = 0$ against the alternative $\xi > 0$.

The second approach is that of Dickey *et al.* (1986). Suppose

$$X_j = (1-\xi)X_{j-1} + Z_j \tag{12}$$

where $\{Z_j\}$ is a stationary process. Assume $\{Z_j\}$ is stationary if $\xi > 0$; otherwise assume it has a zero mean. Again the objective is to test the hypothesis $\xi = 0$ against the alternative $\xi > 0$.

In the first approach the null hypothesis is that the process is stationary and the alternative that the process is integrated. That is, the process appears to be stationary but we want to check whether there is a random drift component.

In the second approach the null hypothesis is that the process is integrated and the alternative is that it is stationary. That is, the process appears to have a random drift component and we want to check whether this drift is biased towards some central value.

The following subsections examine these two approaches. I first look at the situation where the $\{Z_j\}$ is an i.i.d. normal process and then find asymptotic results for $\{Z_j\}$ stationary. Sections 3.1 and 3.2 consider the i.i.d. normal case and section 3.3 extends these to the stationary normal case.

3.1. Sum of a Brownian motion and white noise

Consider the process defined by (1): $X_j = \mu + \xi W_j + \epsilon_j$. Then the $\{F_j : j = 1, \ldots, n-1\}$ defined by (2) are independent normal random variables with variances given by (3). The problem falls into the class described by formula (A.3). Since σ^2 is unknown we will be interested in a scale independent test. Hence the locally optimal test for $\xi = 0$ against $\xi > 0$ is found from (A.30) with ξ^2 as the parameter. It has critical region

$$\left\{ \sum_{j=1}^{n-1} \frac{F_j^2}{4\sin^2(\frac{1}{2}\pi j/n)} > c \sum_{j=1}^{n-1} F_j^2 \right\}. \tag{13}$$

The beta-optimal test is found from (A.28) and has critical region of the form

$$\left\{ \sum_{j=1}^{n-1} \frac{F_j^2}{\xi_0^2 + 4\sin^2(\frac{1}{2}\pi j/n)} > c \sum_{j=1}^{n-1} F_j^2 \right\}. \tag{14}$$

This is similar to (13) except that the first few F_j are given a more equal weighting.

The variances of the first few F_j are asymptotically equal to $\sigma^2 + \xi^2 n^2/(\pi^2 j^2)$ so we are interested in values of $\xi = O(1/n)$ when investigating the power of tests.

It is also of interest to consider the test derived from (14) when we let $\xi \to \infty$. This has critical region

$$\left\{ \sum_{j=1}^{n-1} F_j^2 \cos^2(\tfrac{1}{2}\pi j/n) > c \sum_{j=1}^{n-1} F_j^2 \right\}. \tag{15}$$

Since

$$\sum_{j=1}^{n-1} F_j^2 \cos^2(\tfrac{1}{2}\pi j/n) = \tfrac{1}{2}\sum_{1}^{n}(X_i - \overline{X})^2 + \tfrac{1}{2}\sum_{1}^{n-1}(X_i - \overline{X})(X_{i+1} - \overline{X}) + \tfrac{1}{4}\left(X_1 - \overline{X}\right)^2 + \tfrac{1}{4}\left(X_n - \overline{X}\right)^2$$

and $\sum F_j^2 = \sum \left(X_i - \overline{X}\right)^2$, this test is essentially the same as the test which rejects the hypothesis for large values of the first auto-correlation and so is listed as the correlation test in the tables.

Table 3: Power functions for the Harvey model

	ξ	Test			
		locally optimal	beta optimal	correlation	$F6$
$n = 20$	0.00	0.05	0.05	0.05	0.05
	0.50	0.50	0.55	0.48	0.46
	1.00	0.67	0.82	0.79	0.78
	1.50	0.73	0.90	0.89	0.88
	2.00	0.75	0.93	0.93	0.91
	2.50	0.76	0.94	0.95	0.93
$n = 100$	0.00	0.05	0.05	0.05	0.05
	0.10	0.59	0.65	0.37	0.61
	0.20	0.83	0.93	0.81	0.91
	0.30	0.92	0.98	0.95	0.98
	0.40	0.95	0.99	0.99	0.99
	0.50	0.97	1.00	1.00	1.00
$n = 500$	0.00	0.05	0.05	0.05	0.05
	0.02	0.61	0.67	0.20	0.63
	0.04	0.86	0.94	0.61	0.93
	0.06	0.94	0.99	0.86	0.98
	0.08	0.97	1.00	0.96	1.00
	0.10	0.99	1.00	0.99	1.00

Power functions for $n = 20, 100, 500$ are given in Table 3. Compared with the beta-optimal test the correlation test performs poorly for large n. The locally optimal test is also somewhat less powerful. The table includes a test, $F6$, with critical region

$$\left\{ \sum_{j=1}^{6} F_j^2 > c \sum_{j=7}^{n-1} F_j^2 \right\}. \tag{16}$$

Critical values can be derived from the F distribution. For larger values of n the $F6$ test has power close to that of the beta-optimal test and so is a good practical choice.

3.2. Autoregressive model

Let $\{\epsilon_j\}_{j=1,\ldots,n}$ be independent standard normal random variables. The process

$$X_j = \mu + (1-\xi)X_{j-1} + \epsilon_j, \tag{17}$$

is a first order autoregressive process with parameter $1-\xi$. Assume it is stationary if $\xi > 0$; otherwise suppose $\mu = 0$. Then the $\{F_j\}_{j=1,\ldots,n-1}$ are approximately independently normally distributed with zero mean and variance equal to

$$1/\{\xi^2 + 4(1-\xi)\sin^2(\tfrac{1}{2}\pi j/n)\}.$$

This result is exact if $\xi = 0$ or 1, but is only rough if $\xi = O(1/n)$. Suppose for the moment that it is exact. Then the scale-independent beta-optimal test of $\xi = 0$ against $\xi > 0$ has critical region

$$\left\{\sum_{j=1}^{n-1} F_j^2 \cos^2(\tfrac{1}{2}\pi j/n) < c\sum_{j=1}^{n-1} F_j^2\right\}. \tag{18}$$

This is independent of ξ and so is a uniformly most powerful test. This is the same as the test (15) except now we are rejecting the hypothesis for small values of the left-hand side. As before I refer to this as the correlation test.

To find the tests for the exact distributions, work with the first differences, $Y_j = X_{j+1} - X_j$ for $j = 1,\ldots,n-1$. Under (17) these have second moments given by

$$\operatorname{var} Y_i = 2/(2-\xi), \qquad \operatorname{cov}(Y_i, Y_{i+j}) = -\xi(1-\xi)^{j-1}/(2-\xi) \quad (j > 0).$$

The beta-optimal and locally optimal tests can be found from (A.28) and (A.30). The locally optimal test has critical region $\{\overline{Y}^2 < c\sum_1^{n-1} Y_i^2\}$ and turns out to be a very bad choice.

Power functions for these tests are given in Table 4 for $n = 20, 100, 500$. I also include the power of the test, $\hat{\rho}_\mu$, of Dickey *et al.* (1986, Table 2). This gives similar but very slightly poorer performance than the beta-optimal or correlation tests. Since the test based on $\hat{\rho}_\mu$ is essentially a correlation test it is not surprising that its performance is similar to that of the correlation test.

3.3. Stationary versus integrated processes

This section extends the results of sections 3.1 and 3.2 to the situation where $\{Z_j\}$ in (11) and (12) is a stationary normal process. The results are asymptotic as $n \to \infty$. In general, I do not show the explicit dependence on n. Suppose $\hat{\lambda}_j$ is an estimate of the spectrum λ_j of $\{Z_j\}$ as in (5). This section shows that versions of the tests considered in sections 3.1 and 3.2 remain valid, asymptotically, if F_j is replaced by $F_j/\hat{\lambda}_j$. Suppose that $\max_j |\hat{\lambda}_j - \lambda_j| \to 0$ in probability as $n \to \infty$ and λ_j is bounded away from 0.

Now consider the Harvey model (11) and the locally optimal test (13). Suppose $\xi = O(1/n)$ in (11) and the auto-correlations of $\{Z_j\}$ obey (6). The variance–covariance matrix of F is given by $\operatorname{var}(F) = T_0\Sigma T_0' = \xi^2 B + \Lambda + \Delta$ where B is diagonal with elements given by (A.11), Λ is diagonal with elements given by (5) and, according to Theorem A.5, $n^{-1/2}\|\Delta\|_F \to 0$ and $\max|\Delta_{j,k}| \to 0$.

Table 4: Power functions for the Dickey model

	ξ	Test			
		locally optimal	beta optimal	correlation	Dickey's $\hat{\rho}_\mu$
$n = 20$	0.00	0.05	0.05	0.05	0.05
	0.20	0.10	0.15	0.14	0.14
	0.40	0.14	0.39	0.38	0.35
	0.60	0.17	0.71	0.70	0.67
	0.80	0.20	0.91	0.91	0.90
	1.00	0.22	0.98	0.98	0.98
n=100	0.00	0.05	0.05	0.05	0.05
	0.05	0.11	0.19	0.18	0.17
	0.10	0.16	0.52	0.49	0.46
	0.15	0.19	0.84	0.81	0.78
	0.20	0.22	0.97	0.97	0.95
	0.25	0.25	1.00	1.00	1.00
n=500	0.00	0.05	0.05	0.05	0.05
	0.01	0.11	0.19	0.18	0.17
	0.02	0.16	0.51	0.48	0.45
	0.03	0.19	0.83	0.80	0.77
	0.04	0.22	0.97	0.96	0.95
	0.05	0.25	1.00	1.00	0.99

After replacing F_j by $F_j/\hat{\lambda}_j$ and rescaling, the critical region defined by (13) becomes

$$\left\{ \frac{1}{n^2} \sum_{j=1}^{n-1} \frac{F_j^2}{4\sin^2(\frac{1}{2}\pi j/n)\hat{\lambda}_j} > \frac{c}{n} \sum_{j=1}^{n-1} \frac{F_j^2}{\hat{\lambda}_j} \right\}. \tag{19}$$

The left-hand side tends to c. The right-hand side is dominated by the first few terms and so the Δ term can be ignored. Hence (19), possibly with the right-hand side replaced by a constant, has an asymptotic significance level independent of the spectrum of the $\{Z_j\}$ and power which depends on ξ^2/λ_0. A similar argument can be applied to tests (14) and (16).

Now consider the Dickey model (12) and the test (18). I consider the distribution only under the hypothesis. Rearranging and scaling (18) and substituting F_j by $F_j/\hat{\lambda}_j$, the critical region is where

$$\left\{ \frac{1}{n^2} \sum_{j=1}^{n-1} \frac{F_j^2}{\hat{\lambda}_j} < \frac{c}{n} \sum_{j=1}^{n-1} \frac{4F_j^2 \sin^2(\frac{1}{2}\pi j/n)}{\hat{\lambda}_j} \right\}. \tag{20}$$

According to Theorem A.7 and supposing (A.20) to be satisfied, F has its variance–covariance matrix equal to $T_0 \Sigma T_0' = \Lambda B + \Delta$ where Λ, B and Δ are as in the Harvey model. The right-hand side of (20) tends to c and the left-hand side is dominated by the first few terms and hence, as with the Harvey model, the test (20), possibly with the right-hand side replaced by a constant, has an asymptotic significance level independent of the spectrum of the $\{Z_j\}$.

A. Appendix

A.1. Likelihood ratios

Suppose $X_1, \dots, X_n$, represented by the column vector X, are normal random variables with mean 0. The aim of this section is to find the likelihood ratios for the scale and translation invariant parts of $X_1, \dots, X_n$. Suppose p_1 denotes the probability density of X if the variance–covariance matrix is Σ. Let p_0 denote the probability density if X is standard normal. Suppose Y denotes some function of X and let q_1 and q_0 be the densities of Y corresponding to p_1 and p_0. The following theorem provides formulae for q_1/q_0 for various definitions of Y.

Theorem A.1. *Let*

$$Q = \Sigma^{-1} - \Sigma^{-1} 1_n 1_n' \Sigma^{-1} / (1_n' \Sigma^{-1} 1_n). \tag{A.1}$$

Then the following are values of $q_1(Y)/q_0(Y)$ for various definitions of Y.
(1°) *If $Y = X - \overline{X} 1_n$ then*

$$q_1(Y)/q_0(Y) = n^{1/2} \{1_n' \Sigma^{-1} 1_n \det(\Sigma)\}^{-1/2} \exp\{\tfrac{1}{2} Y'(I_n - Q) Y\}. \tag{A.2}$$

(2°) *If $Y = X/\|X\|$ then*

$$q_1(Y)/q_0(Y) = \{\det(\Sigma)\}^{-1/2} (Y' \Sigma^{-1} Y)^{-n/2}. \tag{A.3}$$

(3°) *If $Y = (X - \overline{X} 1_n)/\|X - \overline{X} 1_n\|$ then*

$$q_1(Y)/q_0(Y) = n^{1/2} \{1_n' \Sigma^{-1} 1_n \det(\Sigma)\}^{-1/2} (Y' Q Y)^{-(n-1)/2}. \tag{A.4}$$

Proof. The proof makes extensive use of the following formula (see e.g. Davies, 1985):

$$q_1(Y)/q_0(Y) = \mathrm{E}_0 \big[p_1(X)/p_0(X) \bigm| Y \big] \tag{A.5}$$

where E_0 denotes expectation with respect to p_0. Details of the proof are left to the reader. □

We need only the scale invariant case (A.3) since the tests are based on the F_j at (2), and these have already been made translation invariant. However, it is the matrix Q in (A.2) and (A.4) that suggests the transform used in this paper. Now we investigate the eigenvalues of Q.

Theorem A.2. *Let $P = I_n - 1_n 1_n'/n$, Q be as in* (A.1),

$$R = P \Sigma P, \tag{A.6}$$

and Λ be an $(n-1) \times (n-1)$ diagonal matrix with elements equal to the reciprocals of the non-zero eigenvalues of Q. Then P, Q and R are simultaneously diagonalisable with non-zero eigenvalues equal to the diagonal elements of I_{n-1}, Λ^{-1} and Λ respectively.

Proof. Matrices Q and R are simultaneously diagonalisable since $QR = P = RQ$. Suppose TQT' and TRT' are diagonal and T orthogonal. Then TPT' is also diagonal with $n-1$ non-zero diagonal elements equal to 1. □

A.2. Orthogonalisation with cosine transform

Now consider the special case where Σ is the variance–covariance matrix of a Brownian motion sampled at equal time intervals.

Theorem A.3. *Suppose the (i,j)-th element of Σ is given by*

$$\Sigma_{i,j} = \min(i,j) \tag{A.7}$$

and T is an $n \times n$ orthogonal matrix with

$$T_{j,k} = (2/n)^{1/2} \cos\{\pi j(k - \tfrac{1}{2})/n\} \tag{A.8}$$

for $j = 1, \ldots, n-1$ and $k = 1, \ldots, n$; $T_{n,k} = n^{-1/2}$. Suppose also that T_0 is a matrix composed of the first $n-1$ rows of T. Suppose P, Q and R are as defined in Theorem A.2 *with Σ as in* (A.7)*. Then Q is tridiagonal with off-diagonal elements equal to -1, and $Q_{i,i} = 1$ if $i = 1$ or n, $= 2$ otherwise. and $Q = T_0' B^{-1} T_0$,*

$$R = T_0' B T_0, \tag{A.9}$$

$$B = T_0 \Sigma T_0', \tag{A.10}$$

where B is diagonal with

$$B_{j,j} = 1/\{4\sin^2(\tfrac{1}{2}\pi j/n)\} \qquad (j = 1, \ldots, n-1). \tag{A.11}$$

Proof. The matrix Σ^{-1} is tridiagonal with off-diagonal elements equal to -1, the last diagonal element equal to 1 and the other diagonal elements equal to 2. It follows that Q is as stated.

$$\begin{aligned}
\sum_{l=1}^{n-1} \frac{T_{l,j} T_{l,k}}{B_{l,l}} &= \frac{4}{n} \sum_{l=-n}^{n-1} \cos\{\pi l(j - \tfrac{1}{2})/n\} \cos\{\pi l(k - \tfrac{1}{2})/n\} \sin^2(\tfrac{1}{2}\pi l/n) \\
&= \frac{1}{n} \sum_{l=-n}^{n-1} [\sin\{\pi l j/n\} - \sin\{\pi l(j-1)/n\}] \\
&\times [\sin\{\pi l k/n\} - \sin\{\pi l(k-1)/n\}] \\
&= \frac{1}{2n} \sum_{l=-n}^{n-1} [2\cos\{\pi l(j-k)/n\} - \cos\{\pi l(j+k)/n\}] \\
&- \cos\{\pi l(j-k+1)/n\} + 2\cos\{\pi l(j+k-1)/n\} \\
&- \cos\{\pi l(j-k-1)/n\} - \cos\{\pi l(j+k-2)/n\}]
\end{aligned}$$

and this is equal to Q. Equation (A.9) follows from Theorem A.2 and (A.10) follows from (A.6) since $P = T_0' T_0$. □

The following theorem is a consequence of (A.10).

Theorem A.4. *Let X denote a Brownian motion with mean* 0 *and variance–covariance as at* (A.7)*, and let T_0 be as in Theorem* A.3*. Then the $n-1$ components of*

$$F = T_0 X \tag{A.12}$$

are independently normally distributed with mean 0 *and variance given by the diagonal elements of B as defined in* (A.11).

A.3. Asymptotic results

The next theorem is an analogue and extension of Davies (1973, lemma 3.1(iv)), and shows that T defined in (A.8) approximately diagonalises a symmetric Toeplitz matrix.

Theorem A.5. *Suppose* $\{a_i : i = 0, 1, \ldots\}$ *satisfy* $\sum |a_j| < \infty$, *and* A *is an* $n \times n$ *matrix with* $A_{j,k} = a_{|j-k|}$. *Let* Λ *be an* $n \times n$ *diagonal matrix with*

$$\Lambda_{j,j} = \begin{cases} a_0 + 2\sum_{k=1}^{\infty} a_k \cos(\pi jk/n) & (j = 1, \ldots, n-1), \\ a_0 + 2\sum_{k=1}^{\infty} a_k & (j = n). \end{cases} \tag{A.13}$$

Then

$$\lim_{n\to\infty} n^{-1/2} \|TAT' - \Lambda\|_F = 0, \tag{A.14}$$

$$\lim_{n\to\infty} \max_{j,k} |(TAT' - \Lambda)_{j,k}| = 0. \tag{A.15}$$

Proof. Consider (A.14). It is sufficient to show $n^{-1/2}\|T'\Lambda T - A\|_F \to 0$. Let $a_{-j} = a_j$.

$$\begin{aligned}(T'\Lambda T)_{p,q} &= \frac{1}{n} \sum_{k=-\infty}^{\infty} \sum_{j=-n}^{n-1} a_k \cos\frac{\pi j(p-\frac{1}{2})}{n} \cos\frac{\pi j(q-\frac{1}{2})}{n} \cos\frac{\pi jk}{n} \\ &= \frac{1}{4n} \sum_{k=-\infty}^{\infty} \sum_{j=-n}^{n-1} a_k \Big[\cos\frac{\pi j(p+q-1+k)}{n} + \cos\frac{\pi j(p+q-1-k)}{n} \\ &\qquad + \cos\frac{\pi j(p-q+k)}{n} + \cos\frac{\pi j(p-q-k)}{n}\Big] \\ &= \sum_{k=-\infty}^{\infty} (a_{p-q+2nk} + a_{p+q-1+2nk}).\end{aligned}$$

Hence

$$\begin{aligned}\frac{1}{n}\sum_{p=1}^{n}\sum_{q=1}^{n} |(T'\Lambda T - A)_{p,q}| &\le \frac{1}{n}\sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{k\neq 0} |a_{p-q+nk}| \\ &= \frac{1}{n}\sum_{p=-n+1}^{n-1} (n-|p|) \sum_{k\neq 0} |a_{p+nk}| = \sum_{p=-\infty}^{\infty} \min\left(\frac{|p|}{n}, 1\right)|a_p| \to 0\end{aligned}$$

as $n \to \infty$. Hence

$$\frac{1}{n}\sum_{p=1}^{n}\sum_{q=1}^{n} (T'\Lambda T - A)_{p,q}^2 \to 0$$

as $n \to \infty$ since the elements of $T'\Lambda T - A$ are bounded.

Now consider (A.15). Suppose $j = k \neq n$. Then

$$\begin{aligned}(TAT')_{j,j} &= \frac{2}{n}\sum_{p=1}^{n}\sum_{q=1}^{n} a_{p-q} \cos\frac{\pi j(p-\frac{1}{2})}{n} \cos\frac{\pi j(q-\frac{1}{2})}{n} \\ &= \frac{1}{n}\sum_{p=1}^{n}\sum_{q=1}^{n} a_{p-q} \left[\cos\frac{\pi j(p-q)}{n} + \cos\frac{\pi j(p+q-1)}{n}\right]\end{aligned}$$

$$= \frac{1}{n}\sum_{p=-n+1}^{n-1} a_p \sum_{q=1}^{n-|p|}\left[\cos\frac{\pi j|p|}{n} + \cos\frac{\pi j(|p|+2q-1)}{n}\right]$$
$$= \frac{1}{n}\sum_{p=-n+1}^{n-1} a_p \left((n-|p|)\cos\frac{\pi jp}{n} - \frac{\sin(\pi j|p|/n)}{\sin(\pi j/n)}\right). \qquad \text{(A.16)}$$

Also,

$$(TAT')_{n,n} = \frac{1}{n}\sum_{p=1}^{n}\sum_{q=1}^{n} a_{p-q} = \frac{1}{n}\sum_{p=-n+1}^{n-1}(n-|p|)a_p.$$

Hence

$$|(TAT')_{j,j} - \Lambda_{j,j}| \le (1+\tfrac{1}{2}\pi)\sum_{p=-\infty}^{\infty}\min(|p|/n,1)|a_p| \to 0 \qquad (n\to\infty).$$

Now suppose j, k and n are all different. Then

$$(TAT')_{j,k} = \frac{1}{n}\sum_{p=0}^{n-1} a_p \sum_{q=1}^{n-p}\left(\cos\frac{\pi[j(p+q-\frac{1}{2})+k(q-\frac{1}{2})]}{n} + \begin{matrix}\text{a similar term with}\\ \text{the sign of } k \text{ reversed}\end{matrix}\right)$$
$$+ \frac{1}{n}\sum_{p=1}^{n-1} a_{-p}\sum_{q=1}^{n-p}\begin{pmatrix}\text{similar terms with}\\ j \text{ and } k \text{ swapped}\end{pmatrix}$$
$$= \sum_{p=0}^{n-1} a_p\left(\frac{\cos[\frac{1}{2}\pi\{p(j-k)+n(j+k)\}/n]\sin\{\frac{1}{2}\pi(j+k)(n-p)/n\}}{n\sin\{\frac{1}{2}\pi(j+k)/n\}}\right.$$
$$\left. + \begin{matrix}\text{a similar term with}\\ \text{the sign of } k \text{ reversed}\end{matrix}\right) + \sum_{p=1}^{n-1} a_{-p}\begin{pmatrix}\text{similar terms with}\\ j \text{ and } k \text{ swapped}\end{pmatrix}$$

Now suppose $j+k$ is even (so $j-k$ is also even). Then

$$(TAT')_{j,k} = -\sum_{p=0}^{n-1} a_p\left(\frac{\cos\{\frac{1}{2}\pi p(j-k)/n\}\sin\{\frac{1}{2}\pi p(j+k)/n\}}{n\sin\{\frac{1}{2}\pi(j+k)/n\}}\right.$$
$$\left. + \begin{matrix}\text{a similar term with}\\ \text{the sign of } k \text{ reversed}\end{matrix}\right) - \sum_{p=1}^{n-1} a_{-p}\begin{pmatrix}\text{similar terms with}\\ j \text{ and } k \text{ swapped}\end{pmatrix}$$
$$= -\sum_{p=-n+1}^{n-1} a_p\left(\frac{\cos\{\frac{1}{2}\pi|p|(j-k)/n\}\sin\{\frac{1}{2}\pi|p|(j+k)/n\}}{n\sin\{\frac{1}{2}\pi(j+k)/n\}}\right.$$
$$\left. + \frac{\cos\{\frac{1}{2}\pi|p|(j+k)/n\}\sin\{\frac{1}{2}\pi|p|(j-k)/n\}}{n\sin\{\frac{1}{2}\pi(j-k)/n\}}\right). \qquad \text{(A.17)}$$

Hence

$$|(TAT')_{j,k}| \le \frac{2\pi}{n}\sum_{p=1}^{n-1} p|a_p| \to 0 \qquad (n\to\infty).$$

Now suppose $j+k$ is odd. Then

$$(TAT')_{j,k} = -\sum_{p=0}^{n-1} a_p \Bigg(\frac{\sin\{\frac{1}{2}\pi p(j-k)/n\}\cos\{\frac{1}{2}\pi p(j+k)/n\}}{n\sin\{\frac{1}{2}\pi(j+k)/n\}} + \begin{matrix}\text{a similar term with}\\ \text{the sign of } k \text{ reversed}\end{matrix} \Bigg) - \sum_{p=1}^{n-1} a_{-p} \left(\begin{matrix}\text{similar terms with}\\ j \text{ and } k \text{ swapped}\end{matrix} \right)$$
$$= 0 \tag{A.18}$$

since $a_p = a_{-p}$ by assumption.

When $j=n$ or $k=n$ use the preceding arguments with $j=0$ or $k=0$. □

Now suppose A is the variance–covariance matrix of a first order autoregressive process: $X_{i+1} = \alpha X_i + \epsilon_i$ where the ϵ_i are uncorrelated and have unit variance.

Theorem A.6. *In the notation of Theorem* A.5 *suppose* $A_{p,q} = \alpha^{|p-q|}/(1-\alpha^2)$. *Then*

$$\Lambda_{j,j} = 1/\{(1-\alpha)^2 + 4\alpha\sin^2(\tfrac{1}{2}\pi j/n)\} \qquad (j = 1,\dots,n-1),$$

and for $j = 1,\dots,n-1$, $k = 1,\dots,n-1$,

$$(TAT'-\Lambda)_{j,k} = \begin{cases} 0 & \text{if } j+k \text{ is odd,} \\ -\dfrac{4\alpha(1-\alpha)\{1-\alpha^n(-1)^j\}\cos(\frac{1}{2}\pi j/n)\cos(\frac{1}{2}\pi k/n)\Lambda_{j,j}\Lambda_{k,k}}{n(1+\alpha)} & \text{if } j+k \text{ is even.} \end{cases}$$

Proof. First consider the case $j=k$. Start from (A.16), and note the identities

$$\sum_{p=1}^{\infty} p\alpha^p \cos 2p\theta = \frac{2\alpha(1-\alpha)^2\cos^2\theta}{[(1-\alpha)^2+4\alpha\sin^2\theta]^2} - \frac{\alpha}{(1-\alpha)^2+4\alpha\sin^2\theta},$$
$$\sum_{p=0}^{\infty} \alpha^p \sin 2p\theta = \frac{\alpha\sin 2\theta}{(1-\alpha)^2+4\alpha\sin^2\theta}.$$

Then $\frac{1}{2}n(1-\alpha^2)(TAT'-\Lambda)_{j,j}$ equals

$$-\left\{\sum_{p=0}^{\infty}\alpha^p\left(p\cos\frac{\pi jp}{n} + \frac{\sin(\pi jp/n)}{\sin(\pi j/n)}\right) - \sum_{p=n}^{\infty}\alpha^p\left((p-n)\cos\frac{\pi jp}{n} + \frac{\sin(\pi jp/n)}{\sin(\pi j/n)}\right)\right\}$$
$$= -[1-\alpha^n(-1)^j]\sum_{p=0}^{\infty}\alpha^p\left(p\cos\frac{\pi jp}{n} + \frac{\sin(\pi jp/n)}{\sin(\pi j/n)}\right)$$

and the result follows.

Now suppose $j \neq k$. If $j+k$ is odd the result follows from (A.18). Suppose $j+k$ is even. Starting from (A.17), the result follows from the identity

$$\sum_{p=1}^{n-1}\alpha^p\left(\frac{\cos\{\frac{1}{2}\pi p(j-k)/n\}\sin\{\frac{1}{2}\pi p(j+k)/n\}}{\sin\{\frac{1}{2}\pi(j+k)/n\}} + \frac{\cos\{\frac{1}{2}\pi p(j+k)/n\}\sin\{\frac{1}{2}\pi p(j-k)/n\}}{\sin\{\frac{1}{2}\pi(j-k)/n\}}\right)$$
$$= 2\alpha(1-\alpha)^2\{1-\alpha^n(-1)^j\}\cos(\tfrac{1}{2}\pi j/n)\cos(\tfrac{1}{2}\pi k/n)/(\Lambda_{j,j}\Lambda_{k,k}).$$ □

The next theorem extends Theorem A.5 to integrated processes.

Theorem A.7. *Suppose $Z_1, \ldots, Z_n$ is a stationary process with auto-covariances given by $\mathrm{cov}(Z_p, Z_q) = c_{|p-q|}$ and $X_1, \ldots, X_n$ is the series of cumulative sums*

$$X_p = \sum_{j=1}^{p} Z_j.$$

Define the spectrum of the Z_j as in (5). *Let T_0 be as in Theorem* A.3, *Σ denote the variance–covariance matrix of $X_1, \ldots, X_n$ and Ψ denote the $(n-1) \times (n-1)$ diagonal matrix with the (j,j)-th element given by*

$$\Psi_{j,j} = \lambda_j / \{4 \sin^2(\tfrac{1}{2}\pi j/n)\}. \tag{A.19}$$

Suppose

$$\sum_{1}^{\infty} k^2 |c_k| < \infty. \tag{A.20}$$

Then

$$\lim_{n\to\infty} n^{-1/2} \|T_0 \Sigma T_0' - \Psi\|_F = 0, \qquad \lim_{n\to\infty} \max_{j,k} |(T_0 \Sigma T_0' - \Psi)_{j,k}| = 0.$$

Proof. Express the elements of Σ in terms of the c_k:

$$\Sigma_{p,q} = \sum_{j=1}^{p} \sum_{k=1}^{q} c_{|j-k|} = \min(p,q) \left\{ c_0 + 2 \sum_{j=1}^{\infty} c_j \right\} - \sum_{j=1}^{\infty} j \big(c_{j+|p-q|} - c_{j+p} - c_{j+q} + c_j \big). \tag{A.21}$$

Only the terms in $\min(p,q)$ and $c_{j+|p-q|}$ contribute to $T_0 \Sigma T_0'$. Hence $T_0 \Sigma T_0' = T_0 \Sigma_1 T_0' + T_0 \Sigma_2 T_0'$ where Σ_1 and Σ_2 have (p,q)-th terms reflecting these two contributions. From (A.10) $T_0 \Sigma_1 T_0'$ is exactly diagonal with jth diagonal element

$$\frac{c_0 + 2 \sum_{k=1}^{\infty} c_k}{4 \sin^2(\frac{1}{2}\pi j/n)}, \tag{A.22}$$

and $T_0 \Sigma_2 T_0'$ is approximately diagonal with the jth diagonal element given by (A.13) with

$$a_k = \sum_{j=1}^{\infty} j c_{j+k} \tag{A.23}$$

and the approximation defined by (A.14) or (A.15). Combining (A.13) and (A.23) gives the approximation to the jth diagonal element of $T_0 \Sigma_2 T_0'$ as

$$\sum_{k=1}^{\infty} c_k \frac{\cos(\pi j k/n) - 1}{2 \sin^2(\frac{1}{2}\pi j/n)} \tag{A.24}$$

Summing (A.22) and (A.24) leads to (A.19). □

A.4. Locally optimal and beta-optimal tests

The tests derived in this paper are locally optimal or beta optimal. This section summarises some of the theory. Suppose we wish to test the hypothesis $\xi = 0$ against the alternative $\xi > 0$. Locally optimal tests maximise the slope of the power function at $\xi = 0$. Beta-optimal tests (also known as point-optimal tests) minimise the value of ξ for which a preassigned power (say 80%) is achieved, see Davies (1969) and King (1988).

Suppose $p_\xi(X)$ is the probability density of the vector of observations, X, and suppose $p_\xi(X)$ is differentiable at 0 in the sense that

$$\lim_{\xi\to 0}\int \left| \frac{\partial p_\eta(X)}{\partial \eta}\Big|_{\eta=0} - \frac{p_\xi(X)-p_0(X)}{\xi}\right| \nu(dX) = 0, \qquad (A.25)$$

where ν is the measure with respect to which the p_ξ are defined. Then the locally optimal test has critical region

$$\left\{ \frac{\partial \log p_\xi(X)}{\partial \xi}\Big|_{\xi=0} > c \right\},$$

where c is chosen to give the desired significance level.

Suppose one wants a locally optimal test based on only part of the data, Y, with density, q_ξ. Then one can use the formula

$$\frac{\partial \log q_\xi(Y)}{\partial \xi}\Big|_{\xi=0} = E_0\left[\frac{\partial \log p_\xi(X)}{\partial \xi}\Big|_{\xi=0} \,\Big|\, Y \right] \qquad (A.26)$$

which follows from (A.5) provided (A.25) is true.

Beta-optimal tests usually have critical region of the form $\{p_\xi(X)/p_0(X) > c\}$ where c and ξ be found by solving

$$P_0\{p_\xi(X)/p_0(X) > c\} = \alpha, \quad P_\xi\{p_\xi(X)/p_0(X) > c\} = \beta. \qquad (A.27)$$

The parameters α and β are the significance level and the power at which we are seeking optimality and P_ξ is the probability measure corresponding to p_ξ. See Davies (1969) for details. If we want beta-optimal tests based on only part of the data then the likelihood ratios should be calculated using (A.5).

Now suppose X has an n-dimensional multivariate normal distribution with $EX_i = 0$ and variance–covariance matrix Σ_ξ and we wish to test $\xi = 0$ against $\xi > 0$.

For a scale-invariant test, the likelihoods under the hypothesis and the alternative can be found from (A.3). The beta-optimal test has critical region

$$\{X'(\Sigma_0^{-1} - c\Sigma_\xi^{-1})X > 0\}. \qquad (A.28)$$

The parameters c and ξ are such that

$$\Pr\{Z'(I_n - c\Lambda_\xi^{-1})Z > 0\} = \alpha, \quad \Pr\{Z'(\Lambda_\xi - cI)Z > 0\} = \beta \qquad (A.29)$$

where Z is an n-dimensional vector of independent standard normal random variables and Λ_ξ is a diagonal matrix with diagonal elements being the eigenvalues of $\Sigma_0^{-1/2}\Sigma_\xi\Sigma_0^{-1/2}$. The probabilities (A.29) can be calculated using, for example, the program of Davies (1980). Fortran and C versions of this program are on the author's website (Davies, 2000).

The locally optimal test has critical region

$$\left\{X'\Sigma_0^{-1}\left(\frac{\partial\,\Sigma_\xi}{\partial\xi}\Big|_{\xi=0}\right)\Sigma_0^{-1}X > cX'\Sigma_0^{-1}X\right\}. \tag{A.30}$$

References

Davies, R. B. (1969). Beta-optimal tests and an application to the summary evaluation of experiments. *J. Roy. Statist. Soc. Ser. B* **31**, 524–538.

Davies, R. B. (1973). Asymptotic inference in stationary Gaussian time-series. *Adv. Appl. Probab.* **5**, 469–497.

Davies, R. B. (1980). The distribution of a linear combination of chi-squared random variables. Algorithm AS155. *Appl. Statist.* **29**, 323–333.

Davies, R. B. (1983). Optimal inference in the frequency domain. In *Handbook of Statistics, Vol 3.* ed. D. R. Brillinger and P. R. Krishnaiah, 73–92. Elsevier, Amsterdam.

Davies, R. B. (1985). Asymptotic inference when the amount of information is random. In *Proceedings of the Berkeley Conference in the Honor of Jerzy Neyman and Jack Kiefer* **2**, ed. L. M. LeCam and R. A. Olshen, 841–864. Wadsworth, Belmont CA.

Davies, R. B. (2000). Linear combination of chi-squared random variables. <http://www.statsresearch.co.nz/robert/QF.htm>.

Dickey, D. A., Bell, W. R. and Miller, R. B. (1986). Unit Roots in Time Series Models: Tests and Implications. *Amer. Statistician* **40**, 12–26.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press, Cambridge.

King, M. L. (1988). Towards a theory of point optimal testing. *Econometric Reviews* **6**, 169–218.

Rao, K. R. and Yip, P. (1990). *Discrete Cosine Transform: Algorithms, Advantages, Applications.* Academic Press, New York.

Van Loan, C. (1992). *Computational Frameworks for the Fast Fourier Transform.* SIAM, Philadelphia.