N73-31552
CR-133986

ICSA

INSTITUTE FOR COMPUTER SERVICES AND APPLICATIONS
RICE UNIVERSITY

ICSA - 275-025-004

POWER SPECTRAL DENSITY ESTIMATION BY SPLINE SMOOTHING IN THE FREQUENCY DOMAIN

BY

R. J. P. de Figueiredo
Mathematics Research Center
The University of Wisconsin

and

J. R. Thompson
Dept. of Mathematical Sciences
Rice University

ABSTRACT

An approach, based on a global averaging procedure, is presented for estimating the power spectrum of a second order stationary zero-mean ergodic stochastic process from a finite length record. This estimate is derived by smoothing, with a cubic smoothing spline, the naive estimate of the spectrum obtained by applying FFT techniques to the raw data. By means of digital computer simulated results, a comparison is made between the features of the present approach and those of more classical techniques of spectral estimation.

Institute for Computer Services & Applications

Rice University

Houston, Texas 77001

September 1972

Research supported under NASA Contract #NAS9-12776

INTRODUCTION

Let $\{X_t :- \infty < t < \infty\}$ denote a second order stationary zero-average ergodic real-valued stochastic process. If the process is Gaussian, then it is completely characterized by the autocovariance function

$$C(\tau) = E[X(t)\, X(t+\tau)] = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x(t+\tau) dt, \qquad (1.\ 1)$$

or, equivalently, by the Fourier transform of $C(\tau)$, referred to as the power spectrum,

$$P(f) = \int_{-\infty}^{\infty} C(\tau)\, e^{-i2\pi f\tau} d\tau$$

$$= \lim_{T\to\infty} \frac{1}{T} \left| \int_{-T/2}^{T/2} x(t)\, e^{-i2\pi ft} dt \right|^2 . \qquad (1.\ 2)$$

Above and henceforth, we denote by $x(t)$ or x_t a realization of $X(t)$ or X_t.

If it is desired to estimate $P(f)$ on the basis of a record of finite length T_n, we may compute the apparent autocovariance function

$$C_{00}(\tau) = \frac{1}{T_n - |\tau|} \int_{-b}^{b} x(t - {}^{\tau}/_{2})\, x\,(t + {}^{\tau}/_{2})\, dt \qquad (1.\ 3)$$

where $b = (T_n - |\tau| / _2)$

and take the Fourier transform of $\tilde{C}_{00}(\tau) = C_{00}(\tau)$ for $|\tau| < T_m < T_n$, $= 0$ otherwise, giving

$$\tilde{P}_{00}(f) = \int_{-\infty}^{\infty} \tilde{C}_{00}(\tau)\, e^{-i2\pi f\tau} d\tau . \qquad (1.\ 4)$$

Unfortunately, in general $\tilde{P}_{00}(f)$ does not converge to $P(f)$, in the L_2-norm, as the length of the time series becomes infinite, i.e., $\tilde{P}_{00}(f)$ is not a consistent estimator for $P(f)$.

The Tukey family of filters(1) remedies this difficulty by an averaging process

$$C_A(\tau) = \left(a + (1-a)\cos\frac{\pi\tau}{T_m}\right) C_{00}(\tau)$$
$$\text{for } |\tau| \le T_m < T_n ,$$
$$= 0 \text{ if } |\tau| > T_m , \qquad (1.\ 5)$$

where $0 < a < 1$, giving

$$P_A(\tau) = \int_{-\infty}^{\infty} C_A(\tau)\, e^{-i2\pi f\tau} d\tau . \qquad (1.\ 6)$$

For the case where the record $\{x(t): t \in (0, T_n)\}$ is sliced at n intervals of length Δt, we first compute the mean lagged products

$$C_r = \frac{1}{n-r} \sum_{q=0}^{q=n-r} x_q\, x_{q+rh} \qquad (1.\ 7)$$

for $r = 0, 1, 2, \ldots, m$ where $m < n$. We then apply a discrete finite cosine series transform to $C_0, \ldots, C_m$ obtaining

$$V_r = \Delta t \left(C_0 + 2 \sum_{q=1}^{m-1} C_q \cos\frac{qr\pi}{m} + C_m \cos r\pi \right). \qquad (1.\ 8)$$

The Tukey filter then gives as estimates of the spectral density for the case where

$$a = \frac{1}{2} \quad \text{(hanning)}$$

$$U_0 = .5V_0 + .5V_1$$

$$U_r = .25V_{r-1} + .5V_r + .25V_{r+1}, \; 1 \le r \le m-1$$

$$U_m = .5V_{m-1} + .5V_m \, . \qquad (1.\ 9)$$

Basically, then, the popular Tukey filter is a local averaging procedure. In this paper we present, as an alternative approach, an estimation procedure based on a global averaging approach. The results given here are of a preliminary nature and are now in the process of being studied in greater detail.

2. DISCUSSION AND RESULTS

Consider the discrete white noise time series

$$X_t = \epsilon_t \; ; \; t = 1, 2, \; . \; . \; . \; , \; n \; , \qquad (2.\ 1)$$

where the ϵ_t are independent Gaussian random variables with mean zero and variance unity. A naive estimate of the power spectrum is given by

$$P_N(f) = \frac{1}{n} \left| \sum_{t=1}^{n} x_t e^{-i2\pi ft} \right|^2 \qquad (2.\ 2)$$

and may be quickly computed by Fast Fourier Transform techniques[(2)]. A plot of this estimate is given for a time series of length $n = 512$ in Figure 1. The inconsistency of this estimate is apparent. We shall attempt to remedy this inconsistency by smoothing $P_N(f)$.

Let us divide the frequency interval [0, .5cps] into k-1 equal subintervals. We seek the solution $P_s(f)$ which minimizes , $\forall\, P_s \epsilon\, C^1$ [0, .5] and piecewise continuous P_s'' (See (3)),

$$\sum_{j=0}^{k-1} a_j \left(P_s\left(\frac{j}{2(k-1)}\right) - P_N\left(\frac{j}{2(k-1)}\right) \right)^2 + g \int_0^{.5} (P_s''(f))^2 df \, , \qquad (2.\ 3)$$

where g is a nonnegative constant which can be selected so as to give appropriate weight to the roughness penalty term, and the a_j are positive constants summing to unity used for weighting the "miss distances" in the first term. In the ensuing discussion we shall take

$$a_j = \frac{1}{k} \, , \quad \text{for } j = 0, 1, 2, \ldots, k-1 \, .$$

Schoenberg has shown[5] that the solution P_s is, in fact, a cubic spline with knots at the subinterval endpoints. The smoothing algorithm employed here is that of Greville[3].

In Figure 2, we show $P_s(f)$ plots for $k = 8$ and $n = 128$ and 512 together with the true power spectrum. We note a dramatic change from the naive estimate and a good approximation to the actual spectral density.

Next, let us consider a spectral estimation problem from Jenkins and Watts' Spectral Analysis and its Applications[4, pp. 263-265]. The spectral density $P(f)/\sigma_x^2$ of the first order autoregressive process

$$X_t = -.4X_{t-1} + \epsilon_t \, ; \; t = 1, 2, \ldots, n \, , \qquad (2.\ 4)$$

where the ϵ_t are independent Gaussian variates with zero mean and unit variance, is given by

$$P(f)/\sigma_x^2 = \frac{1.68}{1.16 + .8 \cos 2\pi f}, \; 0 \leq f \leq \frac{1}{2}. \qquad (2.5)$$

In Figure 3, we demonstrate estimates of the spectral density for series of lengths 128, 256 and 512 using hanning with $m = 32$. In Figure 4, we show $P_s(f)$ for the same time series, 16 knots and $g - 1$. The spline smoothing approach appears to compare favorably with hanning.

The effect of the number of knots is demonstrated in Figure 5 in the estimation of the spectral density of the autoregressive process in (2. 4) for a series of length 1024 using $g = 1$ and $k = 8, 16, 32$. In Figure 6, we demonstrate estimates using the same series with 16 knots and $g = 1, .1$, and .01. Clearly, the selection of g and k is an important consideration and should be based both on the length of the time series and prior feelings as to the shape of the spectral density. We have not yet had opportunity to investigate optimal (g, k) selection procedures but shall approach this problem in the future.

As a final example from Jenkins and Watts [4, pp. 268-272], we consider the second order autoregressive process

$$X_t = X_{t-1} - .5X_{t-2} + \epsilon_t \; ; \; t = 1, 2, \ldots, n, \qquad (2.6)$$

where the ϵ_t are independent Gaussian deviates with mean zero and unit variance. The spectral density here is given by

$$P(f)/\sigma_x^2 = \frac{.834}{2.25 - 3 \cos 2\pi f + \cos 4\pi f}, \; 0 \leq f \leq \frac{1}{2} \qquad (2.7)$$

In Figure 7, we show for a series of length 512, the hanning estimate with $m = 32$ and compare the result with the spline smoothing approach using $g = .01$ and 32 knots.

3. CONCLUSIONS

We have explored in a preliminary fashion the possibility of using spline smoothing techniques, in the frequency domain, for the estimation of power spectra. Although for the examples considered, this technique appears to compare favorably with one of the most popular smoothing techniques-hanning - we still need to study in detail the consistency properties of the proposed estimation procedure. We hope, in addition, to study optimal methods for applying spline smoothing to cross-spectra.

ACKNOWLEDGEMENT

We wish to thank Mr. J. Rod Grisham for programming all routines used in this paper.

REFERENCES

(1) Blackman, R. B. and Tukey, J. W., The Measurement of Power Spectra from the Point of View of Communications Engineering, Dover Publications, New York, 1958.

(2) Cooley, J. W. and Tukey, J. W., "An Algorithm for the Machine Calculation of Fourier Series," Math. Comput., V. 19, 1965 pp. 297-301.

(3) Greville, T. N. E., "Introduction to Spline Functions," Theory and Applications of Spline Functions, Academic Press, New York, 1969, pp. 1-36. .

(4) Jenkins, G. W. and Watts, D. G., Spectral Analysis and its Applications, Holden-Day, San Francisco, 1968.

(5) Schoenberg, I. J., "Spline Functions and the Problem of Graduation," Proc. Nat. Acad. Sci., 1964, pp. 947-950.

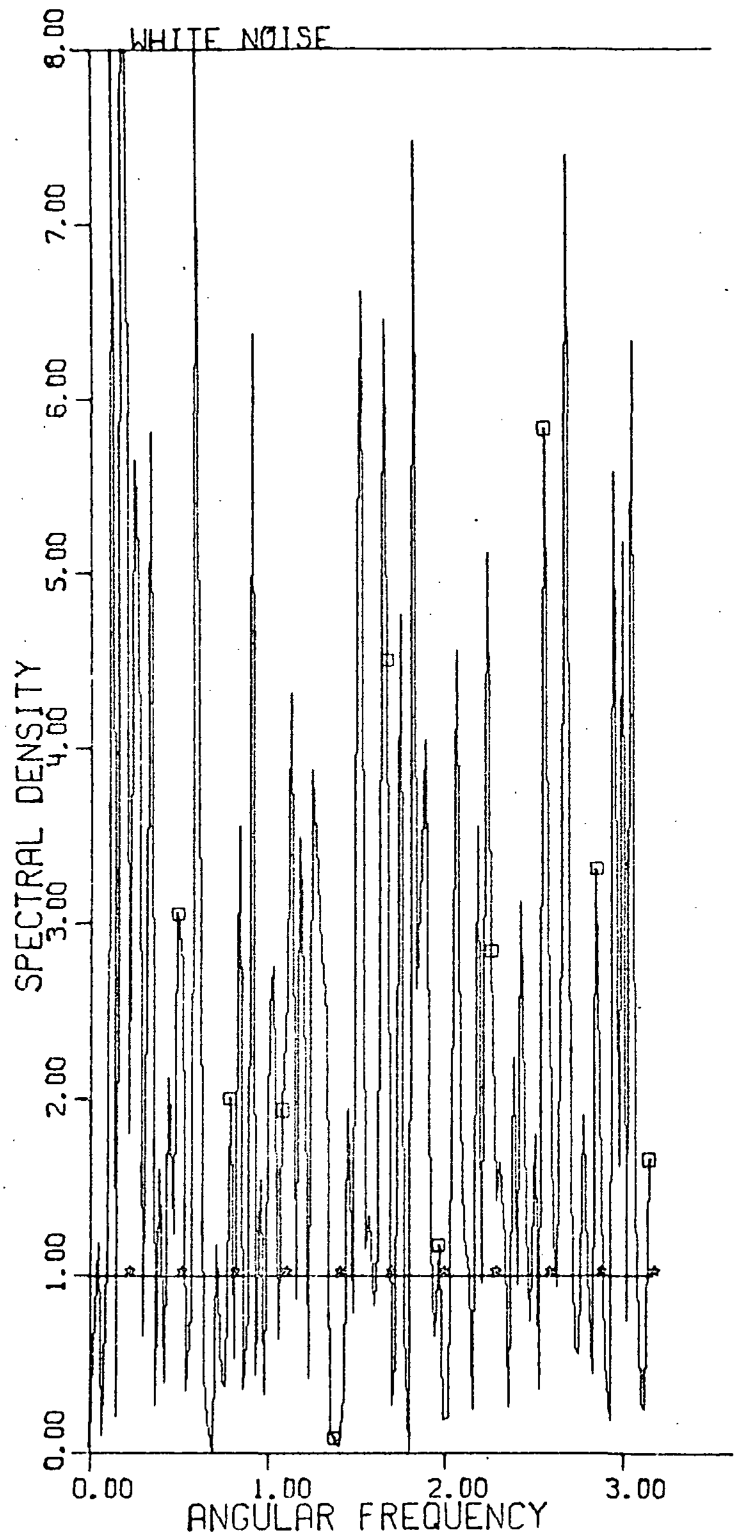

Fig. 1: Naive spectral density estimates for white noise.

* $P(\omega)$, □ $P_N(\omega)$, $\omega = 2\pi f$, $n = 512$

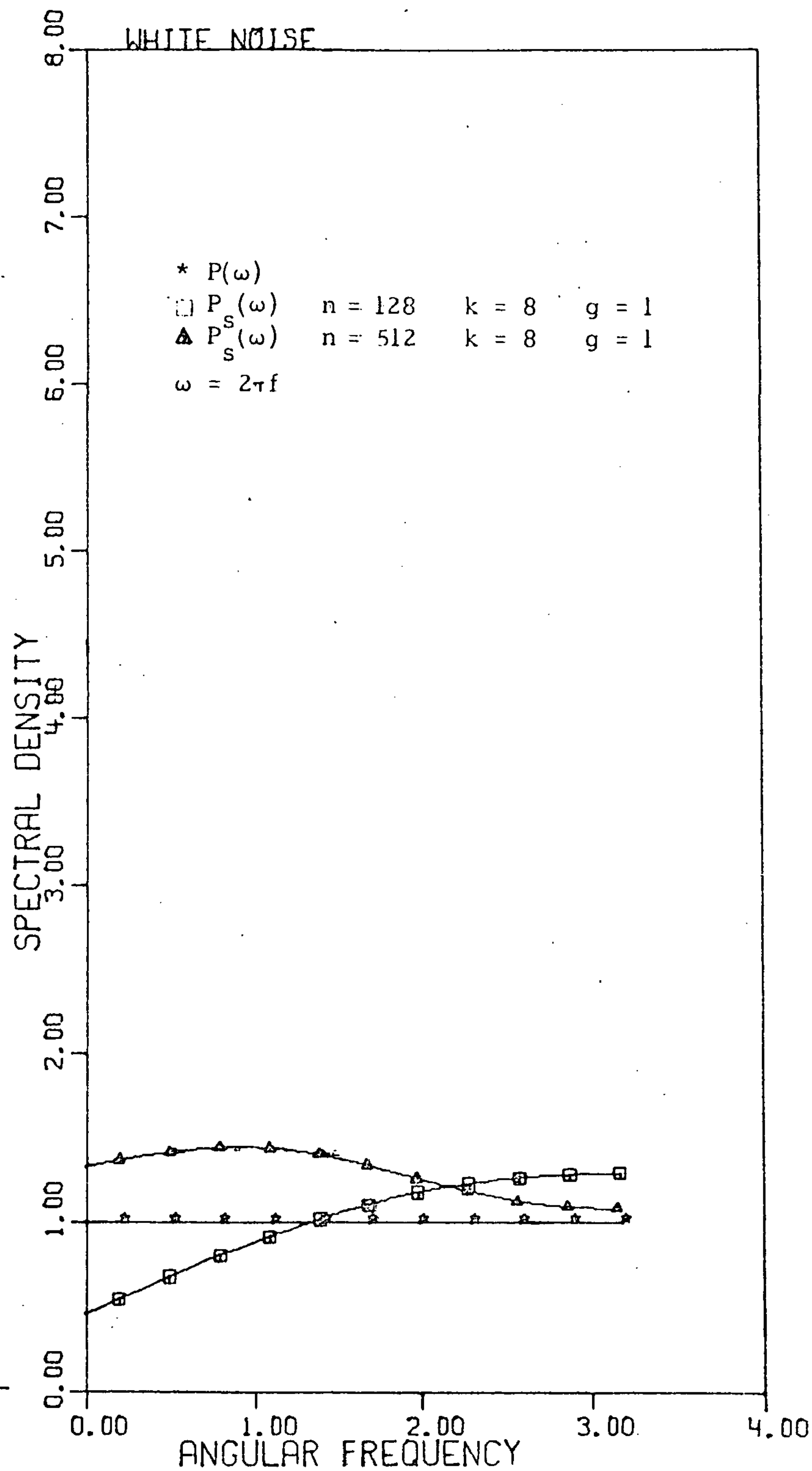

Fig. 2: Spline smoothed spectral density estimates for white noise.

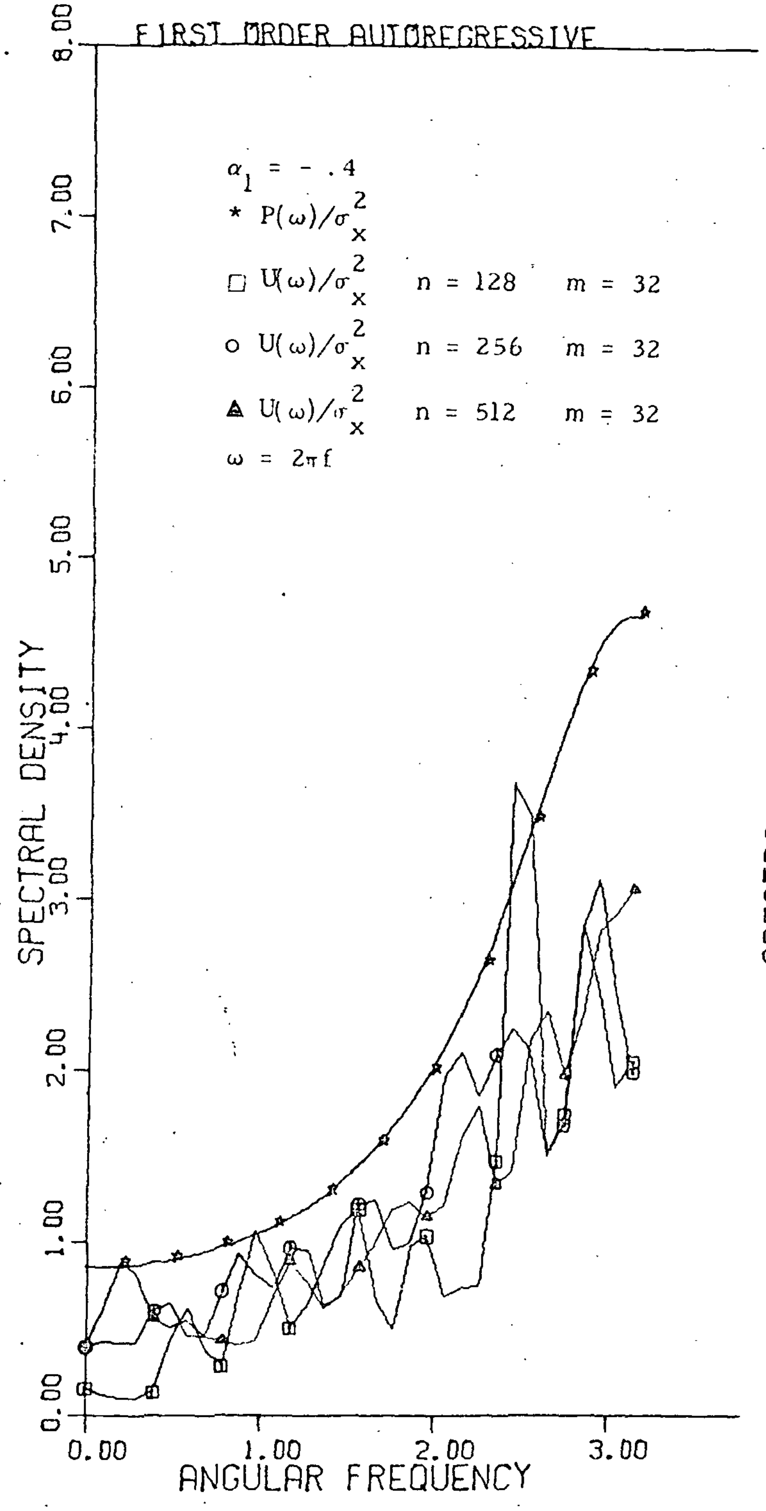

Fig. 3: Hanning smoothed spectral density estimates for a first-order ar process $X_t = \alpha_1 X_{t-1} + \varepsilon_t$.

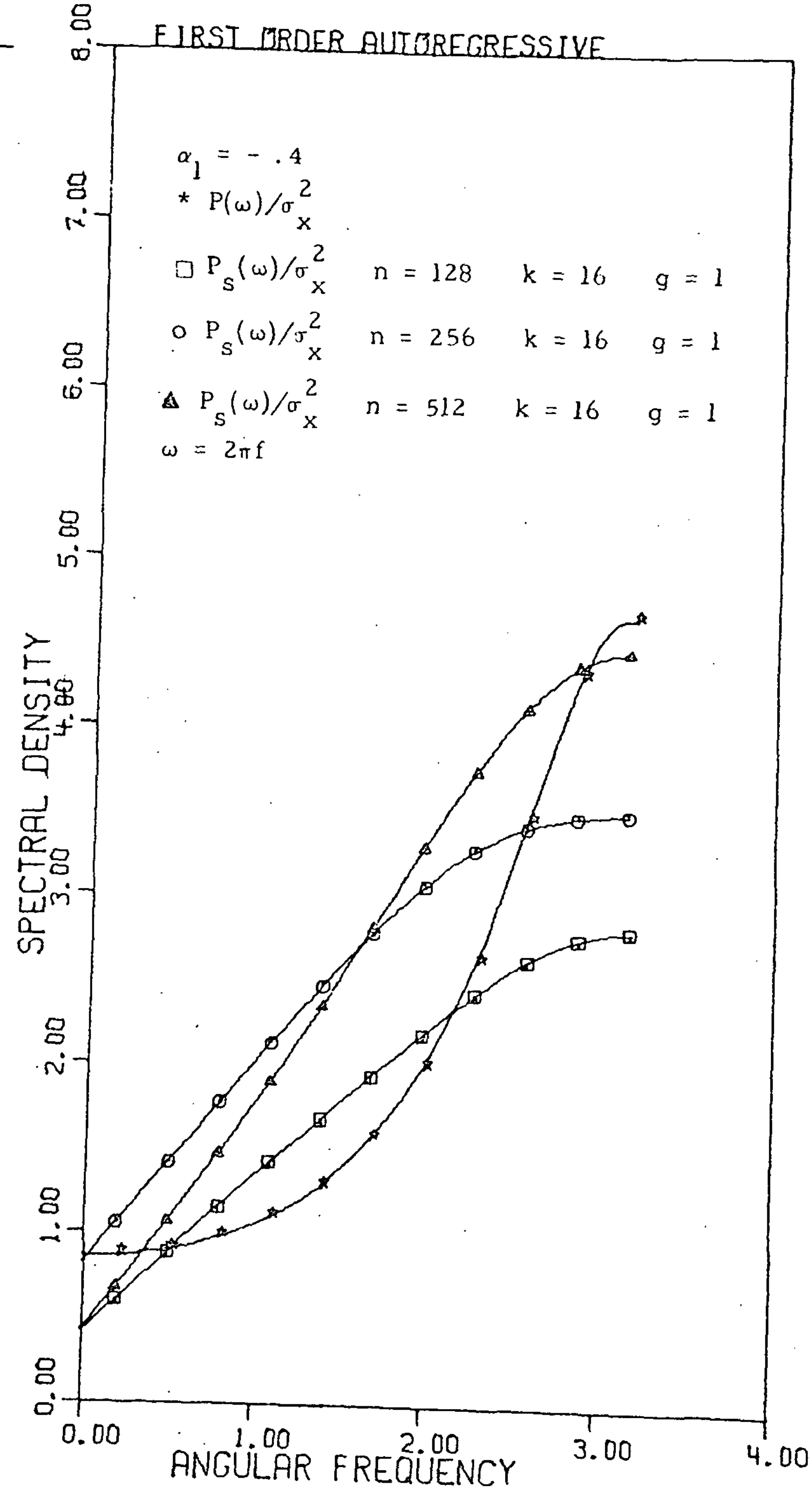

Fig. 4: Spline smoothed spectral density estimates for a first order ar process $X_t = \alpha_1 X_{t-1} + \varepsilon_t$.

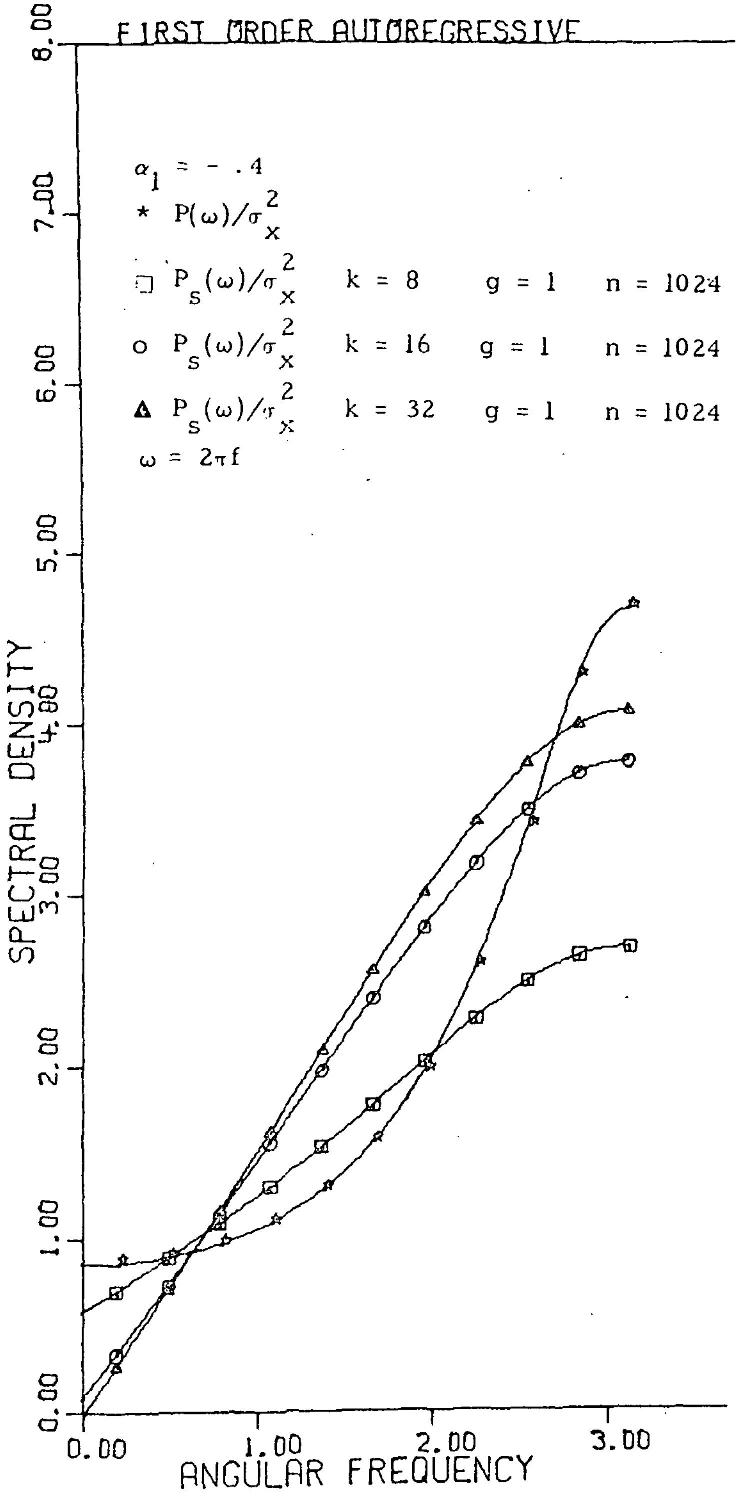

Fig. 5: Spline smoothed spectral density estimates for a first order ar process $X_t = \alpha_1 X_{t-1} + \varepsilon_t$ under different conditions on knots.

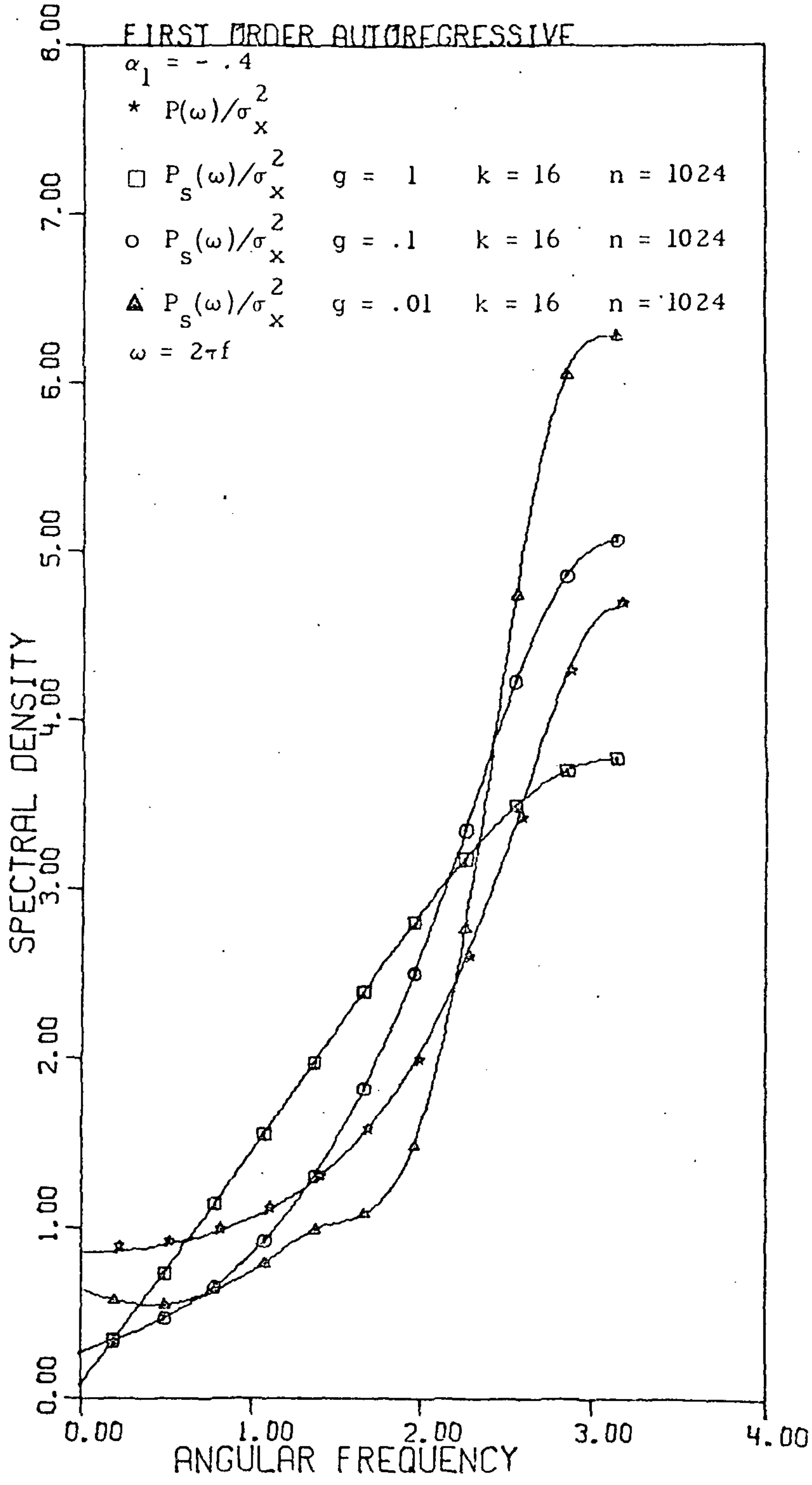

Fig. 6: Spline smoothed spectral density estimates of a first order ar process under different conditions on the smoothing parameter g.

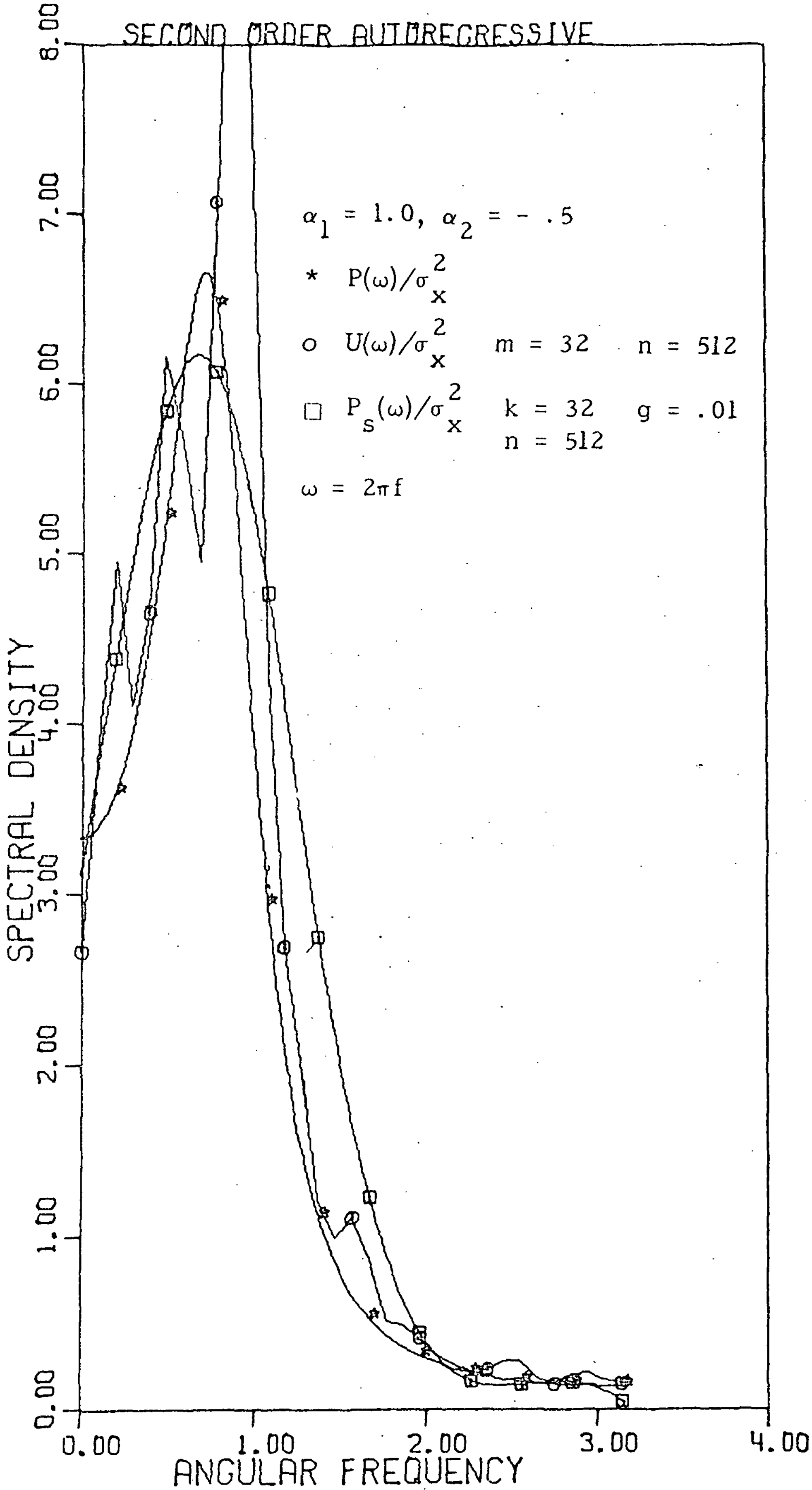

Fig. 7: Hanning and spline smoothed estimates of the spectral density of a second order ar process $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$.