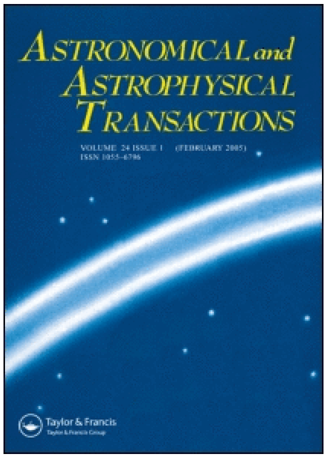

Astronomical & Astrophysical Transactions
The Journal of the Eurasian Astronomical Society
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713453505

Time series analysis of unequally spaced data: The statistical properties of the Schuster periodogram

V. V. Vityazev [a]
[a] Astronomy Department, St. Petersburg University, St. Petersburg, Russia

Online Publication Date: 01 November 1996
To cite this Article: Vityazev, V. V. (1996) 'Time series analysis of unequally spaced data: The statistical properties of the Schuster periodogram', Astronomical & Astrophysical Transactions, 11:2, 159 - 173
To link to this article: DOI: 10.1080/10556799608205462
URL: http://dx.doi.org/10.1080/10556799608205462

Astronomical and Astrophysical Transactions, 1996,
Vol. 11, pp. 159–173
Reprints available directly from the publisher
Photocopying permitted by license only

TIME SERIES ANALYSIS OF UNEQUALLY SPACED DATA: THE STATISTICAL PROPERTIES OF THE SCHUSTER PERIODOGRAM

V. V. VITYAZEV

Astronomy Department, St. Petersburg University, Bibliotechnaya pl. 2, Petrodvorets, St. Petersburg, 198904, Russia

(Received November 9, 1995)

It is generally believed that the classical Schuster periodogram must not be used in the problem of finding periodicities in the irregularly spaced data. This follows from the fact that if the data is pure noise, the statistical distribution of the Schuster periodogram is no longer exponential. From this point of view, the so-called LS-spectra based on the least squares fitting of a sine function to the data are recommended. Nevertheless, in many situations the Schuster periodograms and the LS-spectra are close to being identical. The paper presents a comparative study of the Schuster periodogram and the LS-spectra with respect to their statistical properties. The analytical expression for the probability distribution of the Schuster periodogram when the time series is assumed to be unevenly spaced pure noise is found. It is shown that the probability distribution deviates from the exponential law only at the frequencies ω_j that satisfy the condition $1-W(2\omega_j) \ll 1$, where $W(\omega)$ is the spectral window. The examples of the time points distributions yielding such pathology are given.

KEY WORDS Time series, power spectra

1 INTRODUCTION

Evaluation of the power spectra of unevenly spaced time series is a very important problem especially in astronomy, where irregular observations are often unavoidable. In the task of finding periodicities hidden in the observed data, the calculation of the power spectrum is a useful step which permits us to see the concentrations of power (the spectral peaks) at certain frequencies. All observations are accompanied with noise. To detect a signal in the noisy data, one must know the images of the signal and of the noise in the frequency domain. For regular time series, these images are calculated with the help of the Schuster periodogram (Schuster, 1898). The reason to introduce the Schuster periodogram is that it gives the correlation between the data and a sine function. The properties of this periodogram are well known

(Jenkins and Watts, 1968; Otnes and Enocson, 1978; Marple, 1987; Terebizh, 1992, etc.). In particular, the reflex of a sine wave is described by the spectral window, whereas the reflex of the pure noise is a random variable distributed according to the exponential law. The application of the Schuster periodogram to irregular time series was made by Deeming (1975a, 1975b), Roberts *et al.* (1987). At the same time, Lomb (1976), Scargle (1982), Terebizh (1992) pointed out that the exponential law can not be justified for the periodogram of pure noise when the time series is irregular. To remedy this, Lomb (1976), Ferraz–Mello (1981), Scargle (1982, 1989) proposed alternative procedures to evaluate the power spectra of the time series with missing points. Their approaches are based on the least squares fitting of a sine function to the data (Barning, 1962). The resulting estimators of the power spectrum – the so-called LS-spectra- retain the exponential law, and this is the main reason why the LS-spectra are widely used nowadays. At the same time, the LS-spectra lose several important properties: they cannot be described in terms of the spectral window, they are not rigorously connected with the correlation function, and so on. Moreover, the practice of spectral evaluation shows that in many situations the Schuster periodogram and the LS-spectra are almost identical despite the different theoretical foundations.

In the previous paper (Vityazev, 1996), the intercomparison between the Schuster periodogram and the LS-spectra was made with the goal to find the situations when the said estimators must coincide or be different. It was found that the likeness between them is governed by the properties of the corresponding spectral window $W(\omega)$. The main result is: if the frequency ω_0 of the signal sutisfies the condition $W(2\omega_0) = 0$, then the Schuster periodogram and the LS-spectra are identical, otherwise they are different. Naturally, this simple condition, valid for a pure signal without noise, forces us to continue the intercomparison between the Schuster periodogram and the LS-spectra, this time regarding the time series to be a pure noise – and this is the main purpose of the present paper. In Section 2 the generalised Π-periodogram is introduced and its statistical properties are studied. This allows to clarify the probability distributions and the correlations between various values of the Schuster periodogram and of the LS-spectra. (Sections 3 and 4). The significance tests for all the periodograms are considered in Section 5. The application of the theory to simulated time series is given in Section 6.

2 THE Π-PERIODOGRAM

Assume that the time series x_k, $k = 1, 2, \ldots N$ given at arbitrary set of time points t_k constitutes a random sample from a normally distributed population with zero mean and variance σ_0^2. For this case we have

$$\langle x_p x_q \rangle = \begin{cases} \sigma_0^2, & p = q \\ 0, & p \neq q \end{cases}. \tag{2.1}$$

In order to study various periodograms simulteneously and not to repeat similar calculations, we introduce the generalized Π-periodogram using the following notation:

$$\Pi(\omega) = a^2(\omega)(x,\phi_1)^2 + b^2(\omega)(x,\phi_2)^2, \tag{2.2}$$

where

$$\phi_1(t) = \cos\omega\bar{t}_k, \quad \phi_2(t) = \sin\omega\bar{t}_k, \tag{2.3}$$

$$\bar{t}_k = t_k - \tau(\omega), \tag{2.4}$$

$$\tau(\omega) = \frac{1}{2\omega}\arctan\frac{\Sigma_k \sin 2\omega t_k}{\Sigma_k \cos 2\omega t_k}, \tag{2.5}$$

$$(p,q) = \frac{1}{N}\sum_{k=1}^{N} p(t_k)q(t_k), \quad ||p||^2 = (p,p). \tag{2.6}$$

The reason for using the Π-periodogram is that if we set

$$a(\omega) = 1, \quad b(\omega) = 1, \tag{2.7}$$

then Eq. (2.2) yields the Schuster periodogram:

$$S(\omega) = (x,\phi_1)^2 + (x,\phi_2)^2 = \frac{1}{N^2}|\sum_{k=1}^{N} x_k e^{-i\omega t_k}|^2, \tag{2.8}$$

while with

$$a^{-2} = 2||\phi_1||^2, \quad b^{-2} = 2||\phi_2||^2 \tag{2.9}$$

from Eq. (2.2) we get the Lomb periodogram (Lomb, 1976):

$$L(\omega) = \frac{1}{2}\left[\frac{(x,\phi_1)^2}{||\phi_1||^2} + \frac{(x,\phi_2)^2}{||\phi_2||^2}\right]. \tag{2.10}$$

It was shown (Vityazev, 1996) that the Lomb periodogram is identical to the Barning periodogram (Barning, 1962), and the former coincides with the periodogram based on the orthogonalization of the initial functions $\cos\omega t_k$ and $\sin\omega t_k$ by means of the Gram–Schmidt procedure (Ferraz–Mello, 1981). Thus we see that with the help of Eqs. (2.7) and (2.9) the generalized Π-periodogram gives us either the Schuster periodogram or the LS-spectra. It is important to emphasize that for all the periodograms under consideration one has:

$$(\phi_1,\phi_2) = 0. \tag{2.11}$$

Now we are going to answer two questions: what is the probability distribution of the Π-periodogram and how its values are correlated. In order to do this, we shall use some properties of random variables known in the theory of probability (Ventsel, 1964). Given are two random variables x and y with expectations m_x and m_y and variancies σ_x^2, σ_y^2, respectively. Defining their correlation coefficient as

$$r = \frac{\langle(x-m_x)\rangle\langle(y-m_y)\rangle}{\sigma_x\sigma_y}, \tag{2.12}$$

we assume that the bivariate normal probability function for x and y is

$$f(x,y) = f(x)f(y)\exp\left[\frac{r^2}{1-r^2}\right], \tag{2.13}$$

where $f(x)$ and $f(y)$ are the normal probability functions:

$$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}}\exp\left[-\frac{(x-m_x)^2}{2\sigma_x^2}\right], \tag{2.14}$$

$$f(x) = \frac{1}{\sigma_y\sqrt{2\pi}}\exp\left[-\frac{(y-m_y)^2}{2\sigma_y^2}\right]. \tag{2.15}$$

It is known that two random variables x and y are independent if

$$f(x,y) = f(x)f(y). \tag{2.16}$$

The two variables are said to be *uncorrelated* if

$$r = 0. \tag{2.17}$$

In general, the absence of correlation follows from independence, but the opposite conclusion is false. In particular, when x and y are normally distributed, independence and absence of correlation are all the same.

Now, if x and y are independent and normally distributed, then the variable $z = x^2 + y^2$ is distributed according to (Papoulis, 1965; see also Scargle, 1982):

$$p(z) = \frac{1}{2\sigma_x\sigma_y}\exp\left[-\frac{z}{4}\frac{\sigma_x^2+\sigma_y^2}{\sigma_x^2\sigma_y^2}\right] I_0\left[-\frac{z}{4}\frac{\sigma_x^2-\sigma_y^2}{\sigma_x^2\sigma_y^2}\right], \tag{2.18}$$

where I_0 is the modified Bessel function of the first kind.

Return now to the Π-periodogram, rewriting it in the form

$$\Pi(\omega) = X^2(\omega) + Y^2(\omega), \tag{2.19}$$

where

$$X(\omega) = a(\omega)(x,\phi_1), \quad Y(\omega) = b(\omega)(x,\phi_2). \tag{2.20}$$

It is not difficult to justify that $X(\omega)$ and $Y(\omega)$ are normally distributed random variables for which, from (2.1), one can find:

$$\langle X\rangle = 0, \quad \langle Y\rangle = 0, \tag{2.21}$$

$$\sigma_x^2 = \langle X^2\rangle = a^2(\omega)\frac{\sigma_0^2}{N}\|\phi_1\|^2, \tag{2.22}$$

$$\sigma_y^2 = \langle Y^2\rangle = b^2(\omega)\frac{\sigma_0^2}{N}\|\phi_2\|^2. \tag{2.23}$$

For the correlation coefficient we have

$$r_{xy} = \frac{\langle XY \rangle}{\sigma_x \sigma_y} = 0, \tag{2.24}$$

since from Eq. (2.11) it follows that

$$\langle XY \rangle = a(\omega)b(\omega)\frac{\sigma_0^2}{N}(\phi_1, \phi_2) = 0. \tag{2.25}$$

Thus we see that the random variables $\Pi(\omega)$ are distributed according to Eq. (2.18).

For further study, we consider the correlation moment of two variables $\Pi_1 = \Pi(\omega_1)$ and $\Pi_2 = \Pi(\omega_2)$:

$$K(\omega_1, \omega_2) = \langle \Pi_1 \bar{\Pi} \rangle = \langle \Pi_1 \Pi_2 \rangle - \bar{\Pi}_1 \bar{\Pi}_2, \tag{2.26}$$

where the averaged Π-periodogram is

$$\begin{aligned} \bar{\Pi} &\equiv \langle \Pi(\omega) \rangle = \langle X^2 \rangle + \langle Y^2 \rangle \\ &= \frac{\sigma_0^2}{N}[a^2(\omega)\|\phi_1\|^2 + b^2(\omega)\|\phi_2\|^2]. \end{aligned} \tag{2.27}$$

Adopting notations $X_i = X(\omega_i)$ and $Y_i = Y(\omega_i)$, $i = 1, 2$, from Eq. (2.26) we get

$$K(\omega_1, \omega_2) = \langle X_1^2 X_2^2 \rangle + \langle Y_1^2 Y_2^2 \rangle + \langle X_1 Y_2^2 \rangle + \langle Y_1^2 X_2^2 \rangle - \bar{\Pi}_1 \bar{\Pi}_2. \tag{2.28}$$

For two random variables with the bivariate normal probability function (2.13) one has:

$$\langle x^2 y^2 \rangle = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^2 y^2 f(x, y)\, dx dy = \sigma_x^2 \sigma_y^2 (1 + 2r^2). \tag{2.29}$$

With this result, the final expression for the correlation moment becomes:

$$\begin{aligned} K(\omega_1, \omega_2) &= \frac{2\sigma_0^4}{N^2}[a^2(\omega_1)a^2(\omega_2)(\phi_1, \phi_1)_{12}^2 + b^2(\omega_1)b^2(\omega_2)(\phi_2, \phi_2)_{12}^2 \\ &+ a^2(\omega_1)b^2(\omega_2)(\phi_1, \phi_2)_{12}^2 + a^2(\omega_2)b^2(\omega_1)(\phi_2, \phi_1)_{12}^2, \end{aligned} \tag{2.30}$$

where

$$(\phi_i, \phi_j)_{mn} = \frac{1}{N}\sum_k \phi_i(\omega_m, t_k - \tau(\omega_m)),\ \phi_j(\omega_n, t_k - \tau(\omega_n))),\ i, j, m, n, = 1, 2. \tag{2.31}$$

Now, the correlation coefficient between the values of the Π-periodogram evaluated at two frequencies ω_1 and ω_2 is

$$k(\omega_1, \omega_2) = \frac{K(\omega_1, \omega_2)}{\sqrt{K(\omega_1, \omega_1)K(\omega_2, \omega_2)}}. \tag{2.32}$$

The general properties derived here for the Π-periodogram permit us to deduce the corresponding properties of specific periodograms.

3 THE LS-SPECTRA

From Eqs. (2.9), (2.2) and (2.23) we find:

$$\sigma_x^2 = \sigma_y^2 = \frac{\sigma_0^2}{2N} \equiv \sigma^2. \tag{3.1}$$

Analogously, for the averaged periodogram one has:

$$\bar{\Pi} = \frac{\sigma_0^2}{N}. \tag{3.2}$$

Now we see that, for the normalized Π-periodogram

$$\Pi_N(\omega) = \frac{\Pi(\omega)}{\sigma_0^2/N} \tag{3.3}$$

the distribution function is

$$p(z) = \exp(-z), \quad 0 < z < \infty, \quad \omega > 0. \tag{3.4}$$

As it was pointed out above, this result is valid for the periodograms of Barning, Lomb, and Ferraz–Mello.

The second result concerning the LS-spectra is: the correlation between the heights of a noise spectrum at frequencies ω_1 and ω_2 is equal to the mean height of the periodogram of a sine function of frequency ω_1 at frequency ω_2. This result follows from Eq. (2.30). For the functions ϕ_1 and ϕ_2, defined by Eqs. (2.3)–(2.5) this was found by Lomb (1976). Now we see that, with the corresponding meaning of ϕ_1 and ϕ_2, this result is the same for the Barning and the Ferraz–Mello periodograms. In other words, the values of the LS-spectra are correlated, whence they are not independent.

4 THE SCHUSTER PERIODOGRAM

This time, from Eqs. (2.7), (2.22) and (2.23) one has:

$$\sigma_x^2 = \frac{\sigma_0^2}{2N}[1 + \sqrt{W(2\omega)}], \tag{4.1}$$

$$\sigma_y^2 = \frac{\sigma_0^2}{2N}[-\sqrt{W(2\omega)}], \tag{4.2}$$

where the spectral window is defined as

$$W(\omega) = \frac{1}{N^2}|\sum_{k=1}^{N} e^{-i\omega t_k}|^2. \tag{4.3}$$

It means that, for an arbitrary set of time points, the normalized Schuster periodogram of pure noise

$$S_N(\omega) = \frac{S(\omega)}{\sigma_0^2/N} \tag{4.4}$$

is distributed as

$$p(z) = \exp(-z)D(z,\omega), \quad 0 < z < \infty, \quad \omega > 0, \tag{4.5}$$

where

$$D(z,\omega) = \alpha^{-1/2}\exp[-z(1-\alpha)/\alpha]I_0[-z(1-\alpha)^{-1/2}/\alpha], \tag{4.6}$$

$$\alpha = 1 - W(2\omega). \tag{4.7}$$

Strictly speaking, this distribution function is valid for the Schuster periodogram calculated with the time points $\bar{t}_k$. Nevertheless, it can be applied to the conventional Schuster periodogram defined by the time points t_k, since the squared abcolute values discard the time shift (see Eq. 2.8).

The correlation coefficient between the values of $S(\omega_1)$ and $S(\omega_2)$ is given by the expression

$$k(\omega_1,\omega_2) = \frac{W(\omega_1-\omega_2)+W(\omega_1+\omega_2)}{\sqrt{[1+W(2\omega_1)][1+W(2\omega_2)]}}, \tag{4.8}$$

which follows from Eqs. (2.32) and (2.30) if Eqs. (2.7) are taken into account. It is worthwhile to remark that Eq. (2.30) is valid no matter what time points $\bar{t}_k$ or t_k are used, and the points t_k are to be used to obtain Eq. (4.8), whereas the points $\bar{t}_k$ yield the corresponding correlation for the LS-spectra. We see that, like in the cases of the LS-spectra, the values of the Schuster periodogram evaluated for uneven pure noise are correlated and consequently are not independent.

5 SIGNIFICANCE TESTS

The central question considered in this section is: what is the probability q that a certain peak in the periodogram is generated by noise? If the value q is small, then we conclude that this peak hardly comes from noise, and, consequently, the existence of a harmonic component in the data can be claimed with high probability $1-q$. In this connection, two standard situations we meet in practice can be considered. For the sake of reference they will be called the hypotheses H_1 and H_2.

a) The hypothesis H_1. We have no *a priori* information at what frequency the peak due to a signal can be expected. In this case we may think that it is the highest peak. At the same time, the peaks in the noise periodogram, being the random values, can be sufficiently large due to a chance noise fluctuation. So the problem of finding the distribution function for the highest peak in the periodogram of noisy data must be considered.

b) The hypothesis H_2. We know *a priori* at what frequency the peak due to a signal is to be sought. In this situation, another task must be solved: what is the probability that at the preselected frequency the peak of the noise periodogram can be large.

5.1 The Regular Set of Time Points

It is very instructive to recall solutions of the above-mentioned problems for the case of even time series (Terebizh, 1992). In this case the time points are considered to be

$$t_k = \Delta t(k-1),\ k = 1, 2 \ldots N, \tag{5.1}$$

where Δt is the constant interval of sampling. Consider now the set of *natural frequencies* (we suppose that N is an even number):

$$\omega_j = \frac{2\pi}{N\Delta t} j,\ j = 0, 1, \ldots, \frac{N}{2}. \tag{5.2}$$

It is known that at these frequencies the Schuster periodogram and all the LS-spectra are all the same. At the same time, it is simple to verify that

$$W(\omega_j) = 0,\ j = 1, \ldots, \frac{N}{2} - 1, \tag{5.3}$$

whence, due to Eq. (4.8), the values of the periodogram are not correlated. Moreover, the values of the normalized periodograms (3.3) and (4.4) are distributed according to the exponential law. Now, in the limits of the hypothesis H_1 one has:

a) the probability that each peak of the normalized periodograms (3.3) and (4.4) does not exceed the value $x > 0$ is

$$\int_0^x e^{-z}\, dz = 1 - e^{-x}; \tag{5.4}$$

b) the probability that all peaks $S_N(\omega_j)$, $j = 1, \ldots, \frac{N}{2} - 1$ do not exceed the value $x > 0$ is

$$(1 - e^{-x})^{N/2-1} \tag{5.5}$$

c) the probability that at least one of the $S_N(\omega_j)$ will be above the level $x > 0$ is

$$Q(x) = 1 - (1 - e^{-x})^{N/2-1}. \tag{5.6}$$

The function $Q(x)$ is known as the Walker distribution of the highest peaks in the Schuster periodogram of pure noise (Walker, 1914). If we adopt a value $q < 1$, then the solution of equation

$$Q(X_q) = q \tag{5.7}$$

yields

$$X_q = -\ln[1 - (1-q)^{2/(N-2)}]. \tag{5.8}$$

Thus, the highest peak in the periodogram that satisfies the condition

$$S_{\max} \geq X_q \tag{5.9}$$

can be regarded as a signal with the probability $1 - q$. The value q is known as FAP (*the False Alarm Probability*), whereas the value X_q is called *the detection threshold.*

In the case of the hypothesis H_2, the probability that each peak in the periodogram exceeds the level $x > 0$ is

$$Q(x) = \int_x^\infty e^{-x}\,dx = e^{-x}. \tag{5.10}$$

Now, the detection threshold is detemined according to

$$X_q = -\ln(q), \tag{5.11}$$

and the detection of a signal at a known frequency ω_0 is claimed with the probability $1-q$ if

$$S_N(\omega_0) \geq X_q. \tag{5.12}$$

5.2 The Irregular Set of Time Points

In this case, due to Eq. (3.4), the values of the normalized LS-spectra are distributed exponentially, but they are correlated, as it was pointed out in Section 3. For this reason, the transition from Eq. (5.4) to Eq. (5.5) cannot be justified, and, consequently, it is impossible to obtain analytically the probability distribution of the highest peak in the said periodograms. In some cases it may be worthwhile to establish the probability distribution numerically by calculating the periodograms of different sequences of quasi-random noise. This proposal was made by Lomb (1976), and it was followed by Horne and Baliunas (1986), who introduced the empirical formula

$$Q(x) = 1 - (1 - e^{-x})^{N_i}, \tag{5.13}$$

where N_i (i-independent) designates the number of independent frequencies encorporated in "derivation" of Eq. (5.13). The numerical experiments with different patterns of time points led them to a conclusion that $N_i \approx N$. In this connection Koen (1988) showed that the strict statistical independence appears to be not realizable for unequally spaced data. Now we see that in the case of the Schuster periodogram we are facing the same problem: due to correlation between the peaks of $S_N(\omega)$ (see Eq. 4.8), no analytical expression analogous to the Walker distribution (5.6) can be found, and the only way to evaluate the detection thresholds is numerical experiments with the set of time points which is at our disposal. In Table 1 we show the results of such experiments with the Schuster periodogram for several kinds of time points distributions. Each column of this table was obtained by averaging 10 histograms, each one being derived from 1000 periodograms for quasi-normally distributed random variables.

In the case of the hypothesis H_2, the detection threshold for the LS-spectra is determined by Eq. (5.11) for all distributions of time points (regular and irregular). In general, this is not so for the Schuster periodogram since, in the case of irregular spaced time points, the analog of Eq. (5.10) is

$$Q(x) = \int_x^\infty e^{-z} D(z,\omega)\,dz, \tag{5.14}$$

Table 1. The values of the False Alarm Probability q as a function of the detection threshold X_q: a) an even sequence of 128 points; b) periodical 5-point gaps in the sequence of 115 regularly spaced points; c) two 33-point blocks separated by a 54-point gap; d) a random sample of 57 points from 120 regularly spaced points

X_q	a	b	c	d
4.5–5.0	0.200	0.200	0.210	0.210
5.0–5.5	0.180	0.160	0.170	0.170
5.5–6.0	0.130	0.120	0.120	0.130
6.0–6.5	0.086	0.073	0.072	0.080
6.5–7.0	0.059	0.045	0.043	0.052
7.0–7.5	0.035	0.026	0.024	0.029
7.5–8.0	0.023	0.016	0.017	0.014
8.0–8.5	0.015	0.010	0.010	0.009
8.5–9.0	0.008	0.006	0.006	0.006
9.0–9.5	0.005	0.006	0.006	0.003
9.5–10.0	0.002	0.002	0.002	0.002
10.0–10.5	0.001	0.003	0.001	0.001
10.5–11.0	0.001	0.001	0.001	0.001

whereas the detection threshold X_q is determined from the equation

$$q = \int_{X_q}^{\infty} e^{-z} D(z,\omega)\, dz. \tag{5.15}$$

It is important to note that at frequencies that satisfy the condition

$$W(2\omega) = 0, \tag{5.16}$$

the function $D(z,\omega) \equiv 1$, and Eqs. (5.14) and (5.15) again give Eqs. (5.10) and (5.11). Earlier we have shown (Vityazev, 1996) that Eq. (5.16) determines the set of frequencies at which the Schuster periodogram and the LS-spectra of a sine function are identical for even and uneven time series. Now we see that the same condition yields the identity of the distribution laws of all these periodograms. In other words, we have found the frequencies at which the Schuster periodogram retains the "classical" exponential distribution law for irregularly spaced time points. From Eq. (5.16) it follows that the deviation from the exponential law occurs at the frequencies $\omega \neq 0$ that satisfy the relation

$$2\omega = \bar{\omega}, \tag{5.17}$$

where $\bar{\omega}$ is a frequency at which the spectral window has a strong peak due to irregularity of the data.

6 NUMERICAL EXAMPLES

In Figures 1–3 we show the functions $W(\omega)$, $\alpha(\omega)$, and $D(z,\omega)$ for three models of the time points distribution. The first example demonstrates periodical gaps of

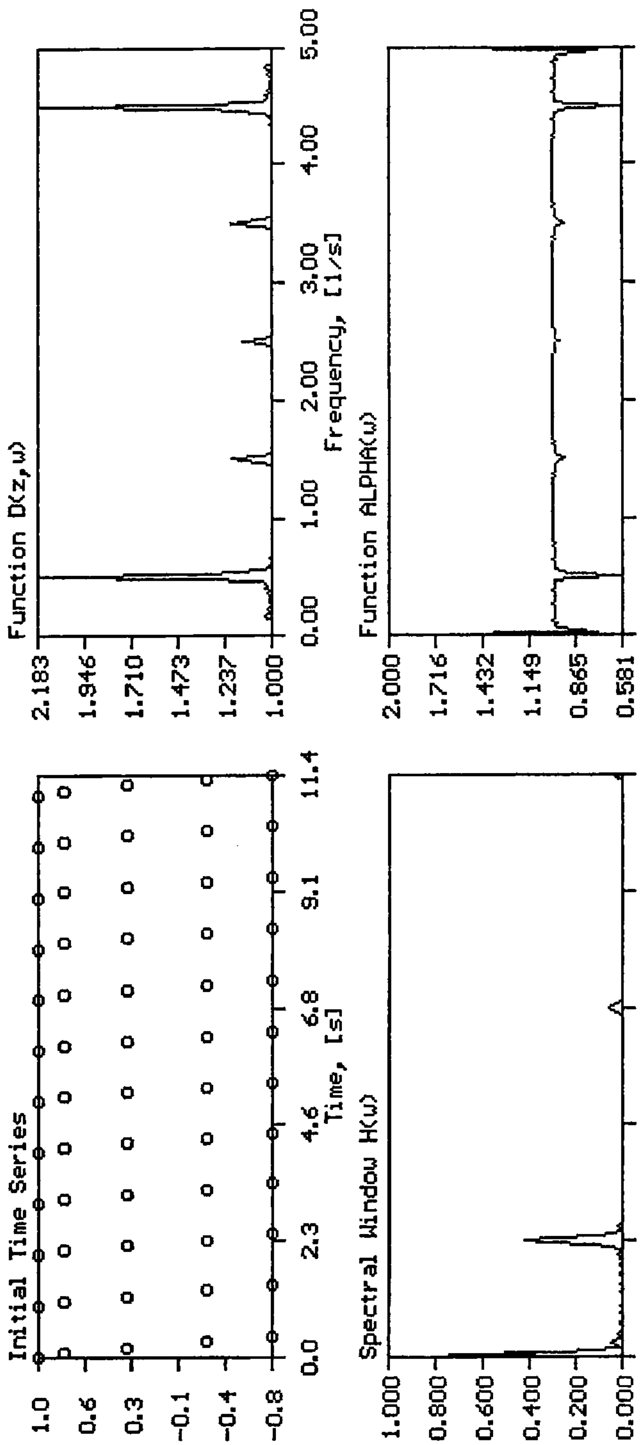

Figure 1 Functions $W(\omega)$, $D(z,\omega)$, $z = 6$, and $\alpha(\omega)$ for the set of time points with periodic gaps ($n = 5$; $p = 5$; $m = 12$).

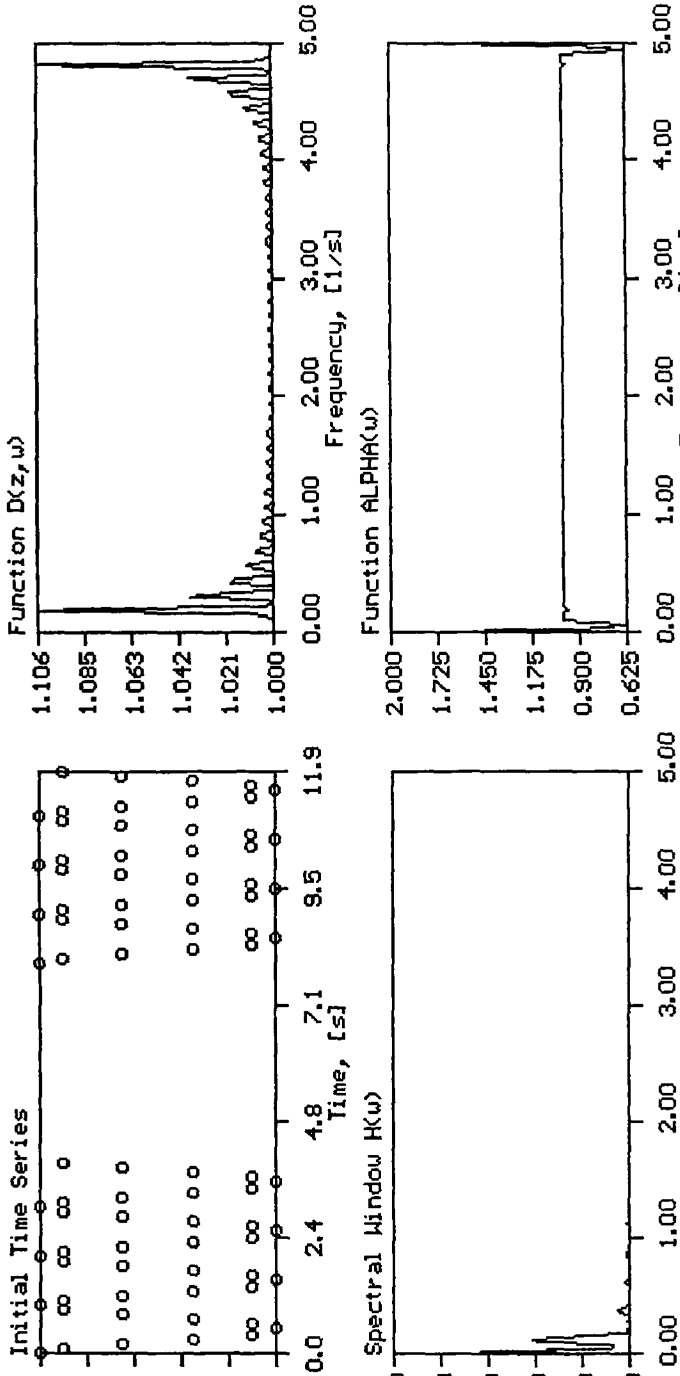

Figure 2 Functions $W(\omega)$, $D(z,\omega)$, $z = 6$, and $\alpha(\omega)$ for the set of time points with a long gap ($n = 40$; $p = 40$).

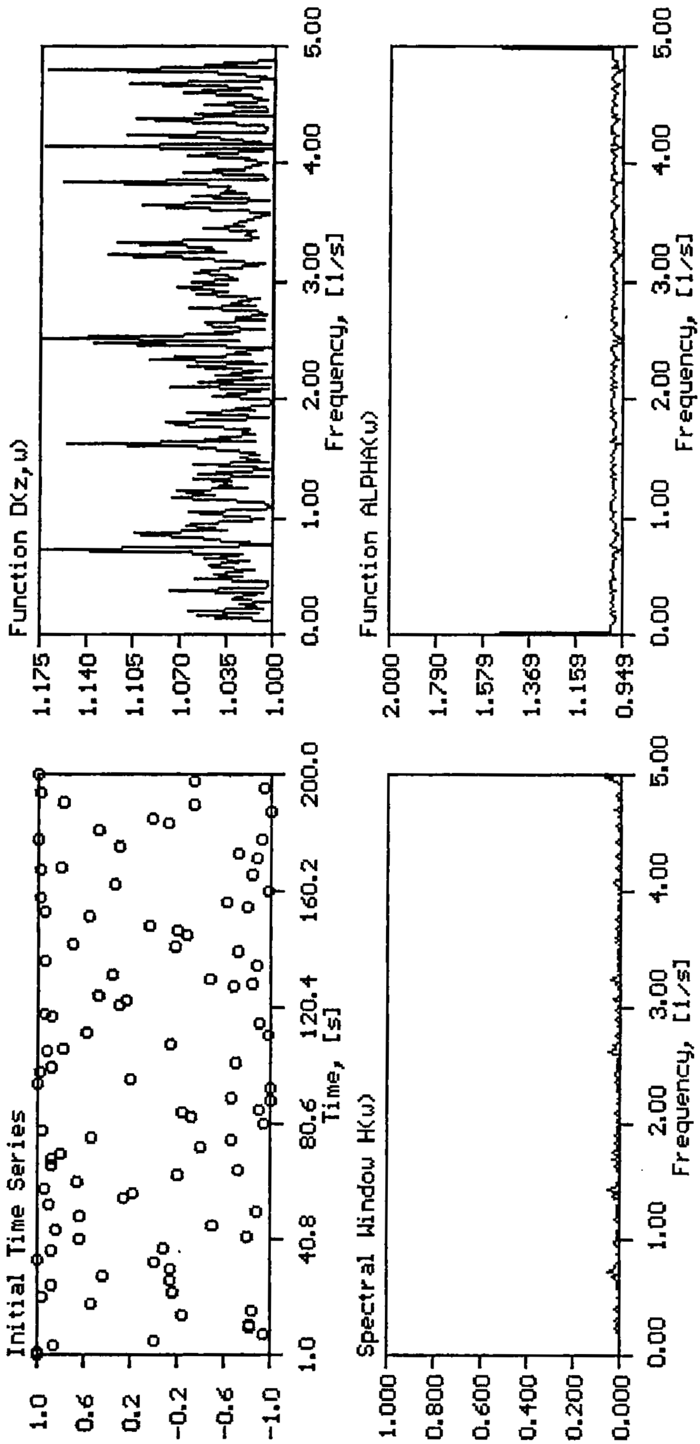

Figure 3 Functions $W(\omega)$, $D(z, \omega)$, $z = 6$, and $\alpha(\omega)$ for the set of 100 irregularly distributed time points.

observations. Here we suppose that at the set of regular points with the constant sampling interval Δt one has n successive observations and p successive missing points, the combinations of $n+p$ points being repeated m times. Figure 1 shows that noticeable deviations from the exponential law are observed only at two frequencies, namely $\bar{\omega}/2$ and $\pi/\Delta t - \bar{\omega}_1/2$, where $\bar{\omega} = 2\pi/m(n+p)\Delta t$ is the frequency of gaps. The second example is a time series that consists of two n-point blocks of observed data separated by a p-point gap, all the points are supposed to be regularly spaced over the time interval $\Delta t = \text{const}$. Similarly to the first case, Figure 2 shows large deviations from the exponential law again at two frequencies $\bar{\omega}/2$ and $\pi/\Delta t - \bar{\omega}/2$, where $\bar{\omega}_1 = 2\pi/(n+p)\Delta t$. In both cases the deviations at all other frequencies do not exceed 20 per cent. In the third example (random shifts from the regular sequence of points, Figure 3) the deviations are of the same order.

7 CONCLUSIONS

Here we summarize the results obtained in this paper:
a) the spectral window $W(\omega)$ is the key in the spectral analysis of irregular time series;
b) at the frequencies that satisfy the relation $W(2\omega) = 0$, the values of the Schuster periodogram, when the time series is uneven pure noise, are distributed exponentially. Otherwise, they are distributed according to the function given by Eqs. (4.5)–(4.7);
c) within limits of the H_2-hypothesis, i.e. when we study the spectrum at a preselected frequency ω_0, no distinction exists between the statistical properties of the Schuster periodogram and of the LS-spectra if the condition $W(2\omega_0) = 0$ is satisfied;
d) within the limits of H_1-hypothesis, i.e. when we study the highest peak in the spectrum, the situation becomes worse for the Schuster periodogram as well as for the LS-spectra due to correlation that exists between the spectral values. For this reason, the exponential law does not save the LS-spectra and the analog of the Walker distribution law must be evaluated numerically for every irregular set of time points, no matter what periodogram is used for the spectral evaluation.

References

Barning, F. J. M., (1962) *B. A. N.* **17**, 22.
Deeming, T. J. (1975a) *Astrophys. Space Sci.* **36**, 137.
Deeming, T. J. (1975b) *Astrophys. Space Sci.* **42**, 257(E).
Ferraz-Mello, S. (1981) *Astron. J.* **86**, 619.
Jenkins, G. M. and Watts, D. G. (1968) *Spectral Analysis and Its Applications*, Holden–Day, San Francisco.
Lomb, N. R. (1976) *Astrophys. Space Sci.* **39**, 447.
Marple, Jr., S. L. (1987) *Digital Spectral Analysis with Applications*, Prentice–Hall, Englewood Cliffs, New Jersey.
Otnes, R. K. and Enocson, L. (1978) *Applied Time Series Analysis*, A Wiley-Interscience Publication, New York.

Roberts, D. H., Lehar, J., and Dreher, J. W. (1987) *Astron. J.* **93**, 968.
Scargle, J. D. (1982) *Astrophys. J.* **263**, 835.
Scargle, J. D. (1989) *Astrophys. J.* **343**, 874.
Schuster, A. (1898) *Terrestrial Magnetism* **3**, 13.
Terebish, V. Yu. (1992) *Time Series Analysis in Astrophysics*, Moscow, Nauka.
Vityasev, V. V. (1994) *Astron. Astrophys. Tr.* **5**, 177.
Vityasev, V. V. and Prudnikova, E. Ya. (1994) *Vestn. St. Petersburg Univ.*, Ser. 1, issue 2, No. 8.
Vityasev, V. V. (1996) *Astron. Astrophys. Tr.* (in press).