FURTHER DEVELOPMENT AND PROPERTIES OF THE SPECTRAL ANALYSIS BY LEAST-SQUARES

PETR VANÍČEK
Surveys and Mapping Branch, Dept. of Energy, Mines and Resources, Ottawa, Ont., Canada†

(Received 7 December, 1970)

Abstract. The concept of spectral analysis using least-squares is further developed to remove any undesired influence on the spectrum. The influence of such a 'systematic noise' can be eliminated without the necessity of knowing the magnitudes of the noise constituents. The technique can be used for irregularly spaced as well as equidistantly spaced data. The response to random noise is found to be constant in the frequency domain and its expected level is derived. Presence of random noise in the analyzed time series is shown to transform the spectrum merely linearly. Examples of applications of the technique are presented.

1. Introduction

The classical methods of spectral analysis are based on the assumption that the analyzed time series is stationary, i.e. that the average of the time series is not affected by any translation of the time origin (see e.g. Blackman and Tukey, 1959, p. 4). Yet the majority of time series encountered in geophysical sciences are known to be non-stationary due to the very nature of the observed phenomena which may include secular trends, long periodic influences and the like. It is understood that the more the analyzed time series departs from stationarity, then the more disturbed the spectral results are likely to be. There are techniques designed to overcome these difficulties, but one feels that these approaches have only a limited power unless the composition of the analyzed empirical function is well known beforehand.

The present method is the result of an attempt to avoid the described hurdles by approaching the problem from a different angle. The method does not claim to be either all powerful or trouble free, although the author would like to think that it possesses some useful properties. Two such properties may be noted at this point. Firstly, the method consists of a prescription for designing a spectrum of the given time series which is insensitive to any particular systematic or random noise; and secondly, a feature of the method is that it can be used for analyzing a time series defined on non-equidistant as well as equidistant time intervals. Unless otherwise stated a general discrete time series is assumed, although the technique can easily be extended to continuous time functions as well.

Throughout our development we shall be using the following notation:

$$\langle F, G\rangle = \sum_{t \in \mathcal{N}_n} (F(t)\, G(t)) \tag{1.1}$$

† Present address: Dept. of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.

Astrophysics and Space Science **12** (1971) 10–33. *All Rights Reserved*

for the scalar product of any two functions F, G defined on $\mathcal{N}_n \equiv \{t_1, t_2, \ldots, t_n\}$,

$$\varrho(F) = \langle F, F\rangle \tag{1.2}$$

for the norm of F, and

$$\varrho(F, G) = \varrho(F - G) \tag{1.3}$$

for the mean-quadratic distance between F and G.

We shall be often referring to the identity

$$\varrho(F \pm G) = \varrho(F) \pm 2\langle F, G\rangle + \varrho(G) \tag{1.4}$$

the proof of which follows immediately by expanding the right-hand side of (1.2) with $F \pm G$ instead of F.

Note a particular case of (1.4)

$$\varrho(F) - \varrho(F, G) = 2\langle F, G\rangle - \varrho(G) \tag{1.5}$$

which is going to play an important role in our problems.

2. Basic Ideas

Suppose we have an observed time series $F(t)$, $t \in \{t_1, t_2, \ldots, t_n\} \equiv \mathcal{N}_n$ where the only fact known is that it contains an arbitrary constant due to the arbitrarily selected datum for the observation. Thus we shall require that the spectrum of F be, at the most, a linear transform of the spectrum $F+C$ for any number C. Further, we must require that the ordinate of the spectrum represents the 'contribution' of the corresponding frequency to the overall variance of F (see e.g. Blackman and Tukey, 1959, p. 7). We shall show that these two requirements are not enough to specify a spectrum.

Let us define

$$\varrho(F, T) = \sum_{t \in \mathcal{N}_n} (F(t) - T(t))^2, \tag{2.1}$$

where T represents a trigonometric term

$$T(t) = c_1 + c_2 \cos \omega t + c_3 \sin \omega t. \tag{2.2}$$

We can see that $\varrho(F, T)$ describes some linear transform of the variance of the residual of F after subtraction of the trigonometric term T. Thus σ defined as follows

$$\sigma(F, T) = \varrho(F) - \varrho(F, T) \tag{2.3}$$

describes the 'contribution' of the frequency ω to $\varrho(F)$.

Because we have not asked for any particular 'contribution', the coefficients c_1, c_2, c_3 are not specified and $\sigma(F, T)$ is not unique. Intuitively, the most natural way to specify them is to ask for the 'maximum possible contribution' of ω, a request which leads immediately to the minimization of $\varrho(F, T)$. But the minimum of $\varrho(F, T)$ is achieved for c_1, c_2, c_3 furnished by the set of normal equations (see e.g. Berezin and

Zhidkov, 1962, I, p. 396 or Handscomb, 1965, p. 27), if they are solvable. Thus $\sigma(F, T)$ for any given F becomes a function of ω only, which may or may not be defined for all $\omega \in \Omega$. Hence, $\sigma(F, T) = \sigma(\omega)|_F$ (with c_1, c_2, c_3 computed from the system of normal equations) equals the maximum contribution of the frequency ω to $\varrho(F)$. It was shown in Vaníček, 1969a, that the location of peaks and dips of σ is invariant in the transformation $F \to F + C$ that satisfies our first requirement. This property is due to the presence of c_1 in Equation (2.2).

Further, it was shown in the same paper that if we subtract from F defined on equidistant $\mathcal{N}_n$ the mean $\bar{F}$ to get ΔF we can write σ in a simple form

$$\sigma|_{\Delta F}(\omega) = p(\omega) \Big(\sum_{t \in \mathcal{N}_n} \Delta F(t) \cos \omega t \Big)^2 + r(\omega) \Big(\sum_{t \in \mathcal{N}_n} \Delta F(t) \sin \omega t \Big)^2, \quad (2.4)$$

which is nothing else but the contribution of the frequency ω to the n-times variance of F. Here

$$p(\omega) = 1 \Big/ \left(\sum_t \cos^2 \omega t - \frac{1}{n} \left(\sum_t \cos \omega t \right)^2 \right), \qquad r(\omega) = 1 \Big/ \sum_t \sin^2 \omega t .$$

It is not difficult to see the similarity of σ given by Equation (2.4) with the Fourier transform

$$R(\omega) = \frac{2}{n} \left(\left(\sum_{t \in \mathcal{N}_n} F(t) \cos \omega t \right)^2 + \left(\sum_{t \in \mathcal{N}_n} F(t) \sin \omega t \right)^2 \right) \quad (2.5)$$

(see e.g. Lanczos, 1957), and it can be shown that $\sigma(\omega) = R(\omega)$ for $\omega = k(n-1)/n$ $k = 0, 1, \ldots$. Thus, the functions p, r in Equation (2.4) may be regarded as corrective terms in the Fourier transform responsible for the mentioned properties.

One more feature of σ is worth pointing out here. If F is a simple sinusoidal wave of say frequency μ, then $\sigma(\mu) = \varrho(F)$ is the only absolute maximum of σ on Ω. This suggests that the spectral image should be more or less correct, even for very low frequencies, if the function F has one predominant periodical constituent in this region. This statement cannot generally be said of the Fourier transform.

Having noticed the consequence of presence of the absolute term c_1 in Equation (2.2), we will now show how it can be further exploited.

3. Known Constituents – Systematic Noise

In geophysical applications we often have to analyze for unknown periodicities of an observed non-stationary time series $F(t)$, when some properties of the series are already known. For example, it may be necessary to deal with a time series containing a theoretically predicted linear (quadratic, exponential, ...) secular trend of unknown magnitude, or to deal with a time series with a well established annual constituent (of unknown magnitude), or again, a time series observed during different periods related to different (unknown) datums. All these properties can be expressed as functions of time

$$c\Phi(t) \tag{3.1}$$

with unknown magnitude c, and can be regarded as undesirable from the point of view of spectral analysis. In other words, these constituents can be considered as 'systematic noise' in the analyzed time series.

As stated in Section 1 there is an obvious danger in removing this noise when the magnitudes are not known.† Regardless of which technique is used, it is not possible to compute the true values of the c's exactly, unless the complete composition of F is known. If the composition of F were known it would not be necessary to analyze it for unknown periodicities. Determination of the c's from the incompletely known F can not only yield quite erroneous values, but can also influence the residue by creating undesired false periodicities. On the other hand, it is usually equally detrimental to leave these constituents unattended, because they may distort the spectral image of F to a considerable degree.

Therefore, it seems sensible to design a formula for the spectrum that would account for this systematic noise without the necessity of determining the c's. Such an extension of the existing least-squares technique, described in the previous paragraph, can be achieved relatively easily.

Let us consider a function F defined on $\mathcal{N}_n$ (or simply a time series $\{F_i, t_i\}$) with m 'known' constituents $\Phi_1, \Phi_2, \ldots, \Phi_m$. To make the notation simpler let us have from now on

$$\Phi_{m+1}(t) = \cos\omega t,\ \Phi_{m+2}(t) = \sin\omega t, \quad \text{and} \quad m+2 = M\,.$$

We can now consider the best fitting generalized polynomial P_M to F,

$$P_M(t) = \sum_{i=1}^{M} c_i \Phi_i(t), \tag{3.2}$$

instead of the simple trigonometric term T given by Equation (2.2), writing for the spectrum

$$\sigma(F, P_M) = \varrho(F) - \varrho(F, P_M). \tag{3.3}$$

The M-tuple c_i, $i=1, 2, \ldots, M$, is for every $\omega \in \Omega$ obviously defined by the set of normal equations

$$\sum_{j=1}^{M} \langle \Phi_i, \Phi_j \rangle\, c_j = \langle F, \Phi_i \rangle = l_i \qquad i = 1, 2, \ldots, M\,. \tag{3.4}$$

Note that assuming $\mathbf{\Phi} \equiv \{\Phi_1, \Phi_2, \ldots, \Phi_M\}$ a subset of a complete orthogonal system, $\mathbf{\Phi}^c$ say, on $\mathcal{N}_n$ we would obtain

$$\langle \Phi_i, \Phi_i \rangle\, c_i = \langle F, \Phi_i \rangle \qquad i = 1, 2, \ldots, M \tag{3.5}$$

† If the magnitudes were known then all the undesired constituents could be subtracted from F beforehand without any difficulty.

where the c's would be the coefficients of the truncated Fourier series

$$\sum_{i=1}^{M} c_i \Phi_i \sim F$$

developed for the base $\mathbf{\Phi}^c$. In our development, however, $\mathbf{\Phi}$ is not assumed to be generally a subset of an orthogonal system so that the Equations (3.4) will not generally collapse into Equations (3.5).

We shall be using a weaker assumption that the functions Φ_i are linearly independent on $\mathcal{N}_n$ for all concerned $\omega \in \Omega \subset [0, \frac{1}{2}(n-1)]$† so that the Equations (3.4) have a unique solution. Whether this assumption, equivalent to the assumption that det $\|\langle \Phi_i, \Phi_j \rangle\| \neq 0$ (Gram's determinant of $\Phi_1, \Phi_2, \ldots, \Phi_M$), is satisfied or not has to be established for each individual case because it depends on the selection of $\mathbf{\Phi}$ as well as $\mathcal{N}_n$.

4. The Generalized Spectrum and Its Basic Property

We can now write for the spectrum given by Equation (3.3) using the identity (1.5):

$$\sigma = 2 \langle F, P_M \rangle - \varrho (P_M) . \tag{4.1}$$

Substituting for P_M from Equation (3.2) and making use of Equations (3.4) we arrive at the well known formula

$$\sigma = \sum_{i=1}^{M} c_i l_i = \mathbf{l}^T \mathbf{c} , \tag{4.2}$$

derived for instance in Munk and Cartwright (1968) and referred to as 'prediction variance'. Let $A = \|\langle \Phi_i, \Phi_j \rangle\|$ be the matrix of the normal Equations (3.4). Then (3.4) become $A\mathbf{c} = \mathbf{l}$, hence $\mathbf{c} = A^{-1}\mathbf{l}$, and we get for the spectrum

$$\sigma = \mathbf{l}^T A^{-1} \mathbf{l} . \tag{4.3}$$

This formula represents the whole family of spectra varying from each other by the choice of $\mathbf{\Phi}^\circ \equiv \{\Phi_1, \Phi_2, \ldots, \Phi_m\}$. We are going to prove that, providing $\mathbf{\Phi}$ is linearly independent on $\mathcal{N}_n$ for all $\omega \in \Omega$, σ acquires only an additive constant when F transforms to $F' = F + L_m$ (not identically zero on $\mathcal{N}_n$) where L_m is any linear combination of the $\Phi_i \in \mathbf{\Phi}^\circ$.

Let

$$L_m = \sum_{i=1}^{m} \lambda_i \Phi_i , \tag{4.4}$$

and let

$$P'_M = \sum_{i=1}^{M} c'_i \Phi_i \tag{4.5}$$

be the best fitting polynomial to $F' = F + L_m$ where the coefficients c'_i, $i = 1, 2, \ldots, M$,

† Here, we are deliberately limiting ourselves to the interval used in the Fourier approach to keep the development as similar as possible.

are given by the set of normal equations

$$\sum_{j=1}^{M} \langle \Phi_i, \Phi_j \rangle c'_j = \langle F', \Phi_i \rangle = l'_i \qquad i = 1, 2, \ldots, M. \tag{4.6}$$

We can write

$$l'_i = \langle F, \Phi_i \rangle + \langle L_m, \Phi_i \rangle = l_i + \sum_{j=1}^{m} \lambda_j \langle \Phi_i, \Phi_j \rangle \qquad i = 1, 2, \ldots, M$$

and Equations (4.6) become

$$\sum_{j=1}^{M} \langle \Phi_i, \Phi_j \rangle c'_j = l_i + \sum_{j=1}^{m} \lambda_j \langle \Phi_i, \Phi_j \rangle \qquad i = 1, 2, \ldots, M,$$

which can be rewritten as

$$\sum_{j=1}^{M} \langle \Phi_i, \Phi_j \rangle c''_j = l_i, \qquad i = 1, 2, \ldots, M, \tag{4.7}$$

where

$$c''_i = \begin{cases} c'_i - \lambda_i & \text{for} \quad i \leqslant m \\ c'_i & \text{for} \quad i > m. \end{cases}$$

But Equations (4.7) are obviously identical with Equations (3.4) so that

$$c_i = c''_i \qquad i = 1, 2, \ldots, M$$

and we have

$$c'_i = \begin{cases} c_i + \lambda_i & \text{for} \quad i \leqslant m \\ c_i & \text{for} \quad i > m. \end{cases}$$

Hence,

$$P'_M = P_M + L_m$$

and

$$\begin{aligned} \sigma|_{F'} &= \sigma(F', P'_M) \\ &= \varrho(F') - \varrho(F', P'_M) \\ &= \varrho(F + L_m) - \varrho(F + L_m, P_M + L_m) \\ &= \varrho(F + L_m) - \varrho(F, P_M). \end{aligned}$$

Using the identity (1.4), we can rewrite

$$\varrho(F + L_m) = \varrho(F) + 2\langle F, L_m \rangle + \varrho(L_m)$$

and we obtain finally

$$\sigma|_{F'} = \sigma|_F + 2\langle F, L_m \rangle + \varrho(L_m), \tag{4.8}$$

where the second and third terms are evidently constant (for constant L_m), and the proof is thus concluded.

5. The Optimum Generalized Spectrum and Its Basic Properties

We can now show that by using the same idea a spectrum with better properties can be derived. It is not difficult to see from Equation (3.3) that $\sigma|_F \in [0, \varrho(F)]$. Let us define a 'normalized' spectrum

$$\tilde{\sigma}|_F = \sigma|_F/\varrho(F) = 1 - \frac{\varrho(F, P_M)}{\varrho(F)}. \tag{5.1}$$

Its values lie in the interval [0, 1] and they show the proportionate contribution of each frequency ω to the overall spectrum of F. We can see that $\tilde{\sigma}$ transforms as

$$\tilde{\sigma}|_{F'} = 1 - \frac{\varrho(F, P_M)}{\varrho(F + L_m)}, \quad (\varrho(F + L_m) \neq 0),$$

when F transforms to $F' = F + L_m$.

This indicates that while the values $\tilde{\sigma}|_{F'} = \tilde{\sigma}'(\omega)$ remain in [0, 1] all the differences in the height of peaks and dips will be emphasized $\varrho(F)/\varrho(F+L_m)$-times, and for any two $\omega_1, \omega_2 \in \Omega$ we get

$$\tilde{\sigma}'(\omega_2) - \tilde{\sigma}'(\omega_1) = \frac{\varrho(F)}{\varrho(F + L_m)}(\tilde{\sigma}(\omega_2) - \tilde{\sigma}(\omega_1)),$$

which obviously represents a linear transformation of $\tilde{\sigma}$ in the frequency domain.

We can obtain the 'optimum' (most pronounced) spectrum by selecting L_m so that the ratio $\varrho(F)/\varrho(F+L_m)$ becomes maximum; remember that we can select L_m in any way without distorting the spectral image horizontally. But this again is equivalent to the requirement that $\varrho(F+L_m)$ be minimum since $\varrho(F)$ is constant. The minimum is achieved for $L_m = -P_m$ where P_m is the best fitting (in the least squares sense) polynomial composed from only the known constituents $\mathbf{\Phi}^\circ$. Thus

$$P_m = \sum_{i=1}^{m} c_i \Phi_i, \tag{5.2}$$

where the c's are given by the following system of normal equations

$$\sum_{j=1}^{m} \langle \Phi_i, \Phi_j \rangle c_j = l_i \qquad i = 1, 2, \ldots, m. \tag{5.3}$$

Denoting the best fitting polynomial to $F^* = F - P_m$ again by P_M, we can write the optimum spectrum as

$$\sigma^*|_F = \tilde{\sigma}|_{F*} = 1 - \frac{\varrho(F - P_m, P_M)}{\varrho(F, P_m)}. \tag{5.4}$$

It is evident that σ^*, besides being the most pronounced spectrum of F, is also completely invariant in any transform $F \to F + L_m$. The implication is that by selecting the appropriate optimum spectrum, i.e. by selecting the right 'known' constituents $\Phi_1, \Phi_2, \ldots, \Phi_m$, we are able to suppress any 'systematic noise' of the form L_m without being obliged to know the magnitudes of its individual constituents. This also allows us to suppress the already known peaks in the spectrum in a search for as yet unknown periodicities.

One more distinct advantage is that we can easily treat data of a geophysical phenomenon which has been observed during several disconnected time intervals and is related to different arbitrary (unknown) datums. This can be achieved by considering for each time interval, an additional function defined as follows:

$$\Phi_i(t) = \begin{cases} 1 & \text{for } t \text{ in the } i\text{-th interval,} \\ 0 & \text{for } t \text{ outside the } i\text{-th interval,} \end{cases}$$

the coefficient of which represents the datum for the i-th interval. By treating such data as one 'continuous' time series, the separability of the long periodic constituents is increased – see example No. 4 accompanying this paper.

6. Development of a Simplified Formula for the Optimum Spectrum

Before showing the influence of random fluctuations in F on the optimum spectrum, let us consider the computational aspects. By realising that all the coefficients of P_m and P_M in Equation (5.4) are to be determined from the appropriate systems of normal equations, we can see that the problem of computing the values $\sigma^*(\omega)$ is not negligible. Fortunately, Equation (5.4) can be considerably simplified.

The coefficients c_i^* of P_M are to be computed from the following system of normal equations

$$\sum_{j=1}^{M} \langle \Phi_i, \Phi_j \rangle \, c_j^* = l_i^*, \qquad i = 1, 2, \ldots, M,$$

where

$$l_i^* = \langle F, \Phi_i \rangle - \langle P_m, \Phi_i \rangle \qquad i = 1, 2, \ldots, M.$$

Substituting for P_m from Equation (5.2) we obtain

$$l_i^* = l_i - \sum_{j=1}^{m} c_j \langle \Phi_i, \Phi_j \rangle \qquad i = 1, 2, \ldots, M$$

and comparison with Equations (5.3) yields

$$l_i^* = \begin{cases} 0 & i \leqslant m \\ \langle F, \Phi_i \rangle - \langle P_m, \Phi_i \rangle = \langle F^*, \Phi_i \rangle & i > m. \end{cases}$$

We can therefore denote the M-tuple $\{l_1^*, l_2^*, \ldots, l_M^*\}$ by $\{0, 0, \ldots, \xi, \eta\}$ and write for the 'plain' spectrum (see formula (4.3)) of F^*

$$\sigma|_{F^*} = \mathbf{l}^{*T} A^{-1} \mathbf{l}^*,$$

where only the terms with $i, j > m$ make any contribution. The optimum spectrum is then given by

$$\sigma^*|_F = \frac{\sigma|_{F*}}{\varrho(F^*)} = \frac{1}{\varrho(F^*)} \mathbf{l}^{*T} A^{-1} \mathbf{l}^* .$$

The process of computing σ^* must involve solving the system (5.3), which for convenience we can rewrite here as

$$\sum_{j=1}^{m} K_{ij} c_j = l_i \qquad i = 1, 2, \ldots, m, \tag{6.1}$$

and computing the residual F^*. It is not difficult to see that A^{-1} becomes a function of ω only. Provided we know the inverse K^{-1} of K we can derive relatively simple formulae for the elements of A^{-1}.

For simplicity let us denote the scalar products involving the variable periodic functions $\Phi_{m+1}(t) = \cos\omega t$, $\Phi_{m+2}(t) = \sin\omega t$ by

$$\langle \Phi_i, \Phi_{m+1} \rangle = u_i, \quad \langle \Phi_i, \Phi_{m+2} \rangle = v_i \qquad i = 1, 2, \ldots, m$$

and

$$\langle \Phi_{m+1}, \Phi_{m+1} \rangle = S_1, \quad \langle \Phi_{m+1}, \Phi_{m+2} \rangle = S_2, \quad \langle \Phi_{m+2}, \Phi_{m+2} \rangle = S_3, \tag{6.2}$$

all functions of ω only. Then we can write the complete matrix A as

$$A \equiv \begin{Vmatrix} K, & \mathbf{u}, & \mathbf{v} \\ \mathbf{u}^T, & S_1, & S_2 \\ \mathbf{v}^T, & S_2, & S_3 \end{Vmatrix} = \begin{Vmatrix} K, & Z \\ Z^T, & S \end{Vmatrix},$$

where K is an $m \times m$ matrix, Z is an $m \times 2$ matrix, and S is a 2×2 matrix. To find the needed inverse of A let us use the 'method of partitioning' (see e.g. Faddeyev and Faddeyeva, 1964, p. 179; Thompson, 1969, pp. 64–6; Hohn, 1960, p. 82) which is particularly handy in our case since the inverse K^{-1} is assumed known from the computation of P_m. We can write

$$A^{-1} = \begin{Vmatrix} C, & B \\ B^T, & D \end{Vmatrix},$$

where we shall further deal with D (2×2 matrix) only, since obviously

$$\mathbf{l}^{*T} A^{-1} \mathbf{l}^* = (\xi, \eta)\, D \begin{pmatrix} \xi \\ \eta \end{pmatrix}.$$

For D we obtain

$$D = \|S - Z^T K^{-1} Z\|^{-1} = \begin{Vmatrix} S_1 - \mathbf{u}^T K^{-1} \mathbf{u}, & S_2 - \mathbf{u}^T K^{-1} \mathbf{v} \\ S_2 - \mathbf{v}^T K^{-1} \mathbf{u}, & S_3 - \mathbf{v}^T K^{-1} \mathbf{v} \end{Vmatrix}^{-1}.$$

Since K is symmetrical we have $\mathbf{u}^T K^{-1} \mathbf{v} = \mathbf{v}^T K^{-1} \mathbf{u}$, and putting

$$\begin{aligned} \mathbf{u}^T K^{-1} \mathbf{u} &= U, \\ \mathbf{u}^T K^{-1} \mathbf{v} &= W, \\ \mathbf{v}^T K^{-1} \mathbf{v} &= V, \end{aligned} \tag{6.3}$$

we have finally

$$D = \frac{1}{(S_1 - U)(S_3 - V) - (S_2 - W)^2} \begin{Vmatrix} S_3 - V, & -S_2 + W \\ -S_2 + W, & S_1 - U \end{Vmatrix}.$$

Hence we have derived the simplified formula for σ^* of the form

$$\sigma^* = p\xi^2 + q\xi\eta + r\eta^2, \tag{6.4}$$

where

$$\begin{aligned} p &= \quad (S_3 - V)/(Q\varrho(F^*)), \\ q &= -2(S_2 - W)/(Q\varrho(F^*)), \\ r &= \quad (S_1 - U)/(Q\varrho(F^*)), \end{aligned} \tag{6.5}$$

and $Q = (S_1 - U)(S_3 - V) - (S_2 - W)^2 > 0$ for $\mathbf{\Phi}$ linearly independent on $\mathcal{N}_n$.

The similarity between σ^* and the Fourier transform is apparent even in this somewhat more complex form. The similarity becomes still more evident for $\mathbf{\Phi}$ all odd or even when q becomes identically zero for all $\omega \in \Omega$, as was the case for instance in Section 2. Thus as noted in Section 2 the functions p, q, r can again be regarded as corrective terms for the Fourier transform responsible for the described properties.

If for any $\omega \in \Omega$, $\mathbf{\Phi}$ is linearly dependent (this occurs, e.g. when we enforce some trigonometric functions as the 'known' constituents with say a frequency μ), $\sigma^*(\mu)$ is not defined and σ^* becomes discontinuous on Ω. However, we can use limits to define an everywhere continuous function σ^{**} since the singularity at $\omega = \mu$ is removable, satisfying $\sigma^{**}(\omega) = \sigma^*(\omega)$ for $\omega \neq \mu$. Note that $\sigma^{**}(\mu)$ is not necessarily zero as the point $(\mu, \sigma^{**}(\mu))$ may lie on the slope of a nearby peak. We shall henceforth assume that this continuation is carried out and write σ^{**} simply as σ^*.

From the computation point of view the evaluation of U, V, W, S_1, S_2, S_3 is the most troublesome part, because it involves determination of the values of the scalar products

$$\sum_{t \in \mathcal{N}_n} \Phi_i(t) \cdot \Phi_j(t), \qquad i = 1, 2, \ldots, M, \qquad j = m+1, m+2.$$

However, if we limit ourselves to an $\mathcal{N}_n$ of equidistant points t, or equidistant points with gaps, we are usually able to express the scalar products in a compact analytical form that is more convenient for numerical evaluation. It is not the aim of this paper to include these details. For some examples the reader is referred to Vaníček (1967), Vaníček (1969a), Vaníček (1969c), and Quraishee and Vaníček (1970).

7. Response to Random Noise

In practical applications we may have to consider the influence of both the 'systematic' noise, and the noise due to random fluctuations in F, on the spectrum. To show how the optimum spectrum 'responds' – a term borrowed from communication theory – to such fluctuations, let us consider a statistically independent random variable $x(t)$, $t \in \mathcal{N}_n$, with the mean $\mathscr{E}(x) = \kappa = \text{const.}$ and the variance $\operatorname{var}(x) = \mathscr{E}(x - \mathscr{E}(x))^2 = \gamma^2 = \text{const.} \neq 0$. (For proper definitions the reader is referred to Wilks (1962, pp. 73, 74)).

It was shown already (see Section 5) that the optimum spectrum is invariant in any transformation $F \to F + L_m$. Taking one of the $\mathbf{\Phi}^\circ$ as a constant function we can force the mean, κ, to be zero without altering the spectrum σ^*, hence $\gamma^2 = \mathscr{E}(x^2)$. This assumption will simplify the forthcoming argument.

To avoid any later confusion, let us put $x^* = x - p_m$, with $p_m = \sum_1^m d_i \Phi_i$ as the best fitting polynomial to x, and $\xi|_x = \langle x^*, \cos\rangle$, $\eta|_x = \langle x^*, \sin\rangle$.† Then we can write for the optimum spectrum of x,

$$\sigma^*|_x = \frac{(S_3 - V)\, \xi|_x^2 - 2(S_2 - W)\, \xi|_x \eta|_x + (S_1 - U)\, \eta|_x^2}{\varrho(x^*)\, Q}, \tag{7.1}$$

and we can prove that under the above assumptions its expected value $\mathscr{E}(\sigma^*|_x)$ is constant.

To prove it let us begin with showing that $\xi|_x$, $\eta|_x$ are linear combinations of the $x(t)$, $t \in \mathscr{N}_n$. We have

$$\begin{aligned} \xi|_x &= \langle x, \cos\rangle - \langle p_m, \cos\rangle \\ &= \langle x, \cos\rangle - \langle \sum_1^m d_i \Phi_i, \cos\rangle \\ &= \langle x, \cos\rangle - \sum_1^m d_i \langle \Phi_i, \cos\rangle , \end{aligned}$$

where, according to 6.2, $\langle \Phi_i, \cos\rangle = u_i$ and the d_i must again satisfy the system of normal equations

$$K\mathbf{d} = \mathbf{l}|_x ,$$

similar to (5.3), where $l_i|_x = \langle x, \Phi_i\rangle$. Hence,

$$\begin{aligned} \xi|_x &= \langle x, \cos\rangle - \mathbf{u}^T K^{-1} \mathbf{l}|_x \\ &= \langle x, \cos\rangle - \mathbf{u}^T K^{-1} \langle x, \mathbf{\Phi}^\circ\rangle \\ &= \langle x, \cos - \mathbf{u}^T K^{-1} \mathbf{\Phi}^\circ\rangle . \end{aligned}$$

Note that $\mathbf{u}^T K^{-1} \mathbf{\Phi}^\circ = P_c$ is nothing else than the best fitting polynomial to $\cos \omega t$ composed of all $\mathbf{\Phi}^\circ \equiv \{\Phi_1, \Phi_2, \ldots, \Phi_m\}$. Therefore $\xi|_x$ is the scalar product of $x(t)$ and the residual $\alpha(t)$ of $\cos \omega t$. Let us write it as

$$\xi|_x = \langle x, \alpha\rangle = \sum_t x(t)\, \alpha(t). \tag{7.2}$$

Similarly we get

$$\eta|_x = \langle x, \sin - \mathbf{v}^T K^{-1} \mathbf{\Phi}^\circ\rangle ,$$

where again $\mathbf{v}^T K^{-1} \mathbf{\Phi}^\circ = P_s$ is the best fitting polynomial to $\sin \omega t$ and the residual can be denoted by β yielding

$$\eta|_x = \langle x, \beta\rangle = \sum_t x(t)\, \beta(t). \tag{7.3}$$

† Note that asterisks in these symbols have nothing to do with the convolution operator. Convolution does not enter explicitly into our argument.

Evidently, $\varrho(x^*)$ in Equation (7.1) is constant in the frequency domain. On the other hand, the rest of the right hand side of (7.1) are all varying with ω. Thus $\varrho(x^*)$ can be regarded as statistically independent of the rest in the frequency domain and we can therefore write for the expected value of $\sigma^*|_x$:

$$\mathscr{E}(\sigma^*|_x) = \frac{(S_3 - V)\,\mathscr{E}(\xi|_x^2) - 2(S_2 - W)\,\mathscr{E}(\xi|_x\eta|_x) + (S_1 - U)\,\mathscr{E}(\eta|_x^2)}{\mathscr{E}(\varrho(x^*))\,Q}. \tag{7.4}$$

According to Wilks (1962, p. 83), we have for the expected values in the numerator – remembering that the $x(t)$, $t \in \mathscr{N}_n$, are themselves assumed statistically independent

$$\begin{aligned} \mathscr{E}(\xi|_x^2) &= \mathscr{E}(x^2)\sum_t \alpha^2(t), \\ \mathscr{E}(\eta|_x^2) &= \mathscr{E}(x^2)\sum_t \beta^2(t), \\ \mathscr{E}(\xi|_x\eta|_x) &= \mathscr{E}(x^2)\sum_t \alpha(t)\,\beta(t). \end{aligned} \tag{7.5}$$

Substituting for α we obtain

$$\sum_t \alpha^2(t) = \sum_t (\cos\omega t - P_c(t))^2 = \varrho(\cos, P_c).$$

By taking, in Equation (3.3), cos instead of F, and m instead of M, we can write

$$\sigma(\cos, P_c) = \varrho(\cos) - \varrho(\cos, P_c).$$

But from Equation (4.3), for m replacing M and writing, therefore, K instead of A we get

$$\begin{aligned} \sigma(\cos, P_c) &= \langle\cos, \boldsymbol{\Phi}^\circ\rangle^T K^{-1} \langle\cos, \boldsymbol{\Phi}^\circ\rangle \\ &= \mathbf{u}^T K^{-1}\mathbf{u}, \end{aligned}$$

and combining the three last results we have finally

$$\sum_t \alpha^2(t) = \varrho(\cos) - \mathbf{u}^T K^{-1}\mathbf{u}.$$

If we use the notation of Section 6 this becomes

$$\sum_t \alpha^2(t) = S_1 - U. \tag{7.6}$$

Similarly, for $\sum_t \beta^2(t)$ we obtain

$$\sum_t \beta^2(t) = S_3 - V. \tag{7.7}$$

We are going to prove that analogously

$$\sum_t \alpha(t)\,\beta(t) = S_2 - W. \tag{7.8}$$

It is

$$\sum_t \alpha(t)\,\beta(t) = \sum_t \left[\left(\cos\omega t - P_c(t)\right)\left(\sin\omega t - P_s(t)\right)\right]$$

$$= \sum_t \cos\sin - \sum_t P_s \cos - \sum_t P_c \sin + \sum_t P_c P_s,$$

where the first term is S_2 and each of the other terms equals W. We have

$$\begin{aligned}\sum_t P_s \cos &= \sum_t \left(\mathbf{v}^T K^{-1} \mathbf{\Phi}^\circ(t)\cos\omega t\right)\\ &= \mathbf{v}^T K^{-1}\langle \mathbf{\Phi}^\circ, \cos\rangle\\ &= \mathbf{v}^T K^{-1}\mathbf{u}\\ &= W\end{aligned}$$

and similarly

$$\sum_t P_c \sin = W\,.$$

Denoting the elements $\langle \Phi_i, \Phi_j\rangle$ of K by K_{ij}, the elements of K^{-1} by K_{ij}^{-1} and making use of the Kronecker δ we can write

$$\begin{aligned}\sum_t P_c P_s &= \sum_t \left[\sum_{i,j=1}^{m} \left(K_{ij}^{-1} u_i \Phi_j(t)\right) \sum_{k,l=1}^{m} \left(K_{kl}^{-1} v_k \Phi_l(t)\right)\right]\\ &= \sum_{ijkl} \left[K_{ij}^{-1} u_i K_{kl}^{-1} v_k \langle \Phi_j, \Phi_l\rangle\right]\\ &= \sum_{ijkl} \left[K_{ij}^{-1} u_i v_k K_{kl}^{-1} K_{jl}\right]\\ &= \sum_{ijk} \left[K_{ij}^{-1} u_i v_k \delta_{kj}\right]\\ &= \sum_{ij} \left[K_{ij}^{-1} u_i v_j\right]\\ &= W,\end{aligned}$$

which concludes the proof.

Substituting (7.6), (7.7), (7.8) in (7.5) and (7.5) in (7.4) we obtain

$$\mathscr{E}(\sigma^*|_x) = \frac{2\mathscr{E}(x^2)}{\mathscr{E}\left(\varrho(x^*)\right)}, \tag{7.9}$$

which is obviously constant for all ω and we have thus proved that the expected response to the purely random noise is constant.

To estimate the magnitude of this response let us establish the value of the denominator. It is

$$\begin{aligned}\mathscr{E}\left(\varrho(x^*)\right) &= \mathscr{E}\left(\varrho(x - p_m)\right)\\ &= \mathscr{E}\left(\varrho(x) - 2\langle x, p_m\rangle + \varrho(p_m)\right)\\ &= \mathscr{E}\left(\varrho(x)\right) + \mathscr{E}\left(-2\langle x, p_m\rangle + \varrho(p_m)\right)\\ &= \mathscr{E}(\mathrm{I}) + \mathscr{E}(\mathrm{II})\,.\end{aligned}$$

The first term $\mathscr{E}(\mathrm{I})$ evidently becomes

$$\mathscr{E}(\varrho(x)) = n\mathscr{E}(x^2). \tag{7.10}$$

We shall show that the second term, $\mathscr{E}(\mathrm{II})$ equals $-m\mathscr{E}(x^2)$, with m being the number of 'known' constituents. To prove this, we can use an argument similar to the development of (7.6). We can write

$$\begin{aligned} \mathrm{II} = -2\langle x, p_m\rangle + \varrho(p_m) &= \varrho(x, p_m) - \varrho(x) \\ &= -\sigma(x, p_m) \\ &= -\langle x, \mathbf{\Phi}^\circ\rangle^T K^{-1}\langle x, \mathbf{\Phi}^\circ\rangle . \end{aligned}$$

Rewriting this in a more elementary form we obtain

$$\begin{aligned} \mathrm{II} &= -\sum_{i=1}^{m}\left[\langle x, \Phi_i\rangle \sum_{j=1}^{m} K_{ij}^{-1}\langle x, \Phi_j\rangle\right] \\ &= -\sum_{i,j=1}^{m}\left[K_{ij}^{-1}\langle x, \Phi_i\rangle\langle x, \Phi_j\rangle\right] \\ &= -\sum_{i,j=1}^{m}\left[K_{ij}^{-1}\sum_{k=1}^{n}\left(x(t_k)\,\Phi_i(t_k)\right)\sum_{l=1}^{n}\left(x(t_l)\,\Phi_j(t_l)\right)\right] \\ &= -\sum_{k,l=1}^{n}\left[x(t_k)\,x(t_l)\sum_{i,j=1}^{m}\left(K_{ij}^{-1}\Phi_i(t_k)\,\Phi_j(t_l)\right)\right]. \end{aligned}$$

The expected value of II thus becomes

$$\mathscr{E}(\mathrm{II}) = -\sum_{kl}\left[\mathscr{E}\left(x(t_k)\,x(t_l)\right)\sum_{ij}\left(K_{ij}^{-1}\Phi_i(t_k)\,\Phi(t_l)\right)\right],$$

where

$$\mathscr{E}\left(x(t_k)\,x(t_l)\right) = \mathscr{E}(x^2)\,\delta_{kl}$$

since $x(t)$ are assumed statistically independent. Hence, we have

$$\begin{aligned} \mathscr{E}(\mathrm{II}) &= -\sum_{k=1}^{n}\left[\mathscr{E}(x^2)\sum_{ij}\left(K_{ij}^{-1}\Phi_i(t_k)\,\Phi_j(t_k)\right)\right] \\ &= -\mathscr{E}(x^2)\sum_{t}\sum_{ij}\left[K_{ij}^{-1}\Phi_i(t)\,\Phi_j(t)\right] \\ &= -\mathscr{E}(x^2)\sum_{ij}\left[K_{ij}^{-1}\langle\Phi_i, \Phi_j\rangle\right] \\ &= -\mathscr{E}(x^2)\sum_{ij}K_{ij}^{-1}K_{ij} \\ &= -\mathscr{E}(x^2)\sum_{i=1}^{m}1 \\ &= -\mathscr{E}(x^2)\,m, \end{aligned} \tag{7.11}$$

which was to be proved.

Substituting (7.10) and (7.11) in (7.9) we get

$$\mathscr{E}(\sigma^*|_x) = 2/(n-m); \tag{7.12}$$

and the expected level of random noise in the spectrum, which corresponds to the random noise in the analyzed time series, is constant and is inversely proportional to the number of statistical degrees of freedom $n-m$ of the analyzed time series $x-p_m$. From our earlier discussions concerning the optimized spectrum, it is obvious that the same spectrum belongs to the noise $X=L_m+x$.

It is felt that the investigation could be extended to establish the probabilities attached to particular magnitudes of peaks in the spectrum in terms of the probability distribution of the noise. Thus, statistical criteria could be derived for testing the statistical significance of individual peaks in a way similar to that used for the classical methods. However, to indulge in such argument is beyond the scope of this paper. The author believes that the examples attached to this paper should be sufficient at this stage to show the behaviour of the spectrum from the statistical point of view.

Remark – According to the argument in this section, we can interpret the formula (6.4) in such a way that p, q, r can be regarded as functions of the residuals $\alpha(t)$, $\beta(t)$ of $\cos\omega t$ and $\sin\omega t$. We can write

$$\begin{aligned} p &= \quad \varrho(\beta)/(\varrho(F^*)\,Q),\\ q &= -2\langle\alpha,\beta\rangle/(\varrho(F^*)\,Q),\\ r &= \quad \varrho(\alpha)/(\varrho(F^*)\,Q),\\ Q &= \quad \varrho(\alpha)\,\varrho(\beta) - \langle\alpha,\beta\rangle^2. \end{aligned}$$

If the pair of trigonometric functions $\cos\mu t$, $\sin\mu t$ is among the 'known' constituents $\mathbf{\Phi}^\circ$ then obviously the residuals α, β go to zero for $\omega\to\mu$ for all $t\in\mathcal{N}_n$. Hence $\varrho(\alpha)$, $\varrho(\beta)$, $\langle\alpha,\beta\rangle\to 0$ for $\omega\to\mu$. At the same time even $\xi|_F$ and $\eta|_F$ go to zero.

8. Influence of Noise on the Spectrum of a Time Series

It remains to be seen whether the spectral image of a time series $F(t)$ is spoiled by the presence of the noise X. Let us consider a time series F with a noise x superimposed. Let us assume that the noise x has the same statistical properties as in Section 7. The spectrum of such a time series $F'=F+x$ will be given by

$$\sigma^*|_F = p\,(\xi|_F + \xi|_x)^2 + q\,(\xi|_F + \xi|_x)\,(\eta|_F + \eta|_x) + r\,(\eta|_F + \eta|_x)^2, \qquad (8.1)$$

where again $\xi|_F=\sum_t F^*(t)\cos\omega t$, analogously $\xi|_x$, $\eta|_F$, $\eta|_x$, and p, q, r are given by

$$\begin{aligned} p &= \quad (S_3 - V)/(\varrho(F'^*)\,Q),\\ q &= -2\,(S_2 - W)/(\varrho(F'^*)\,Q),\\ r &= \quad (S_1 - U)/(\varrho(F'^*)\,Q). \end{aligned}$$

We can write

$$\begin{aligned} \sigma^*|_{F'} = {} & \frac{\varrho(F,P_m)}{\varrho(F',P'_m)}\,\sigma^*|_F + \frac{\varrho(x,p_m)}{\varrho(F',P'_m)}\,\sigma^*|_x \\ & + (2p\xi|_F + q\eta|_F)\,\xi|_x + (2r\eta|_F + q\xi|_F)\,\eta|_x, \qquad (8.2) \end{aligned}$$

having denoted by P_m, P'_m, p_m the best fitting polynomials to F, F' and x respectively. If one of the 'known' functions in $\Phi°$ is constant (which we can always ensure) we can again assume without any loss of generality that $\mathscr{E}(x)=0$. Since the last two terms in Equation (8.2) represent only a linear combination of x's, they disappear when we take the average of $\sigma^*|_{F'}$. We obtain

$$\mathscr{E}(\sigma^*|_{F'}) = \frac{\varrho(F, P_m)}{\mathscr{E}(\varrho(F', P'_m))}\,\sigma^*|_F + \frac{\mathscr{E}(\varrho(x, p_m)\,\sigma^*|_x)}{\mathscr{E}(\varrho(F', P'_m))}. \tag{8.3}$$

It is not difficult to show – using an argument similar to the one used in Section 4 – that $P'_m = P_m + p_m$. Further, making use of Equation (7.9) we have

$$\mathscr{E}(\varrho(x, p_m)\,\sigma^*|_x) = 2\mathscr{E}(x^2) = 2\gamma^2.$$

For the denominator of Equation (8.3) we can write (using the identity (1.4))

$$\mathscr{E}(\varrho(F', P'_m)) = \mathscr{E}[\varrho(F, P_m) + 2\langle F - P_m, x - p_m\rangle + \varrho(x, p_m)],$$

where the first term remains $\varrho(F, P_m)$, the third was shown to equal $(n-m)\gamma^2$ (see the development of Equation (7.9)), and the second can be rewritten as follows:

$$2\mathscr{E}(\langle F - P_m, x - p_m\rangle) = 2\sum_t ((F - P_m)\cdot\mathscr{E}(x - p_m)),$$

which goes to zero for $\mathscr{E}(x)=0$, since $\mathscr{E}(x)=0$ implies $\mathscr{E}(p_m)=0$, and, hence $\mathscr{E}(x-p_m)=0$. Thus we can write Equation (8.3) in the form

$$\mathscr{E}(\sigma^*|_{F'}) = \frac{\varrho(F, P_m)}{\varrho(F, P_m) + (n - m)\gamma^2}\,\sigma^*|_F + \frac{2\gamma^2}{\varrho(F, P_m) + (n - m)\gamma^2}. \tag{8.4}$$

Now, the function $F-P_m$, i.e. the residue of F after subtracting the best fitting polynomial composed of the 'known' functions $\Phi°$, can be intuitively regarded as a 'useful signal'. In other words $F-P_m=F^*$ is the function in which we are interested from the point of view of spectral analysis – bearing in mind that P_m cannot completely eliminate the undesired constituents because the whole structure of F is not known. It becomes natural to denote $\varrho(F, P_m)$ by $(n-m)\Gamma^2$ and call Γ^2 the 'variance of the useful signal'. Substituting this into Equation (8.4) we obtain

$$\mathscr{E}(\sigma^*|_{F'}) = \frac{\Gamma^2}{\Gamma^2 + \gamma^2}\,\sigma^*|_F + \frac{2\gamma^2}{(\Gamma^2 + \gamma^2)(n - m)}. \tag{8.5}$$

As a check, we may ask what would happen if F itself were also a random noise, say x'. Then $\sigma^*|_{F'}$ would become $2/(n-m)$ and we would get

$$\mathscr{E}(\sigma^*|_{x+x'}) = \frac{2(\Gamma^2 + \gamma^2)}{(\Gamma^2 + \gamma^2)(n - m)} = \frac{2}{n - m},$$

the expected answer.

Introducing the 'signal to noise ratio'

$$J = \Gamma/\gamma, \tag{8.6}$$

we can rewrite Equation (8.5) as

$$\mathscr{E}(\sigma^*|_{F'}) = \frac{J^2}{J^2+1}\,\sigma^*|_F + \frac{2}{(J^2+1)(n-m)}. \tag{8.7}$$

The result implies that as the signal to noise ratio increases then the shift of the spectrum of the signal is reduced. Hence the optimum spectrum behaves reasonably, and by considering the invariance under the transformation $F \to F + L_m$ we can confirm the quotation from Section 1 of this paper that the spectrum is not only insensitive to 'white noise' (for definition of white noise see e.g. Blackman and Tukey, 1959, p. 173) but is also insensitive to any noise which can be expressed in the form $L_m + x$.

Let us point out, for comparison, that if we apply a similar argument to the Fourier transform R (see Equation (2.5)) we shall get

$$\mathscr{E}(R|_x) = 2\mathscr{E}(x^2). \tag{8.8}$$

For the 'normalized Fourier transform' $\tilde{R} = R/\varrho(x)$ we have

$$\mathscr{E}(\tilde{R}|_x) = 2/n,$$

as compared to our Equation (7.12). If the presence of any 'systematic noise', L_m say, is suspected and its elimination is carried out, then this would introduce some linear dependence of the treated x' and $\mathscr{E}(x_i', x_j')$ would no longer equal to const. $\times\, \delta_{ij}$. In deriving Equation (8.8) we should have to write more generally

$$\mathscr{E}(x_i', x_j') = \operatorname{cov}(x_i', x_j') + \kappa^2,$$

(see e.g. Wilks, 1962, p. 78) and $\mathscr{E}(R|_{x'})$ would no longer be constant. An attempt to eliminate a 'systematic noise' will thus generally introduce false periodicities as we have already mentioned. This results in a need to correct the spectrum while any uncertainty in the elimination of 'systematic noise' remains.

9. Examples

To illustrate some of the points discussed in previous sections the following numerical examples were computed on a Univac 1108 using a flexible program designed for the technique.

Example No. 1 – 500 normally distributed pseudo-random numbers x_i with mean $\mathscr{E}(x) = 1.5$ and var $(x) = 1$ were generated using the existing subroutines. Their spectrum σ^* is plotted in Figure 1. The best fitted mean and trend, i.e. the coefficients to the enforced constituents $\Phi_1(t) = 1$, $\Phi_2(t) = t$, yielded values of 1.487 and 0.0005. Within the used frequency span only one peak (for the frequency 12.9) surpasses 1%. The whole spectrum looks fairly uniform and apparently has random fluctuations. Comparison with the expected mean level, 0.401% in our case, seems to support the theoretical conclusions. The computation took approximately 15 sec.

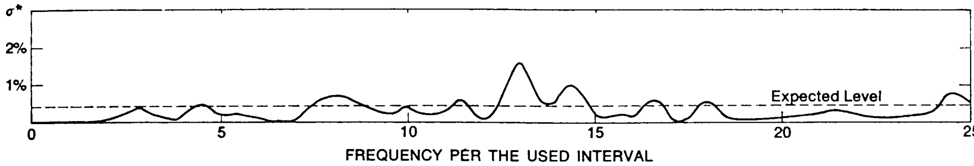

Fig. 1.

Example No. 2 – A time series of the following form

$$F(t) = 20 + 3t + 0.6\cos(1.4(2\pi/50)t) - 1.8\sin(1.4(2\pi/50)t)$$

was generated for $t=0.1, 0.2, \ldots, 50.0$ yr (500 values). The frequency 1.4 corresponds thus to the period 35.7143 yr and the linear trend to 3 units per year. The graph of the function is shown in Figure 2.

It was again decided to enforce $\Phi_1(t)=1$, $\Phi_2(t)=t$ and their coefficients came out to within 0.6% and 1.1% of the generated values:

$$c_1 = 19.87, \quad c_2 = 2.967.$$

The optimum spectrum σ^* of F is given in Figure 3. It can be seen that the main peak is located as close to the period 35.7143 yr as the reading precision allows. This had to be expected from the theory, even though the peak corresponds to a very low frequency. The inevitable 'side lobes', familiar to the users of any spectral analysis

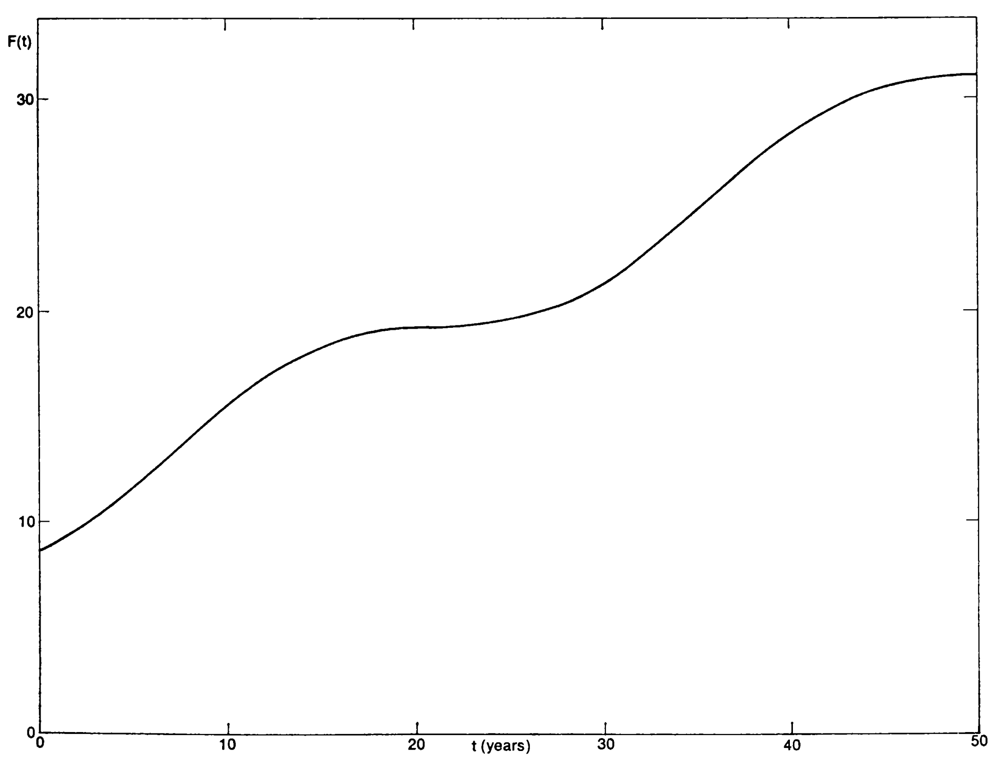

Fig. 2.

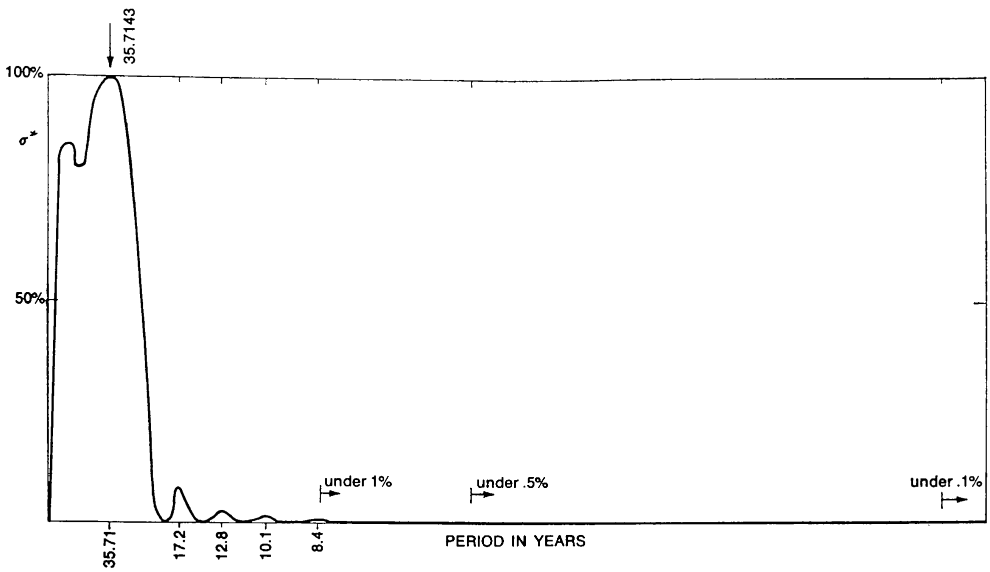

Fig. 3.

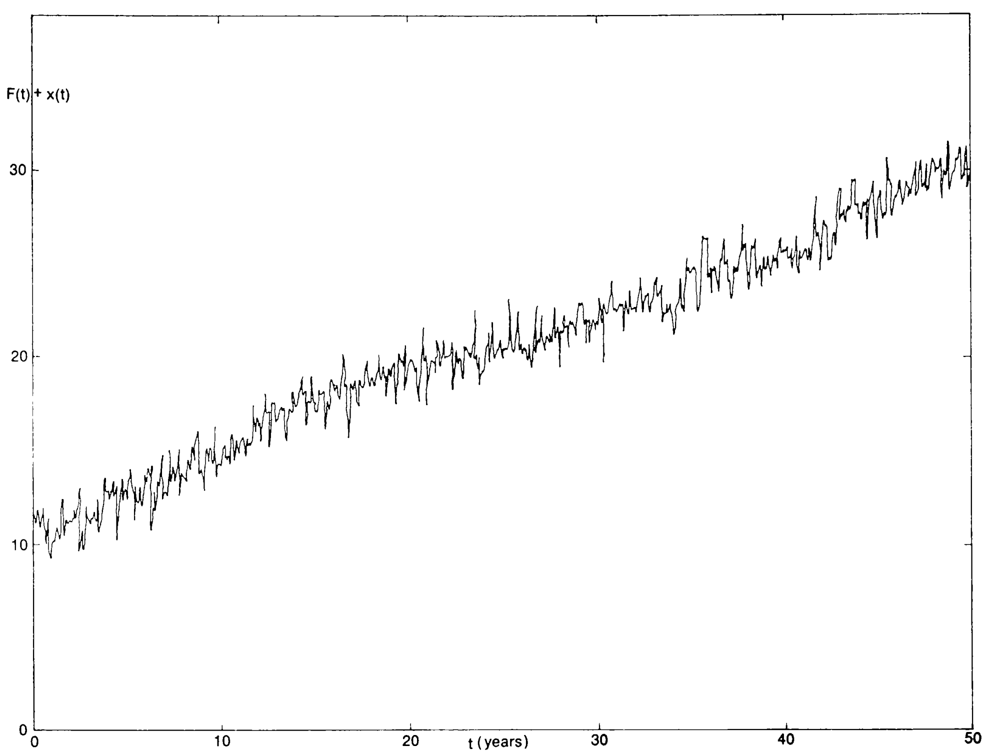

Fig. 4.

technique are, indeed, present. This phenomenon may be understood here – without going into any detail – by realizing that certain frequencies must exist that are liable to contribute more to the variance of a term '$a \cos \mu t + b \sin \mu t$' than other frequencies. More will be said about this point after Example No. 4. The computing time was approximately 15 sec.

Example No. 3 – To illustrate the influence of a random noise on the spectrum of a time series we have generated a time series

$$F'(t) = 19 + 3t + 0.2\cos(1.4(2\pi/50)t) - 0.6\sin(1.4(2\pi/50)t),$$
$$t = 0.1, 0.2, \ldots, 50.0 \text{ yr},$$

very similar to the one used in Example No. 2 and added to it the pseudo-random noise from Example No. 1. This 'noisy' time series is presented in Figure 4. If the theory holds the spectrum σ^* of F' should be similar to the spectrum of F from Example No. 2. Due to the presence of noise we have (according to Equation (8.7)) to expect the spectrum of F' to be diminished by the factor $J^2/(J^2+1)$, where $J = \Gamma/\gamma$. Γ^2 in our case is given by

$$\Gamma^2 = \frac{1}{498}\sum_{t=1}^{500}(0.2\cos 1.4(2\pi/500)t - 0.6\sin 1.4(2\pi/500)t)^2 \approx 0.2158$$

and γ^2 equals var (x) from Example No. 1 – i.e., 1. We thus get

$$J^2 \approx 0.2158$$

and the diminishing factor becomes approximately 0.1775. Hence we should have

$$\sigma^*|_{F'+x} = 0.1775\,\sigma^*|_F + 0.0033\,.$$

The values of the mean and trend were established as

$$c_1 = 20.444, \quad c_2 = 3.030$$

within 2.2% and 1.0% of the true values. The actual spectrum $\sigma^*|_{F'+x}$ is plotted in Figure 5. We see that the main peak is again located, within the reading precision, at the period of 35.7143 yr and its magnitude is rather close to the 17.75% as predicted. The side lobes are relatively higher. Otherwise, the spectrum looks very much like

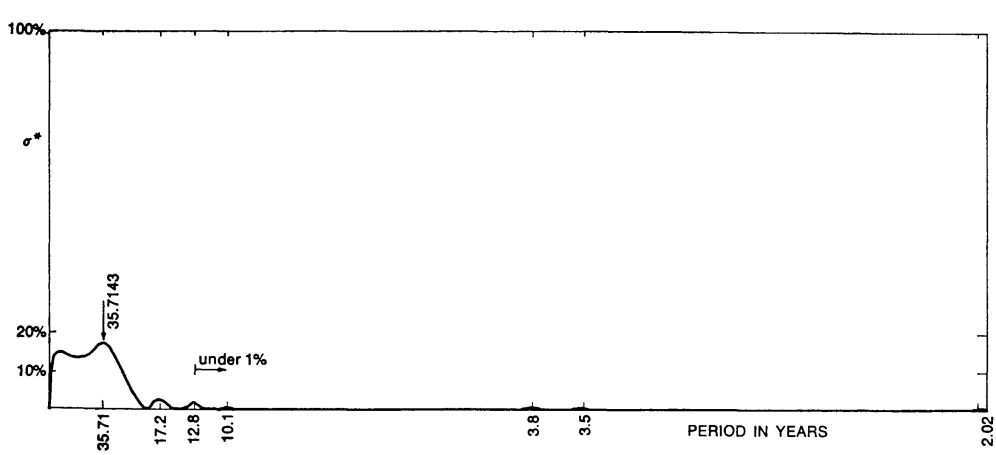

Fig. 5.

the one from Example No. 2. Except for three minor peaks reaching 0.9%, 0.7%, 0.7% that correspond to the similar peaks found in Example No. 1, the spectrum is well below 0.5%. The persistence of the three minor peaks introduces the suspicion that there may be some hidden periodicities in the generated series of pseudo-random numbers. Computing took again approximately 15 sec.

Example No. 4 – In this example a rather more complex time series

$$F(t) = C_i + 0.01t + \sum_{j=1}^{5} (a_j \cos \mu_j t + b_j \sin \mu_j t)$$

has been generated for $t \in \mathcal{N}_{270} \equiv \bigcup_{i=1}^{3} \mathcal{M}_i$ with

$$\mathcal{M}_1 \equiv \{0.1, 0.2, \ldots, 10.0\} \text{ yr},$$
$$\mathcal{M}_2 \equiv \{20.1, 20.2, \ldots, 25.0\} \text{ yr},$$
$$\mathcal{M}_3 \equiv \{28.1, 28.2, \ldots, 40.0\} \text{ yr},$$

$$C_1 = 1, \quad C_2 = -1, \quad C_3 = 3,$$

$$\begin{array}{ll} a_1 = 0.5, & b_1 = -0.5, \\ a_2 = -0.25, & b_2 = 0.0 \\ a_3 = 0.0 & b_3 = 1.0, \\ a_4 = 1.0, & b_4 = 0.5, \\ a_5 = 0.5, & b_5 = 1.0, \end{array}$$

and μ_i's corresponding to 40, 16, 5.714, 3.636, 2.759 yr respectively. The graph of $F(t)$ is shown in Figure 6.

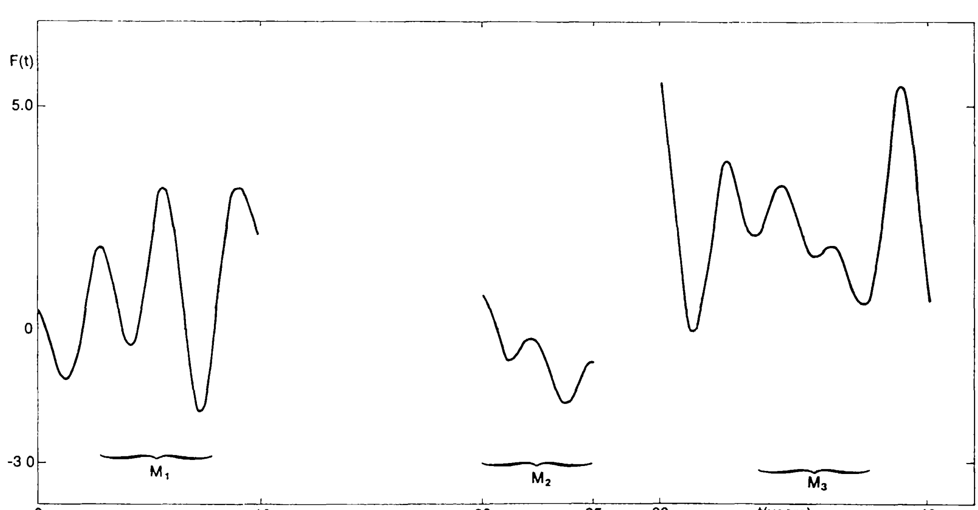

Fig. 6.

Ten 'known' constituents were selected as follows:

$$\Phi_1(t) = \begin{cases} 1 & t \in \mathcal{M}_1 \\ 0 & t \notin \mathcal{M}_1, \end{cases}$$

$$\Phi_2(t) = \begin{cases} 1 & t \in \mathcal{M}_2 \\ 0 & t \notin \mathcal{M}_2, \end{cases}$$

$$\Phi_3(t) = \begin{cases} 1 & t \in \mathcal{M}_3 \\ 0 & t \notin \mathcal{M}_3, \end{cases}$$

$$\Phi_4(t) = t,$$

$$\Phi_5(t) = \cos \mu_1 t, \quad \Phi_6(t) = \sin \mu_1 t,$$

$$\Phi_7(t) = \cos \mu_3 t, \quad \Phi_8(t) = \sin \mu_3 t,$$

$$\Phi_9(t) = \cos \mu_5 t, \quad \Phi_{10}(t) = \sin \mu_5 t, \quad t \in \mathcal{N}_{270},$$

i.e. the datums for the three individual intervals, the linear trend and three of the five generated trigonemetric terms. Their magnitudes computed by the least-squares technique, were as follows (in brackets the differences from the true values in percentages):

$$\begin{aligned} c_1 &= 2.312\ (131.2\%), \\ c_2 &= -1.600\ (-60.0\%), \\ c_3 &= 2.129\ (-29.0\%), \\ c_4 &= 0.064\ (540\%), \\ c_5 &= 0.983\ (96.6\%), \\ c_6 &= -0.229\ (54.2\%), \\ c_7 &= 0.107\ (\rightarrow \infty), \\ c_8 &= 0.912\ (-8.8\%), \\ c_9 &= 0.514\ (2.8\%), \\ c_{10} &= -0.923\ (7.7\%). \end{aligned}$$

The results are interesting because they show the order of inexactitude one can encounter when trying to eliminate a systematic influence from an incompletely known time series by using the least-squares technique. We have already pointed out the danger of doing so in the introduction to this paper. It is worth noting that the agreement is somewhat better for higher frequencies.

The optimum spectrum is drawn in Figure 7. We can notice that the dominant peak is located within the reading precision limits of the period 3.636 yr showing a contribution to the variance of the order of 98.5%. The peak corresponding to the second period (16 yr) has been shifted to 14.6 yr, i.e. by approx. 8.8%, indicating a

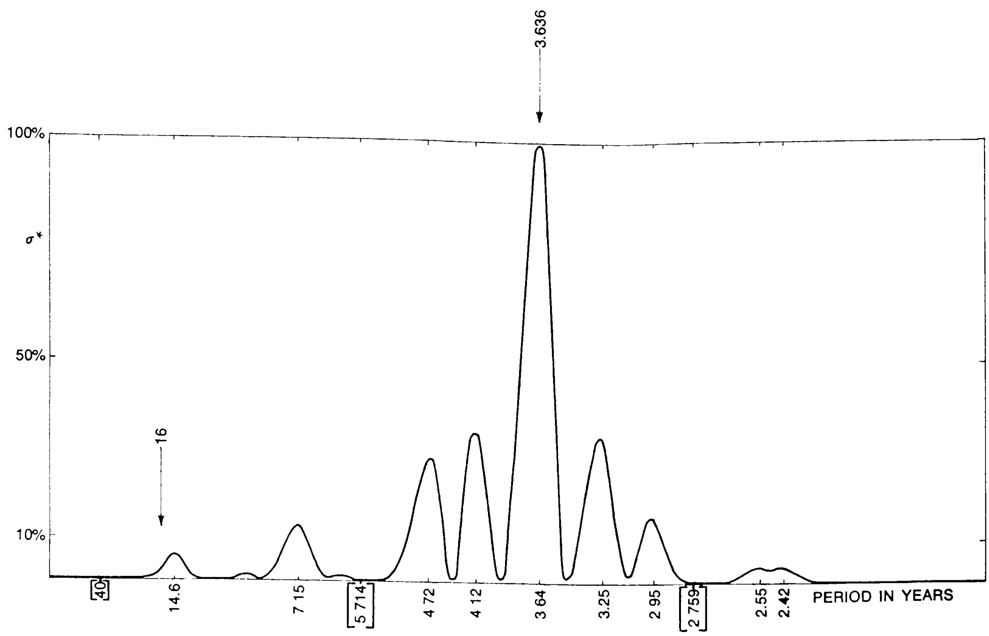

Fig. 7.

contribution of about 5.5%, and making hence the ratio of the two contributions close to 18. This corresponds roughly to the ratio of the squares of the amplitudes of the constituents: $(1+0.5^2)/0.25^2=20$. Further, we can see that the contributions of the remaining periods (40, 5.714 and 2.759 yr) are close to zero.

On the other hand, the side lobes are more pronounced than those found in Example No. 2. This is presumed to be due to the much more adverse shape of the analyzed function, and therefore, to the more adverse structure of the transformation (spectrum) needed. Besides, since there are several periodic terms involved we have to expect the presence of peaks due to the mutual interference of such terms.

The author is inclined to believe that the difficulties with side lobes and interference could be overcome to some extent by using following approach. Whenever a dominant peak shows up on spectrum that could be associated with a known geophysical phenomenon, and therefore attributed some physical meaning, the corresponding two trigonometric constituents can be subsequently taken as two additional "known" constituents and the process of analysis repeated. Thus we should eliminate the direct influence, as well as the side lobes and interference terms arising from the presence of the periodic term. The possibility cannot be ruled out that it might be feasible with the method to use a technique similar to 'hanning' or 'hamming' (see Blackman and Tukey, 1959, pp. 14, 98, 171), to smooth the spectrum by removing the side lobes to some degree. This again, however, is beyond the scope of this paper.

Acknowledgements

I wish to acknowledge that the research program was carried out under the kind sponsorship of the N.R.C. of Canada. I am particularly grateful to Mr T. Wray for his very thorough check of the mathematical formulae and many useful suggestions. My thanks go also to Mr W. A. Williamson for the editorial comments and to Mrs B. L. Murphy for her expert typing.

References

Berezin, I. S. and Zhidkov, N. P.: 1962, *Metody Vychisleniy*, Tom I, GIFML, Moscow.
Blackman, R.B. and Tukey, J.W.: 1959, *The Measurement of Power Spectra*, Dover, New York.
Faddeyev, D. K. and Faddeyeva, V. N.: 1964, *Vychislitel'nye Metody Lineynoy Algebry* (Czech. translation), SNTL, Prague.
Handscomb, D. C.: 1965, *Methods of Numerical Approximation*, Pergamon Press.
Hohn, F. E.: 1960, *Elementary Matrix Algebra*, MacMillan, New York.
Lanczos, C.: 1957, *Applied Analysis*, Pitman, London.
Munk, W. E. and Cartwright, D. E.: 1968, *Tidal Spectroscopy*, Philosophical Transactions of the Royal Society.
Quraishee, G. S.: 1969, 'Numerical Methods in Tidal Analysis and Prediction', Ph.D. Thesis, The University of Liverpool.
Quraishee, G. S. and Vaníček, P.: 1970, 'A Search for Low Frequencies in Residual Tide and Mean Sea Level Observations by Means of the Least Squares Spectral Analysis', Proceedings of the 1st International Symposium on Coastal Geodesy, Munich, in press.
Thompson, E. H.: 1969, *Introduction to the Algebra of Matrices with some Applications*, The University of Toronto Press.
Tidal Institute and Observatory, The University of Liverpool, *Annual Report 1968–69*, pp. 15–7.
Vaníček, P.: 1967, 'Anharmonic Analysis and Its Applications in Geophysics', C.Sc., Thesis, Czechoslovak Academy of Sciences, Prague.
Vaníček, P.: 1969a, 'Approximate Spectral Analysis by the Least Squares Fit', *Astrophys. Space Sci.* **4**, 387–91.
Vaníček, P.: 1969b, 'New Analysis of the Earth Pole Wobble', *Studia Geophys. Geodaet.* **13**, 225–30.
Vaníček, P.: 1969c, 'An Analytical Technique to Minimize Noise in a Search for Lines in the Low Frequency Spectra', Proceedings of the 6th International Symposium on Earth Tides, Strasbourg.
Wilks, S. S.: 1962, *Mathematical Statistics*, J. Wiley and Sons, New York.