Scanning Chapter Request

libweb@mcmaster.ca on behalf of Amanjot Bhullar [bhullas3@mcmaster.ca]
Sent: Monday, September 11, 2017 2:47 PM
To: Scanning, Journal

Submitted on Monday, September 11, 2017 - 14:47

Submitted values are:

Library : Thode Library
Book Title: The Study of variable stars using small telescopes
Chapter Title & Number: Searching for periodicity in astronomical data (p. 201 - 218)
Chapter page range: 201 - 218
Book Call Number: QB 835 .S84 1986
Your Name: Amanjot Bhullar
Your email address: bhullas3@mcmaster.ca
I am a McMaster : Undergraduate student
Faculty: are you uploading this item to Avenue to Learn (course reading)? No
Barcode (14 digit number on your McMaster ID): 29005012269045

McMaster University LIBRARIES

This copy has been made solely for use by a student, staff or faculty member at McMaster University for research, private study, review, criticism or new reporting purposes only, under the University's Fair Dealing Policy.

More information about the Fair Dealing Policy can be found at http://www.copyright.mcmaster.ca/content/fair_dealing_policy

Further copying, scanning, transmitting or otherwise making and distributing copies of this item without a proper license or the express consent of the copyright owner is not permitted.

If you still have questions about the use of this material please contact copyright@mcmaster.ca

SEARCHING FOR PERIODICITY IN ASTRONOMICAL DATA

A. W. Fullerton
David Dunlap Observatory
Department of Astronomy
University of Toronto
Toronto, Ontario
Canada M5S 1A1

ABSTRACT

The goals and principles of astronomical period searching are reviewed. Particular emphasis is placed on the distinction between the detection of signals and the estimation of their properties. Several of the most commonly used algorithms are described in general terms, and their respective merits and deficiencies are discussed. Difficulties which hinder all period analyses are illustrated and some remedies are prescribed.

I. INTRODUCTION

The tricks and techniques employed to extract "hidden" periodicities from a series of observations rank among the most important tools used by observers and analyzers of variable stars. The importance of these calculations arises from the bias towards periodic phenomena shared by both observers and theoreticians, a bias so strong that "variable" is often used to mean "periodically variable". The observer appreciates periodic signals because they lessen his work load: having established the properties of a periodic variation, the star doesn't have to be observed intensely because its behaviour is (more or less!) prescribed. For the theoretician, both extrinsic and intrinsic periodic phenomena betray the existence of an underlying regularity which may be used to interpret the star's behaviour or to probe its structure.

This paper reviews the most popular methods of period searching in use today. The proliferation of digital computers has been a tremendous boon to the development of these algorithms and has greatly facilitated the computations they entail. Unfortunately, computers are not a panacea! The results of all period search programs are subject to a variety of difficulties arising from the inherent imprecision of the observations, the characteristics of the sampling of the data, and the "noise" generated by the algorithm itself. These difficulties and ambiguities are discussed in the last two sections of the paper. The statistical properties of the period finding programs offer the only protection against the detection of spurious periods.

Consequently, the need for rigorous statistical analysis is stressed throughout this paper.

II. PRELIMINARY CONSIDERATIONS

II.1 Input Data

The input data consists of a series of N_o observations of a dependent variable, $f(t_i)$, obtained at times t_i where i ranges from 1 to N_o. The times t_i are not necessarily equally spaced. In astronomical applications the dependent variable $f(t)$ is often the differential magnitude of an object in a particular filter or the radial velocity of an object. Each observation is contaminated by some random uncertainty ("noise") due to measuring errors, atmospheric fluctuations, guiding errors, and so on. This random variable is usually assumed to be from a Gaussian distribution characterized by a mean value of 0 and a standard deviation of σ_o. Noise processes are also assumed to be independent of frequency or "white", although this model is not always the most appropriate one.

II.2 Goals

There are generally two distinct goals associated with period analysis. In the first instance the input data must be searched for periods, and the statistical significance of each of the possibilities encountered must be rigorously tested. This is the problem of *detection*, and may be characterized by the question "Is there a statistically significant periodic signal in this data set?". In this context, periodic signals with a low probability (usually 1 per cent or less) of arising by chance are taken to be significant. The statistical properties of any period search method, especially those involved in the assessment of significance, are of crucial importance to the detection problem. For, as Morbey (1978) states: "only very rarely will there be an appropriately scaled assemblage of data where no periodicity at all is indicated. Even in apparently random data some periods will be found."

Presuming that a statistically significant signal is detected in a data set, the second goal of the analysis is to estimate the properties of the signal (e.g. period, amplitude) and to attach uncertainties to these estimates. This second goal is one of *estimation* and is typified by the question "What are the parameters associated with the signal which has been detected?". The determination of the best estimates of the signal parameters and their uncertainties will also involve the statistical properties of the period search method employed.

These two steps are not always separated by researchers undertaking period analysis. The distinction between "detection" and "estimation" is slight, but the questions which characterize each of the goals are quite distinct. More importantly, the statistics and frequency space sampling which are appropriate to "detection" are quite different from those appropriate to "estimation" (see II.3 below). Of course, not every period search requires the application of detection statistics. The existence of a periodicity may be known *a priori* from either independent observations or theoretical considerations. For example, once a periodic spectroscopic variation has been detected in an object, *any* photometric variation associated with this period is *automatically* significant. Similarly, if a star is known to be a Cepheid variable, the existence of a periodicity does not have to be demonstrated; only an "estimation" period search is required to establish the parameters of the light curve. In general, however, the two separate goals must be distinguished and the appropriate statistical analysis performed in order to prevent the proliferation of spurious periods.

II.3 Period Searching in General

To locate "hidden" periodicities, the input data are manipulated and tested by a function which is "sensitive" to periods in a manner described in the next section. At trial frequency (or period) grid points, the test function is evaluated. It attains an extremum (either a maximum or a minimum depending on the nature of the function) when a sympathetic trial frequency is encountered. A plot of the period search function against trial frequency indicates periods which may be present in the input data. Such a plot is often termed a "periodogram", although in this paper these diagrams will be called "generalized periodograms" to distinguish them from the class of functions described in section III.3). Once a generalized periodogram has been obtained, further analysis is required to determine whether any of the extrema are significant (*i.e.*, whether any significant periods are detected) or to estimate the properties of a known signal. In principle these procedures will work only for strictly periodic signals (see following article by Willson). Consequently, any changes in the period which occur must do so over intervals which are very long compared with the duration of the observations. In practice, however, these techniques remain useful for signals with periods which vary slowly and slightly.

The characteristics of the frequency (period) grid used to evaluate the period search function depend on the purpose of the analysis. For period searches designed to detect unknown frequencies which may or may not be present, it is crucially important that the frequency grid points be independent, or at least "quasi-independent" of each other (Scargle 1982). This property is required to ensure the validity of the statistical test used to assess the significances of extrema. In this regard, a generalized periodogram calculation is nothing more than a numerical experiment. In all experiments, whether physical or numerical, factors are varied independently of each other to determine relationships and causes. If the manipulations are not arranged so that each trial is independent of the others, then no conclusions can be reached concerning the effects of the factors on the system. Similarly, if neighbouring periodogram frequency grid points are not independent of each other, then no conclusions can be reached about the effect of a particular trial frequency on the system defined by the input time series. However, independence of the grid points is not a serious consideration if the periodogram is designed to estimate properties of a known signal. In this case, "oversampling" in frequency space is equivalent to interpolation, which is certainly permitted when best estimates of parameters are sought (Black & Scargle 1982; Scargle 1982). Thus an arbitrarily fine frequency grid *is* permitted for estimation purposes, but is *not* permitted for detection purposes.

There are further considerations concerning the frequency information present in a particular data set, both of which are determined by the temporal sampling characteristics of the data. It is intuitively clear that the highest frequency (*i.e.*, shortest period) information contained in a data set is determined by the intervals between successive measurements (Kurtz 1983). In the case of equally spaced data, 3 points are necessary to define a sine curve, hence the minimum period (maximum frequency) which can be detected is $2\Delta t$. The corresponding frequency is known as the Nyquist frequency: $\nu_{Ny} = \nu_{Max} = (2\Delta t)^{-1}$. However, the Nyquist frequency is not well defined for input data obtained at unequally spaced time intervals, the usual case in astronomy (Scargle 1982). Theoretically such a data set contains information on periodicities down to $\Delta t = \min(t_i - t_{i-1})$. In practice, however, a pseudo-Nyquist frequency may be defined by averaging $\Delta t = (t_i - t_{i-1})$, where large, uncharacteristic temporal gaps are avoided. Alternatively, the harmonic mean of all Δt may be used. The result is that a useful pseudo-Nyquist frequency may be defined by $\nu_{Ny} = (2\langle\Delta t\rangle)^{-1}$.

The final consideration concerns the longest period (smallest frequency) information present in a particular data set. Clearly this time is given by the total time spanned by the data, T. The corresponding minimum frequency is given by $\nu_{min} = (T)^{-1}$. This is the minimum frequency grid increment that any "detection" periodogram could utilize, providing that all the grid points are quasi-independent of each other. In the Deeming (1975) and Scargle (1982) Fourier-type periodogram techniques the "window function" contains information about the degree of coupling between grid points. Quasi-independent grid points are typically found to be separated by ν_{min}.

III. AN OVERVIEW OF PERIOD SEARCH TECHNIQUES

Three broadly defined classes of techniques are used to search for periodicities and to estimate their properties. These are a) string length methods, b) phase binning methods, and c) Fourier transform (or equivalently least squares or periodogram) techniques. Several variations usually exist within each class but they all function in much the same way.

III.1 String Length Methods

String length methods seek to minimize the total length of the line segments (termed the "string") which join adjacent observations in phase space (in essence, to maximize the "smoothness" of the curve). For a trial period P the phases ϕ for the input data are calculated from

$$\phi = FRC[(t_i - t_o)/P]$$

where t_o is the adopted epoch and where FRC indicates that only the fractional part of the expression is kept. Next the observations are put in order within the phase space associated with the trial period, and the string length is calculated according to the formula given by Dworetsky (1983):

$$l(P) = \Sigma[(f_j - f_{j-1})^2 + (\phi_j - \phi_{j-1})^2]^{1/2} + [(f_1 - f_n)^2 + (\phi_1 - \phi_n + 1)^2]^{1/2}$$

Clearly, the string length is calculated by repeated application of the Pythagorean Theorem in phase space. The function $l(P)$ is calculated for the complete range of trial periods, and plotted against P. The smallest value of l occurs for the true period, when all the observations lie along a smooth curve. The underlying curve may be of arbitrary shape.

There are two variations on the basic method given by Dworetsky (1983). The Lafler & Kinman (1964) method was the first to utilize modern digital computers. However, it only calculates the one-dimensional string length corresponding to the vertical distance between adjacent data points in phase space. Neglect of the contribution of the horizontal component of the string length prevents biasing by those regions which happen to be more densely populated, but it also increases the noisy appearance of the generalized periodogram. Renson's (1978) formulation includes a weighting function which allows for the unequal distribution of the observations in phase space.

The basic advantage of these methods is their simplicity. The major drawback they suffer is that for each trial period the phased observations must be sorted and placed in increasing order by phase. Although efficient sorting algorithms exist this procedure remains computationally intensive, especially for large input data sets.

III.2 Phase Binning Methods

Techniques which "fold" the input data using a trial frequency or period and examine the contents of the bins in phase space are frequently used in astronomical studies. Stellingwerf's (1978) Phase Dispersion Minimization (PDM) technique is particularly popular. The methods described by Jurkevich (1971) and Morbey (1973, 1978) are conceptually similar to it.

In the PDM method, phase space (0.0,1.0) is divided into a number N_b of compartments ("bins"), typically 5, each of equal extent. The input data are phased with a trial frequency and are distributed among the bins accordingly. Usually a number, N_c, of overlapping bins are considered. The various sets of bins ("covers") are designed so that the outcome of the calculation will not depend on the distribution of the input data in the bins. A typical bin structure is characterized by $(N_b, N_c) = (5, 2)$; this structure is illustrated schematically in Figure 1. Each of the input data points falls into N_c bins.

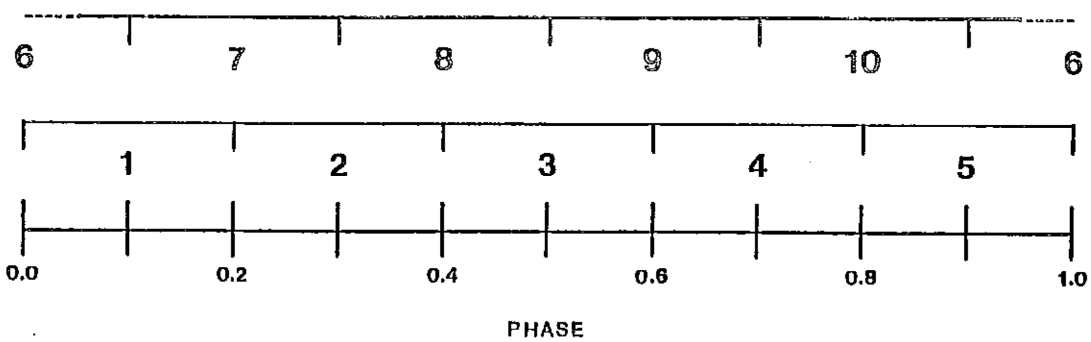

Figure 1. Schematic of PDM (5,2) bin structure. Phase space is divided into 5 equal bins, each extending for intervals of 0.2 in phase. There are 2 sets of "covers"; each set is offset from the other by 0.1 in phase. The 10 resulting bins are labelled. Clearly each data point falls into 2 bins.

For a particular trial frequency, the variances of the data assembled in each of the $N_b N_c$ bins are calculated and summed. The sum of the bin variances is compared with the global variance of the input data. Most trial frequencies will produce a random distribution of data points in phase space and so the ratio of the total bin variances to the global variance will be approximately unity. However, for the trial frequency which corresponds to a true period the scatter within the phase bins will be reduced dramatically and the ratio of the two variances will be a small number. In a plot of the ratio of the two variances *vs.* frequency the frequency of the true periodicity is indicated by a deep minimum. The depth is an indication of the significance of the fit for *it reveals the degree to which the scatter about the mean light curve is reduced.*

Like the string length methods, phase bin techniques make no assumptions about the shape of the variation being studied. However, the phase bin methods are not usually as computationally intensive as the string length methods, primarily because the data do not have to be ordered within the phase bins.

III.3 Fourier Transform, Periodogram, and Least Squares Methods

The Fourier transform and related techniques are all based on the discrete Fourier transform and employ the elegant concepts embodied in harmonic analysis. Although apparently distinct, all of these period search methods are conceptually similar. Fourier Transform and periodogram techniques are very closely related, for the latter are essentially the squared modulus of the former, suitably normalized (Deeming 1975; Scargle 1982). Least squares fitting techniques are formally equivalent to periodogram analysis under certain circumstances (Scargle 1982).

The current popularity of these techniques is largely due to the computational efficiency of the Fast Fourier Transform (FFT) routine. Unfortunately this algorithm requires input data which are equally spaced in time, an unusual circumstance in optical astronomy. Even in cases where the FFT is applicable (*e.g.*, Middleditch & Nelson 1973) the periodogram or "power spectrum" is usually calculated, at the expense of some phase information.

An extremely useful formulation of the periodogram function has been developed by Scargle (1982):

$$P_X(\omega) = \frac{1}{2}[C^2(\omega) + S^2(\omega)]$$

where

$$C(\omega) = A(\omega)\Sigma X(t_j)\cos(\omega[t_j - \tau])$$

$$S(\omega) = B(\omega)\Sigma X(t_j)\sin(\omega[t_j - \tau])$$

$$A(\omega) = (\Sigma\cos^2(\omega[t_j - \tau]))^{-1/2}$$

$$B(\omega) = (\Sigma\sin^2(\omega[t_j - \tau]))^{-1/2}$$

and where $\tau(\omega)$ is defined by

$$\tan(2\omega\tau) = (\Sigma \sin 2\omega t_j)/(\Sigma \cos 2\omega t_j)$$

In these expressions, ω is the angular frequency, $\omega = 2\pi\nu = 2\pi(P)^{-1}, X_j$ are the input data obtained at arbitrarily spaced times t_j, $j = 1$ to N_o, and $\tau(\omega)$ is a phase function designed to make this periodogram formally equivalent to least squares fitting techniques. To understand how this function is sensitive to frequencies present in the data, suppose that the data set X_j contains a sinusoidal signal of angular frequency ω_o. When the trial frequency is not close to ω_o the input data X_j and the $\cos(\omega[t_j - \tau])$ and $\sin(\omega[t_j - \tau])$ terms are out of phase. The terms in the summations have random signs with respect to each other and they largely cancel out. However, when $\omega \simeq \omega_o$ the data and the sinusoidal factors are nearly in phase and the terms in the summations add coherently so that P_X attains a relatively large value. Thus the periodogram reveals hidden periodicities by "resonating" with them. The Scargle formulation is somewhat more complicated than its predecessor, the famous Deeming algorithm (Deeming 1975), but it possesses very desirable statistical properties which make its use preferable (Horne & Baliunas 1985). An example of a periodogram obtained by this method is contained in the paper by Stagg, elsewhere in this volume.

By design, the calculation of a Scargle periodogram is equivalent to fitting sinusoids to the data by the method of linear least squares. Least squares techniques are conceptually straightforward; descriptions are given by Vanicek (1969, 1971), Lomb (1976), and Ponman (1981). At each of the trial frequencies ω a linear combination of sine and cosine functions with argument $(\omega t - j)$ are fit to the data by least squares. The sum of the squared amplitudes from both of these terms is plotted against frequency to form a graph which closely resembles a periodogram (*cf.* Lomb 1976, Figures 1 and 2). Peaks in the graph indicate periods; the heights of the peaks indicate significances through their relationship with the F statistic.

In addition to the linear least squares approach (where the frequency enters as a parameter), nonlinear least squares calculations have been used to solve iteratively for the period(s) and amplitude(s) which best fit the data (*e.g.*, Wizinowich & Percy 1979). These methods require good initial estimates for the period to prevent convergence to a spurious period located in a local rather than the global minimum in the parameter space.

Fourier techniques and their relatives differ from the other two classes of methods in that the signal is assumed to have a definite functional form (sinusoidal). This would appear to limit the usefulness of these techniques for the analysis of nonsinusoidal variations (*e.g.*, Cepheid light curves). In principle Fourier methods are applicable to arbitrary waveforms through the Fourier Decomposition Theorem which states that any periodic signal can be decomposed into a (possibly infinite) summation of sinusoids. The frequency of each of these sinusoids will be an integral multiple of the fundamental frequency ν_o, *e.g.*, $2\nu_o$, $3\nu_o$, $4\nu_o$, etc. For example, adequate representations of Cepheid light curves usually require summation over the first 4 to 8 harmonics (Moffett & Barnes 1985). Many more Fourier components are required to reconstruct the light curve of a typical eclipsing binary.

However, some practical difficulties arise owing to the effects of aliasing and noise (see section V) and there is considerable risk that the higher harmonics may not be recovered. Under these circumstances the other "non-parametric" methods provide better estimates of the properties of the signal. However, for detection purposes Fourier methods remain valid even for significantly nonsinusoidal signals (Black & Scargle 1982). Despite the distribution of the signal over more than one Fourier component, most of the "power" in curves typical of Keplerian orbits, for example, remains at the fundamental frequency (Jensen & Ulrych 1973). Thus the presence of at least the fundamental frequency should be detectable in any data set of adequate quality (with the exception of many eclipsing binary light curves) although other techniques may be required to estimate the harmonic content of the waveform.

IV. TESTS AND COMPARISONS

From the descriptions of their operation, it is clear that each of the three classes of period search techniques is able to estimate the properties of periodic signals. Series of tests performed by Perez de la Blanca & Garrido (1980) and by Heck, Manfroid, & Mersch (1985) on various subsets of techniques confirm this capability. The tests performed by Heck, Manfroid, & Mersch are particularly impressive. These researchers concluded that the non-parametric PDM and "Renson" methods are slightly (probably insignificantly) preferable to the other methods though they also comment that non-parametric and Fourier methods are complementary to some degree. Similar tests performed by various workers at the University of Toronto give comparable results, although a preference for periodogram techniques has been established. Heck, Manfroid, & Mersch find all the methods to be roughly equal in computational efficiency with the exception of the PDM method which is slower because of its complicated bin structure.

However, the equivalence among these techniques vanishes when the problem of detection is considered, primarily because of the lack of appropriate statistics for all methods except the Scargle periodogram (and, by extension, least squares fitting). The factor which the other methods fail to account for is the "statistical penalty" exacted for examining many frequencies (Scargle 1982). The origin of this penalty is best understood by considering a periodogram (now taken in the general sense) to be a numerical experiment. Since the "experiments" a periodogram performs are inherently noisy, large random fluctuations are expected to occur. Clearly if the experiment is repeated often enough (*i.e.*, if enough trial frequencies are searched) the chance of encountering a noise spike of any specified height increases and approaches certainty as the number of experiments tends to infinity. If one large periodogram peak is uncovered among the N frequencies examined, the fact that there were $(N-1)$ non-detections must be accounted for when the significance of the peak is assessed.

Within the context of Scargle's modified periodogram the statistical penalty for searching many frequencies is embodied in the threshold height which a peak must attain to be considered significant. This height is usually expressed in units of the noise, *i.e.*, it is a power signal-to-noise (S/N) ratio. As the number of frequencies searched increases so must the height of a peak in order to achieve a specified degree of statistical significance. Thus, if z is the power S/N threshold for 1 per cent significance, then z is given by

$$z(N;0.01) = 4.6 + \ln(N)$$

If the detection periodogram is calculated at 100 frequency points the power S/N threshold for 1 per cent significance is 9.2 (corresponding to an amplitude S/N ratio of 3.0); if 500 frequencies are searched this number becomes 10.8 (amplitude S/N ratio of 3.3). Of course each of the frequency grid points must be quasi-independent in order for this test to be valid (section II.3).

Only the statistics of the Scargle periodogram account for the statistical penalty. Heck, Manfroid & Mersch (1985) have also recognized that the other methods are inappropriate for detection purposes, for they state that "these techniques should then stay in an essentially descriptive scheme. When ν has to be estimated none of these criteria has a known distribution, not even Fourier's". The last comment refers to the Deeming periodogram; Scargle's modifications remedy this difficulty.

The reasons for these shortcomings lie in the complex, essentially undetermined, statistical properties of the non-parametric methods. These properties are exceedingly difficult to evaluate analytically. Consider, for example, the PDM method. Individual data points occupy more than one bin, consequently the various bins are not independent of each other! Rigorous detection statistics are probably not derivable under these circumstances. The only approach available for the non-parametric functions is to undertake extensive computer simulations with random data (obtained by shuffling the temporal order of the actual data, for example) to derive "empirical" significance estimates (Nemec & Nemec 1985). An advantage of these simulations is that they make no assumption about the distribution of the noise. However, they can be quite costly in computer time, particularly when large data sets are involved.

V. COMPLICATIONS

Unfortunately, all the procedures for performing period analyses described in the preceeding sections suffer from several complications. These difficulties arise from the imprecision inherent in the input data, and from interactions with the sampling "window" through which the data are obtained. They may be overcome under many circumstances.

V.1 Aliasing and Pseudo-Aliasing

The temporal sampling characteristics of the input data give rise to insidious difficulties known as aliasing and pseudo-aliasing. As the term suggests, "aliasing" inhibits and sometimes precludes the identification of the correct signal frequency. Instead a related "alias" frequency is detected. This phenomenon arises when the input data are sampled at equally spaced time intervals, Δt. This temporal regularity permits sinusoids of frequencies $\nu_o \pm n(\Delta t)^{-1}$ to fit the data equally well (Figures 2a) and 2b)); in this expression ν_o is the true signal frequency and n is an integer. Fortunately, astronomical observations are seldomly spaced equally in time, usually because of the follies of the weather. Even small deviations from equal sampling are sufficient to make aliasing unimportant (Figure 3). As a result, aliasing is not usually a problem for astronomical time series.

However, a similar phenomenon termed "pseudo-aliasing" by Scargle (1982) (though most people do not distinguish the two phenomena) is almost always present in astronomical observations. Pseudo-aliasing causes the same sort of confusion over the identity of the true period as genuine aliasing although it originates with the cycle-count ambiguities which exist in gapped data. Put another way, pseudo-aliasing is caused by the presence of a regularity in the overall observing pattern. For example, the periodogram of observations obtained on consecutive nights will show strong ± 1 d^{-1} pseudo-aliases; observations made on an annual basis possess ± 1 yr^{-1} pseudo-aliases, and so on. Without resorting to gap-filling algorithms (see V.2), the only ways of avoiding pseudo-aliasing consist of restricting attention to observations of rapid phenomena on a single night or arranging for essentially continuous unequally spaced observations from a satellite, from telescopes situated at a variety of longitudes, or from a telescope located at one of the Earth's poles (during the appropriate season, of course!).

Although pseudo-aliasing can make period analysis impossible, there are usually ways of dealing with it. This is particularly true of Fourier techniques which

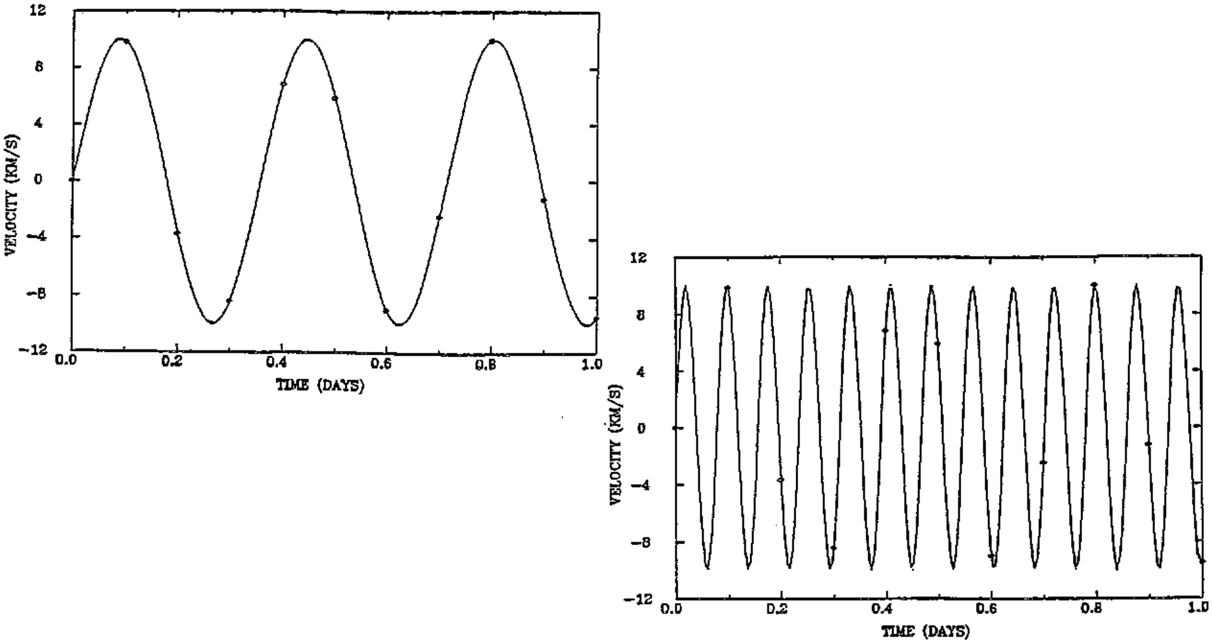

Figure 2. An illustration of aliasing. In this simulation, the data points (◇) were obtained at intervals separated by exactly 0.1 day, with "infinite" precision. (a) A sinusoid with amplitude 10 kms^{-1} and period 0.36 day (frequency 2.8 d^{-1}) fits the "observations" quite well. (b). However, the same data are fit equally well by a sinusoid with amplitude 10 kms^{-1} and period 0.08 day (frequency 12.8 d^{-1}). Thus the actual periodicity present in the data cannot be determined uniquely. The two possible periods are "aliases" of each other.

permit the calculation of a "window function" (*i.e.*, a periodogram of the independent variable). All information regarding the temporal sampling of a particular data set is contained in the window function (Deeming 1975; Scargle 1982). Comparison of the window function peaks with the pattern of pseudo-aliases in the periodogram of the data can aid identification of the true frequency, which is usually, though not always, the highest peak in the periodogram. Gray & Desikachary (1973) have formalized this procedure through their method of cross-correlating the window function with the periodogram. This approach utilizes all the information contained in the pattern of the pseudo-aliases, but it does not guarantee the successful identification of the genuine period. In most situations "prewhitening" can be used to determine which of a group of peaks corresponds to the physical period (see VI.1).

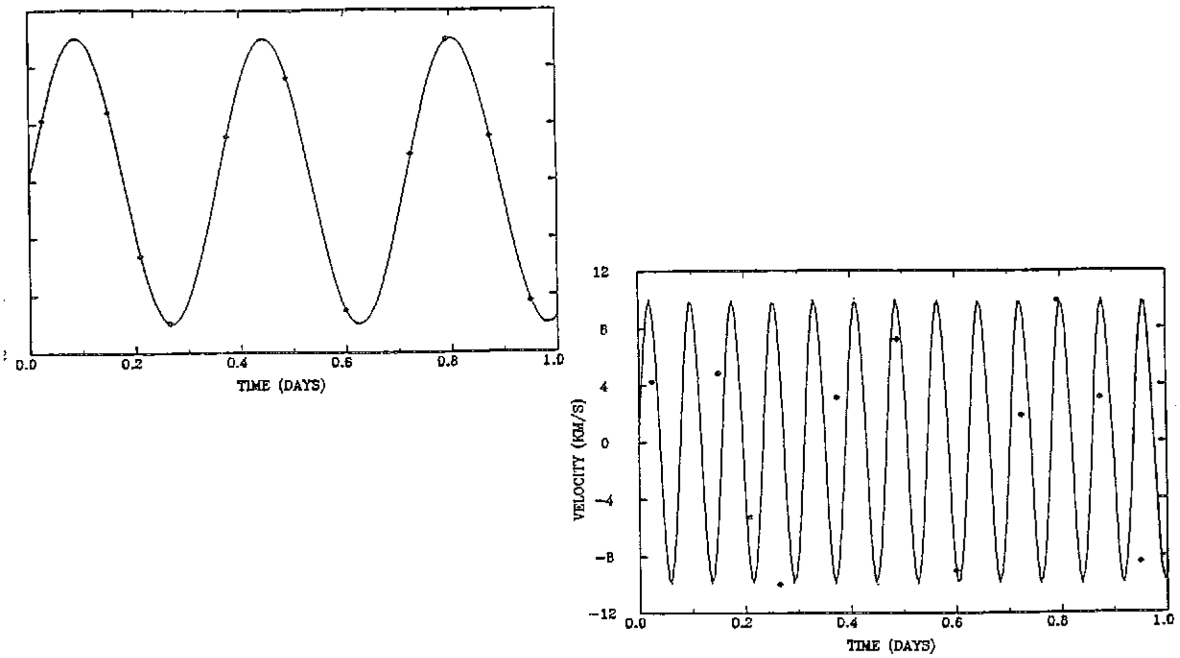

Figure 3. The same as Figure 2, only this time the data were sampled at unequally spaced times. The period of the variation is no longer ambiguous. Apparently the longer of the two periods illustrated in Figure 2 is the correct one.

V.2 Filling Gaps in the Data

Geophysicists frequently use the "maximum entropy method" (MEM) to fill temporal gaps in a data set, thereby preventing the occurrence of pseudo-aliasing. This technique uses the observations from one night as a template to link successive nights in a continuous and "maximally noncommittal way". Fahlman & Ulrych (1982) have demonstrated the power of this technique for astronomical applications. However, the usefulness of MEM suffers from its inability to utilize unequally spaced data as input. Nevertheless, MEM has very desirable frequency resolution characteristics, so even a single night's observations can yield good period estimates for stars which vary sufficiently rapidly (Percy 1977).

V.3 Noise Estimation

As discussed in section 4, the only protection against "detecting" spurious periods lies in rigorous statistical analysis. Accurate estimation of the noise level present in a data set is of considerable importance since the S/N ratio is the quantity by which the significance of a periodogram peak is assessed.

There are a variety of approaches for estimating the noise in a data set. The external precision of the observations (*e.g.*, the standard deviation of (Check-Comparison) observations in photometry or the mean error of a radial velocity measurement) provides a straightforward estimate. In principle the noise may be determined directly from a Fourier-type periodogram, under the assumption that the noise is locally white. This strong assumption is necessary because photometric observations suffer low frequency fluctuations due to hourly and nightly changes in atmospheric conditions. By averaging the periodogram over a region near the peak of interest (and preferably free of pseudo-aliases since these features are not random fluctuations) a good estimate of the mean square noise, σ_o^2, is usually obtained. Horne & Baliunas (1986) demonstrate that the total variance of the input data (in the time domain) is the *only* acceptable noise estimator for use with the Scargle periodogram if the valuable statistical properties of this algorithm are to be retained.

VI. SECOND ORDER COMPLICATIONS

In addition to the complications which plague attempts to recover single periodicities from a data set, further difficulties are associated with the analysis of multiperiodic or quasi-periodic phenomena.

VI.1 Prewhitening and the Detection of Multiple Periods

Prewhitening is a term used to indicate the removal of a signal at a particular frequency from a data set. Usually this is accomplished by fitting sinusoids of the appropriate frequency to the data by least squares. The residuals from the fit form the prewhitened data.

Prewhitening has two uses in time series analysis. The first, alluded to in section V.3), is to aid in the identification of a (single) "true" periodicity under conditions of unfavourable pseudo-aliasing. To decide which of the statistically significant peaks corresponds to the real frequency, the data are prewhitened for each of the aliased frequencies in turn and periodograms of the residuals are examined. When a pseudo-alias frequency has been removed the other peaks remain since a "physical" signal is still present. Often the periodogram of the residuals contains new peaks and the overall noise level rises. In contrast, prewhitening with the true signal frequency removes all the pseudo-alias peaks from the periodogram of the residuals; the overall noise level usually decreases. Thus the true frequency may be distinguished from the pseudo-aliases in many cases.

More frequently prewhitening is used to search a data set for more than one periodicity. This procedure involves the detection of a period through periodogram analysis, its subsequent removal by prewhitening, and reanalysis of the residual data for more periods. In principle this step-by-step approach is valid for analyzing a linear combination of sinusoidal, multi-periodic signals. In practice the detection of more than one period is very difficult, primarily because the prewhitening procedure destroys the independence of the input data. This is particularly true if the prewhitening fits possess large uncertainties. As a result, the statistical significances of the peaks in the periodogram of the prewhitened residuals are impossible to evaluate analytically.

This bleak outlook is corroborated by a series of tests on synthetic data performed by Walker, Pike, & Hartley (1984). Their conclusions are worth repeating:

1) "It is well known that any noisy data set can be adequately (*i.e.*, down to the noise level) modelled by a power spectrum of sufficient complexity. Typical large data sets on a variety of stars often require 10 frequencies.

2) The power spectrum for typical data sets is not unique and if the wrong frequency is used at any stage in the complex fitting and prewhitening procedure then a different frequency set will be found which will fit the data equally well.

3) A typical data set is so poorly constrained that the total number of frequency sets which will adequately represent the data is large and may be infinite. Alternately we might summarize by saying that generally one needs a lot more data than one originally thought...."

Often, no more than one periodicity can be reliably extracted from a data set.

VI.2 Detection of quasi-periods (timescales)

The period search techniques discussed in this paper are valid for "strictly" periodic signals only (see section II.3). However, there are numerous phenomena in astronomy which are not periodic but which do change over characteristic time intervals (*e.g.*, photometric variations of supergiants, Be stars, spotted rotating stars, quasars). These timescales convey useful physical information about the origins of the variability, and are therefore worthy of detailed study.

Autocorrelation methods have been used successfully to estimate timescales (Burki, Maeder, & Rufener 1978; Percy, Jakate, & Matthews 1981; Baliunas *et al.* 1983). These techniques differ from those discussed in section III in that the analysis occurs in the time domain. All autocorrelation methods seek to determine the time lag required for the observations to recognize or report themselves. This is accomplished by introducing a series of discrete time shifts to the data. At each step in the procedure the shifted data are compared with the unshifted data by means of a "correlation function". A simple example of such a function is just the sum of the squared differences between the parts of the shifted and unshifted data which overlap. A plot of the correlation function *vs.* time delay has a large dip when the data sets are in approximate registration, *i.e.*, when the shifted data "recognizes" itself in the unshifted data. This value of the time delay corresponds to the timescale, or at least to a harmonic or multiple of the actual timescale of the variation.

As with the frequency domain analysis of strictly periodic phenomena, statistical significances must be calculated for the dips observed in autocorrelation diagrams. Baliunas *et al.* (1983) discuss a method of calculating statistical significances for their autocorrelation technique. Heck, Manfroid, & Mersch (1985) examined the performance of an autocorrelation algorithm in their extensive series of tests of period search methods. As expected, they found that it was not as reliable as the other methods for estimating strictly periodic signals. However, they stress that the real utility of the autocorrelation method lies with the problem of determining quasi-periods.

VII. Concluding Remarks

In their famous treatise on mathematical methods, Jeffreys & Jeffreys (1953) make the following remarks on period searching: "Without some such precaution periodicities found by harmonic analysis and not predicted by previous theoretical considerations should be mistrusted, as many complications are capable of giving rise to spurious periods; not more than a tenth of those that have been asserted will bear a proper statistical examination."

The theme of this paper has been to echo these sentiments. In particular, the distinction between detection ("periodicities found by harmonic analysis") and estimation ("predicted by theoretical considerations", or by independent observations) has been emphasized. The importance of rigorous statistics, particularly for the detection of suspected signals, has also been stressed, primarily to improve upon the ten per cent success rate suggested by Jeffreys & Jeffreys!

Even with rigorous statistical analysis, the complexities associated with period detection and estimation can sometimes confound a researcher. In such cases only a thorough knowledge of the behaviour of the period search method used will prevent the proliferation of spurious periods in the literature. Thus, every method should be tested comprehensively using both simulated and real data. Careful, conservative analysis of "detection" periodograms will eliminate a great many spurious periods, but, in the final analysis, further observations will always tell the tale!

Acknowledgements

It is with great pleasure that I record my gratitude to Drs. C. T. Bolton and J. R. Percy for their encouragement and instruction in learning the black art of period searching. Many "periodic" discussions with Dr. D. R. Gies, C. R. Stagg, and M. F. Bietenholz have been extremely valuable to the formulation of my thoughts, as have comments from D. L. Welch and L. M. Oattes. Finally, financial support from the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged.

REFERENCES

Baliunas, S. L. *et al.* . (1983) *Ap. J.*, **275**, 752.
Black, D. C. & Scargle, J. D. (1982). *Ap. J.*, **263**, 854.
Burki, G., Maeder, A., & Rufener, F. (1978). *Astr. Ap.*, **65**, 363.
Deeming, T. J. (1975). *Ap. Space Sci.*, **36**, 137.
Dworetsky, M. M. (1983). *M. N. R. A. S.*, **203**, 917.
Fahlman, G. G. & Ulrych, T. J. (1982). *M. N. R. A. S.*, **199**, 53.
Gray, D. F. & Desikachary, K. (1973). *Ap. J.*, **181**, 523.
Heck, A., Manfroid, J., & Mersch, G. (1985). *Astr. Ap. Suppl.*, **59**, 63.
Horne, J. H. & Baliunas, S. L. (1986). *Ap. J.*, **302**, in press.
Jeffreys, H. & Jeffreys, B. S. (1953). *Methods of Mathematical Physics*, 3rd edition. p. 452. London: Cambridge University.
Jensen, O. G. & Ulrych, T. J. (1973). *A. J.*, **78**, 1104.
Jurkevich, I. (1971). *Ap. Space Sci.*, **13**, 154.
Kurtz, D. W. (1983). *Inf. Bull. Var. Stars*, # 2285.
Lafler, J. & Kinman, T. D. (1964). *Ap. J. Suppl.*, **11**, 216.
Lomb, N. R. (1976). *Ap. Space Sci.*, **39**, 447.
Middleditch, J. & Nelson, J. (1973). *Astrophys. Lett.*, **14**, 124.
Moffett, T. J. & Barnes, T. G. (1985). *Ap. J. Suppl.*, **58**, 843.
Morbey, C. L. (1973). *Pub. D. A. O.*, **14**, 185.
Morbey, C. L. (1978). *Pub. D. A. O.*, **15**, 105.
Nemec, A. F. L. & Nemec, J. M. (1985). *A. J.*, **90**, 2317.
Percy, J. R. (1977). *M. N. R. A. S.*, **181**, 647.
Percy, J. R., Jakate, S. M., & Matthews, J. M. (1981). *A. J.*, **86**, 53.
Perez de la Blanca, N. & Garrido, R. (1981). In *Workshop on Pulsating B Stars*, eds. G.E.V.O.N. & C. Sterken, p. 285. Nice: Nice Observatory.
Ponman, T. (1981). *M. N. R. A. S.*, **196**, 583.
Priestly, M. B. (1981). *Spectral Analysis and Time Series*, Chapter 6, London: Academic.
Renson, P. (1978). *Astr. Ap.*, **63**, 125.
Scargle, J. D. (1982). *Ap. J.*, **263**, 835.

Stellingwerf, R. F. (1978). *Ap. J.*, **224**, 953.

Vanicek, P. (1969). *Ap. Space Sci.*, **4**, 387.

Vanicek, P. (1971). *Ap. Space Sci.*, **12**, 10.

Walker, E. N., Pike, C. D., & Hartley, K. F. (1984). In *Space Research Prospects in Stellar Activity and Variability*, eds. A. Mangeney and F. Praderie, p. 151. Paris: Observatoire de Paris-Meudon.

Wizinowich, P. & Percy, J. R. (1979). *Pub. A. S. P.*, **91**, 53.