Multiperiodicity, modulations and flip-flops in variable star light curves

I. Carrier fit method

J. Pelt1, N. Olsper1, M. J. Mantere2, and I. Tuominen2,*

1 Tartu Observatory, 61602 Tõravere, Estonia
e-mail: pelt@aai.ee
2 Department of Physics, Gustaf Hällströminkatu 2a (PO Box 64), 00014 University of Helsinki, Finland

Received 14 March 2011 / Accepted 13 September 2011

ABSTRACT

Context. The light curves of variable stars are commonly described using simple trigonometric models, that make use of the assumption that the model parameters are constant in time. This assumption, however, is often violated, and consequently, time series models with components that vary slowly in time are of great interest.

Aims. In this paper we introduce a class of data analysis and visualization methods which can be applied in many different contexts of variable star research, for example spotted stars, variables showing the Blazhko effect, and the spin-down of rapid rotators. The methods proposed are of explorative type, and can be of significant aid when performing a more thorough data analysis and interpretation with a more conventional method.

Methods. Our methods are based on a straightforward decomposition of the input time series into a fast “clocking” periodicity and smooth modulating curves. The fast frequency, referred to as the carrier frequency, can be obtained from earlier observations (for instance in the case of photometric data the period can be obtained from independently measured radial velocities), postulated using some simple physical principles (Keplerian rotation laws in accretion disks), or estimated from the data as a certain mean frequency. The smooth modulating curves are described by trigonometric polynomials or splines. The data approximation procedures are based on standard computational packages implementing simple or constrained least-squares fit-type algorithms.

Results. Using both artificially generated data sets and observed data, we demonstrate the utility of the proposed methods. Our interest is mainly focused on cases where multiperiodicity, trends or abrupt changes take place in the variable star light curves.

Conclusions. The presented examples show that the proposed methods significantly enrich the traditional toolbox for variable star researchers. Applications of the methods to solve problems of astrophysical interest will be presented in the next papers of the series.

Key words. stars: variables: general – methods: data analysis – methods: statistical

1. Introduction

The light curves of variable stars can display very different modes of behaviour. They can be quite regular, often even periodic, or relatively fast changes can be detected. The methods to deal with such observed data can roughly be divided into two classes. The first class encompasses tools that are based on Fourier analysis methods, assuming that the light curve can be decomposed into a set of periodic (harmonic) components. The components themselves can then be interpreted as manifestations of, e.g., different pulsational modes (as in asteroseismology), rotation, orbiting, and so forth. The local behaviour of such multicomponent curves can be quite complex. For instance, it is quite easy to build simple examples where destructive interference between two nearby components with nearly equal frequencies allows some third component to dominate the observed curve. The detection of such local regularities often lead to the usage of the second type of analysis tools, usually referred to as the time-frequency analysis methods. Instead of a global model, a certain set of local variability elements is postulated, and then their change in time is followed. Probably the best known method of this kind is the wavelet transform.

For continuously observed signals the Fourier transform-based methods and different types of time-frequency transforms are a well-developed part of the traditional data processing work flow (see e.g. Poularikas 2010). The situation is different in the case of astronomical data. The irregular observing patterns and long time gaps in the measured series significantly complicate the analysis (see e.g. the review by Schwarzenberg-Czerny 2003, and references therein; or Chapter 5 in Aerts et al. 2010). For the first type of methods, the computed frequency spectra often show spurious peaks due to the regularities in the gap structure of the input data. The second class of methods, on the other hand, can only be used for the short, more or less continuous, subsets of data.

In this paper we propose a relatively new method, at least in the astronomical context, to deal with irregularly spaced data. We start from the simple physical assumption that even if an object is observed to exhibit irregular or semi-regular variability, some regular process can be identified in the background. The best example of this is stellar rotation, normally quite stable over long periods of time, the semi-regularity of the observed light curve being caused by inhomogeneities, such as dark spots, on the stellar surface. Another relevant setting is provided by accretion disks, where the Keplerian rotation of the accretion elements is more regular than the observed outflow of radiation.
Using this assumption, we decompose the observed light curves into two components: a rapidly changing carrier tracing the regular part of the signal, and its slowly changing modulation. In this way, we construct a method combining both aspects of the two classes of methods presented above. The proposed method is certainly not a new method to seek hidden periods from data with gaps and irregularities. Instead, we try to build continuous models for irregularly observed data using a priori known frequency. Proper visualization of the model can then help us to interpret the data at hand and to reveal possible physical effects. The method is of explorative nature and can be used as a starting point before building more sophisticated statistical, numerical or dynamical models.

The layout of the paper is somewhat non-traditional. Instead of commencing with a lengthy introduction and comparison of the new proposed method to various time series analysis methods that have been used in variable star research, we start with introducing the basic elements of the analysis method in Sect. 2, where we also give a detailed description of the algorithms used. Section 3 is dedicated to demonstrate the suitability of the proposed method in various astronomical settings, using both artificially generated and observational data sets. This allows the reader to get appreciation of the new method at work. In Sect. 4 we will give an overview of the related developments, discuss the new method in the light of previous work, and in Sect. 5 discuss the domains of its applicability. The full-scale application of the method to longer sets of observational data will be presented in the forthcoming papers in the series.

2. Carrier fit method

The notion of the carrier frequency is well known from communication theory. Very often this is just a single harmonic tone with a fixed frequency \( f_0 \). The carrier is modulated by a signal of interest and transmitted to receiver. In the receiver the modulated signal is demodulated and useful information is recovered. The simplest model for such a process is just a waveform

\[
    f(t) = a(t) \cos(2\pi f_0 t) + b(t) \sin(2\pi f_0 t),
\]

(1)

where the carrier wave is modulated by two smooth low-frequency signal components \( a(t) \) and \( b(t) \). In some cases the signal can be represented in an analytic complex form

\[
    f_a(t) = A(t)e^{i\varphi(t)},
\]

(2)

which allows the explicit tracking of the time dependent amplitude \( A(t) \), the phase \( \varphi(t) \), and most importantly, the instantaneous frequency

\[
    \nu(t) = \frac{d\varphi(t)}{dt}.
\]

(3)

The real astronomical signals very often contain also higher harmonics of the basic frequency. Due to this we introduce, from the very beginning, a more complex model of the form

\[
    f(t) = a_0(t) + \sum_{k=1}^{K} (a_k(t) \cos(2\pi k f_0 t) + b_k(t) \sin(2\pi k f_0 t)).
\]

(4)

By including the term \( a_0(t) \), we allow the mean level of the signal to be different from zero. We also assume that the total number of harmonics, \( K \), is not very large.

2.1. The algorithm

To describe the details of the carrier fit method we start from the simplest model of a variable star light curve:

\[
    f(t) = a(t) \cos \left( \frac{2\pi t}{P_0} \right) + b(t) \sin \left( \frac{2\pi t}{P_0} \right),
\]

(5)

where \( P_0 \) is the period (measured in days) corresponding to the carrier frequency \( \nu_0 = 1/P_0 \) (measured in cycles per day), and the two functions \( a(t) \) and \( b(t) \) describe slow changes of the curve. Here we define a slow process as one for which the characteristic changes of the functions \( a \) and \( b \) occur on time-scales significantly longer than the period of the carrier oscillation, \( P_0 \).

To build approximation curves for the observed data we firstly need to have a proper value for the carrier period \( P_0 \), the determination of which is discussed in detail in Sect. 2.5. Secondly, certain analytical or numerical models for the both slow modulating functions are required. These models themselves depend on certain sets of parameters, as discussed in Sects. 2.2 and 2.3 for the chosen two representations. After reliably determining \( P_0 \) and formulating a suitable model for the modulators, the remaining task is to find the model that best fits the observations, and retrieve the corresponding parameters. This data fitting process itself can be formulated as a standard least-square approximation procedure. This procedure can, in general, involve both linear and non-linear parameters. Because of the overall complexity of non-linear fitting, we usually restrict our analysis to the linear regime.

Below we demonstrate the carrier fit tool using two different methods for smooth approximation of the modulators: trigonometric polynomials and splines. There are certainly many other possibilities – essentially every known method of data smoothing from literature can also be used in this context.

2.2. Trigonometric approximation

We start from the relatively transparent method of trigonometric approximation. This allows us to explicitly control the frequency domain behaviour of our algorithm.

Let the time interval \([t_{min}, t_{max}]\) be the full span of our input data. Then we can introduce a certain period \( D = C \times (t_{max} - t_{min}) \) for which the coverage factor \( C \) is larger than unity (typically \( C = 1.1 - 1.5 \)). Using the corresponding frequency, \( \nu_D = 1/D \), we can now build a trigonometric (truncated) series of the type:

\[
    a(t) = c^a_0 + \sum_{l=1}^{L} \left( c^a_l \cos(2\pi l \nu_D t) + s^a_l \sin(2\pi l \nu_D t) \right),
\]

(6)

and

\[
    b(t) = c^b_0 + \sum_{l=1}^{L} \left( c^b_l \cos(2\pi l \nu_D t) + s^b_l \sin(2\pi l \nu_D t) \right),
\]

(7)

to be used as models for smooth modulating curves. The coverage factor \( C \) must be larger than the data span to avoid artefacts due to the built-in periodicity of our trigonometric models. Too large coverage factor leads to redundant parameterization. Very often the building of the trigonometric model starts from the rounding of the full span period, \( D \), and checking that the coverage factor remains in reasonable limits. According to our definition of a slow process, the period \( D \) must be significantly longer than the carrier period \( P_0 \).
Using basic trigonometric identities it is now easy to show that the full expansion of the carrier fit model will contain cosine and sine terms of the form \( \cos \left( 2\pi (\nu_0 + lv) \right) \) and \( \sin \left( 2\pi (\nu_0 + lv) \right) \) where \( l = 0, \ldots, L \). Proper expansion coefficient estimates must be computed for every term in the series for the fixed carrier frequency \( \nu_0 \) and “data frequency” \( \nu_d \); this is a standard linear estimation procedure and can be implemented using standard mathematical (statistical) packages.

In the beginning of our discussion about the algorithm, we started off with the simplest light curve model, Eq. (5), including only the lowest harmonic \( k = 1 \). To generalise to the full case, Eq. (4), including also higher harmonics, we will construct trigonometric representations \( a_k(t) \) and \( b_k(t) \) for each overtone of the carrier \( k \nu_0, \ k = 1, \ldots, K \). The time dependence of \( a_k(t) \) is also modelled using special trigonometric polynomial. If all polynomials use the same number of harmonics \( L \) and we approximate separate cycles by \( K \)-harmonic model, then the overall count of linear parameters to be fitted is \( N = (2 \times L + 1) (2 \times K + 1) \).

The actual choice of the representative parameters \( K \) and \( L \) depends on the particular object we are working with. The number of tones, \( K \), depends on the complexity of the phase curves. The choice of \( L \) is constrained by the longest gaps in the time series. In principle, it is also possible to compute certain formal constraints using standard statistical regression theory (see Draper & Smith 1998). Very often the inspection of phase plots which are computed with different parameter sets can also be useful.

2.3. Spline approximation

The trigonometric models of the modulation curves are the most useful when the actual behaviour of the modulations is relatively continuous, i.e. without sharp transients. The spline approximations are more flexible from this point of view.

To build spline approximations we need to divide the full data time span into, say \( L \), intervals. For every interval with index \( l = 1, \ldots, L \), and for each overtone of the carrier \( k = 1, \ldots, K \), we define two local sets of cubic curves

\[
a_{kl}(t) = \sum_{m=0}^{3} a_{klm} t^m,
\]

and

\[
b_{kl}(t) = \sum_{m=0}^{3} b_{klm} t^m,
\]

to be used as local modulators for the different tones of the cosine and sine components of the carrier wave. The complete analytical model for the data will then consist of cosine components

\[
a_{klm} \cos(2\pi k t \nu_0),
\]

and sine components

\[
b_{klm} \sin(2\pi k t \nu_0).
\]

There will altogether be \( 2 \times 4KL \) such components. To take into account the general level of the signal, we need to include also the unit component, after which the total number of modes to be fitted into data will be \( 8KL + 1 \). The splines are not just a sample of local polynomials, but must obey certain restrictions. In our simple case, we demand that the approximation curves themselves, their first and second derivatives should be continuous. As a result we need to complement our set of fitting equations with \( 2 \times 3K(L - 1) \) linear constraints. The total number of free parameters will then be \( N = 2KL + 6K + 1 \).

To perform the constrained least-squares fitting we used routines from the AS 164 -package written by Stirling (1981). The implemented algorithms are based on the so-called Givens rotations (see Gentleman 1973) and allow to solve weighted least-squares equations. The constraints are effectively taken into account by assigning infinite weights to the constraining equations.

2.4. Construction of an analytic signal

One of the useful properties of the carrier fit method is that it allows easily to compute Hilbert transforms for the model waveforms. Using the Bedrosian theorem (Bedrosian 1963), we can easily compute for term

\[
f(t) = a(t) \cos(2\pi \nu_0 t) + b(t) \sin(2\pi \nu_0 t),
\]

its Hilbert transform

\[
h(t) = a(t) \sin(2\pi \nu_0 t) - b(t) \cos(2\pi \nu_0 t),
\]

By definition, the modulators \( a(t) \) and \( b(t) \) are smooth functions, or in spectral terms, their spectra are concentrated around the zero frequency. The carrier frequency is significantly higher, and therefore the modulators can be removed from the transform. The complex analytic signal (Gabor 1946) can now be constructed as

\[
u(t) = f(t) + jh(t),
\]

its amplitude as function of time

\[
A(t) = \sqrt{f^2(t) + h^2(t)},
\]

and the instantaneous frequency

\[
\nu(t) = \frac{h'f - f'h}{2\pi (f^2 + h^2)},
\]

can be computed. For the details we refer the reader to the paper by Vakman (1996).

Unfortunately the analytic signal approach is useful only for waveforms which are either completely or nearly harmonic. For more general curves the instantaneous values for amplitude, phase and frequency can be formally computed, but their interpretation is not straightforward.

2.5. Estimation of the carrier frequency

The most important input parameter for the carrier fit method is naturally the carrier frequency \( \nu_0 \), or the corresponding period \( P_0 = 1/\nu_0 \) itself. Very often a useful value is already known from previous data analysis steps. Here are some typical cases:

- estimate from previous literature is at hand;
- orbiting period is known from eclipses;
- rotation period is computed from radial velocity data (say, in fully synchronized binaries);
- the period is estimated by averaging its seasonal values (Blazhko effect stars, spotted stars);
the central frequency of the band in the Fourier spectrum can be used (Quasi-Periodic Objects).

In most cases, therefore, an initial value can be easily determined. The first trial can be improved later if the phase plots show significant upward or downward trends.

In the case when we do not have a proper value for carrier frequency we need to estimate it from the data. Next we formulate a method to compare different carrier frequencies, and describe how a rough initial value can be improved to eliminate any trends in phase plots.

The most natural choice to measure the quality of the carrier frequency value is to inspect the stability of the resulting light curve – with the optimal carrier frequency, the most stable behaviour should be obtained. For this purpose we can use the interpolated carrier fit trial waveform. Even if built using a very rough preliminary estimate for the frequency \( \nu_0 \), it approximates the input waveform in its entirety and is a continuous function. Let us fix a certain trial frequency \( \nu_T \) somewhere around the preliminary carrier frequency \( \nu_0 \). We can divide the interpolated waveform into intervals of the length \( P_T = 1/\nu_T \), and construct local phase curves

\[
\phi_i(\phi) = \tilde{f}(t_i + P_T \phi),
\]

where \( t_i = t_0 + iP_T, \ i = 0, 1, \ldots, N_T - 1 \) is a sequence of starting points of the different intervals of the interpolated curve \( \tilde{f}(t) \), phase \( \phi \) runs from 0 to 1, and \( N_T \) is the total number of the intervals for a particular trial period. The phase curves change if we move along time (later we will plot such time-dependent curves as phase diagrams). We can now measure the rate of change of the phases along time using a statistic

\[
D^2(\nu_T) = \frac{1}{N_T - 1} \sum_{i=0}^{N_T - 2} \left( \frac{\phi_i(\phi) - \phi_{i+1}(\phi)}{\nu_T} \right)^2 d\phi.
\]

Loosely speaking, we estimate the mean square of the first derivative of the phase diagram along time. The optimal value for carrier frequency should minimize our statistics (see the discussion in Sect. 3.9 and Fig. 22).

After computing the value for an optimal carrier frequency, we can recompute the curve estimates and repeat the optimal frequency search, and thereby obtain a refined value for the optimal carrier frequency. In actual computations such a procedure takes only a couple of iterations to converge. Using this approach, uncertainties in the initial value of the carrier frequency do not matter, as the method ultimately finds the optimal and consistent carrier. However, one problem can be identified with this procedure. The optimal carrier frequency, as defined above, is essentially only a representative device. Sometimes it may be hard to find a physical interpretation to the resulting carrier value computed from the data. Only if we can assume that behind the seemingly erratic variation of oscillation parameters there is a certain mean flow which can be characterized by a certain mean frequency, we can accept the solution as physically meaningful. Fortunately, in most practical cases, the carrier frequency is available from the very beginning or physical considerations lead to a proper interpretation easily.

2.6. Visualization

We use the following scheme to visualize our results. First we calculate a continuous curve estimate, \( \tilde{f}(t) \), from the randomly spaced and gapped data set, using a carrier-based least-squares fitting scheme. This approximation is continuous and does not contain gaps, and therefore allows us to get a smooth picture of the long-term behaviour. Below we will discuss why and when we can trust such a gap-filling scheme. Next we divide this continuous curve into segments with a length of the carrier period \( P_0 = \frac{1}{\nu_0} \). Throughout the paper, we measure periods in the units of days, and frequencies in the units of cycles per day. We then normalize each segment so that the approximating values span the standard range of \([-1, 1]\). After normalization, we stack the segments along the time axis. To enhance the obtained plot, we somewhat extend every segment along phases, so that the actual display is wider (along phases) than a single period. The normalization is a relevant part of our procedure because it helps to grasp phase information we are interested in (trends, drifts, flip-flops etc.). In the lower part of the plots we often depict the distribution of the actual observational time moments in the form of “bar code”. The black stripes correspond to the periods during which at least one observation is available. The bright stripes correspond to gaps. This method of visualization allows to verify that the model fits into the data and not into the gaps. If the number of nodes or harmonics \( L \) for the modulation model curves is chosen properly then the phase plots do not reveal underlying timing structure.

2.7. Local time-frequency analysis

The obtained continuous, “de-gapped” curve \( \tilde{f}(t) \) obtained by the carrier fit method can be used as an input to different additional analysis procedures. Now that a fully sampled data set has been obtained, the standard approaches are applicable, due to which many different choices are possible. For instance, we can compute local Fourier spectra for sliding time windows, or compute various wavelet transforms (see Poularikas 2010, for additional possibilities). The corresponding time-frequency plots can then be combined with phase diagrams to get full grasp of the changes in the underlying processes involved.

In particular, we use stacks of normalized, local in time, power spectra to reveal trends in frequency structure of artificial or observed curves.

3. Examples and applications

The carrier fit method is quite straightforward in its formulation and uses rather standard tools for its implementation. However, its full utility can be appreciated only by looking at some actual applications. To build up the required intuition, we first illustrate the method applying it to very simple analytical light curve models which are sampled at time points taken from real observations, and after that, to real astronomical data sets. In the current context, however, we totally ignore the physical background of the effects observed in each example star – our goal here is to show the versatility of the method.

3.1. Timing sequences and data sets

To get realistic long-term photometric timing sequences we use a compilation of the V-band photometry for the variable star LQ Hya (Berdyugina et al. 2002). The used subset of 1627 points (displayed in Fig. 1) covers the interval JD 2 447 881–2 452 053 and contains rather long gaps (up to 284 days). In the most of the examples below we use only time points from this data set and calculate artificial time-dependent process values using analytical formulae. To model observational errors we add to each
analytical time series value a noise component $\varepsilon(t)$, which we generate as a normally distributed random variable with standard deviation of 5% from the overall series amplitude. This level of scatter is quite characteristic for reasonably good photometry, corresponding roughly to the signal-to-noise ratio of 20.

To demonstrate the suitability of the new methods to analyse the so-called Blazhko effect, we use the RR Lyr-type variable star MW Lyr. The data and detailed photometric solutions for the extensive photometry can be found in Jurcsik et al. (2008). As an input for reanalysis we used a well-sampled subset of the observations spanning over the interval JD 2 453 920–2 454 052, containing 3216 points (see Fig. 2).

The usability of the method in the context of the flip-flop phenomenon (e.g. Jetsu et al. 1993) is demonstrated using a sub-set of the extensive photometry of the rapidly rotating variable star FK Com, described in Korhonen et al. (2007) and depicted in Fig. 3. The data set contains 968 points and spans over the time interval JD 2 451 144–2 451 380.

The periods used as analytical model parameters are also taken to match some values that can be derived from observations; we return to their actual deduction from the data in the next papers of the series.

### 3.2. Mismatched carrier frequency

Very often the appropriate carrier frequency $\nu_0$ is obtained from information which is available before the actual data processing. Being it rotation, orbiting or previously determined frequency, we need to confirm or re-evaluate it. To do this, a model with smoothly varying coefficients is fitted into the data, using the procedure described in Sect. 2.5. If the actual frequency is stable, but differs from the used carrier frequency we will see it from the time-dependent phase diagram. In the following example we use artificially generated data, but use the time-points from actual observations of LQ Hya described above. For instance, if the input model is just a single harmonic with a frequency $\nu = 0.149009 = 1/6.711001$, magnitude off-set and an added noise term $\varepsilon(t)$ (see Sect. 3.1)

$$f(t) = 12.0 + \cos(2\pi t) + \varepsilon(t),$$

then using a carrier value $\nu_0 = 0.148714 = 1/6.724333$, we will get a time-dependent phase diagram depicted in Fig. 4. To stress that the actual time points used in the analysis are taken from real observations with long gaps, we often plot a stripe-like illustration of the observing times on the lower part of the phase diagrams, as explained in detail in Sect. 2.6. With the refined-search method for the carrier frequency, we can recover the actual period in the input data set, after which the tilted stripes in the figure become strictly horizontal.

To analyse time-dependent frequency spectra, we use the same continuous estimate, $\hat{f}(t)$, but now we use it as an input to a standard time-frequency software\(^1\). In our particular case we compute frequency spectra for sliding time windows. After

---

\(^1\) ISDA – Irregularly Spaced Data Analysis, available at [http://www.aai.ee/~pelt/soft.htm](http://www.aai.ee/~pelt/soft.htm)
computing each local in time spectrum we again normalize it to span the standard range [0, 1]. In this case the normalization helps us to track minor shifts in frequencies even if the amplitudes of the separate components change drastically. In this way our visualization method enhances the aspects of the variability which we are looking for. The result of this analysis is shown in Fig. 5.

In principle it is possible to compute local frequency spectra for the original (not interpolated) data set. In Fig. 6 we depict the results of this kind of transform. From this plot it is evident how the gaps in the data tend to distort the overall image. The changes in the spectra can be (mistakenly) counted as real modulations.

### 3.3. Abrupt change in the period

The next interesting feature in variable star light curves is a possible jump in the frequency. Applying the same time points from observations as in the previous example, such a process can be described with a simple time series model

\[
f(t) = \begin{cases} 
\cos(2\pi P_1) + \varepsilon(t), & t < 3000, \\
\cos(2\pi P_2) + \varepsilon(t), & t \geq 3000,
\end{cases}
\]

where \( P_1 = 6.747017, P_2 = 6.701800, \varepsilon(t) \) is noise term described above (Sect. 3.1) We choose the carrier frequency, \( \nu_0 \), as a mean of the two model frequencies, and apply the carrier fit method to build a continuous curve \( \hat{f}(t) \); the results are depicted in Fig. 7. As a comparison, local time-frequency analysis is made, the results of which are plotted in Fig. 8. The smooth appearance of the two plots is a result of the proper approximation scheme. The gaps of the input data are no longer visible, as they are properly filled in. This facilitates the visibility of the continuous patterns in the changing phase and period. However, if to compare the two plots, we see quite a relevant difference: the transition time from one period to the other is much longer for the time-dependent spectra if compared with the phase plot. To explain the difference, we need to take into account the fact that the time resolution of the transients differs for the two analysis methods. The smoothness of the phase plots is determined by the smoothness of the modulating coefficients \( a(t) \) and \( b(t) \), and can be adjusted using different fitting schemes. For time-dependent spectra the smoothness is controlled by the width of the sliding transform window. To obtain sufficient resolution in frequency domain we lose resolution in time domain. This is a well-known and ubiquitous uncertainty relation, and is common for all time-frequency analysis methods.

### 3.4. Two beating waves

Next we consider an example, where an interference between two nearby frequencies occurs. Again, the same time points with
gaps as in the previous examples are used, and an artificial time series model of the form

\[ f(t) = \cos(2\pi t/P_1) + \cos(2\pi t/P_2) + \epsilon(t), \]  

(20)

where \( P_1 = 6.747017 \), \( P_2 = 6.701800 \), and \( \epsilon(t) \) is noise term (see Sect. 3.1), is analysed using the carrier fit method to build a phase diagram, plotted in Fig. 9, where a clear chessboard pattern can be seen. This can be understood by considering the trigonometric identity

\[ \cos(A) + \cos(B) = 2 \cos\left(\frac{A + B}{2}\right) \cos\left(\frac{A - B}{2}\right). \]  

(21)

The interference between the two nearby frequencies can be looked upon as a certain high-frequency main tone modulated by a low-frequency beat tone. In our case the beat period is \( \approx 2000 \) days. Because we are dealing here with 100% modulation, the sign of the beat wave changes after every 1000 days, and the overall image of the time-dependent phase looks like a chessboard. These abrupt changes in the phase are somewhat unexpected if we take into account that the model contains only two simple harmonics. The carrier fit method and the used visualization technique allows us to reveal this natural flip-flop in the clearest way.

The continuous approximation curve \( \hat{f}(t) \) can be further analysed by using the local time-frequency analysis method. The results of such an analysis with two different sliding windows are shown in Figs. 10 and 11. The first of these two figures, in which a sliding window of the length 600 days was used, demonstrates a very serious problem of the time-frequency analysis. From the form of the input model, we expect that two parallel lines along time are recovered. Instead, we see that the plot is dominated by a single maximum at a period which corresponds to the mean of the two model frequencies, and only around the locations, where the modulating beat wave changes sign, the spectrum splits into two separate frequencies. Even then the local maxima are not placed at the correct periods. Such a weird picture is due to the usage of a relatively short sliding time window (600 days) if to compare with beat period (2000 days). Even if we compute the local spectra for a four times longer time window (as in Fig. 11), there is still a slow wave seen in the positions of the two frequencies. This example illustrates a typical error made in time series analysis, where the original data is divided into segments, and each segment is analysed separately, after which conclusions are made according to the obtained local periods. Were the local spectra computed without a preceding continuous interpolation, the results could be even more misleading. In this particular case we consider the time-dependent phase plot a more appropriate analysis and visualization method.

3.5. Linear trend in the frequency

As our next example of the carrier fit method, we define the input data as follows

\[ f(t) = \cos(2\pi t((1 - \lambda(t)) \times 0.148214 + \lambda(t) \times 0.149214)) + \epsilon(t), \]  

(22)

where \( \lambda(t) = \frac{t - t_{\min}}{t_{\max} - t_{\min}} \) moves linearly from 0 to 1 during the time span of the input data \([t_{\min}, t_{\max}]\) and \( \epsilon(t) \) is a noise term (see Sect. 3.1).

As seen from Fig. 12, the phase diagram for a curve with a changing frequency is rather complex, and it is not easy to interpret the actual form of the variability in the input data. In this case, however, the local time-frequency analysis reveals quite explicitly what is going on, as shown in Fig. 13.

For this model, we can also apply an instantaneous frequency estimation scheme based on the Hilbert transform to construct an analytic signal, as described in Sect. 2.4. In Fig. 14, the computed run of the instantaneous frequency is displayed. Due to the noise component and gaps in the data set, the recovery is
Parameters for the carrier fit model: ν₀ = 0.148714, K = 1, and L = 8.

It is quite clear that for the modelling of the process we need higher carrier harmonics. In Fig. 15 we show the results of the carrier fit with K = 3. Again, a regular pattern of minima and maxima is visible. If we compare with the simple beat pattern of Fig. 9, the most important difference is the occasional bi-modality of the phase curves. This is a result of an interplay between phases of the two "activity zones". If the zones occur at phases nearby each other, we have curves with a single maximum; in case of roughly 180° separation, two maxima appear. For more complex waveforms, the interpretation of the phase plots can be even more convoluted.

3.7. Blazhko effect

As an example of an application of the carrier fit algorithm, based on trigonometric modulation curves, to real observational data, we re-analyse a subset of the photometry for the RR Lyr-type star MW Lyr, collected by Jurcsik et al. (2008). This pulsating star is known to exhibit a modulation of its light curve in addition to the variability due to the pulsations; this effect is known as the Blazhko effect. Jurcsik et al. (2008) proposed a bi-periodic model with the very accurate pulsational period of
$P_0 = 0.397680$ days modulated by a significantly longer period of $P_b = 16.5462$ days.

We start from a similar bi-periodic model with the frequencies $k\nu_0 + l\nu_b$, where $k = 1, \ldots, 5$ and $l = -10, \ldots, 10$, and $\nu_0 = 1/P_0$ and $\nu_b = 1/P_b$. The main difference to the original analysis by Jurcsik et al. (2008) is the usage of a smoother form of the basic waveforms ($|k| \leq 5$ instead of $|k| \leq 13$), and a more detailed modulation model ($|l| \leq 10$ instead of $|l| \leq 4$). The resulting normalized phase diagram is presented in Fig. 16.

To apply the carrier fit method to the same input data sequence, we fixed the "data period", $D$, to be 150 days with the coverage factor of $C \approx 1.14$, so that the trigonometric modulation curves cover slightly more than the full data span. It is important to state that the resulting phase plot is not very sensitive to this number. We also set the number of harmonics $K = 5$ for the base frequency, and $L = 10$ for the modulations. In this way the system of linear equations to be fitted into the data was equal to the number of equations in the bi-periodic fit. The results are presented in a phase diagram of Fig. 17, showing a striking similarity of the phase patterns to the ones plotted in Fig. 16. An important point here is that the characteristic modulation pattern is revealed without pre-set (or precomputed) value for the modulation period. The only difference between the two diagrams is the somewhat larger number of small details for the bi-periodic fit and its exact periodicity. The phase diagram for the carrier fit solution is somewhat smoother and reveals minor fluctuations in the general periodic pattern. Whether these are fluctuations resulting from the observational noise, method artefacts, or sampling irregularities, remains currently open.

As this example clearly demonstrates, it is possible to recover the overall bi-periodic phase structure even without knowing the value of one of the periods.

### 3.8. Flip-flop

To demonstrate the spline method in action, we performed both a trigonometric and spline-based fit into a subset of photometry of the star FK Com. A flip-flop phenomenon has earlier been detected in this very subset of data (Jetsu et al. 1993) covering one single observing season, during which a sharp jump of roughly 0.5 in phases occurs. The results of our analysis are shown in Figs. 18 and 19; for both the trigonometric and spline fits, the general picture is quite similar, and the abrupt jump in phase is clearly revealed. There are, however, some subtle differences. The trigonometric method is in a sense more "global", due to which sharp features in one location can show up as oscillations somewhere far away. The spline method is essentially more "local", and as a result the picture is somewhat smoother. We deliberately chose our parametrization here so that the total number of free parameters is equal for both methods.

---

**Fig. 16.** Blazhko effect, phase diagram, bi-periodic fit with $P_0 = 0.397680 \ (K = 5), P_b = 16.5462 \ (L = 10)$. See Sect. 3.7.

**Fig. 17.** Blazhko effect, phase diagram, carrier fit with $P_0 = 0.397680, D = 150, K = 5, \ and \ L = 10$. To be compared with Fig. 16.

**Fig. 18.** FK Com flip-flop, phase diagram, carrier fit with trigonometric modulations, carrier $\nu_0 = 2.40025, K = 3, L = 3, N = 49$. See Sect. 3.8.

**Fig. 19.** FK Com flip-flop, phase diagram, carrier fit with spline modulations, carrier $\nu_0 = 2.40025, K = 3, L = 5, N = 49$. To be compared with Fig. 18.
3.9. Computation of the carrier frequency

Finally, we illustrate the method to estimate the carrier frequency. Our input data model is the same as used in Sect. 3.3, Eq. (19), containing an abrupt change in the period. The noise added to the input data through the term $\epsilon(t)$ is now increased to 20% level. A fragment of a standard phase-dispersion spectrum (Stellingwerf 1978) for this data set is depicted in Fig. 20. The strongest peak is at a period $P_0 = 6.7470$, coinciding closely with $P_1$ in the input data model. This is to be expected, as the input data consists of a long fragment with this particular period. We proceed by using it as a carrier period value to build a time-dependent phase diagram depicted in Fig. 21. This carrier value produces a nearly horizontal stripe with some oscillatory behaviour for the part of the data having a period close to the used carrier, and a strongly tilted pattern for the part with a shorter period, already familiar from Sect. 3.2 for the case of a mismatched carrier frequency. The situation can be improved by refining the carrier frequency with a procedure described in Sect. 2.5 using the statistic $D^2$ defined by Eq. (17), measuring the speed of changes in the phase diagram. For this example calculation, we plot this statistics in Fig. 22, showing a well-defined minimum at the frequency $\nu_0 = 0.14850 = 1/6.7334$. Using this frequency as the carrier, a time-dependent phase plot with minimum changes can be obtained, and is depicted in Fig. 23. Now the decreasing and increasing fragments of the phase diagram are well balanced.

Here we again note that the optimum carrier value is just a formal construct, and not based on any physical principle or insight. However, the procedure can be used as the last resort if no a priori information is available.

4. Related work

Next we discuss the proposed algorithms in the context of the other well-known analysis methods for variable star light curves.

4.1. The O–C method

In the classical O–C (observed versus computed) method (see for instance Sterken 2005), the input light curve is divided into short fragments, so that for each fragment it is possible to securely determine a certain light curve event, most often a local minimum or maximum. For each event, a difference between the observed (O) and the computed (C) time moments is computed. The differences O–C are then plotted for a full time span of the input data, and a proper model for the deviations is then sought for. Most often polynomial models are used.

In the carrier fit method proposed above, we are not very far off from this traditional scheme. The carrier period, $P_0 = \nu_0^{-1}$, is just an analogue of the period $P$ in the equation for computed
time point values. The fitting procedure itself tries to tightly model local fragments. In doing so, the local phase properties of the carrier wave indicate peculiarities of the original curve (minima and maxima among others). And finally, the smooth modulation curves are just analogues of the traditional polynomial models for O–C differences. In principle, it is even possible to modify the carrier fit method so that instead of trigonometric or spline models, simple polynomials are used as modulation curves. 

The O–C method is often used to refine the already roughly estimated periods. As shown above, a mismatch in the carrier frequency reveals itself as a linear trend in the phase diagram. Therefore, to estimate the needed correction to the carrier period, the linear model for carrier modifications is considered sufficient.

The carrier fit method even shares the well-known problems of the O–C method. Most importantly – the wrong (or not exact enough) base period $P$ can totally spoil the O–C analysis. This is also true when the choice of the proper carrier frequency is considered.

### 4.2. Fourier analysis

Methods based on, or connected to, standard Fourier analysis are certainly the most popular in variable star research. The classical series of papers, starting from Barning (1963), and ending with Koen (1990) and Frescura et al. (2008), gives a statistically sound treatment for the case when the input data can be described by a simple harmonic. The generalizations for non-harmonic (but periodic) and multiperiodic cases are obvious (see the chapter “Frequency Analysis” and references in the monograph by Aerts et al. 2010). For us the important point is that the classical methods are based on the assumption that the Fourier spectra of the underlying processes are discrete, and form a sparse set of frequencies. The sparseness of the spectrum makes it possible to recover the constituent waveforms from data sets with long gaps. Unfortunately, the periodicities caused by the gaps can be a source of serious problems. Along with every discrete peak in the original spectrum there will be a set of spurious peaks, and a perfect recovery can be a challenge.

Another method of frequency analysis is based on the idea of inversion. In this type of methods, the full discretized Fourier spectrum is fitted into data (see for instance Kuhn 1982, or more recently Nygrén & Ulich 2010). For a sensible inversion, the input data needs to be sampled quite densely. It must also be noted that inversion techniques involve huge matrix computations, and for this reason these methods fail to work for long input data sets.

The carrier fit algorithm described above combines properties of the discrete and continuous spectrum methods. First, we have a carrier frequency $v_0$ and its overtones $k v_0$, which constitutes the discrete aspect. If the modulating waveforms $a_k(t)$ and $b_k(t)$ are smooth enough, the resulting spectrum will consist of $K$ narrow but continuous frequency bands, accounting for the continuous aspect. By combining some good properties of the standard methods, we can achieve a compromise, where the discreteness of the midpoints of the narrow bands allow to “bridge” long gaps in time domain, and the continuity of the local spectra around the mid-frequencies allows to describe smooth time-dependent changes. Our method can essentially be regarded as a quite straightforward combination of the two classical approaches.

### 4.3. Local frequency analysis

There is another set of methods popular among variable star researchers. These methods are based on a division of the input data into separate segments, and then using some standard method to analyse them. The segmentation can be based on natural fragmentation of the input data or by formal division of the full data span. In some methods the segments can be overlapping in time (see recent detailed implementation by Lehtinen et al. 2011). It is assumed that the changes in the local Fourier spectra or other type statistics computed for different segments reveal actual trends in frequency structure of the underlying signal.

A typical example – the δ Scuti – type star θ Tuc was observed during six nights (Stobie & Shobbrook 1976), and data for each night was Fourier analysed separately. The conclusion of the authors was

Photometric observations of θ Tucanae were Fourier analysed for their component frequencies. No coherent frequency could be found spanning the complete data set. It is shown that both the frequencies and amplitudes present in θ Tucanae change on time scale as short as 24 h.

However, more recent analysis of more extended data sets by other authors (for example Kurtz 1980; Paparó et al. 1996) gave another interpretation to the local variability of θ Tuc. It occurred that the varying local spectra result from constructive and destructive interferences of different coherent oscillation modes. The interpretation of the local spectra is further complicated by their low resolution. Above we gave an illustration of this effect (see Fig. 10) in the context of straightforward local Fourier analysis. This example shows that it is a good practice to use time-dependent phase diagrams (built by the carrier fit method) and time-dependent frequency analysis together to get a full understanding of the underlying process.

Localisation is carried to the limit in the method proposed by Tsantilas & Rovithis-Livaniou (2008). They use a very similar light curve model to Eq. (1)

\[
 f(t) = a(t) \sin(b(t) \times t + c(t)),
\]

where observed data is decomposed into three continuous curves: amplitude $a(t)$, frequency $b(t)$ and phase $c(t)$. The point estimates for the three components are then obtained by performing local non-linear least squares fits. Unfortunately, the presentation of the light curve in this form is not unique (for instance frequency changes can be compensated by phase changes) and consequently the obtained numerical results contain large amount of arbitrariness. In our method the carrier is fixed and modulations are constrained to be smooth. This allows to get more useful presentations for the light curves.

### 4.4. Time-frequency analysis

Another set of methods worthwhile to mention in this context are different time-frequency transforms. They are all based on a decomposition of the input data into set of sub-waves. Every sub-wave (in some methods called wavelet) is localized in time as well as in frequency. For a large set of this kind of decompositions we refer to a recent work by Blackman (2010) and references therein. In the context of variable star research, the most important problem with the standard time-frequency methods is related to the gaps in the input data. Usually the set of sub-waves is mathematically complete, but this leads to unrealistic “filling of gaps”. The sub-waves which are localized in short time intervals inside observing gaps will get zero values and as a result of
that the structure of gaps will be seen through in the final two-dimensional spectrum. A very nice illustration of this point can be found in Szatmary (1994), where a large catalogue of model wavelet decompositions is given.

To overcome problems with gaps it is possible to modify the original wavelet method by making its resolution to depend on time. This can be achieved by the use of certain weighting schemes (see Foster 1996 for details) or adaptive wavelets, as done by Frick et al. (1997).

In the carrier fit method we use a priori information to constrain our sub-waves to be localised in fixed frequency strips around carrier and possibly around its harmonics. The shortcoming of our formulation is that our system of sub-waves is not complete. Nevertheless, with our algorithm, the gaps can be properly filled in, which enables to get a continuous phase or frequency diagrams that are more easy to interpret.

5. Discussion and conclusions

5.1. Utility of the carrier fit

All the simple examples above demonstrate that the carrier fit procedure itself is not sufficient to perform proper data analysis and interpretation. It is just a particular method to interpolate data sets with gaps to obtain a relatively smooth estimate for the variable star light curve. Interpolation itself has some value only when the real observed process and its observed values have certain properties. Most importantly, the process must be governed by some relatively stationary inherent clocking mechanism whose parameters change slowly. Fortunately this is the case for many variable stars.

5.2. Domain of applicability

By formulation and by algorithmic implementation the proposed utility can be considered as automatic. However, the obtained results can be very often misleading. This is why the complete understanding of the involved ideas is very important. Let us now reiterate major requirements for successful application of the carrier fit method.

Most importantly, the data to be analysed must come from an object with a certain internal clocking mechanism, being it rotation, orbiting or pulsation. This allows to fix, at least approximately, a proper value for a carrier frequency. The modulating curves of the basic cycle must be relatively slow so that the spectral representations of the modulations and carrier frequency are located in different parts of the spectrum, i.e. the Bedrosian theorem is valid. It is also important that models for modulations must be smooth enough to allow "bridging the gaps" in observational sequences.

There must be enough of high quality observational points to allow adequate interpolation. To avoid over fit and under fit, standard methods from regression analysis should be used (Draper & Smith 1998).

The user of the methods must have a certain expertise to interpret the obtained phase diagrams. The proper and required intuition for that can be developed by using model computations with known input components. As a starting point the examples provided above can be useful.

Because the method involves inversion of large matrices, certain numerical problems may occur. The methods based on row-wise updating (Given rotations) to solve the least-squares computational problems are very useful in this context; for details see Lawson & Hanson (1987).

The carrier fit method belongs to the class of data analysis methods called explorative. After getting an intuitively pleasing and convincing phase diagrams the researcher must quantify the results using more precise models (with proper estimates of errors etc.).

Another important aspect of the new method is the focus on phase behaviour. For that purpose, the visualization of normalized light curves is of great importance. This facilitates the detection of relevant trends and abrupt changes in light curves. This is just the aspect which usually remains hidden when standard multicomponent fit into the data is used.

In this paper we have presented a relatively new but simple class of data analysis methods which are useful in different astro-physical contexts. The best way to promote the methods would be their successful application to actual data sets. We are undertaking this effort in the next papers of the series.

The developed software is a part of the general time series package (Irregularly Spaced Data Analysis) which is freely available2.

Acknowledgements. Ilkka Tuominen sadly passed away on March 19, 2011. We wish to express our respect for his importance for the research on magnetically active stars. Financial support from the Academy of Finland grants 141017 and 218159 is acknowledged. We also thank the anonymous referee for useful comments on the manuscript.

References

Aerts, C., Christensen-Dalsgaard, J., & Kurtz, D. W. 2010, Asteroseismology (Springer)
Barning, F. J. M. 1963, B.A.N., 17, 22
Bedrosian, E. 1963, Proc. IEEE, 51, 368
Lawson, L. C., & Hanson, R. J. 1987, Solving Least Squares Problems. Society for Industrial Mathematics

2 ISDA, http://www.aai.ee/~pelt/soft.htm, details by e-mail pelt@aai.ee