Inconsistencies in the harmonic analysis of the time series of pulsating stars observed by space missions

J. Pascual-Granado,* R. Garrido,† and J. C. Suárez‡
Instituto de Astrofísica de Andalucía (CSIC), 18008, Granada, Spain
(Dated: July 24, 2013)

Purpose: The current paradigm of asteroseismology establishes the harmonic analysis of the time series of a pulsating star as the method to find unambiguously its frequency content. On the light of the challenges raised by the ultra-precise photometry supplied by space missions this assumption is here reconsidered. In the present work we test the analyticity of the underlying functions describing the observations.

Method: The reliability of the paradigm is studied using a method based on the Stone-Weierstrass theorem. More specifically we check the correspondence of some properties derived from the spectral components found on the stellar light curves with classical Fourier algorithms. This is done by examining a fine-scale property of the time series, the connectivity, which is intrinsically related with the differentiability of the underlying functions. We performed a connectivity analysis to a multiperiodic pulsating δ Scuti star HD 174936 observed by *CoRoT*. The connectivities of the *CoRoT* data were compared to those obtained from an analytic model typically representative of the frequency content of the star. In order to discard any possible contribution of the instrumentation to the results the method was applied also to the hybrid δ Scuti-γ Dor star KIC 006187665 observed by the *Kepler* space mission.

Results: The comparison of the connectivities of *CoRoT* data and the analytic model yielded very different results demonstrating that the light curve cannot be represented by the analytic model. In addition, the correlated residuals resulting from the connectivity analysis for *CoRoT* and *Kepler* data demonstrate that the functions underlying the light curves are not analytic. Instrumental effects are discarded, since the method was applied to stars observed by two different satellites of different characteristics (precision, sampling, etc.).

Conclusions: The dramatic consequence is that the periodogram is not a consistent estimator of the spectral density of the time series. Therefore, the light curves of these pulsating stars cannot be expanded in Fourier series.

Keywords: methods: data analysis — methods: statistical — stars: oscillations — *Kepler*

INTRODUCTION

Since *SoHO* started the era of asteroseismology from space with the Sun observations [1], outstanding advances in the knowledge of the solar structure have been made. However, some of the objectives of this space mission remains unsolved: the non-detection of solar g-modes [2] and the physical origin of the near surface frequency effects [3] are good examples. More recently Hanasoge et al. [4] using data from NASA satellite SDO (Solar Dynamics Observatory) found a velocity field on the solar atmosphere two orders of magnitude lower than theoretically predicted, which lead them to prompt the following question: "(sic) What mechanism transports the heat flux of a solar luminosity?".

Thanks to *CoRoT* [5] and later *Kepler* [6] missions, ultra-high quality asteroseismic data from space were also available for other stars than the Sun [see the impact of these missions into the physics of stellar structure models in 7], eliminating ground-based observational windows and removing the noise produced by the atmospheric transparency variations. Analysis of such precise time series of multi periodic pulsating stars [e.g. 8–10] has revealed a new puzzle not yet understood: the detection of a huge number of frequencies together with non-gaussian residuals of the light curve fitting.

The unprecedented quality of the data obtained from the above mentioned space missions allows the improvement of the theoretical pulsation models and therein to infer a better description of the stellar structures. But, for a correct understanding of the nature of the observed physical processes, the mathematical description used to interpret the data has to be able to reproduce the complete data set while only leaving as residuals an independent random series. This rarely happens in practice where the residuals, small as they are, keep some structure, which makes them far different from an independent random sequence. This is especially the case in some δ Scuti stars [10, 11] and γ Dor stars [12] observed by *CoRoT* .

These challenges lead us to reconsider some of the basic foundations of the harmonic analysis of data in asteroseismology, where Fourier series (i.e. Lomb-Scargle periodogram, [13]) have been used widely as mathematical description.

In this paper we check the correspondence of some properties derived from the spectral components found on the stellar light curve of HD 174936 using classical Fourier algorithms with properties derived from the original data. With this aim, a fine-scale property called connectivity, which is intrinsically related with the differentiability of the function describing the observations, has been examined in detail.

This allows us to check the identity between the obtained spectral representation and the data, and at the same time to test the validity of the assumption of the square integrability of the signal, which is a necessary condition for the Fourier expansion to be convergent.

The paper begins in Sec. 2 with some esential considerations concerning the reliability of the expansion of a function in Fourier series depending on its analyticity. In Sec. 3 the definition of connectivity is given, which is, basically, a function describing how smoothly connected are the data points of a given sampled function. Thus defined, connectivity gives an adequate approximation of the differentiability for discrete time series, provided that the function is band-limited and sufficiently sampled. In Sec. 4, two approximations are used to calculate the connectivities: one is based on a closed-form mathematical model, the cubic spline polynomials, and the other one is based on a non-closed-form model given by autoregressive techniques (ARMA). The results of the tests on real data obtained from the asteroseismic space mission *CoRoT* are presented in Sec. 5, where we try to reproduce the results obtained by fitting a model based on the assumption that the time series consisted on a discrete set of individual frequencies with white noise added. We compared the results obtained for both real and fitted light curves showing large differences. In order to discard eventual instrumental effects, the same test is performed on a time series obtained by another space instrument, *Kepler* (Sec. 6). Finally, in Sec. 7, a discussion of the main results of the paper is presented together with the relevant consequences derived from our findings and summarized in the Conclusions.

ANALYTICITY

Fourier analysis is a useful technique for frequency detection but it has a variety of difficulties that are not always considered in the literature. If the sequence analyzed belongs to a nonperiodic or quasiperiodic data set, the number of spectral components needed to represent the signal can be so large that the physical information provided by its Fourier representation could be no longer useful. Constraining the number of spectral components used for the harmonic analysis only increases the probability of misleading information as a result. In any case, and more importantly, in order to apply Fourier techniques to the function, Parseval's theorem must hold. This theorem states that the integral of the squared modulus of a function is equal to the integral of the squared modulus of its Fourier transform. For this to be true the function must be square integrable

$$\int_{-\infty}^{+\infty} |F(x)^2| dt < \infty \tag{1}$$

The function is square integrable if the Fourier series of the function converge uniformly pointwise. Unlike the Fourier series, which is an infinite sum, the Discrete Fourier Transform (DFT) and its inverse, the IDFT, which is the analogous to the Fourier series, always converges if the sequence is finite, because it is a finite sum. But when the Fourier series of the function is divergent, or not convergent to the value of the function at this point, the DFT no longer provides a physical description of the function based on its frequency components. That is, a frequency is only a well-defined physical observable when the Fourier series converge.

Therefore, before a Fourier analysis can be applied to a time series, the convergence of the Fourier series expansion of the function must be demonstrated.

On the other hand, the analyticity of a function is a sufficient condition for the square integrability. A function is said to be analytic if it is infinitely differentiable so that its Taylor expansion is convergent. If this condition is not met the convergence of the Fourier series is not guaranteed. That is the reason why it is so important to evaluate the analyticity of the function. In this paper the term analyticity refer to the property that a given function can be expanded in Fourier series. To study the analyticity of a function we give an approximation of the differentiability for discrete time series, the connectivity.

CONNECTIVITY

Time dependent astronomical data are given as a discrete sequence of samples from a function, which is the representation of an observable satisfying the Lebesgue measure definition[25]. In the case of photometric data, such a discrete sequence is called light curve.

Let us consider here a discrete series as a finitely close sequence of data points and the function being sampled, a set of infinitely close datapoints. The first question arising is: can a discrete sequence of values fully and uniquely determine the properties of the function? This is the well-known problem of sampling. The Nyquist-Shannon theorem states that the function can be reconstructed perfectly well by means of a regular sampling of the function if, and only if, it is band-limited to less than half the sampling rate, i.e. the Nyquist frequency.

When the conditions above mentioned are satisfied, then all the properties of the function can be studied from a discrete series, even those concerning the pointwise limits of the function. For our purposes this means that the differentiability of the function can be evaluated adequately.

Concept

Here we are interested on how smoothly the discrete sequence can be connected in order to reproduce exactly the function. We call this property: the connectivity of the function at a given point. This property is closely related to the differentiability.

In general it can be said that a discrete sequence of data points generated from a function is composed of a stochastic sequence plus a deterministic contribution that can have or not a wave-like structure. In fact, the Wold decomposition theorem [14] states that every stationary process can be represented always as an stochastic process with a purely deterministic component. The random component of the stochastic process appears to complicate the study of a time series when the connectivity of the data is being analysed but a very robust mathematical result, the Kolmogorov continuity theorem, which applies for stochastic time series, allows us to fully characterise the properties of the function [15]. This theorem permits to extend properties of stochastic processes, like differentiability, which would require infinitely close data points, to a function, given a finite number of samples. This is a consequence of the topological separability of the function, and it is applicable if certain constraints on the moments describing the data differences are satisfied. Basically, the condition states that the moments of the data differences change at most with a power law. Therefore, it is not expected that the random components of the data difficult seriously the characterization of the function by means of its connectivities.

Definition

It is usual to study the continuity of a function at a given point by calculating both one-sided limits in order to check whether they converge to the same value and coincide with the value of the function at that point. In that case the function can be said to be continuous at that point.

In a similar way, given a datapoint from a discrete sequence, a forward and backward extrapolation from a subset bracketing the selected datapoint can be computed to check whether they both converge to the same value and coincide with the value of the selected datapoint of the sequence.

The above theoretical framework permits the following self-consistent definition of the connectivity $\mathcal{C}_n$ of a datapoint x_n of a discrete sequence as sampled from the function $F(x)$ as

$$C_n = \epsilon_n^f - \epsilon_n^b \,, \tag{2}$$

where ϵ_n^f, ϵ_n^b, are the deviations of the forward and backward extrapolations from the datapoint x_n.

$$\begin{aligned} \epsilon_n^f &= x_n^f - x_n \\ \epsilon_n^b &= x_n^b - x_n \,. \end{aligned} \tag{3}$$

are the deviations of the forward and backward extrapolations from the datapoint x_n, represented by x_n^f and x_n^b, respectively.

A discrete approximation for the pointwise derivability condition for $F(x_n)$ can be:

$$\frac{x_n^b - x_{n+1}}{\Delta t} = \frac{x_n^f - x_{n-1}}{\Delta t} \tag{4}$$

where Δt is the sampling rate of the sequence. It is evident from these equations that connectivity is closely related to derivability. The numerical approximation of the point derivative $\mathcal{D}_n$ at x_n can be expressed as:

$$\mathcal{D}_n = \frac{\mathcal{C}_n + x_{n+1} - x_{n-1}}{2\Delta t} \tag{5}$$

which reduces to the typical point derivative for discrete data when $\mathcal{C}_n = 0$. When connectivities are not zero but independent randomly distributed values, the derivative is still well-defined. In that case, the connectivities can be considered simply as deviations from the true derivative at this point. On the contrary, when the connectivities are correlated with the data the derivability condition is not fulfilled.

METHODS

The forward and backward extrapolations necessary to obtain the connectivities have to be calculated numerically. Two numerical approaches, conceptually very different, have been considered. The first one is based on a representation given by a closed-form function, the cubic spline polynomials, the second one is based on an autoregressive model, which is not described by any closed-form function. The detailed description of these calculations and the properties of these mathematical tools are given in the next sections.

Cubic splines

The Stone-Weierstrass theorem [16] states that if a function is uniformly continuous in a closed interval it can be approximated arbitrarily well, i.e. as closely as desired, by a polynomial of degree n, with n a natural number. Then, the entire function can be expressed as a piecewise polynomials function of finite degree. But piecewise polynomials functions have a drawback, the Runge's phenomenon, resulting in oscillations at the edges of the approximated function, which is the equivalent to the Gibbs phenomenon in the Fourier description.

Another crucial point here is that piecewise polynomials are continuous functions but not necessarily analytic because they can show instabilities at high order derivatives. Then, instead of choosing any polynomials to model the function, piecewise smooth polynomials were chosen. Adding smoothness as an additional constraint guarantees analyticity and then the convergence of Fourier series is assured, also giving a natural solution to the Runge's phenomenon. That is the reason why cubic splines are selected as the mathematical technique to calculate forward and backward extrapolations using an analytic approach.

Formulated in this very general way any particular property can be extended to any analytic function. That is, if a continuous function can be fitted arbitrarily well with a given cubic spline parametric model then the analyticity condition is automatically satisfied, otherwise the convergence of the Fourier series is not guaranteed for that particular function.

The approximation using cubic splines gives us key information about local features of the function which generates the observed discrete sequence. Other non-smooth approaches, such as the Hermite polynomials, are able to capture more features from the function when it is non-analytic. This is the case for the time series analyzed here as will be discussed later in Sec. VII. However, we are more interested now in another approximation using autoregressive methods, because they can fit a non-analytic function also and no closed-form formula is needed for that.

ARMA

Fourier decomposition is useful for describing periodic processes with small non-periodic contributions. That is because the expansion in Fourier series of a periodic function has a finite number of terms. In contrast, when every feature of a function requires independent modelling and there are non-periodic features whose Fourier series expansion might be infinite, the autoregressive methods can be more appropiate for the description.

In the present work, we used the well-known autoregressive method ARMA (Auto Regressive Moving Average) [17] for the connectivity analysis, which is particularly interesting for our purposes because it can fit fairly well the data even when the function is non-analytic and no closed-form formula is necessary (see below Eq. 6)

ARMA models have been used to calculate a forward and backward extrapolation from a subset bracketing every datapoint of the dataset. The ARMA model is composed of two statistical processes: the autoregressive (AR) and a moving average (MA) contribution:

$$x_n + \sum_{i=1}^{p} a_i x_{n-i} = \epsilon_n + \sum_{j=1}^{q} b_j \epsilon_{n-j} \qquad (6)$$

where the left hand side represents the autoregressive process being x_n the data points and the right hand side corresponds to the moving average process being ϵ_n the terms of a white noise process. The a_i, b_j parameters are numerical coefficients containing the correlation between one term of a given sequence and the previous one.

Just like cubic splines modelling, the accuracy of the predictions with ARMA models depends on the number of data points of the segment to be modelled in order to calculate forward and backward extrapolations as defined in Sec. III. In addition to this, ARMA extrapolations depend on two parameters: the orders p and q, which are the number of autoregressive coefficients and the number of moving average coefficients, respectively. Hereafter, ARMA descriptions of time series will be denoted by its coefficients as ARMA(p, q).

In general, an ARMA model is able to reproduce a sinus wave with only three terms. Accordingly, the number of terms increases with the frequency content of the signal.

The ARMA modelling involves three steps:

1. Identification of the order of the starting model. This affects principally the velocity of convergence but does not modify the result.

2. Calculation of the parameters a_i, b_j using a Steiglitz-McBride algorithm [18] given the orders for the AR and MA contributions.

3. Evaluation of the validity of the model based on the goodness of the fit in backward and forward extrapolations. The algorithm starts again with increased orders until a minimum in the residuals is reached.

In the following section both techniques will be applied, cubic splines and ARMA, to our case of study: the time series corresponding to light variations of a star given as the integrated flux as measured by the photometric instrument onboard a space satellite, where the satellites are *CoRoT* and *Kepler*, and the stars are the δ Scuti HD 174936 and the hybrid pulsating star KIC 006187665.

COROT TARGET: HD 174936

The light curve of the δ Scuti star HD 174936 has been analyzed using classical Fourier analysis in García-Hernández et al. [11]. They found 422 significant frequencies in the power spectrum of this star. The authors follow the conservative limit given by Breger et al. [19] to consider a given peak in the power spectrum as a statistically significant frequency as the Fisher's test for periodicity establishes.

However, the 422 frequencies are distributed in such a large range of periods that none of the stellar theoretical models studied in the paper to represent the internal structure of this star is able to reproduce.

The physical origin of the excitation of the peaks interpreted as frequencies in the power spectra of this kind

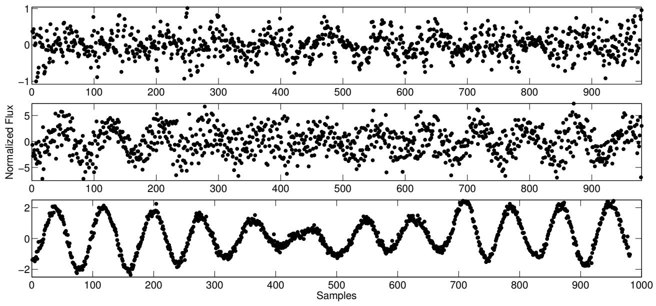

FIG. 1. Connectivities of the *CoRoT* data for the δ Scuti star HD 174936. Upper panel shows ARMA connectivities, middle panel shows splines connectivities, and lower panel corresponds to the CoRot time series light curve.

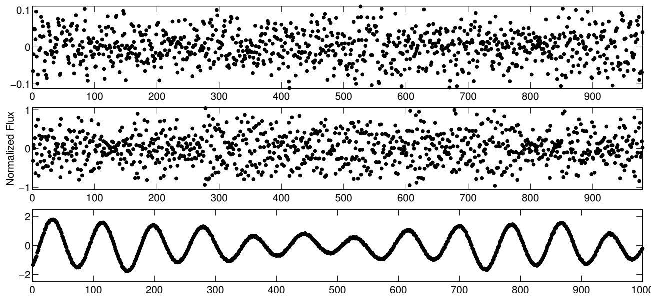

FIG. 2. Connectivities of the analytic model for the δ Scuti star HD 174936 based on the first 422 frequencies detected using classical Fourier techniques. Upper panel shows ARMA connectivities, middle panel shows splines connectivities, and lower panel corresponds to the analytical model time series.

of pulsating stars have been explained until now by the well understood κ driving mechanism. Although some hypothesis have been worked out for such a discrepancy between the theoretical predictions and these new ranges of variability put in evidence by space photometry, no definitive explanation is accepted yet.

One of these hypotheses [20] stated that most of the peaks in the power spectrum of HD 174936 are the signature of a fully developed turbulence giving rise to a non-white granulation background noise[26]

Given the difficulties with the interpretation of the power spectrum of this star, it was selected as a case of study for our tests.

Connectivity analysis

We started using ARMA modelling assuming initial tentative orders. Then we iterated (see previous section) until minimal residuals randomly distributed were obtained so fixing optimal orders p and q. For this particular case it was found that $p = 20$ and $q = 1$ were optimal orders for the light curve of HD 174936. Cubic spline and ARMA(20,1) connectivities were then calculated for every point in a segment of 1000 samples of the light curve using subsets of 80 datapoints (40 forward and 40 backward extrapolations). The segment was previously normalized to have zero mean and unit standard deviation $(0, 1)$.

The same test was applied to a model fitted to the light curve of the δ Scuti HD 174936. A realization of this fitted model is given by constructing a time series using amplitudes and phases of the first 422 frequencies listed in García-Hernández et al.. A white noise component was added with an amplitude equal to the average of the noise amplitudes corresponding to every frequency as a simulation of the expected noise. The realization of the model, in this way constructed, is an analytic function as derived by the approximation expressed in Eq. 5.

In both cases the connectivities based on the cubic splines approximation have much more dispersion than the corresponding ARMA ones. Nevertheless, the most striking effect is seen only in the *CoRoT* data: spline connectivities are not randomly distributed but they are strongly correlated with the signal suggesting that these data are not sufficiently well described by the analytic model. This peculiar behaviour is of considerable interest if we want to understand the physical processes causing the phenomenon either in terms of noise components or any other type of unknown dependence from which the data have to be corrected for.

The same effect seems to be present, although around 10 times smaller in amplitude, when the ARMA plot is examined in detail. In conclusion, and for this particular application, ARMA processes fits clearly better than the analytic approximations.

In general, the goodness of the fit, which is measured as the distance to the real observed values, depends on the number of data points used for that particular model. In Fig. 3 this dependence is plotted as the sum of the squared connectivities over the segment, understood as deviations from the true derivatives. These can be interpreted as the sum of the squared errors SSE versus the number of data points N_p.

Connectivities tend to have an almost zero value when sufficient data are given to the ARMA(20,1) model for *CoRoT* data. Surprisingly, a model with so few parameters (20+1) appears to be a very good model for observed data, yielding much better results than cubic splines models. On the other hand, the results for the analytic model are worse, as could be expected based on theoretical considerations. Actually an ARMA(20,1) model could only fits poorly a function having 422 independent frequencies.

Interestingly, the SSE calculated using the cubic splines approach tends to almost zero for the analytical model, the reason being that, provided that the differentiability is preserved, a function can be represented arbitrarily well. But, as depicted in lower panel of Fig. 3, the

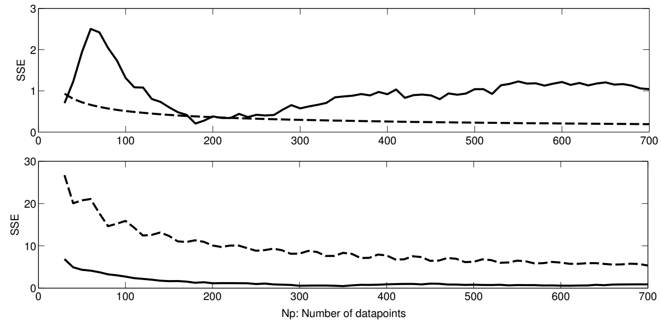

FIG. 3. Sum of the square errors (SSE) vs. the number of datapoints. The upper panel shows the results for the analytical model and lower panel shows the same calculations for the *CoRoT* data of the δ Scuti star HD 174936. ARMA connectivities are shown with a solid line and spline connectivities with a dashed line. Note the rapid convergence of ARMA results for satellite data, and the poor performance of the cubic splines calculations.

SSE of *CoRoT* data decrease much more slowly. Besides, the ripples observed in the SSE in that case indicate that part of the time series exhibits a cyclic structure that is not fully captured by the model.

The difference between the analytic model and real data has been usually interpreted as correlated noise caused by the effects of the turbulence present in the envelopes of stars [21]. In order to check if the convection is at the origin of this phenomenon, some additional tests were performed by adding non-white noise to the analytic model following the paper by Kallinger & Matthews. A power law with two components (36 ppm and 19 ppm of amplitudes) were added in the frequency domain to the analytic model constructed with the 422 independent frequencies mentioned above. The time series is obtained through an inverse Fourier transform using normally distributed random phases. The light curve so constructed is supposed to be originated by the normal discrete oscillation modes of the pulsating star plus granulation noise mimicking a fully developed turbulence regime.

Similar results were obtained when the test was applied to the new time series. The addition of the non-white noise described above does not improve the results and even larger differences were found between the observations and the analytic model. This is a coherent and expected result as far as a non-white noise, showing an inverse frequency dependence, varies in time much more smoothly than the white noise does. Consequently, the effect on the differentiability of the function is expected to be much lower than the corresponding to the white noise. On the other hand, adding more complexity to the analytic model can only make the fitting by the ARMA(20,1) approximation worse, giving as result even more different connectivities to those corresponding to real data.

Definitively, a very important result is emerging when

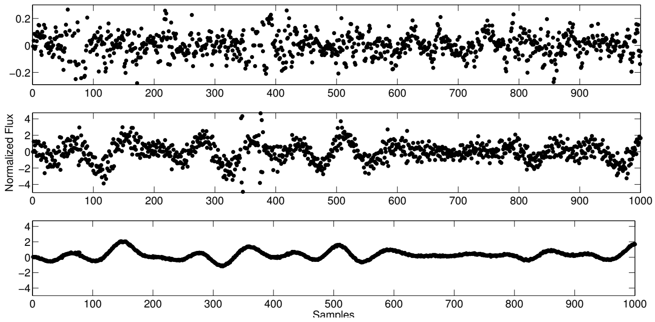

FIG. 4. Connectivities of the *Kepler* data for the hybrid star KIC 006187665. Upper panel shows ARMA connectivities, middle panel splines connectivities, and lower panel the original *Kepler* light curve for comparison.

the connectivity of the *CoRoT* time series is studied in detail: the lack of analyticity. In other words, when a time series, like the ultra-precise photometry given for many targets observed by the asteroseismic camera in *CoRoT*, is almost free of poissonian noise, then a clear difference always exists between the analytic model and the real data. Hence, a working hypothesis can be that the physical real function as observed by *CoRoT* has not the mathematical properties required to be an infinitely differentiable function (C^{∞}).

However, this is only one among the many cases of time series corresponding to *CoRoT* and *Kepler* targets where difficulties in the interpretation of their spectrum have been found.

KEPLER TARGET: KIC 006187665

In order to discard possible instrumental effects specifically linked to a given instrument, a similar test was applied to the time series from another star KIC 006187665, obtained by using a different instrument, the *Kepler* satellite.

The selected star was classified as a hybrid Gdor/Dscut star by Uytterhoeven et al.. The classical Fourier analysis[27], yielded 659 significant peaks from which 184 were in the γ Dor and δ Scuti frequency range.

In order to perform a direct comparison with the previous case the same orders as those used in the *CoRoT* data were used for the ARMA model of the Kepler stellar light curve. The results for the connectivity analysis are shown in Fig. 4. The connectivities of the *Kepler* target show a very similar behaviour to the *CoRoT* target: when calculated using cubic splines they are more dispersed than the corresponding to the ARMA model and strongly correlated with the original *Kepler* light curve, too.

If the connectivity at a point is interpreted as the deviation from the true derivative (Eq. 5), then the analysis

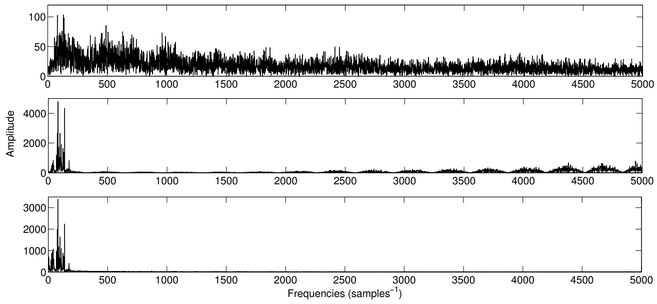

FIG. 5. Amplitude spectrum of the ARMA connectivities (upper panel), cubic spline connectivities (middle panel) and the original *Kepler* light curve of KIC 006187665. Power spectra are calculated from a set of 10000 data points. We use arbitrary units for the frequencies and for the amplitudes..

of its time variations allow us to correct the original time series in order to guarantee differentiability, in this way solving the problems put in evidence in this paper.

To give more insight into these deviations, our working hypothesis will be that the connectivities as they are calculated, fulfill the differentiability required to satisfy the Parseval's theorem, finally allowing Fourier analysis techniques. This hypothesis is being evaluated for the light curves of other targets and will be the subject of a forthcoming paper.

In Figure 5 the power spectrum of the connectivities is shown. The low frequency range for splines connectivities is very similar to that found in the power spectrum of the original data. However a periodic pattern can be seen along the frequency axis increasing in amplitude with frequencies. When the connectivities are calculated with an ARMA model some structure is visible at low frequencies whereas an almost flat spectrum is obtained at higher frequencies. In any case the amplitudes of the ARMA connectivities are 1-2 orders of magnitudes lower than those calculated with cubic splines.

Summarizing, the puzzling features observed in HD 174936 appear also in this hybrid star observed by *Kepler*, so discarding that instrumental effects are involved in this phenomenon. The similarity between the low frequency part of the spline connectivities amplitude spectrum and the corresponding to the original data can be conjectured as a kind of self-similarity. On the other hand, the periodic pattern observed at high frequencies could be an effect related to the sampling rate. However, we do not know the exact origin of this pattern and more tests are being done in order to disentangle this striking effect.

As conclusion we can assert that the effect of the lack of differentiability of the time series analyzed in this work can be put in evidence through the concept of connectivity. This effect has been shown to be independent of the instrument, observed target and sampling. Furthermore, we can also assert that the connectivities, when calculated using analytic models, are 1-2 orders of magnitudes larger than those corresponding to a non-analytic model (ARMA).

DISCUSSION

As mentioned in Sec. II, before a Fourier analysis can be applied to a time series, the convergence of the Fourier series expansion of the function must be demonstrated. In this paper we have introduced the study of the analyticity of the function underlying a discrete sequence to demonstrate the convergence of the Fourier series prior any further analysis.

Provided a sufficiently good sampling and a sufficiently large amount of data, a given function can be fully characterised even if some properties like the differentiability depend on an infinite number of values of the function. As Revuz & Yor [15] demonstrated by means of the Kolmogorov continuity theorem, this is possible even for stochastic signals provided that certain constraints on the moments of the data differences are fulfilled assuring the separability of the function.

Once satisfied the necessary conditions to fully characterise the function, we have studied a property identified as connectivity which is closely related to the differentiability of the function. The definition of this property (Eqs. 2-5) gives the numerical approximation that permit us to characterise the derivability of the function.

In the ideal case that the function is differentiable and an infinite amount of data is available for fitting a cubic splines model, the connectivities would be zero. However, with a finite number of points we demand only that the connectivities have no structure at all, i.e. the autocorrelation is null. This is the case when the connectivities are a strictly white noise process.

Now, at the light of our results, we can state that the huge number of frequencies detected in the light curve of the δ Scuti star HD 174936 cannot reproduce the properties of the original light curves.

The fact that an ARMA process with a significant smaller number of parameters can model the data quite accurately indicates that the real number of frequencies for HD 174936 might be an order of magnitude lower than currently obtained with the classical Fourier analysis without introducing any granulation background noise.

Note the SSE of the spline connectivities tends to a finite value (Fig. 2), which is significantly greater than zero. This test was also repeated for the hybrid star KIC 006187665 showing that in this case, it is not possible to fit arbitrarily well the function using cubic splines either.

Finally, we used Hermite polynomials instead of cubic splines obtaining a much better results. Indeed, the SSE rapidly decreased with increasing the number of data points. As piecewise Hermite polynomials are only C^1

functions, this suggests that the observable functions of the studied stars are not smooth and, therefore, the convergence of their Fourier series is not guaranteed.

CONCLUSIONS

It is known from the general spectral analysis theory of time series that a square summable gaussian process can be represented always by a linear model [23]. When this happens the expansion in Fourier series is convergent and hence the periodogram is a consistent unbiased estimator of the spectral density function of the signal. When this is not the case, as for the light curves studied in this paper, the periodogram is an inconsistent estimator of the spectrum of the signal. On the other side, given that the time series cannot be represented by a linear model and because by definition, every linear combination within a normal distribution must be necessarily normal, the distribution of that time series must be non-gaussian. This explains the reason why the light curves studied here have a non-gaussian distribution and in addition, it can be shown that their discrete time series are derived from a non-analytic function, therefore they are not square summable series.

The tests carried out here for HD 174936 show that the description based in Fourier frequencies is not consistent with the Parseval's theorem. Analyses of different observations from different satellites and instruments discarded this phenomenon to be originated by instrumental effects.

Furthermore, it is demonstrated that this effect is not originated by the surface convection because the granulation noise as given by a Harvey model [24] has no impact in our results.

It has been our aim here to pay attention to the fundamental concept related with the definition of frequency as defined in the Parseval's theorem. Classical time series techniques are based on the reliability of this theorem and our analyses have shown here that when the observable corresponds to the measurement of the flux over the visible disk of a given pulsating, the necessary condition of differentiability, here introduced by means of the connectivity property, is not satisfied.

As demonstrated in the analysis of the SSE for different types of stars, the lack of differentiability is intrinsically related with the original time series. This points to a reconsideration of what is called an unambiguous detection of a physical frequency, usually associated to a statistically significant peak in the power spectrum. In the case of asteroseismology this is understood as solutions in terms of harmonic series of the linear differential equation describing the time evolution of a given stellar oscillating model.

A possible correction preserving the frequency definition and at the same time obtaining a time series which is numerically differentiable is being explored. In this sense, in a forthcoming paper we will investigate the ubiquitous character of this phenomenon in all kind of pulsating stars and different instrumentation.

The ultra-precise data supplied by the asteroseismic missions demand more and more a new and different approximation for the time series analysis, posing real challenges not only for asteroseismology but for any other related field with similar characteristics. The connectivity property as defined in this paper, together with the numerical tool developed to test standard time series, could open a new window for exploring not fully understood properties of physical systems studied by means of time series analysis.

The *CoRoT* space mission, launched on December 27th 2006, has been developed and is operated by CNES, with the contribution of Austria, Belgium, Brazil, ESA (RSSD and Science Programme), Germany and Spain. Funding for the *Kepler* Discovery mission is provided by NASAs Science Mission Directorate. The authors acknowledge support from MINECO and FEDER funds through the Astronomy and Astrophysics National Plan under number AYA2012-39346-C02-01. J.P-G. acknowledges support from MINECO through the FPI grant number BES-2008-008252. JCS acknowledges support by the European SPACE program project SPACEINN.

* javier@iaa.es
† garrido@iaa.es
‡ jcsuarez@iaa.es

[1] Domingo, V., Fleck, B., Poland, A. I. 1995, Solar Physics, 162, Issue 1-2, pp. 1-37
[2] Appourchaux, T., Belkacem, K., Broomhall, A.-M., et al. 2010, A&A Review, 18, Issue 1-2, 197-277
[3] Kjeldsen, H., Bedding, T. R., Christensen-Dalsgaard, J. 2008, Astrophys. J., 683, L175-L178
[4] Hanasoge, S. M., Duvall, T. L., Sreenivasan, K. R. 2012, PNAS, 109, 11928-11932
[5] Auvergne, M., et al. 2009, A&A, 506, 411-424
[6] Borucki, W. J. et al. 2010, *Science*, 327, 977
[7] Suárez, J. C., Garrido R., Balona, L. A., Christensen-Dalsgaard, J. (Eds.) 2013, Impact of New Instrumentation and New Insights, Stellar Pulsations, Ap&SS Procs. 31
[8] Balona, L., Pigulski, A., Cat, P. De, et al. 2011, MNRAS, 413(4), 2403-242
[9] Mantegazza, L., Poretti, E., Michel, E., et al. 2012, A&A, 542, A24
[10] Poretti, E., Michel, E., Garrido, et al. 2009, A&A, 506, 85-93
[11] García Hernández, A., Moya, A., Michel, E., et al. 2009, A&A, 506, 79-83
[12] Chapellier, E., Rodríguez, E., Auvergne, M., et al. 2011, A&A, 525, A23
[13] Scargle, J. D. 1982, Astrophys. J., 263, 835-853
[14] Wold, H. 1938, A Study in the Analysis of Stationary Time Series, (2nd ed.; Uppsala: Almqvist and Wiksell)

[15] Revuz, D., Yor, M. 1999, Continuous Martingales and Brownian Motion, Vol. 293, A Series of Comprehensive Studies in Mathematics (Springer)
[16] Royden, H. L. 1988, Real Analysis (Prentice-Hall)
[17] Box, George E. P., Jenkins,, Gwilym M. 1976, Time Series Analysis, Forecasting and Control (Holden-Day Series in Time Series Analysis)
[18] Steiglitz, K., McBride, L. E. 1965, IEEE Trans. Automatic Control, Vol. AC-10, 461-464
[19] Breger, M., Stich, J., Garrido, R., et al. 1993, A&A, 271, 482
[20] Kallinger, T., Matthews, J. 2010, Astrophys. J., 711, L35-L39
[21] Chaplin, W. J., Elsworth, Y., Howe, R., et al. 1997, MNRAS, 287, 51-56
[22] Uytterhoeven, K., Moya, A., Grigahcène, A., et al. 2011, A&A, 534, A125, 70
[23] Priestley, M. B. 1981, Spectral analysis and time series, Vols. I and II (Probability and Mathematical Statistics; Z. W. Birnbaum, E. Lukacs; Academic Press)
[24] Harvey, J., 1985, In ESA Future Missions in Solar, Heliospheric and Space Plasma Physics, 199-208
[25] Note that an observable of a given physical property is not necessarily coincident with the function which represents it.
[26] After removal of turbulence peaks, only 70 frequencies remain in the resulting power spectrum.
[27] This analysis was performed to the light curve supplied by the satellite at quarter Q2.2, observed with a sampling rate of 60 seconds (short cadence regime)