Distortion of the p-mode peak profiles by the solar-cycle frequency shifts: do we need to worry?

W. J. Chaplin,1⋆ Y. Elsworth,1 R. New2 and T. Toutain1

1School of Physics and Astronomy, The University of Birmingham, Edgbaston, Birmingham B15 2TT
2Faculty of Arts, Computing, Engineering and Sciences, Sheffield Hallam University, Sheffield S1 1WB

Accepted 2007 December 6. Received 2007 November 22; in original form 2007 October 11

ABSTRACT
We seek to address whether solar-cycle frequency shifts of the Sun’s low-l p modes ‘distort’ the underlying shapes of the mode peaks, when those peaks are observed in power frequency spectra made from data spanning large fractions, or more, of the cycle period. We present analytical descriptions of the expected profiles, and validate the predictions through use of artificial seismic time-series data, in which temporal variations of the oscillator frequencies are introduced. Our main finding is that for the Sun-like frequency shifts the distortion of the asymmetrical Lorentzian-like profiles is very small, but also just detectible. Our analysis indicates that by fitting modes to the usual Lorentzian-like models – which do not allow for the distortion – rather than new models we derive, there is a bias in the mode height and linewidth parameters that is comparable in size to the observational uncertainties given by multi-year data sets. Bias in the frequency parameter gives much less cause for worry, being over an order of magnitude smaller than the corresponding frequency uncertainties. The distortion discussed in this paper may need to be considered when multiyear Sun-like asteroseismic data sets are analysed on stars showing strong activity cycles.

Key words: methods: data analysis – methods: statistical – Sun: helioseismology.

1 INTRODUCTION

High-quality observations of the solar p modes are now available, which for some instruments cover almost three complete 11-yr cycles of activity (Chaplin et al. 2007a). Accurate and precise mode parameter data are a vital prerequisite for making accurate inference on the solar interior, be that on the hydrostatic or dynamic structure. There are clear advantages to be gained by extracting estimates of the mode parameters from power frequency spectra made from several, sometimes many, years of contiguous observations. The excellent resolution in frequency, and excellent height-to-background ratios observed in the mode peaks, then allows mode parameters to be extracted to extremely high precision. Subtle phenomena, such as asymmetry of mode peaks (a diagnostic of the stochastic excitation and granulation) and asymmetry of mode frequency splittings (a diagnostic of the surface activity) may then also be extracted reliably from the data, and the weakly damped p modes at very low frequencies become visible and amenable to measurement and study.

The observations may then span a sizeable fraction, or more, of an 11-yr solar activity cycle period. The question arises: what effect do the well-known solar-cycle shifts in frequency through a long time-series have on the underlying shapes of the mode peaks, when those peaks are observed in power frequency spectra made from the full time-series? Are the shapes so distorted from the Lorentzian-like form that the peak-bagging codes – with their Lorentzian-like fitting models – inevitably return biased estimates of the parameters? In this paper, we seek to answer this question for the low-degree (low-l) core-penetrating solar p modes. We seek an analytical description of the underlying peak shapes expected from time-series in which the frequencies of modes are known to vary. We then look at whether the form given is significantly different from the assumed Lorentzian-like profiles. We also use simulations of artificial time-series data, in which temporal variations of the oscillator frequencies have been introduced, to validate use of the analytical expressions.

The frequency-shift regime of interest is illustrated in Fig. 1. The left-hand panel plots the sizes of the minimum-to-maximum solar-cycle frequency shifts of the low-l modes (here, averaged over l = 0 to 3), as a function of mode frequency (e.g. see Chaplin et al. 2004). The average shift amounts to about 0.4 μHz for a mode at ~3000 μHz. While modes at higher frequency suffer a bigger shift (e.g. about 1 μHz at frequency 4000 μHz) their linewidths are then also about an order of magnitude larger than at 3000 μHz, and it is the shift-to-linewidth ratio that is of relevance for determining the impact of any distortion. This ratio is plotted in the right-hand panel of Fig. 1: clearly, modes at the centre of the low-l p-mode spectrum are potentially most susceptible to the distortion effect.
\[ \nu = \frac{H}{\Gamma} \]

where the angled brackets indicate an average over time, and \( H \) and \( \Gamma \) are the mode height (maximum power spectral density) and linewidth, respectively.

In what follows we will discuss the profiles given for two functions describing the frequency shifts in time: first, the simplest possible function, this being a linear variation over time; and secondly, a cosinusoidal variation to mimic the full solar activity cycle. We have neglected the effects of the solar-cycle variations in height and width, which will be included in future work.

3 ANALYTICAL MODE PROFILES

3.1 Linear variation in time

We begin by assuming a simple linear variation in time, i.e.

\[ \nu(t) = \nu_0 + \Delta \nu \frac{t}{T}, \]

where \( \nu_0 \) is the unperturbed frequency, and the shift is calibrated so that the total shift from the start \((t = 0)\) to the end \((t = T)\) of the time-series is \( \Delta \nu \). Substitution of equation (2) into equation (1), followed by solution of the integral, gives the predicted mode profile:

\[ \langle P(\nu) \rangle = \frac{H}{2\epsilon} \text{atan} \left( \frac{2\epsilon}{1 - \epsilon^2 + \chi^2} \right), \]

where

\[ \epsilon = \frac{\Delta \nu}{\Gamma} \]

is the frequency shift in units of the mode linewidth and

\[ \chi = \nu - \frac{\nu_0 + (\Delta \nu/2)}{\Gamma/2}. \]

Fig. 2 shows profiles given by equation (3). The unperturbed profile (solid line) is for a mode having an unperturbed frequency of \( \nu_0 = 3000 \muHz \), an unperturbed linewidth of \( \Gamma = 1 \muHz \) and an unperturbed height of \( H = 100 \) units. The other curves show the profiles that result when the frequency shift, \( \Delta \nu \), is: 0.15 \( \muHz \) (dotted line); 0.40 \( \muHz \) (dashed line); 1.50 \( \muHz \) (dot–dashed line) and 3.0 \( \muHz \) (dot–dot–dot–dashed line). Since \( \Gamma = 1 \muHz \), the \( \Delta \nu \) also correspond to the shift-to-linewidth ratios, \( \epsilon \). To put the values in context, low-\( l \) modes at \( 3000 \muHz \), which also have width \( \approx 1 \muHz \), suffer a total fractional shift of approximately 0.40 \( \muHz \) from the minimum to the maximum of the solar activity cycle.

With reference to Fig. 2, it is evident that only at the two largest shifts (dot–dashed and dot–dot–dot–dashed lines) do the profiles depart appreciably from the Lorentzian form. However, these shifts are somewhat larger than those suffered by the real \( p \) modes. Closer inspection of the profiles does reveal some modest distortion at the two, smaller, Sun-like shifts. These have \( \epsilon = 0.15 \) and 0.40, respectively. We discuss the implications of this distortion for introducing bias in results of the usual peak fitting in Section 5 below.

We also tested the predictions of equation (3) with Monte Carlo simulations of artificial data. The Laplace transform solution of the equation of a forced, damped harmonic oscillator was used to...
generate mode components at a 40-s cadence in the time domain, in the manner described by Chaplin et al. (1997). Components were re-excited independently at each sample with small ‘kicks’ drawn from a Gaussian distribution. The simulations gave modes having Lorentzian limit shapes in the power frequency spectrum. The underlying frequency of the oscillator was then varied over the course of the time-series to give the required shifts. We averaged many thousands of independent realizations, for computations made at each of the shifts indicated above, to recover estimates of the underlying profiles that agreed excellently with the analytical predictions.

At first glance it would seem that the profile given by equation (3) is quite different from a Lorentzian. However, for \( \epsilon \ll 1 \) the function simplifies to give a Lorentzian. In this case, by solving analytically the likelihood maximization of the profile described by equation (3), it is possible to show that \( H^* \) and \( \Gamma^* \), the height and linewidth of this simplified (Lorentzian) function are, to the first non-null order, related to the original \( H \) and \( \Gamma \) by

\[
H^* = H \left( 1 - \epsilon^2 \frac{\pi}{6W} \right)
\]

and

\[
\Gamma^* = \Gamma \left( 1 + \epsilon^2 \frac{\pi}{6W} \right).
\]

Here, \( W \) is the width in frequency (in units of the linewidth \( \Gamma \)) over which the fit is made in the power frequency spectrum. From these equations, we can say that the linewidth will be slightly overestimated, and the height slightly underestimated, if a simple Lorentzian fitting model is used to fit the mode peaks. It is interesting to note that the estimates \( H^* \) and \( \Gamma^* \) tend towards \( H \) and \( \Gamma \) as \( W \) increases. This means that ideally with a sufficiently large frequency window it should be possible to recover the original parameters of the mode to good accuracy.

Finally in this section, we note that the observed solar low-\( l \) p-mode peaks are slightly asymmetric in shape, albeit at the level of only a few per cent at most. It is also possible to derive a version of equation (3) based on an unperturbed profile that is asymmetric, for only a few per cent at most. It is also possible to derive a version of equation (3) based on an unperturbed profile that is asymmetric, for only a few per cent at most. It is also possible to derive a version of equation (3) based on an unperturbed profile that is asymmetric, for only a few per cent at most.
We then apply the technique to real low-solar activity cycle; however, our point here is to test the principle of complete solar activity cycle, and so we therefore use fitting models the frequency shifts, since those shifts may give rise to measurable
mode peaks in the peak-bagging codes it may be possible to estimate which equations (12) and (13) are based (e.g. see Chaplin et al.

The observed time variations of the mode frequencies have extracting estimates of mode frequencies from short time-series, USING NEW FORMALISM

4 ESTIMATION OF FREQUENCY SHIFT USING NEW FORMALISM

Solar-cycle frequency shifts of p modes are usually estimated by extracting estimates of mode frequencies from short time-series, distributed at different epochs along the solar activity cycle. In this section, our aim is to see whether it is instead possible to use the new analytical expressions to extract estimates of the frequency shifts. The observed time variations of the mode frequencies have a qualitatively similar form to the cosinusoidal time variation on which equations (12) and (13) are based (e.g. see Chaplin et al. 2007a). The idea is that by using the new equations to model the mode peaks in the peak-bagging codes it may be possible to estimate the frequency shifts, since those shifts may give rise to measurable distortions.

In our tests we assume the observations span approximately one complete solar activity cycle, and so we therefore use fitting models based on the cosinusoidal model. A cosinusoidal variation is, of course, not an ideal match to the real-observed time variation of the solar activity cycle; however, our point here is to test the principle of the technique. We begin with tests on artificial data that do have an underlying variation of the frequencies that is cosinusoidal in time. We then apply the technique to real low-l time-series, which have lengths spanning approximately one 11-yr solar activity cycle.

4.1 Application to artificial data

We first made 1000 realizations of an artificial time-series comprising an l = 0/2 mode pair, in which the modes were given underlying parameters expected for low-l modes at ~3000 μHz. Frequencies of the artificial modes were varied over time in a cosinusoidal manner. The time-series were each T = 11 yr long – corresponding to the length of the artificial cycle, P365 – and all modes were given a total frequency shift, from the minimum to the maximum of the cycle, of Δν = 0.4 μHz. The artificial mode pairs were then fitted in power frequency spectra of the time-series to fitting models based on equation (13). The results showed it was possible to extract estimates of the frequency shift given to the modes, to a typical precision of ~0.15 μHz.

In 50 cases out of the total of 1000 simulations (i.e. in 5 per cent of the realizations), the fits ‘locked on to’ a null, or zero-valued, estimate of the frequency shift. This is a recurrent problem when fits are made for parameters which are too sensitive to the realization noise (e.g. estimation of component frequency splittings of blended modes). In such cases, the maximum of the likelihood function can be far enough from the solution that it lies outside the range accessible to the parameter. The fit is therefore ‘stopped’ by the hard limit, which is zero in the case of the frequency shift.

As noted in Section 4 above, the classic approach to estimation of the shifts involves dividing the full time-series into shorter sub-series, which are then fitted to yield time-dependent estimates of the frequencies. Analysis of the resulting set of frequency estimates yields an estimate of the total frequency shift. When we applied this classic approach to the 1000 artificial time-series, we found it was possible to estimate the shift to a precision of ~0.07 μHz. The precision is clearly superior to that given by fitting the new equations to the peak profiles. This is not surprising: we showed in Section 3.2 that for realistic low-l frequency shifts, the distortion of the mode profiles is very modest. This makes it hard to measure the distortion, and extract a robust estimate of the frequency shift, using the new technique.

We then extended the Monte-Carlo tests to simulate a range of mode pairs across the low-l power frequency spectrum. We adopted the strategy of Toutain, Elsworth & Chaplin (2005), whereby the artificial underlying limit power frequency spectrum was computed and then fitted. This strategy saves on computing time, since one does not have to generate, and then fit, a large number of independent time-series to give useable statistics. Equation (13) was used to make the mode profiles. The best-fitting uncertainties on fits made to these artificial data gave direct estimates of the precision expected from a single time-series realization of the same length as that used to compute the underlying limit power frequency spectrum. (Note the fits recover the input frequencies accurately, since the fitting model was based on the equations that were used to describe the artificial profiles.) The results, which are plotted in Fig. 4, show that the precision in the estimates is quite modest, particularly at lower frequencies where the input frequency shifts are smallest in size.

4.2 Application to real data

Next, we made use of four time-series of real disc-integrated Doppler velocity observations of the Sun. Two time-series comprised Sun-as-a-star observations: one made of data collected by the ground-based BiSON between 1993 and 2003; and one made of data collected by

Figure 4. Fitted frequency shifts (symbols, with associated 1σ error envelope shown by dashed lines) and input frequency shifts (solid line) for modes in artificial l = 0/2 mode pairs. The frequency shifts were estimated by fitting the artificial mode peaks to fitting models based on equation (13).
the GOLF instrument on board the ESA/NASA SOHO spacecraft between 1996 and 2004. The other two time-series comprised resolved Sun observations that were spatially averaged to give a Sun-as-a-star proxy signal: one made of data collected by the ground-based GONG between 1995 and 2004 and one made of data collected by the MDI instrument on SOHO between 1996 and 2006. Each of the four time-series covers more or less one 11-yr cycle of solar activity.

The power frequency spectrum of each time-series was computed, and low-\(l\) mode pairs were fitted to models based on equation (13). We used equation (13), rather than equation (12), because the real low-\(l\) peaks show small amounts of asymmetry. The \(l = 0/2\) pairs were fitted assuming all constituent components had the same frequency shift. For the \(l = 1/3\) pairs, our fitting results demonstrated that the relative weakness of the \(l = 3\) peaks meant the distortion of their profiles could not be fitted reliably. At \(l = 1\), the blending of the constituent components of each mode caused some problems for the fitting, which was manifested by cross-talk between the fitted frequency shift and frequency splitting parameters.

Fig. 5 shows the estimated frequency shifts as a function of mode frequency, averaged over \(l = 0, 1\) and 2 for each time-series (see the caption). As we would have expected, given the simulation results discussed in the previous section, it was not always possible to extract accurate estimates of the frequency shifts, and there were therefore several occasions on which the fits locked on to zero-valued estimates. Nevertheless, the non-zero best-fitting values follow the expected trend in frequency, giving estimates of the shifts that are consistent with previous frequency-dependent estimates for the low-\(l\) modes (e.g. Chaplin et al. 2001; Jiménez-Reyes et al. 2001; Gelly et al. 2002), apart from the highest frequencies, where the shifts are somewhat larger.

5 BIAS ON MODE PARAMETERS

In the previous sections, we have seen that mode peaks of the low-\(l\) solar p modes should be slightly distorted, with respect to the basic Lorentzian-like profiles, when they are observed in power frequency spectra made from observations spanning a substantial fraction or more of the 11-yr solar activity cycle. This distortion is caused by variation of the mode frequencies over time. Even though the distortion is modest, we have demonstrated that it is possible to measure the distortion (and estimate frequency shifts) by using new fitting models.

This result raises an important question: if the mode profiles are indeed distorted, what bias might we expect in the best-fitting mode parameters? We use results on artificial (Section 4.1) and real (Section 4.2) solar p-mode data to seek an answer to this question.

Let us consider first results on artificial data. We again adopted the strategy of Toutain et al. (2005), creating artificial underlying limit power frequency spectra corresponding to 11 yr of observations. Equation (13) was used to make the mode profiles, with the assumed input frequency shifts having the same sizes as those shown in Fig. 4. The low-\(l\) pairs were then fitted to fitting models made with the Nigam–Kosovichev formalism. Comparison of the fitted and input parameters gave the bias estimates plotted as solid lines in the left-(for linewidth) and right-hand (for frequency) panels of Fig. 6.

The dotted line in the left-hand panel shows the full best-fitting linewidth uncertainty (which is of similar magnitude to the height uncertainty). The overestimation of the linewidth – a similar-sized underestimation of the height is also observed – is seen to be of a size comparable to the estimated uncertainties. The dotted line in the right-hand panel is one-tenth the size of the best-fitting mode frequency uncertainty. Bias in the mode frequencies is evidently very small indeed, in contrast to the linewidth parameter. That said, we need to interpret the frequency result with a little care.

When the observations span either a complete cycle (as here), or one-half of a cycle, the frequency parameter should, in principle,
not be biased because the distortion is symmetric in frequency. Evidently, the small residual bias exhibited in the right-hand panel of Fig. 6 is due to other effects (e.g. from parameter cross-talk in the fitting, and the impact of the asymmetric shapes of the peaks). We would, however, expect there to be a bias from the distortion effect when observations span, say, one-quarter of a cycle. Under these circumstances, the distortion will not be symmetric in frequency.

To test this case, and other intermediate cases, we made further artificial data sets. Artificial power frequency spectra were made to mimic observations ranging in length from 1 to 11 yr. Furthermore, the observations were assumed to start anywhere from the beginning of the activity cycle up to 10 yr of the cycle (in 1 yr steps). Estimates of the frequency bias of modes at 3000 µHz, as a function of the simulated length of the observations, are plotted in Fig. 7. The various curves show the bias – as a percentage of the formal frequency uncertainty – for different starting points along the 11-yr cycle. The magnitude of the bias never rises above a few per cent of the frequency uncertainty. We may therefore conclude that frequency bias is not a major cause for concern.

Next, we fitted low- l modes in the BiSON, GOLF, GONG and MDI power frequency spectra (Section 4.2) to fitting models made with the Nigam–Kosovichev formalism. By taking differences between these results and those from fits to models based on equation (13), we had another means of judging the bias. The results (symbols for different time-series) are shown in the two panels of Fig. 8. The solid line in the left-hand panel shows the full best-fitting linewidth uncertainty; while the solid lines in the right-hand panel correspond to plus and minus one-tenth of the formal frequency uncertainty.

The results are shown to be in a reasonable agreement with those of the artificial data (Fig. 6). Overestimation of the linewidths (left-hand panel of Fig. 8) is on average larger than the overestimation implied by the results of the simulations. We believe this largely reflects the fact that in the real data it is not only the frequencies that vary over the solar cycle, but also the heights, linewidths (Chaplin et al. 2000) and peak asymmetries (Jiménez-Reyes et al. 2007). In deriving the equations (e.g. equation 13), and constructing the artificial data for the simulations, we assumed only the frequencies varied in time. Proper allowance will be made for the height, linewidth and asymmetry variations in the next phase of this work.

Bias in the frequencies (right-hand panel of Fig. 8) is again shown to be but a small fraction of the size of the frequency uncertainties.

6 CONCLUSION

We have studied the impact of the solar-cycle frequency shifts on the underlying shapes of p-mode peaks seen in power frequency spectra made from data spanning large fractions of the cycle period. Analytical descriptions of the resulting mode profiles were presented for two functions describing the shifts in time: a simple linear variation and a cosinusoidal variation to mimic the full solar activity cycle.

We presented plots of the profiles expected for shifts similar in size to, and also larger than, those observed for the low- l solar p modes. The analytical predictions were also validated by Monte Carlo simulations with artificial data. In summary, we showed that while any distortion of the Lorentzian-like profiles of the solar p modes is very modest, it is nevertheless just detectable. Furthermore, our analysis indicates that by fitting models to the usual Lorentzian-like models – which do not allow for the distortion – rather than new models we derive, overestimation (underestimation) of the linewidth (height) parameter results. This bias is estimated to be of size comparable to the observational uncertainties given by data sets of length several years. Bias in the frequency parameter is much less of an issue, being over an order of magnitude smaller than the frequency uncertainties.

The distortion discussed in this paper may, of course, be an issue for analysis of multiyear asteroseismic data sets on some stars that...
show Sun-like oscillations. The effect will be most important in those stars for which the ratio of the stellar-cycle frequency shifts to the mode linewidths is larger than for the Sun. Indeed, visible distortion of the mode profiles in asteroseismic data may provide an initial diagnostic of strong stellar-cycle signatures over the duration of the observations (Chaplin et al. 2007b; Metcalfe et al. 2007).

We finish by offering an answer to the question posed by the title of this paper. As far as the low-\(l\) solar p-mode frequencies are concerned, there is probably no need to worry about the distortion introduced by the frequency-shift effect. However, in the case of the height and linewidth parameters, a systematic bias of \(1\sigma\) or more means we do need to worry if we wish to obtain accurate estimates of these parameters.

ACKNOWLEDGMENTS

BiSON is funded by the UK Science and Technology Facilities Council (STFC). We would like to thank all those who are, or have been, involved in BiSON. GOLF and MDI are the result of the cooperative endeavours of several institutes, to whom we are deeply indebted. SOHO is a mission of international cooperation between ESA and NASA. This work also utilizes data obtained by the Global Oscillation Network Group (GONG) Program, managed by the National Solar Observatory, which is operated by AURA, Inc. under a cooperative agreement with the National Science Foundation. The data were acquired by instruments operated by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísica de Canarias and Cerro Tololo Interamerican Observatory. The authors acknowledge the support of STFC.

REFERENCES

Chaplin W. J., Elsworth Y., Isaak G. R., McLeod C. P., Miller B. A., New R., 1997, MNRAS, 287, 51
Chaplin W. J., Elsworth Y., Isaak G. R., Miller B. A., New R., 2000, MNRAS, 313, 32
Chaplin W. J., Appourchaux T., Elsworth Y., Isaak G. R., New R., 2001, MNRAS, 324, 910
Chaplin W. J., Elsworth Y., Isaak G. R., Miller B. A., New R., 2004, MNRAS, 352, 1102
Chaplin W. J., Elsworth Y., Miller B. A., New R., Verner G. A., 2007a, ApJ, 659, 1749
Chaplin W. J., Houdek G., Elsworth Y., New R., 2007b, MNRAS, 377, 17
Gelly B. et al., 2002, A&A, 394, 285
Jiménez-Reyes S. J., Corbard T., Pallé P. L., Roca Cortés T., Tomczyk S., 2001, A&A, 379, 622
Jiménez-Reyes S. J., Chaplin W. J., Elsworth Y., Garcia R. A., Howe R., Socas-Navarro H. Toutain T., 2007, ApJ, 654, 1135
Metcalf T. S., Dziembowski W. A., Judge P. G., Snow M., 2007, MNRAS, 379, L16
Nigam R., Kosovichev A. G., 1998, ApJ, 505, L51
Toutain T., Elsworth Y., Chaplin W. J., 2005, A&A, 433, 713

This paper has been typeset from a \TeX/LaTeX file prepared by the author.