

On the backwards difference filter (Research Note)

R. A. García^{1,2} and J. Ballot³

¹ DSM/DAPNIA/Service d'Astrophysique, CEA/Saclay, 91191 Gif-sur-Yvette Cedex, France
e-mail: rafael.garcia@cea.fr

² AIM – Unité mixte de recherche CEA - CNRS - Université Paris VII – UMR n°7158, 91191 Gif-sur-Yvette Cedex, France

³ Max Planck Institut für Astrophysik, Karl-Schwarzschild-Str. 1, Postfach 1317, 85741 Garching, Germany
e-mail: jballot@mpa-garching.mpg.de

Received 2007; accepted

ABSTRACT

Context. It is usual in helioseismology to remove unwanted instrumental low-frequency trends by applying high-pass filters to the time series. However, the choice of the filter is very important because it is suitable to keep the periodic signals in all the spectrum. At the same time, these filters should not introduce any spurious effects on the remaining signal that would modify the periodic signatures in the Fourier domain.

Aims. One of the most used filters is the so-called backwards difference filter that can be applied when the time series are regularly sampled. The problem of this filter is that the amplitudes **and the phases** of the periodic signals in the Fourier domain are modified and, therefore, their extracted amplitudes are biased when the frequency of the periodic signals decreases. The objective of this research note is to give a correction that could be applied to the power spectrum to get rid of this problem.

Methods. We properly derive the transfer function of the backwards difference filter and we show how this can be applied to the power spectrum to correct the amplitudes. **For the amplitude spectrum and the time series, we also derive a correction for the phase.**

Results. The amplitudes of the periodic signals in the resultant power spectrum density are corrected from the cut-off frequency down to zero.

Key words. Methods: data analysis – Methods: filters – Stars: oscillations – Sun: helioseismology

1. Introduction

One of the most common problems to be faced when analyzing oscillatory components of an observed signal is the presence of low-frequency drifts which can be of instrumental or natural origin. These low-frequency trends introduce a background on the Fourier domain that reduces the signal-to-noise ratio (S/N) of searched periodic signals. To avoid this, high-pass filters are commonly used. The ideal filter should be the one that reduces the effect of the trend while preserving the periodic signals even at low frequencies. Moreover, it is important to keep all the parameters of periodic signals untouched, including their amplitudes.

In this research note we study a very easy-to-use filter: the backwards difference filter which is commonly used in helioseismology (see for example Christensen-Dalsgaard 2002, for a complete review on this research topic) to reduce the low-frequency unwanted stochastic contribution of regularly sampled observations. We also deduce the transfer function of this filter. This transfer function allows to correct the filtered power spectrum in order to retrieve the correct amplitudes of the periodic signals in the spectrum. **Finally, we determine a correction for the resultant phase for the case we are interested in the amplitude spectrum or in the time series.**

2. Backwards difference filter

2.1. Definition

The backwards difference filter (hereafter called BDf) can be obtained by substituting the value of every data point of a regularly sampled time series by the difference of two consecutive points. Let $\{f_n\}$ be a set of regularly sampled measurements of a function $f(t)$, such that $f_n = f(t_n)$ with $t_n = t_0 + n\Delta t$, where Δt is the sampling rate. By definition, the filtered data set $\{\delta f_n\}$ is such that

$$\delta f_n = f_{n+1} - f_n \quad (1)$$

what can just be seen as the discretization of the function

$$\delta f(t) = f(t + \Delta t) - f(t) \quad (2)$$

Thus, for a given Δt , the BDf can be defined as a linear operator, δ , that can be applied to any function $f(t)$ using Eq. 2.

2.2. Transfer function of the backwards difference filter

Denoting $\tilde{f}(\omega) = \mathcal{F}[f(t)]$ the Fourier transform of the function $f(t)$, we can write that:

$$\begin{aligned} \tilde{\delta f}(\omega) &= \mathcal{F}[f(t + \Delta t) - f(t)] \\ &= [\exp(i\omega\Delta t) - 1] \tilde{f}(\omega) \\ &= 2 \sin\left(\frac{\omega\Delta t}{2}\right) \tilde{f}(\omega) \exp\left(i\frac{\omega\Delta t + \pi}{2}\right) \end{aligned} \quad (3)$$

We deduce that the power spectrum of the original function $f(t)$ and the filtered one, $\delta f(t)$, are linked by:

$$|\widetilde{\delta f}(\omega)|^2 = Q(\omega) |\widetilde{f}(\omega)|^2 \quad (4)$$

where $Q(\omega)$ is the transfer function of the BDf. It can be defined from the previous expressions by:

$$Q(\omega) = \left[2 \sin \left(\frac{\omega \Delta t}{2} \right) \right]^2 = \left[2 \sin \left(\frac{\pi \omega}{2 \omega_c} \right) \right]^2 \quad (5)$$

where we have introduced the cut-off frequency $\omega_c = \pi/\Delta t$.

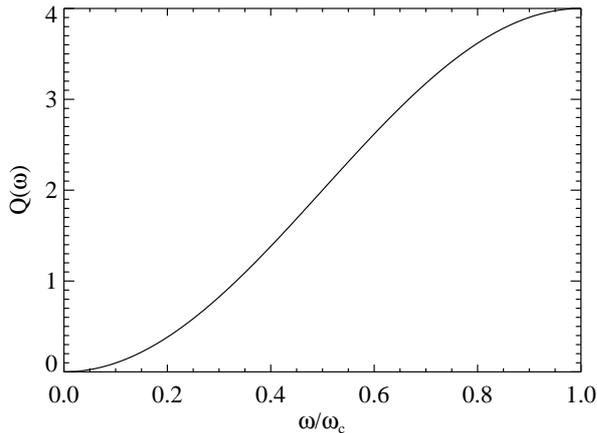


Fig. 1. Transfer function $Q(\omega)$ (defined in Eq. 5) of the backwards difference filter. The frequency ω has been normalized to the cut-off frequency $\omega_c = \pi/\Delta t$.

This transfer function shows the effect of the filter on the amplitude of the power spectrum (see Fig. 1). Indeed, at lower frequencies, the amplitudes tend towards zero. In order to recover the correct amplitudes in the filtered power spectrum, we only need to divide it by $Q(\omega)$:

$$|\mathcal{F}[f(t)]|^2 = Q(\omega)^{-1} |\mathcal{F}[\delta f(t)]|^2 \quad (6)$$

3. Example

To illustrate the behavior of the BDf and to show the advantage of such a filter, we have simulated a simple observation containing low-frequency drifts and periodic signals (see Fig. 2). The length of the simulated series is $T = 150\,000$ s with a constant sampling rate, $\Delta t = 1$ s. We have simulated three sine waves with periods of 1 000, 5 000 and 10 000 s, corresponding to frequencies of 10^{-3} , 2×10^{-4} and 10^{-4} Hz, in a way to cover a large range of frequencies. Their amplitudes are all equal to 1 000 arbitrary units. A low-frequency trend has been added to mimic the behavior of real observations. Thus, we have used a 7th-order polynomial function fitted to the light curve of the Procyon data recovered by the MOST¹ satellite. These data set can be found in the MOST public data archive² and it is described in Matthews et al. (2004).

¹ A Canadian Space Agency mission, jointly operated by Dynacon Inc., the University of Toronto Institute for Aerospace Studies and the University of British Columbia, with the assistance of the University of Vienna

² <http://www.astro.ubc.ca/MOST/data/data.html>

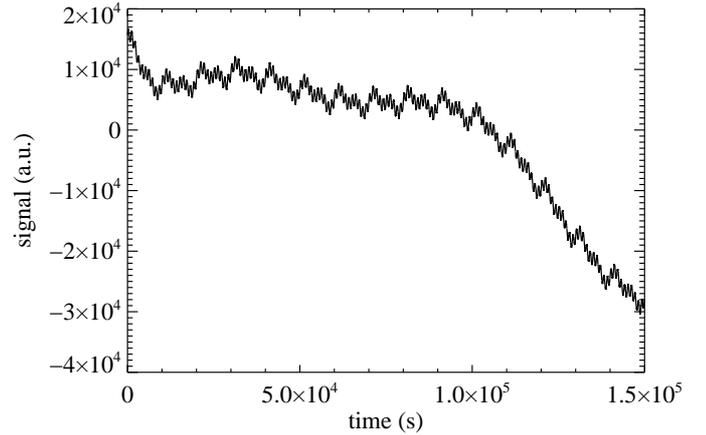


Fig. 2. Simulated time series of the sum of three sine waves and a background that mimics a standard asteroseismic observed background.

The power spectral density (PSD) – calibrated as a single-sided spectrum, see Press et al. (1992) – of the simulated time series is plotted in Fig. 3a. The low-frequency trend introduces a background in the power spectrum reducing the S/N of the sine waves as the frequency decreases when it is compared to the PSD of the three sine waves computed without any trend (cf. Fig. 3b). The application of the BDf reduces the effect of the unwanted background (cf. Fig. 3c). Indeed the S/N ratio increase comparing to the raw PSD. However, the amplitudes of the sine waves are modified by the transfer function of the filter. Thus, when we divide this PSD by the $Q(\omega)$ factor, we retrieve the correct amplitudes (cf. Fig. 3d).

Another way sometimes used to filter slow trends is to remove a smoothing of time series. To make a comparison, we have also plotted the PSD of our artificial data after applying such a filter. In this example, we have smoothed our time series by a boxcar function with a width of $T/10$ and subtracted it to the data. Figure 3e clearly shows that, in this last case, the background of the PSD is dominated by wiggles which modify the background shape. Such wiggles can easily introduce artifacts.

In Table 1 we have summarized the S/N ratio of the three simulated sine waves. The S/N ratio has been defined as the ratio of the peak amplitude to the local background. It does not take into account any effects of stochastic noise in time series. The increase in the S/N after applying the BDf is clearly shown. For the data filtered by smoothing, the variation of S/N ratio are a direct consequence of the wiggles, and is very sensitive to the size of the boxcar which has been chosen.

Table 1. S/N ratio in PSD of the three simulated sine waves at frequencies $\nu_1 = 10^{-4}$, $\nu_2 = 2 \times 10^{-4}$ and $\nu_3 = 10^{-3}$ Hz.

Applied filter	S/N		
	ν_1	ν_2	ν_3
No (raw data)	4	10	140
BDf (corrected)	314	5254	3×10^6
Smoothing	165	268	98

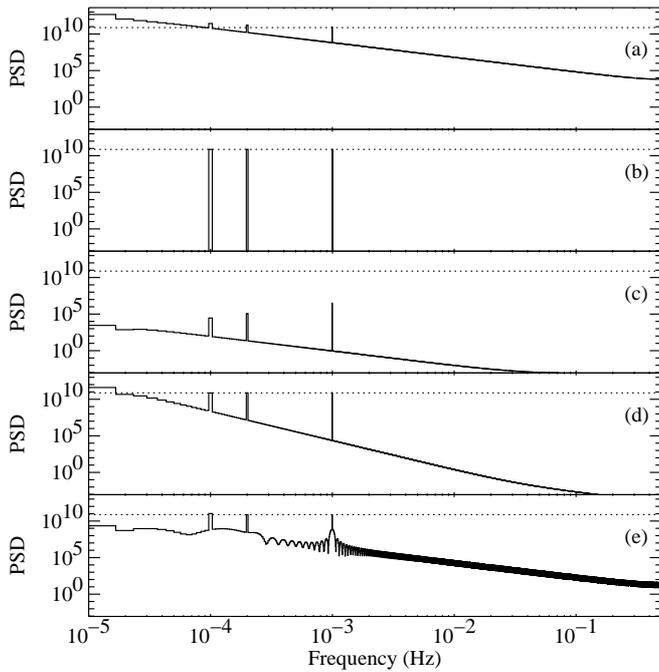


Fig. 3. (a) Power spectrum density of the raw time series; (b) PSD of the three simulated sine waves without the low-frequency trend; (c) PSD after a direct (uncorrected) application of the BDF; (d) the same spectrum corrected by the $Q(\omega)$ factor; (e) PSD of time series filtered by removing a smoothing. The horizontal dotted lines correspond to the power of the simulated sine waves.

4. Correction to the phase

We have seen that the BDF provides the right frequencies of the oscillatory signals and the right amplitudes after correcting them using the $Q(\omega)$ factor. However, the BDF also modifies the phase of the oscillations with a phase-shift depending on ω . If we denote ϕ the real phase and ψ the phase observed in δf , we can deduce a relationship between the two variables from equation (3):

$$\phi = \psi - \frac{\omega\Delta t + \pi}{2} = \psi - \frac{\pi}{2} \left(1 + \frac{\omega}{\omega_c} \right). \quad (7)$$

The phase correction varies linearly from $-\pi/2$ to $-\pi$, over the whole frequency range, from 0 to the cut-off frequency, ω_c .

The modification of the phase induces a modification of the shape of the total signal. Figure 4 illustrates this phenomenon. The first plot shows the beating of two oscillations with close frequencies (0.07 and 0.08 Hz), that can be seen as a modulated carrier wave of frequency 0.075 Hz. In the filtered series, the amplitude is, of course, affected: it is multiplied by a factor $2 \sin(0.075\pi) \approx 0.47$. Furthermore the wave has a phase lead of about $\pi/2$ (if we consider $\omega \ll \omega_c$). The second plot shows what happens for a mode ($\nu_0 = 0.07$ Hz) and its first overtone ($\nu_1 = 0.14$ Hz). We clearly see that the resulting pattern is different. This strong modification appears because the phases have been modified, even if the introduced phase-shift for both oscillations is almost the same ($\psi_1 \approx 1.8$ and $\psi_2 \approx 2.0$).

5. Limitations in the application of the BDF

As we have seen in the previous sections, the BDF can be a powerful detrender but it must not be overused. If there is no drift in the original signal, one has to avoid to apply this filter. Indeed,

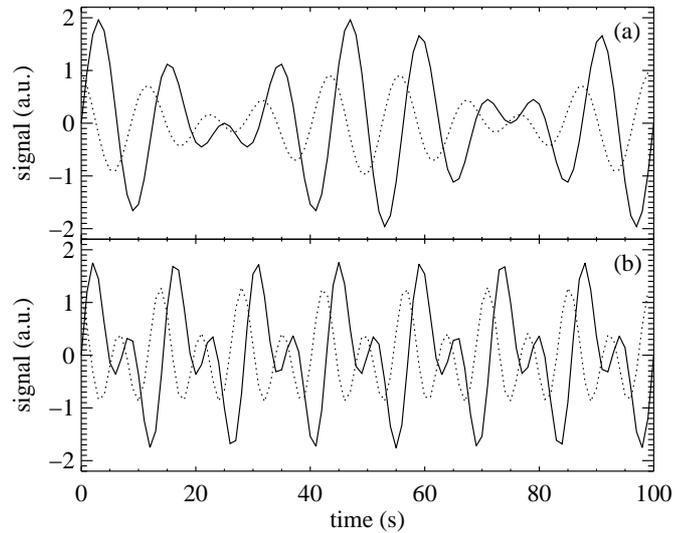


Fig. 4. Examples of time series before (solid lines) and after (dotted lines) applying the BDF: (a) The signal contains two purely sinusoidal oscillations with close frequencies $\nu_0 = 0.07$ Hz and $\nu_1 = 0.08$ Hz; (b) The signal contains a purely sinusoidal oscillation with a frequency $\nu_0 = 0.07$ Hz and its first overtone ($\nu_1 = 0.14$ Hz). The sample rate is $\Delta t = 1$ s.

there is always a small residual background in $\widetilde{\delta f}(\omega)$ that is amplified when we correct the amplitudes by dividing the original power spectrum by the $Q(\omega)$ factor. This effect is magnified at very low frequencies.

When the BDF is applied to any time series – by construction – we lose one point of the signal (the last one) that is usually left to zero to maintain the same number of points as in the initial series. Another problem could arise when data gaps are present in the original time series. When the BDF is applied to data containing N gaps, we lost N additional points as the consequence of the edges of those gaps. Therefore, if N is small compared to the total number of points, then the effect can be neglected. However if it is not the case, it has to be taken into account and if the number is too big, another kind of filter should be used.

6. Conclusion

In this research note we have shown the advantage of removing low-frequency trends by applying the backwards difference filter. The background in the PSD produced by this kind of trends is reduced and the S/N ratio of the periodic signals are improved. Moreover, the transfer function of this filter, $Q(\omega)$, has been deduced in order to correct – by dividing the spectrum by this quantity – the power of the periodic signals in the PSD. The resulting amplitudes inferred from the PSD are the original ones.

Filtering with a smoothing is often used, because it is a powerful detrending tool. However, it must be carefully applied and the choice of the boxcar size needs a preliminary study, according to frequencies of searched signals. The appearance of the wiggles must be taken into account, ideally by correcting the spectrum with the transfer function of this filter.

Even though this presentation does not deal with effects of stochastic noise, it emphasizes the advantages of the BDF and explains the manner to use properly this filter in a way very easy to implement.

Acknowledgements. The authors want to thank the members of the asteroFLAG group³ present at the first ISSI (International Space Science Institute) meeting for triggering this discussion. **The authors want to thank the anonymous referee for his useful comments that have improved the final version of this paper.**

References

- Christensen-Dalsgaard, J. 2002, *Reviews of Modern Physics*, 74, 1073
Matthews, J. M., Kusching, R., Guenther, D. B., et al. 2004, *Nature*, 430, 51
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P.
1992, *Numerical recipes in FORTRAN. The art of scientific computing*
(Cambridge: University Press, 2nd ed.)

³ <http://www.issi.unibe.ch/teams/Astflag/>