# Chapter 4

# Variability filters

## 4.1 Introduction

We have seen that stellar micro-variability will be an important noise source for space-based planetary transit searches, and that it is vital to reduce its impact for *Kepler* and *Eddington* to achieve their stated goals. The basic idea behind the variability filters developed in the present chapter is the following: it is possible to disentangle the planetary transit signal from other types of temporal variability if the two have sufficiently different temporal characteristics.

The micro-variability simulator introduced in Chapter 3 provides us with the means of illustrating and testing the effect of the filters we develop as solutions to this problem. As an example, we use throughout the present section a light curve simulated according to the planned characteristics of the *Eddington* mission, containing stellar variability, planetary transits and photon noise. The light curve lasts 3 years and has a sampling time of 10 min. The transits were simulated using the Universal Transit Modeller (UTM) software of Deeg et al. (2001), while the photon noise was simulated as Gaussian distributed noise with a standard deviation equal the square-root of the expected photon count per integration given the collecting area (0.764 $m^2$) and throughput of the December 2003 *Eddington* baseline design[1]. The light curve contains transits of a 2 $R_\oplus$ planet orbiting a G2V star ($R_\star$ = 1.03 $R_\odot$), i.e. a radius ratio of 0.018, leading to a relative transit depth of $3.24 \times 10^{-4}$. The planet's orbital period is 1 year, and its orbital distance 1 AU, leading to a transit duration of $\sim$ 13 hours. The epoch of the first transit is 1.5 day. The star's age is 4.5 Gyr and its apparent magnitude $V$ = 13, leading to a photon count rate of $8.4 \times 10^7$ 10 min integration. In this regime, the photon noise in each integration is well approximated by a Gaussian distribution with a standard deviation of $1.09 \times 10^{-4}$.

[1] `http://astro.estec.esa.nl/Eddington/Tempo/eddiconfig.html`

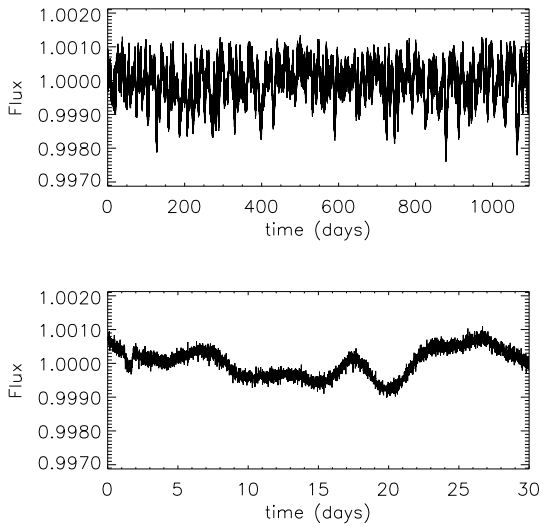


*Figure 4.1: Simulated Eddington light curve for a V =13 solar-age G2V star orbited by a 2$R_\oplus$ planet with a period of 1 year. Top panel: entire light curve. Bottom panel: first 30 days, with a transit 1.5 day after the start. The flux values shown have been normalised to a mean of 1.*

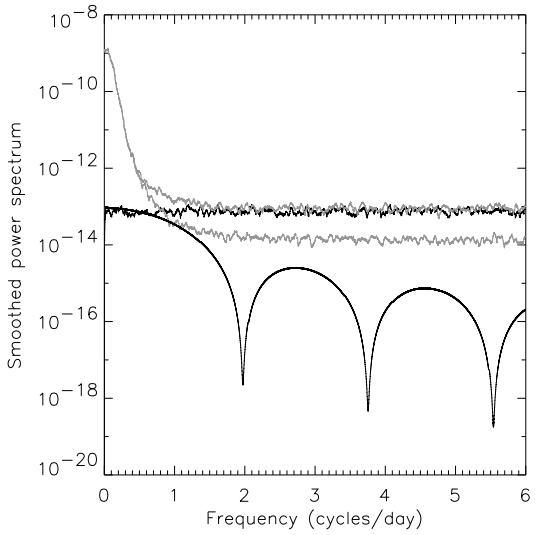


*Figure 4.2: Power spectrum of the light curve shown in Figure 4.1 (upper grey line). Lower grey line: stellar variability only. Lower black line: transits only (3 transits). Upper black line: photon noise. The power spectrum is dominated by stellar variability at low frequencies and by photon noise at high frequencies.*

The power spectra of the different components of the light curve mentioned above are shown in Figure 4.2. Although the power contained in the transit signal is small compared to both stellar and photon noise components (and would be even smaller for the case of an Earth-sized planet), it retains significant power for frequencies higher than $\sim 1\,\mu$Hz, where the stellar signal starts to drop off steeply. As long as the aforementioned condition that the planetary and stellar signal be sufficiently well separated in the frequency domain is fulfilled (i.e. if the stellar variability occurs on sufficiently long timescales), one should be able to separate and detect the transits. Furthermore, in the case of multiple transits, the regular period of the transits also helps constrain the Fourier space occupancy of the transit signal with respect to the stellar signal.

In the spirit of modularity adhered to throughout this thesis, the filters are developed as pre-processing tools, the output of which can be fed to a transit search algorithm. This differs from previous publications on the topic (Defaÿ et al. 2001; Jenkins 2002), which concern transit search algorithms specifically designed to detect signals buried in non-white noise. One advantage of our approach is that the filtered light curves can be searched for any kind of short-timescale event, not only transits.

The work presented in this chapter is the continuation of work carried out on this issue at ESTEC (Carpano et al. 2003). The latter article was a detailed exploration of a pre-whitened matched filter which we show in Section 4.2 to be closely related to a Wiener filter. This filter is then generalised to be applicable to data with gaps and/or irregular sampling in Section 4.3, while an alternative, the iterative non-linear filter, is presented in Section 4.4. This marks the difference between our approach and that of Defaÿ et al. (2001); Jenkins (2002) and Carpano et al. (2003): the filters are required to be *directly* applicable to data with gaps and irregular sampling, a problem that will affect any real dataset to some extent, whether ground- or space-based. Note that Jenkins (2002) does address this issue, but in a different way: by developing a method to effectively regularise the sampling before applying the algorithm. The characteristics of the light curves after application of the two filters are compared in Section 4.5, and their performance, particularities and potential improvements are discussed in Section 4.6.

## 4.2 Wiener or matched filtering approach

Carpano et al. (2003) demonstrated how use of an optimal filter can simultaneously pre-whiten and enhance the visibility of transits in data dominated by stellar variability. The Fourier-based method presented there is also closely related to a minimum mean square error (MMSE) Weiner filter. However, even for space-based missions uneven sampling of the data will occur. In these real-life cases, standard Fourier methods are no longer directly applicable and a more general technique is required.

To gain some insight to the problem consider the general case of intrinsic stellar variability, with the received signal $x(t)$ is composed of the three components:

$$x(t) = s(t) + r(t) + n(t) \tag{4.1}$$

where $s(t)$ is the intrinsic time variable stellar light curve, $r(t)$ is the transiting planet signal, and $n(t)$ denotes the measurement plus photon noise, which we can take to be random (and Gaussian in the cases of interest here)[2]. Each component is statistically independent, hence the expected power spectrum $\Phi(\omega)$ of the received signal is simply given by:

$$\Phi(\omega) = \left\langle |S(\omega)|^2 \right\rangle + \left\langle |R(\omega)|^2 \right\rangle + \left\langle |N(\omega)|^2 \right\rangle \tag{4.2}$$

and in the case of random, or white, noise $\left\langle |N(\omega)|^2 \right\rangle$ is a constant, hence guarantee-

[2]Strictly speaking, the 1st two terms in Equation (4.1) should be multiplicative, but in the limit of low amplitude variability and shallow transits, an additive combination is a very good approximation.

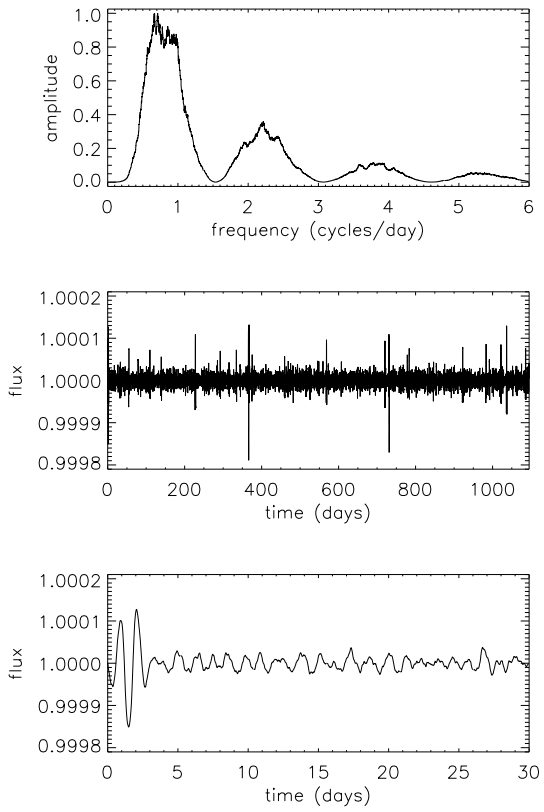


*Figure 4.3: Top panel: Wiener filter constructed using the light curve shown in Figure 4.1 and a reference box-shaped transit of duration 0.65 day . Middle panel: filtered light curve. Bottom panel: idem, 1st 30 days, with a transit 1.5 day after the start.*

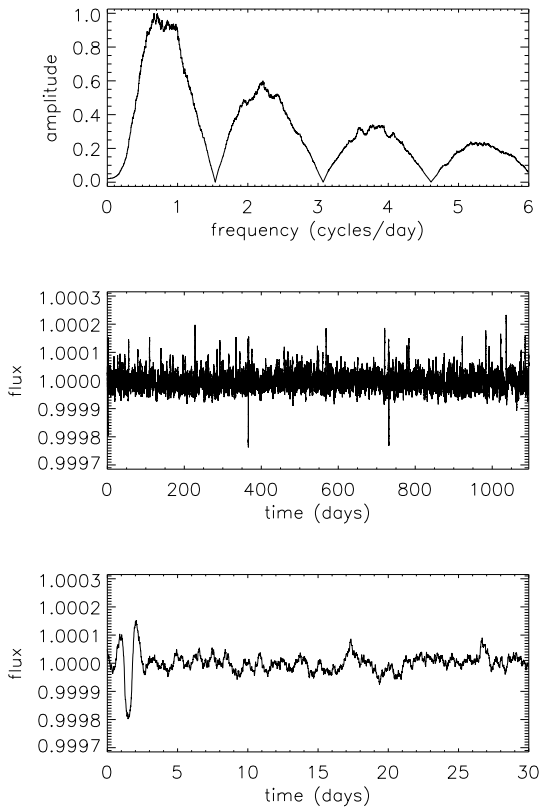


*Figure 4.4: Top panel: matched filter constructed using the light curve shown in Figure 4.1 and a reference box-shaped transit of duration 0.65 day. Middle panel: filtered light curve. Bottom panel: idem, 1st 30 days, with a transit 1.5 day after the start.*

ing positivity of the right hand term. This also highlights in a natural way a justification for the somewhat arbitrary constant in Equation (6) in Carpano et al. (2003) and how its value is related to the expected noise properties (although it would be more natural to implement it as a lower bound). However, as outlined below there is a simpler way to implement their technique without the need for the additional constant.

A standard MMSE Wiener filter attempts to maximise the signal-to-noise in the component of interest, in this case $r(t)$, by convolving the data with a filter, $h(t)$, constructed from the ratio of the cross-spectral energy densities between observation and target, such that:

$$x'(t) = h(t) \otimes x(t) \qquad X'(\omega) = H(\omega)\, X(\omega) \tag{4.3}$$

and (using * to denote complex conjugate):

$$H(\omega) = \frac{\langle R(\omega)R(\omega)^* \rangle}{\langle X(\omega)X(\omega)^* \rangle} = \frac{\langle |R(\omega)|^2 \rangle}{\langle |X(\omega)|^2 \rangle} \tag{4.4}$$

for a long enough run (a fair sample) of observations. In practice the only example we have of $x(t)$ is often singular, implying that the best estimate of the denominator is simply the observed power spectrum $\Phi(\omega)$, subject to the constraint of positivity imposed by the implicit $\langle |N(\omega)|^2 \rangle$ term. Such a filter is illustrated in Figure 4.3: the top panel shows the filter, constructed using the Fourier transform of the light curve shown in Figure 4.1 and a box-shaped reference transit of duration 0.65 day, and the bottom two panels show the filtered light curve. Note that this filtering method does modify the transit shape. In particular, it induces positive deviations either side of the transit, which is effectively equivalent to removing some of the transit signal as well as the stellar and noise signal. However, the transit signal-to-noise ratio is obviously enhanced, and it becomes easily discernible even by eye.

This should be contrasted with the pre-whitened matched detection filter employed by Carpano et al. (2003), illustrated in Figure 4.4 (using the same layout as Figure 4.3), and which can be written in the form:

$$X'(\omega) = H(\omega)\, X(\omega) = \frac{X(\omega)}{\langle |X(\omega)| \rangle}\ \langle |R(\omega)| \rangle \tag{4.5}$$

and hence is equivalent to reconstructing the data using just the phase of the input signal Fourier transform modulated by the amplitude spectrum from the expected transit shape (see Figure 4.5). Viewing the problem in this way removes the need for the additional constant in their Equation (6) and emphasises the two stage nature of the filtering. The pre-whitening suppresses the stellar variability component, while the matched filter is directly equivalent to the transit search algorithm case presented in Section 2.2.2 with $n = 1$.

In practice, transit searching can be based directly on the output of the filtering, or preprocessing can be used to decouple the stellar variation estimation from the transit search phase, which then proceeds using algorithms optimised for white Gaussian noise. (In either case, detailed investigation of the transit depth and shape involves phase folding, unfiltered data, and local modelling.)

Either of these preprocessing filters works well in the case of regularly sampled data with no gaps and with a reasonable separation between the signatures of the Fourier components of the transits and the stellar variability. In Figures 4.3, 4.4 & 4.5, the transits are distinctly visible in the filtered light curve. The results in terms of transit detection performance using either method are very similar. For simplicity, the

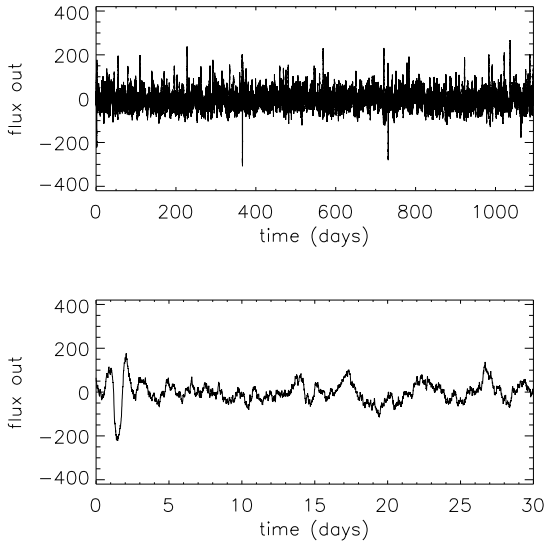


*Figure 4.5: As Figure 4.4, but the filtered light curve was obtained by modulating the phase of the Fourier transform of the data by the amplitude spectrum of the reference transit signal. The filter was omitted as it is effectively identical to that shown in Figure 4.4. Comparing, visually, the amplitude, shape and timescale of the variations in the filtered data with the bottom two panels of Figure 4.4 confirms that this gives very similar results to the matched filter approach.*

matched filter approach, rather than the Wiener filter, is used in the remainder of this paper.

However, real data, even space-based, suffers from irregular sampling and the presence of significant gaps. Fourier domain methods cannot be directly applied to irregularly sampled data, but it is possible to treat regularly sampled data with gaps as a series of $n$ independent time series, and to filter them separately. To test this, four arbitrarily chosen sections were removed from the light curve shown in Figure 4.1 (see Figure 4.6). The matched filter was then applied to the five unbroken intervals separately, and the results are shown in Figure 4.7. Though the filtering is effective on relatively long sections of data (bottom panel) it is not successful for short intervals (middle panel), even if they are significantly longer than the transit duration. This is because the power spectrum of the stellar noise is estimated from the data in order to construct the filter. For this to be successful, the data segment needs to be at least twice as long as the longest significant timescale in the star's variability, which is either the rotation period or the long end of the starspot lifetime distribution

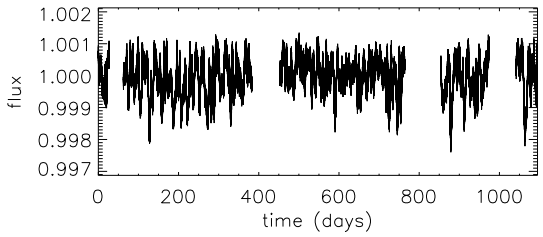


*Figure 4.6: Simulated light curve with data gaps. Four arbitrarily chosen sections were removed from the light curve shown in Figure 4.1. Note that the gaps were chosen to avoid the transit regions.*

(Aigrain et al. 2004). In the case of the G2V star used in the simulations, the minimum data segment length for which the filtering was successful was $\sim 60$ days (last data segment in Figure 4.7), consistent with a rotation period of $\sim 30$ days for such a star.

It is therefore necessary to find other means of coping with this additional complexity. We have investigated two alternative approaches: one based on a least-squares generalisation of the Fourier filtering approach; the other based on a general purpose iteratively clipped non-linear filter. In both cases we use the preprocessing to attempt to remove the stellar signature, as much as possible, prior to invoking the transit search algorithms developed in Chapter 2.

## 4.3 Least-squares fitting

For a long run of regularly sampled data, a discrete Fourier transform asymptotically approaches a least-squares fit of individual sine and cosine components (see e.g. Bretthorst 1988). This naturally suggests an extension of the approach described in Section 4.2 to the case of irregularly sampled data. An analogous situation occurs in the generalisation of the periodogram method to Fourier estimation of periodicity; using generic least-squares sine curve fitting is a more flexible alternative (Brault & White 1971). This allows the case of gaps in the data, or more generally irregular sampling, to be dealt with in a consistent and simple manner.

The procedure is basically identical to that employed for the Wiener or matched filters described in the previous section, but the calculation of the Fourier transform, or power spectrum, of the received signal is replaced by an orthogonal decomposition of this signal into sine components whose amplitude, phase and zero-point are fitted by least-squares. Each of the components has the form:

$$\psi_k(t) = a_k \sin(2\pi k t/T + \phi_k) \tag{4.6}$$

where $T$ is the time range spanned by the data. The number of components to fit can be chosen such than the maximum frequency fitted is equal to some fraction of the Nyquist frequency, but for this one must define an equivalent sampling time $\delta t$. In the case of regular sampling with gaps, $\delta t$ is simply the time sampling outside the gaps. In the case of irregularly sampled data the definition of $\delta t$ is more open ended. However, provided that the sampling is close to regular, a good approximation will be the average time step between consecutive data points – keeping in mind that any significant gaps should be excluded from the calculation of this average. The potentially highest frequency component should then have frequency $\approx 1/(2\delta t)$, although in practice a much lower frequency cutoff for the components is all that is

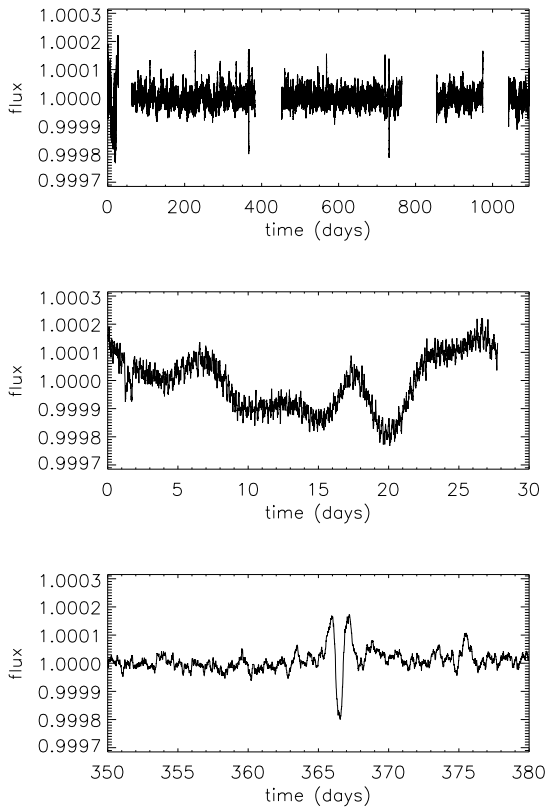


*Figure 4.7: Results of applying the matched filter independently to the 5 unbroken intervals of the light curve shown in Figure 4.6. Top panel: entire filtered light curve. Middle panel: 1st 30 days. Bottom panel: another 30 day section centred on the second transit (at 366.5 days). See text for an explanation.*

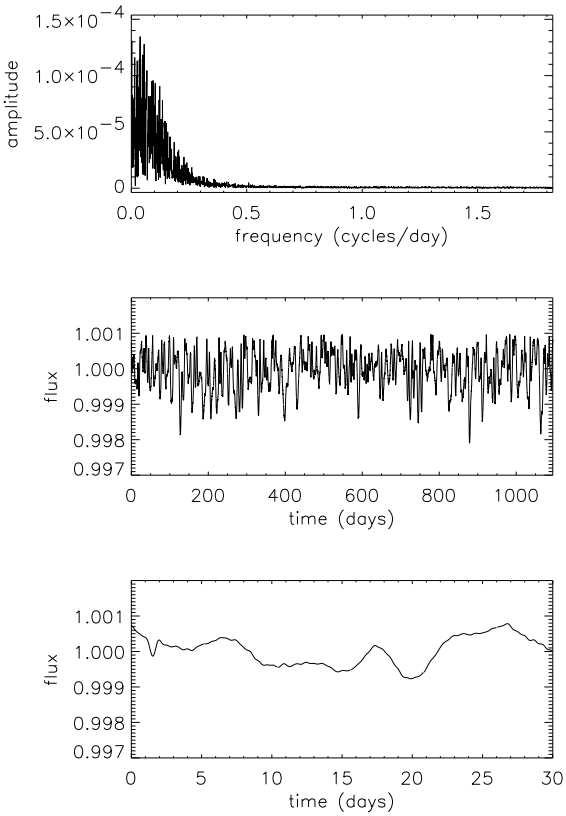


*Figure 4.8: Top panel: "power spectrum" (i.e. coefficients $a_k$ versus frequency) obtained by the least-squares fitting method for the light curve shown in Figure 4.1. Middle panel: reconstructed light curve, obtained by summing over the fitted sine-curves up to a frequency of $\sim$1.8 cycles/day. Bottom panel: 1st 30 days of the reconstructed light curve.*

required.

Note that the first (zero-frequency) component is effectively the mean data value $\langle x(t) \rangle$ (which can be pre-estimated and removed in a robust way e.g. by taking a clipped median). The presence of gaps in the data provides us with a natural way of obtaining several independent estimates of $\langle X_{ls}(w) \rangle$ by measuring it separately in each interval between gaps, or alternatively provides a natural boundary for doing independent light curve decompositions.

Figure 4.8 illustrates this least-squares fitting method, as applied to the light curve shown in Figure 4.1. The top panel shows the "power spectrum", i.e. the coefficients $a_k$ versus frequency, while the bottom two panels show the light curve re-

constructed by summing the fitted sine-curves. Note that high frequency variations are not reconstructed as only the first 2000 sine components were fitted (well below the Nyquist limit, but amply sufficient for the purposes of following the long timescale stellar variability).

The decomposition of the reference (transit) signal can usually be well approximated analytically. For example if a simple box-shaped transit of duration $d$ is adopted as reference signal, the $k^{\text{th}}$ coefficient is given by:

$$r_k = \frac{\sin(\pi k d/\delta t)}{\pi k d/\delta t} \tag{4.7}$$

However, this decomposition can also be performed in the same way as for the received data, for a reference signal of any given shape. The sets of coefficients $a_k$ and $r_k$ then define the filter $h_k$, which is equivalent to the filters of the previous section:

$$h_k = \frac{\langle |r_k|^2 \rangle}{\langle |a_k|^2 \rangle} \qquad \text{or} \qquad h_k = \frac{\langle |r_k| \rangle}{\langle |a_k| \rangle} \tag{4.8}$$

where the first expression corresponds to the standard Wiener filter, and the second to the pre-whitened matched filter used in Carpano et al. (2003).

Figure 4.9 illustrates this filtering method. Using the second expression in Equation (4.8) (equivalent to Equation 4.5), a 'matched filter' $h_k$ is constructed from the coefficients $a_k$ and $r_k$ (the latter computed according to Equation 4.7). The filtered light curve, obtained by multiplying the $a_k$ by $h_k$ and reversing the 'transform', is shown in the middle panel, with a zoom on the first 30 days in the bottom panel.

Figure 4.10 shows the results of the matched filter constructed using the least-squares fitting method when the light curve contains gaps (as in Figure 4.6). The performance of the filter is generally not affected by the gaps, though artifacts near gap boundaries can sometimes be introduced.

The case of irregular sampling is not illustrated here, for practical reasons: if the sampling was allowed to vary, say, by $\pm 10\,\%$ of the normal sampling time in a random fashion, the effect is not visible in plots of such long light curves. In any case, we have found it to have negligible effect on the the least-squares filtering.

Note that the combination of the least-squares fitting method to construct power spectra with the pre-whitened matched filter (RHS of Equation 4.8) will, for conciseness, be referred to hereafter simply as 'the least-squares filtering method'.

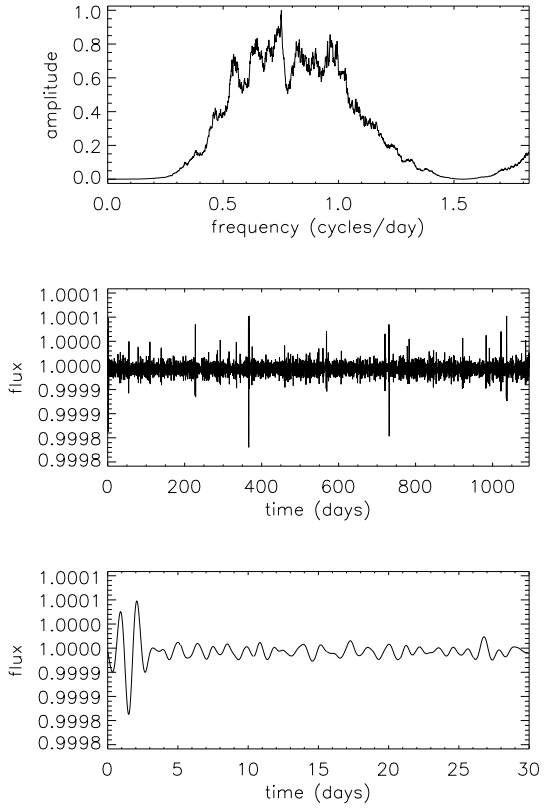


*Figure 4.9: Top panel: equivalent matched filter constructed according to the 2nd expression in Equation (4.8), using the light curve shown in Figure 4.1 and a reference box-shaped transit of duration 0.65 day. Middle panel: filtered light curve. Bottom panel: idem, 1st 30 days, with a transit 1.5 day after the start.*

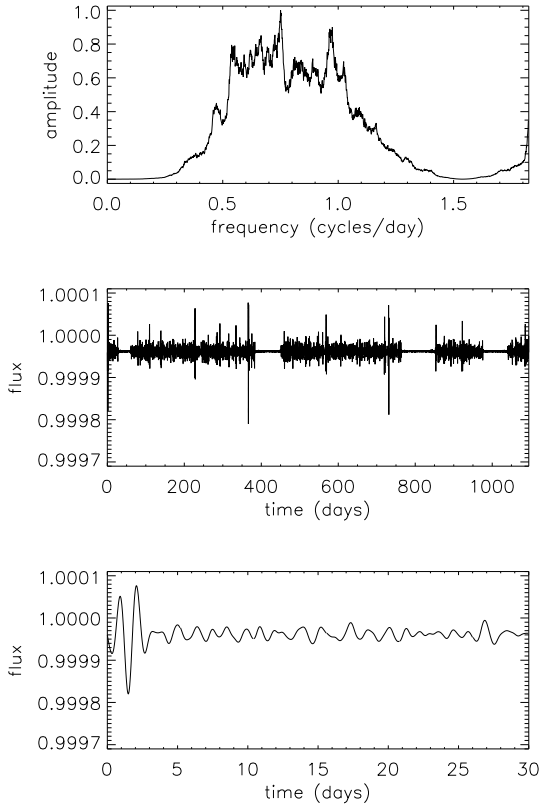


*Figure 4.10: As Figure 4.9, but the input light curve is that shown in Figure 4.6, with 4 significant data gaps.*

## 4.4 Non-linear filtering

If the timescale of the transits is shorter than that of the dominant stellar variations, iterative non-linear time domain filters can pick out short timescale events. A standard median filter is a good starting point for this type of approach.

The data are first, if necessary, split into segments, using any significant gaps in temporal coverage to define the split points. These gaps, defined as missing or bad data points, or instances where two observations are separated in time by more than a certain duration, can be automatically detected.

Each segment of data is then iteratively filtered using a median filter of window $\sim$ 2 to 3 times the transit duration, followed by a (small window) box-car filter

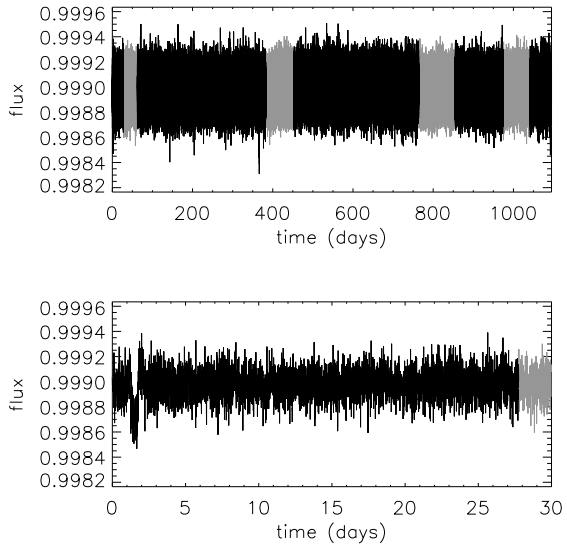


*Figure 4.11: Light curve with data gaps filtered using the non-linear technique (black curve). The input data was the light curve shown in Figure 4.6. The window of the iterative median filter used was* $3 \times 0.65$ *days. The grey curve shows the same data with the residual noise level after filtering measured and artificial data with Gaussian distributed noise of the same standard deviation generated to fill the gaps. This illustrates the fact that, after non-linear filtering, the light curve (outside the transits) is well approximated by a constant level plus white noise.*

to suppress level quantisation, which avoids excessive sensitivity to low-level dips (i.e. potential transits). Irregularities in sampling within a given segment are ignored, the filtering window being defined simply in terms of a number of data points. The difference between the filtered signal and the original is used to compute the (robust) MAD-estimated scatter (sigma) of the residuals. The original data segments are then $k$-sigma clipped (with $k$ = 3) and the filtering repeated, with small gaps and subsequent clipped values flagged and ignored during the median filtering operation. The procedure converges after only a few iterations.

While applying this filter in the context of the COROT blind experiment (see Chapter 6), we found that the results were improved by the addition of a pre-filtering step, which consists in smoothing each interval using a median filter with a width of 2 or 3 data points before applying the iterative non-linear filter to construct the continuum (though the continuum is still subtracted from the original, unsmoothed data to give the filtered light curve). Because it smooths the sharp edges of the transits, this pre-filtering reduces the amount of transit signal (which we want to preserve) that is removed by the filter.

Break points and/or edges are dealt with using the standard technique of edge reflection to artificially construct temporary data extensions. This enables filtering to proceed out to the edges of all the data windows. The continuum obtained in this manner is then subtracted from the original to obtain the filtered light curve.

The main advantage of using a non-linear filter is that the exact shape of the transit is irrelevant and the only free parameter is the typical scale size of the duration of the transit events. The main drawback is that the temporal information in the

segments is essentially ignored. However, providing the sampling within segments is not grossly irregular this has little impact in practice. This filter is also relatively fast due to its simplicity: with the 512 MB RAM, 1.2 GHz processor laptop previously used to test the box-shaped transit finder (see Chapter 2), the running time for a transit duration of $\sim$ 0.5 day is 4 seconds per light curve, about the same as the time required for the standard Wiener filter. The least-squares filtering method was significantly slower (requiring approximately 30 s when 1500 frequencies were fitted).

Figure 4.11 illustrates the non-linear filter as applied to the light curve with gaps shown in Figure 4.6. As with the indirect least-squares filtering, the high frequency noise remains, but this does not impede transit detection. Given the simplicity of this method and its good performance in the presence of data gaps, it appears to be the most promising, as long as the sampling remains relatively regular (if the sampling is significantly irregular, the least-squares filtering method, which takes the time of each observation into account directly, is likely to perform better).

## 4.5 Light curve characteristics after filtering

Several factors are to be taken into account when assessing the performance of the filters:

- How noisy is the filtered light curve?
- How Gaussian is the noise distribution in the filtered light curve?
- Has the transit shape and depth changed in the filtering process and how much?

Comparing Figures 4.10 and 4.11, one can readily see that, while both filtering methods suppress low frequency variations, the least-squares filtering method enhances any variations on the timescale of the reference transit (including the real transit) while suppressing high frequency noise. It also changes the shape of the transit significantly, as well as enhancing its contrast. On the other hand, the non-linear filter does not affect any variations on timescales shorter than two or three times the transit duration. This implies that, although the transit is more obvious to the naked eye after application of the least-squares filtering method, the noise distribution is close to Gaussian after non-linear filtering.

This is illustrated in a quantitative manner in Figure 4.12, which shows distributions of the deviations from the median before and after filtering (note that the transits were excluded from these distributions). In each case, a 1-D Gaussian was fitted to the distribution. The respective half-widths of the fitted Gaussian were $\sim 6 \times 10^{-4}$, $1 \times 10^{-5}$ and $1 \times 10^{-4}$. The least-squares filtering method therefore removes more

noise. It also reduces the transit depth, though by a lesser factor: the approximate transit depth in each case was $3.2 \times 10^{-4}$, $1.4 \times 10^{-4}$ and $3.2 \times 10^{-4}$. The ratio of the transit depth to the Gaussian width is thus higher after least-squares filtering than after non-linear filtering. However, the distribution after non-linear filtering is much closer to a Gaussian.

## 4.6 Discussion

The two filtering methods presented here share some advantages – both can be applied to data with gaps – but they also have different properties.

The least-squares filtering method is capable of making use of the time information in data with irregular sampling. It also allows a theoretically optimal filter (i.e. the Wiener or matched filter) to be combined with a pre-whitening filter, although from the point of view of detection, the matched filter is the main active component of any maximum likelihood-based detection algorithm. It is designed to be the method with the highest performance in terms of enhancement of the transit depth to noise ratio. As a by product of the filtering, the stellar signal can also be reconstructed. However, this is computationally intensive, particularly if one wishes to fit higher frequencies. Its performance also depends quite critically on concordance between the duration of the reference transit and that of any true transit. Its primary use in the context of space-based transit searches will therefore be the detailed investigation of borderline candidates, originally identified in light curves treated with the non-linear filter: there will be few of these so that the computing time requirements do not matter, and one will already have an idea of the approximate transit depth.

On the other hand, iterative non-linear filtering is simple to implement and fast, and produces nearly Gaussian residuals. This is a very important point because transit search algorithms in general – and in particular those developed in Chapter 2 – are optimised for white Gaussian noise. It is also less sensitive to the choice of reference transit duration, because it simply removes any signal on timescales longer than two or three times this duration, rather than applying a Fourier domain filter which has a complex structure over a wide range of frequencies. The non-linear filter is thus our filter of choice for space-based transit searches, where the time sampling is regular apart from the occasional data gaps. However, it ignores any local time information (except for the long gaps which are detected automatically). This means that its performance is likely to degrade if the sampling is seriously irregular, e.g. ground based transit searches, where one will have to resort to the least-squares filtering method.

Whatever the method used, there is a fundamental limit to what can be filtered

out. Stellar variability can only be filtered out if an orthogonal decomposition of the transit and stellar signal is possible, e.g. if the two signatures in the frequency domain do not overlap too much. Therefore, very rapidly rotating stars where the rotation period is close to the transit duration, or stars showing much more power than the Sun on timescales of minutes to hours (e.g. higher meso- or super-granulation) will be problematic targets – although perfectly periodic stellar signals, even if they have large amplitudes and periods close to transit timescales, are easily removed using e.g. the sine-fitting technique discussed in Chapter 6, Section 6.3.1.3.

Of course, these filters will not be used in isolation. The quantity we are really interested in is the performance of the transit search algorithms when applied to their output. This is the subject of the next chapter, where the filters are coupled with the box-shaped transit finder from Chapter 2 and applied to the output of the simulator from Chapter 3 for a wide range of input parameters, given the observational characteristics of the *Eddington*, COROT and *Kepler* missions. The same combination is also tested on COROT data simulated on the basis of inputs from a number of European groups as part of the blind experiment described in Chapter 6.

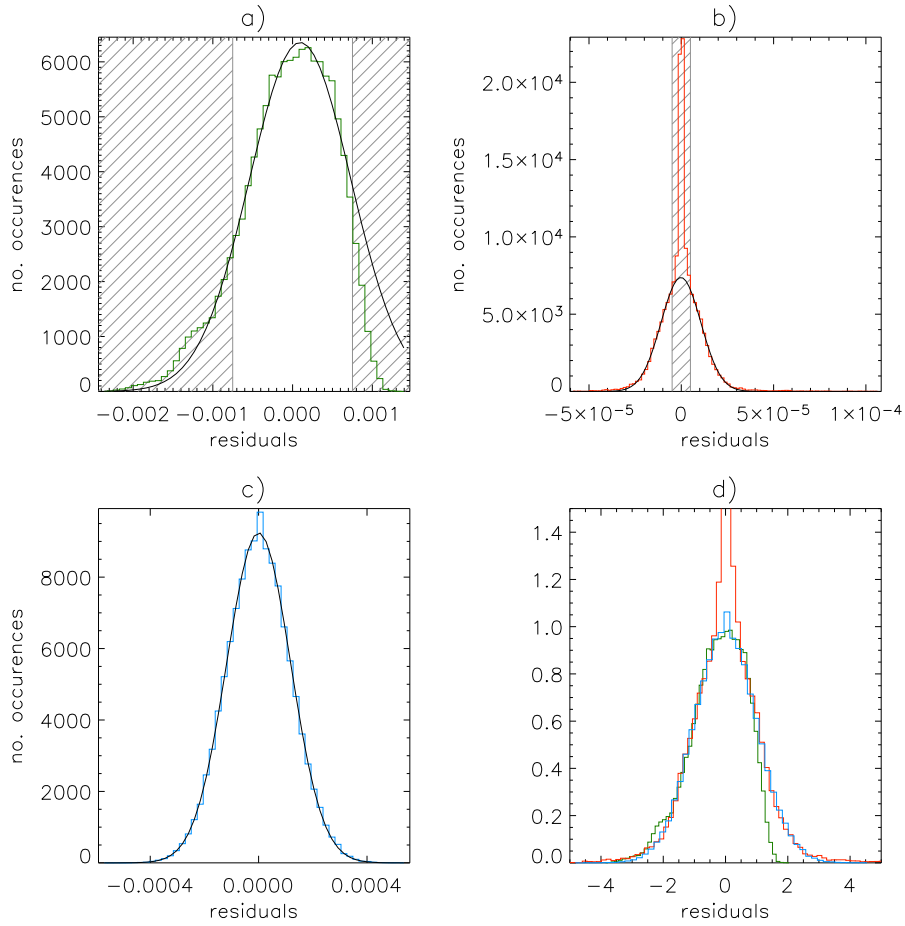


*Figure 4.12: Distributions of the deviations from the median* **a)** *before filtering (in green),* **b)** *after applying the least-squares filtering method (in red) and* **c)** *after applying the non-linear filter (in blue). A Gaussian fit is shown in each case (black line). Obvious departures from Gaussianity were excluded from the fits (grey hashed regions).* **c)** *shows all three distributions, scaled and shifted so that each fit corresponds to a zero mean, unit variance, unit amplitude Gaussian.*