Instrument Science Report

No. OTA 18

Title: Focus monitoring and recommendation for secondary mirror move

Author: Stefano Casertano, SIB

Date: January 3, 1995

Abstract

Phase retrieval on well-exposed PC images shows that shrinkage of the OTA continues at a rate of about $0.90 \pm 0.15\,\mu$m/month, close to that determined by FOC measurements. A movement of $5.0\,\mu$m of the secondary mirror away from the primary is recommended for mid-January, to compensate for the shrinkage occurred since the last mirror movement on June 29, 1994. A program of focus monitoring based on phase retrieval solutions is recommended.

1. Introduction

This memo summarizes the results of a systematic study of the characteristics of WFPC2 images over the period March-December 1994, undertaken with the primary goal of monitoring small focus changes in the OTA.

In the first three years after launch, the focus position changed significantly, probably because of desorption-induced OTA shrinkage; see Hasan, Burrows and Schroeder (1993) for a detailed history of the HST focus and of the measures taken to keep it under control. Since mid-1993, the focus changes have been relatively predictable, and systematic movements of the secondary mirror have been executed to keep it within a few μm of the optimal focus. Emphasis was to maintain a focus position within 5–10 μm of the optimal value. (All measurements of the focus position, regardless of their cause, are expressed here in terms of the equivalent secondary mirror position; zero corresponds to best focus, and negative values indicate secondary mirror too close to the primary.)

After correction of the spherical aberration for both cameras, the demands for an accurate focus have increased. On the basis of pre-refurbishment simulations, Bèly & Hasan (1993) and Hasan & Bèly (1994) have derived a goal of $\pm 2\,\mu$m for the deviation of the secondary mirror from its optimal position in order to prevent significant image degradation in the UV. This criterion is indeed adequate for WFPC2, and it appears that some measurable degradation of the FOC images at the shortest wavelengths is visible even with somewhat smaller deviations (Jedrzejewski, private communication).

The results of the early monitoring program indicate that a feasible goal is maintaining the secondary mirror within 2–3 μm of its optimal position with two movements of the secondary mirror per year, to compensate for the continuing shrinkage of the OTA. The somewhat more stringent requirements of the FOC can be met with additional movements of its own internal focusing mechanism, to be performed when necessary - of order of four movements per year, two of which are to be coordinated with the secondary mirror movements.

One potential difficulty with this program is that it requires frequent measurements of the OTA shrinkage with accuracy of order of 1 μm. This accuracy is not easy to achieve with WFPC2 images, which are not strongly affected by changes of this magnitude in the focus position for near-focus images (see Hasan 1993, 1994 and references therein).

Nonetheless, it is highly desirable to monitor the focus position accurately with the WFPC2, for a number of reasons. First, accurate knowledge of the PSF as a function of time is important in a number of projects that attempt to measure faint extended sources near bright point sources. Second, although the Faint Object Camera is more strongly affected by changes in the focus position at the 1 μm level, focus monitoring with the FOC is quite expensive in terms of telescope resources. Third, focus monitoring with WFPC2 images can be achieved with essentially no additional cost in telescope time, since—as will be seen in the following—the images acquired as a part of the WFPC2 calibration suite suffice to measure at least the long-term trends in focus position. Finally, monitoring the long-term variation in focus position is a significant quality control issue for WFPC2, which lacks an internal focus mechanism.

Three methods have been used to try to follow focus changes. The first uses the sharpness parameter, with a definition similar to that of Hasan (1994), but with some modifications to take into account the noise contribution. The second is direct, qualitative inspection of PC images of the spectrophotometric standard GRW +70D5824, observed as part of the photometric monitoring program (currently Proposal 5563, PI Mark Clampin). The third, and most successful, uses the phase retrieval code developed by Krist and Burrows (1995) to reproduce the detailed shape of the image of GRW +70D5824, and therefore determine the relative focus position. A simplified correction for telescope breathing has also been attempted.

The results are discussed in the framework of a possible mevement of the secondary mirror, for which January 15, 1995 has been used as a target date; the results can easily be adapted for other target dates, if necessary.

2.1 The sharpness measure

The sharpness Σ of an image is defined as

$$\Sigma = \big(\sum_{i,j} w_{ij} I_{ij}\big)^2 / \sum_{i,j} \big(w_{ij} I_{ij}\big)^2$$

where $I_{i,j}$ is the intensity - in arbitrary units - at pixel (i,j). The w_{ij} are suitable weighting factors, which in practice have been set to unity. The sharpness is defined in a box of fixed size; it is advisable to keep the box size small (7×7 or 9×9 at the largest) to minimize the contribution of noise.

The sharpness parameter has several desirable properties: it is fast and easy to measure, it is model-independent, it is in principle independent of the signal in an image (but see below), and it measures directly a desirable quality of an image. Furthermore, useful calibration curves of the sharpness as a fucntion of focus position at different wavelengths have been determined from the focus sweep carried out after WFPC2 was installed.

On the other hand, the sharpness parameter suffers from a few crucial weaknesses for the current purpose. First, the sharpness depends on the location of the image center within a pixel, especially for undersampled images. Second, any noise in the image will introduce both a bias and significant scatter in the value of the sharpness. The bias is in the sense of noisier images having a larger sharpness; thus, faint stars will have systematically larger sharpness than bright stars. The amount of the bias can be predicted and subtracted, but at the cost of introducing additional scatter in the measurement. Third, any image defect (bad/hot pixels, cosmic rays etc) affects the image sharpness; because of its non-parametric definition, it is not easy to correct for missing pixels. Fourth, sharpness is quadratic in focus position near best focus, and thus it is not a sensitive function of focus position for in-focus images; it is also symmetric, and thus it cannot in principle discriminate - for small shifts - between 'too near' and 'too far', although this latter limitation can be overcome by using information from all four cameras, which have known small offsets from the individual optimal focus position.

Finally, there is a practical issue related to the availability of the necessary data to carry out the sharpness measurement. Since the sharpness parameter cannot be reliably measured from a single stellar image, calibration stellar fields - such as ω Cen and NGC 6752 - have been used for such measurements. However, the number of stars in the PC part of such observations is not very large (of order 20), and the stars are not very well exposed for the purpose of a sharpness measurement.

I have carried out sharpness measurements for about 20 stars in the PC, using two different filters and two different epochs. I have corrected for the magnitude dependence of the measured sharpness either by using an analytic estimate of the expected bias or by selecting a subsample in a narrow magnitude range. In order to remove the dependence on position within the central pixel, I have looked for correlations between the corrected sharpness and the distance of the image centroid from a pixel center. When no statistically significant correlation could be found, I tried using the median sharpness of the sample. None of these methods could give an overall sharpness measurement with fractional error better than 15%, which translates - near focus - to an error of 5 μm, much larger than the desired accuracy.

2.2 Qualitative inspection of stellar images

As a consequence of the small amount of astigmatism present in the PC, the appearance of well-exposed stellar images changes as the instrument moves through its ideal focus. Specifically, images taken at a negative focus value (secondary closer to primary than optimal) appear slightly elongated along the positive diagonal, and images taken at a positive focus position appear elongated along the negative diagonal (Krist, private communication).

A composite of images taken at monthly intervals is shown in Figure 1. With careful analysis, it can be seen that most images are slightly elongated; the change in orientation of the major axis is most evident when comparing the June 14 and the July 4 images, which were taken just before and after the 5 μm move of the secondary mirror away from the primary, on June 29, 1994. (Detecting the image elongation is made easier by changing the contrast on the screen, and it may be tricky in the printed images.) The interpretation is that the focus position was negative just before the move (image elongated towards PA -45), and became positive by roughly the same amount after the move (image elongated towards PA $+45$).

The trend continues after July, with the image elongation progressively decreasing until September, when it is essentially round. This corresponds to the nominal zero position for the focus. In October, the image begins to show an elongation in the perpendicular direction, which grows through November and December. These observations, although qualitative, provide strong support for a continuing shrinking trend of the OTA, in an amount that can be estimated in approximately 1 μm per month.

I have attempted to turn these considerations into a more quantitative criterion, based on combinations of image moments; this has not been very successful, because of a combination of image noise and subpixel positioning. For a truly quantitative study, a more powerful tool was needed, and is discussed in the next section.

2.3 Phase retrieval solutions

The qualitative study of the PSF uses only a small part of the information present in its detailed shape. Clearly, more information is present, but retrieving it requires a sophisticated model of the telescope and WFPC2 optics. Fortunately, such a tool is available in the phase retrieval program developed by Chris Burrows and John Krist, which has kindly been made available for this study. The phase retrieval program tries to reproduce an observed, well-exposed image of a point source, by varying the position of the source at the subpixel level, as well as the focus, coma, astigmatism, and other telescope and image parameters. 'Features' such as spreading in signal between pixels are also taken into account. A detailed description of the algorithm, its properties and limitations can be found in Krist and Burrows (1995).

For this focus analysis, I have been able to build on the considerable experience accrued by Burrows and Krist on WFPC2 images, and I needed to optimize only a few of the many adjustable parameters available in the program. For the other parameters, I have used the best values previously determined by the study of a large number of WFPC2 images at different locations in the various chips. The parameters I actually varied are the sub-pixel location in x and y, the two components of coma, and of course the focus.

The phase retrieval program has been used to fit the PC image of the photometric standard star GRW +70D5824, the same used in the previous section, in two filters: F439W and F814W. The fit is carried out on a 35×35 pixel subimage, centered on the star image; the actual location of the image within the PC is also retained and used in the fit.

The resulting best-fit focus position is shown in Figure 2 vs. the date of observation: filled squares represent the F439W data, open circles the F814W. It is immediately apparent that the focus position shows an overall downward trend, consistent with the expected shrinkage of the OTA, with an obvious jump before Day 185 (July 4): this jump corresponds to the secondary mirror movement of $5\,\mu$m away from the primary which was carried out on Day 180 (June 29) to compensate for the OTA shrinkage until that point. There is, however, considerable scatter with respect to the overall shrinking trend: this is partly due to variations in the breathing of the telescope, which can amount to $2\,\mu$m (occasionally more) in any given observation, and partly to the intrinsic uncertainty in the parameters derived by the phase retrieval program. The scatter appears much larger in the F814W data, pointing to larger errors in the phase retrieval solution; this is to be expected, because the F814W image has a larger diffraction size, and therefore is less affected by small changes in focus.

A quantitative confirmation of the validity of the phase retrieval analysis can be obtained by fitting separate linear trends to the focus position before and after Day 180. The resulting piecewise-straight lines (dashed for F439W, dotted for F814W) are drawn on Figure 2. The derived value of the jump at day 180 for F439W is $4.4\,\mu$m, in very good agreement with the expected value of $5\,\mu$m. The F814W fits are very steep, especially before the move, and are probably influenced unduly an abnormal low value at Day 121, when the focus position is estimated at $-9.8\,\mu$m; yet the measured value of the jump is $7.5\,\mu$m, also in acceptable agreement with the expected value. The fit wihtout this anomalous point (long-dashed line) indicates a jump of $6.6\,\mu$m. Note that no information on the actual movement, except for its date, is used in this measurement. The typical rms deviation between the data points and the best-fit linear regression is $1.0\,\mu$m in F439W, and $1.4\,\mu$m in F814W.

The best-fit slope of the linear regression is a measure of the rate at which the OTA is shrinking. The rate measured from the independent fits before and after the move ranges from 0.7 to 1.8 μm/month, with the large variation due to the small number of data points and their relatively large individual errors. For a better definition of the variation of focus position with time and of the overall desorption trend, it is desirable to fit for the best-fit slope *assuming* a 5 μm move on Day 180 and forcing the slope to be constant. This solution is shown in Figure 3, again with the dashed line for F439W and the dotted line for F814W. The slopes correspond to 0.75 μm/month and 0.92 μm/month respectively, and in both cases best-focus was traversed on Day 264, September 21 (this coincidence is most likely purely accidental). If the anomalous F814W point at Day 121 is excluded, the slope increases to 1.07 μm/month, and best-focus is achieved at day 268 (dot-dashed line in Fig. 3); this solution has an rms scatter of 1.6 μm, compared with 2.6 μm for the solution that uses all the F814W points.

An extrapolation of each of these solutions to January 15, 1995 (Day 380 in Figure 3) yields a predicted the focus position somewhere between $-2.8\,\mu$m and $-3.9\,\mu$m. A combined linear regression to both data sets, excluding the anomalous point, gives $-3.4\,\mu$m at Day 380; we take this as the nominal value of the linearly extrapolated focus position on that date—but see Section 3 for a cautionary remark.

2.4 Correction for breathing

One of the elements that contribute to the error in the focus measurement from a small set of images is the well-known 'breathing' of the OTA, which translates into changes in the effective focus position over a time scale of an orbit. Such changes, probably of thermal origin, can amount to $\pm 2\,\mu$m or more during a single orbit, and thus may represent a significant contribution to the scatter in the phase retrieval solutions.

An estimate of the position of the mirror within the breathing cicle can be obtained, following Hasan & Bèly (1994), by using the average $T_{\rm LS}$ of the temperatures measured by the four sensors in the light shield around the secondary mirror. They suggest that the secondary mirror moves by an amount

$$\Delta \mathrm{SM} = 0.7(T_{\rm LS} - \langle T_{\rm LS} \rangle) + K$$

where the average $\langle T_{\rm LS} \rangle$ is taken over the orbit preceding the observation, ΔSM is meanured in μm, and K is an arbitrary constant. The relevant temperature data have been obtained with the kind help of Bruce Toth and Richard Bouchard. However, it should be kept in mind that the above correction is really intended to track variations of focus along each orbit; any terms on a longer time scale would be absorbed into a variation of the constant K, for which I am not aware of a specific prescription.

Individual focus measurements have been corrected by an amount proportional to the measured temperature difference $\Delta T = T_{\rm LS} - \langle T_{\rm LS} \rangle$; in order to evaluate the improvement obtained with the breathing correction, a range of values has been used for the proportionality constant, besides the value of 0.7 μm/K given by Hasan & Bèly (1994). For the

F814W data points, a slight improvement is indeed obtained with the breathing correction, in the sense that the linear regression has a significantly smaller scatter—1.0 μm rms, compared to 1.6 μm for the uncorrected solution (in both cases, the anomalous point at day 121 is removed). This solution corresponds to a slightly lower shrinkage rate of 1.01 μm/month, with best focus at day 249; the predicted focus position on January 15, 1995 is $-4.37\,\mu$m.

On the other hand, for the F439W data the breathing correction yields a significantly worse linear regression than for the uncorrected data. Of course, this should not be interpreted as evidence against the breathing correction itself, which is meant to track short-term focus changes only. Also, individual data points have significant noise, and focus position need not be exactly linear with time. However, it does put in question the validity of the improvement achieved for F814W with exactly the same method. Therefore I prefer to forego the breathing correction altogether in the current analysis, and use only the uncorrected results in my conclusions. Note that, in any event, the change in expected focus position at the target date is only about 0.4 μm, comparable with the intrinsic accuracy in the prediction.

3. Discussion and recommendations

The available data are consistent with a continuing shrinkage of the OTA at a rate of $0.90 \pm 0.15\,\mu$m/month, with zero focus being traversed around day 265. This rate of shrinkage is consistent with that measured by the FOC (Jedrzejewski, private communication), but is significantly higher than that predicted by the extrapolation of the double-exponential trend observed through mid-1993 (Hasan et al 1993; Hasan, private communication).

Linear extrapolation to mid-January 1995 indicates a likely focus position of $-3.4 \pm 0.6\,\mu$m at that date. A movement of 5.5 μm performed on that date would ensure that the average focus position remains within 2 μm of the optimal value for the following 4–5 months, barring unforeseeable changes in the desorption trend or other abnormal events. If the tolerance is to be increased to 3 μm, then a movement of 6.5 μm would ensure that the focus position remains within the desired tolerance for approximately 7 months. If the current desorption trend continues, a schedule of regular moves twice a year is compatible with a tolerance of 2.5 μm.

As a cautionary note, it must be noted that these estimates are based on a linear regression applied to the measured focus position. However, the most recent measurements (November and December 1994) do not indicate as large an offset as the linear regressions predict. This may be due to random errors, or it may be an indication that the rate at which the OTA shrinks has changed somewhat in the last few months. Therefore, it may be appropriate to reduce the amount of the movement of the secondary mirror to a more conservative 5.0 μm.

Because of this uncertainty, and of the possible change in the rate of shrinkage of the telescope in the future, it is recommended that systematic focus monitoring be continued on at least a monthly basis. Fortunately, the tools exist to carry out this monitoring using other calibration data, and thus without any increase in demands on telescope time. Possible improvements on the monitoring program with respect to what is described here include extending the study to more wavelengths, hopefully with the effect of reducing the uncertainties, and new attempts to include the effects of telescope breathing.

Acknowledgements

This work was carried out with substantial help from all those who have worked previously on focus monitoring and image quality of HST, and who have shared liberally their expertise and the numerous tools they had developed for this task. I am especially grateful to John Krist and Chris Burrows, who have put their phase retrieval program at my disposal and have taught me how to use it; to Hashima Hasan, who has helped me understand many of the intricacies of focus monitoring; to Robert Jedrzejevski, Antonella Nota, and George Hartig, for several discussions on the effect of focus changes on other instruments; to Pierre Bèly, Bruce Toth and Richard Bouchard, for help with understanding telescope breathing and for retrieving the relevant data; and to Sylvia Baggett and Christine Ritchie, for their help in retrieving and studying the relevant WFPC2 images.

References

Bèly, P. Y., & Hasan, H., 1993, Instrument Science Report, OTA–11

Hasan, H., 1993, Instrument Science Report, OTA–13

Hasan, H., 1994, Instrument Science Report, OTA–17

Hasan, H., & Bèly, P. Y., 1994, in *The restoration of HST Images and Spectra II*, R. J. Hanisch and R. L. White, eds (Baltimore: STScI), p. 157

Hasan, H., Burrows, C. J., & Schroeder, D. J., 1993, PASP 105, 1184 and Instrument Science Report, OTA–12

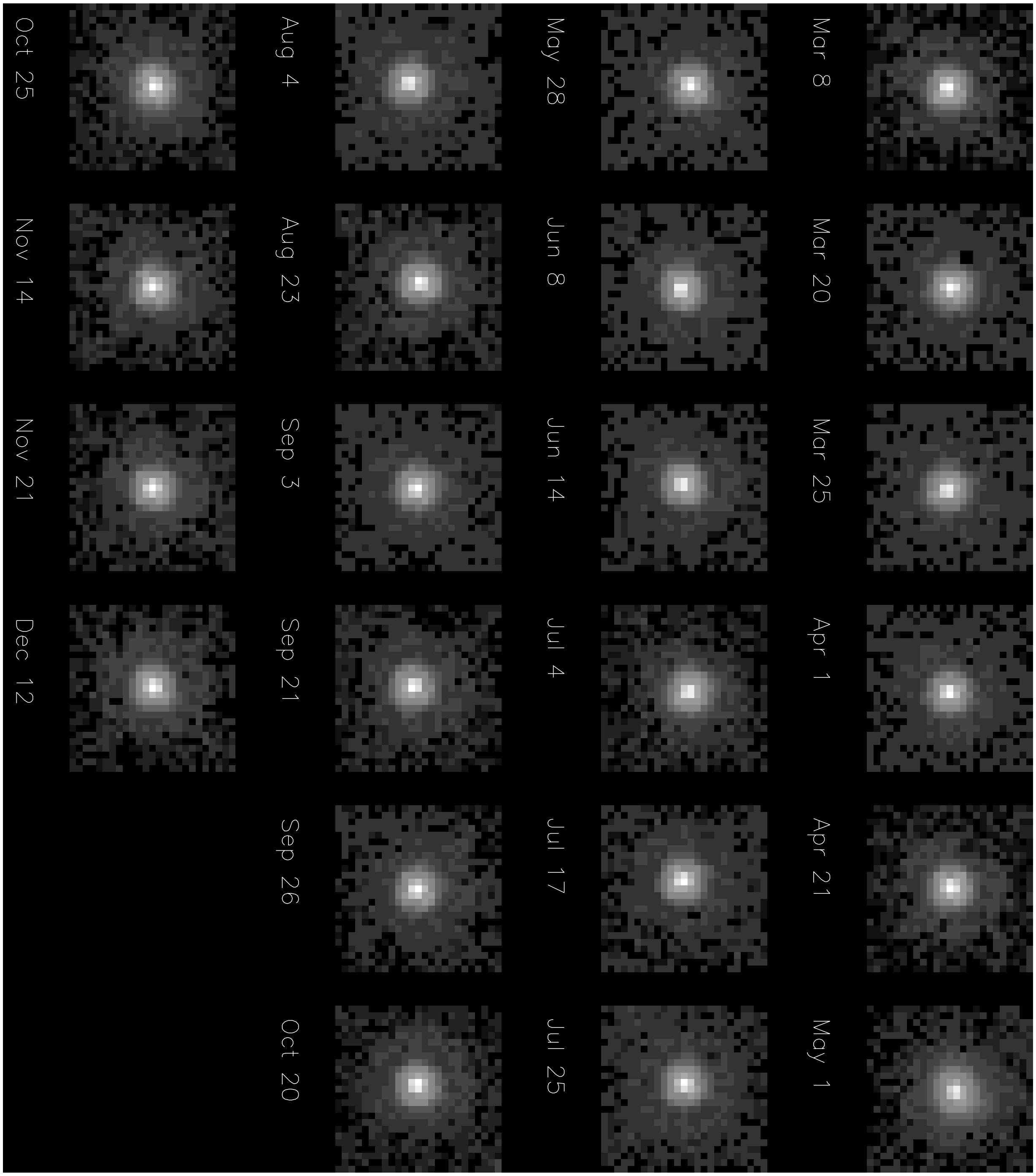
Krist, J., & Burrows, C. J., 1995, Applied Optics, in press

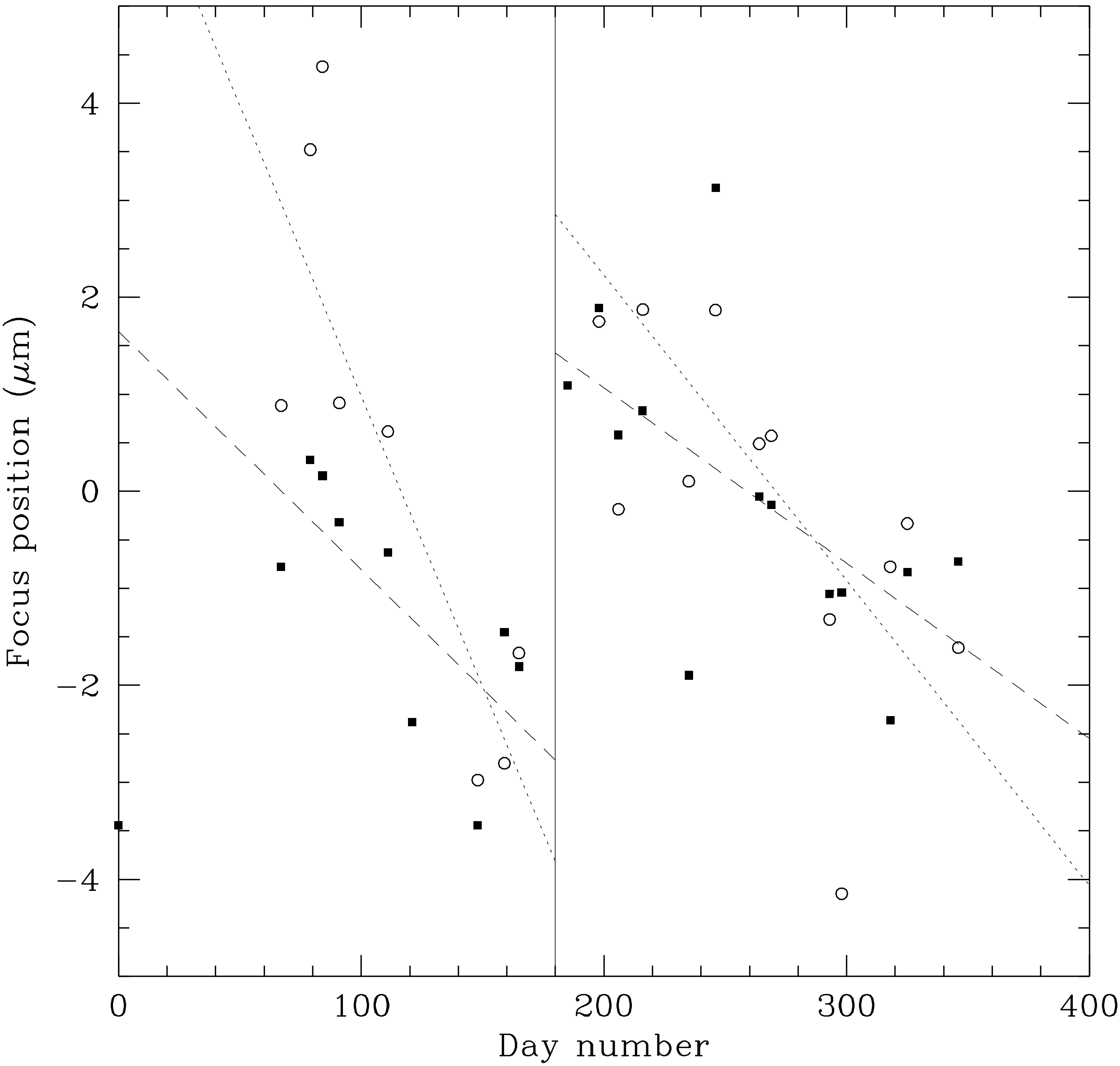
Figure Captions

Figure 1: Grayscale representation of the PC image of the spectrophotometric standard star GRW +70D5824, taken in the filter F814W at various dates throughout 1994.

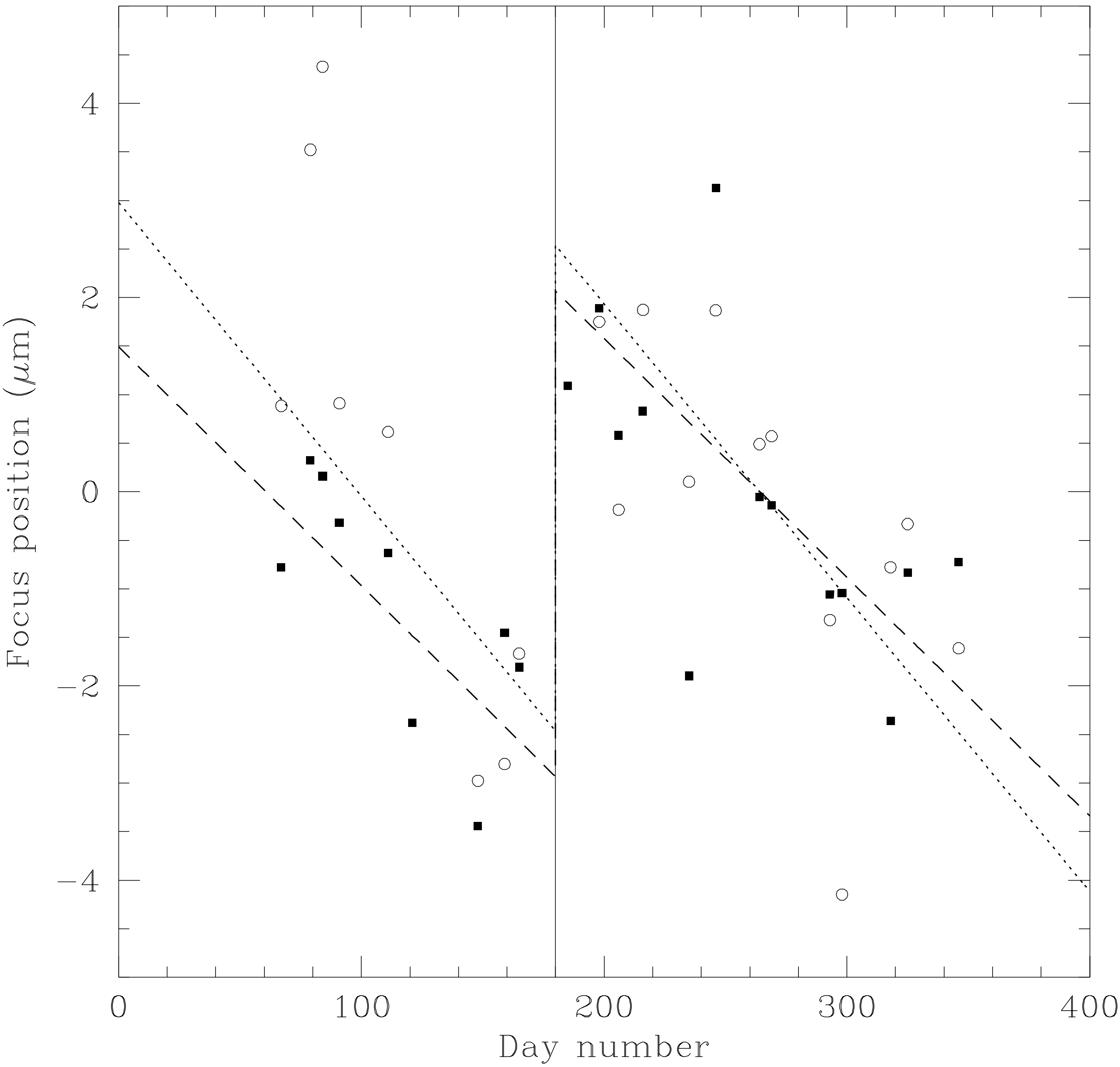
Figure 2: Focus position measured using the phase retrieval method on PC images of GRW +70D5824, both in F439W (filled squares) and in F814W (open circles). The vertical line represents Day 180 (June 29), when the secondary mirror was moved by $5\,\mu$m. The other lines are best-fit linear relations for the focus position in F439W (dashed) and F814W (dotted); the fits have been carried out separately for pre- and post-movement data. The very low point for F814W at day 121 appears anomalous; the long-dashed line represents the best-fit for F814W pre-movement data if that point is excluded.

Figure 3: Same as Figure 2, except that the straight lines represent now a single fit to pre- and post-movement data, with the change at Day 180 constrained to be $5\,\mu$m. The slopes of the two lines represent a rate of change in the focus position of -0.75 and $-0.92\,\mu$m/month for F439W and F814W, respectively. The long-dashed line is the F814W fit with the anomalous data point at day 121 excluded, with a slope of $-1.07\,\mu$m/month.


Mar 8
Mar 20
Mar 25
Apr 1
Apr 21
May 1
May 28
Jun 8
Jun 14
Jul 4
Jul 17
Jul 25
Aug 4
Aug 23
Sep 3
Sep 21
Sep 26
Oct 20
Oct 25
Nov 14
Nov 21
Dec 12



Focus position (μm)
4
2
0
−2
−4
0
100
200
300
400
Day number



Focus position (μm)
4
2
0
−2
−4
0
100
200
300
400
Day number