ASTRONOMY & ASTROPHYSICS
SUPPLEMENT SERIES

Astron. Astrophys. Suppl. Ser. **111,** 183-193 (1995)

Model-mismatch errors in least-squares 1-D centering

M. David and W. Verschueren

Astrophysics Research Group, University of Antwerp (RUCA), Groenenborgerlaan 171, B-2020 Antwerpen, Belgium
email: astro@ruca.ua.ac.be

Received July 9; accepted November 26, 1994

Abstract. — The nature of "pixelation errors" in 1-D centering by means of least-squares fits is analyzed in detail. For the case of intrinsically symmetric data we propose an improved position estimator, based on a combination of the fit results obtained with an even and an odd number of fitpoints. Numerical tests on synthetic data confirm that the combined estimator reduces the pixelation errors – which might be called more aptly *model mismatch errors* – by roughly an order of magnitude. The influence of weak asymmetries in a feature is discussed and we indicate how, in practice, the importance of model mismatch errors (and the relevance of a correction thereof) with respect to noise-induced errors, can be assessed.

Key words: methods: data analysis — methods: numerical — techniques: spectroscopic

1. Introduction

The problem of accurately determining the position of some feature in a digitized image arises frequently in many branches of observational astronomy, e.g. CCD astrometry, radial velocity measurements, wavelength calibration in spectroscopy, pointing and tracking applications, etc. All major astronomical software packages offer several algorithms (with varying degree of sophistication) to handle this task, leaving the decision on their suitability for a specific purpose to the users. Faster algorithms are often obtained by making simplifying assumptions about the data (e.g. that the calibration line profile is a parabola, or that the background around a feature is negligible) which may cause the accuracy of the result to be dominated by induced systematic errors rather than by random noise.

In the case of pointing and tracking applications using centroiding, the different sources of systematic errors have been discussed in some detail by Dick et al. (1989). The detection of systematic errors in wavelength calibrations has been discussed by Hensberge & Verschueren (1989) who demonstrated in particular that the so-called 'gravity' method (a pseudo-centroid method using only the differences of the highest two pixels with respect to the third highest pixel in a feature) is quite inadequate for the purpose at hand. The accuracy of radial velocity measurements by means of cross correlation or from individual line positions has been discussed by several authors, e.g. Tonry & Davis (1979), Brown (1989), Verschueren (1991).

On the other hand, a recent comparative study of several spectral line fitting techniques, on calibration data, by Gulliver & Hill (1990) shows that the understanding of the mechanisms leading to systematic errors and of the proper diagnostics to detect them, is not yet common knowledge in the astronomical community. In the present paper we therefore present an analysis of the possible systematic error in positions obtained by fitting a model for the feature to the observed data.

There are several possible choices (e.g. L_1, L_2, ...) for the metric to be used in calculating the "distance" between the model and the feature; in particular, according to the quality of the data, it may be appropriate to weight the contributions from different data points. Among all these possibilities we shall limit our discussion to the case of the L_2 metric with equal weights for all data points, for the following two reasons. Firstly, we are concerned with the problem of maximizing sub-pixel precision in a position measurement, which requires, above all, well sampled high S/N data that have been subjected to a thorough cosmetic treatment eliminating the effects of CCD-blemishes, cosmic rays, etc., so that there will seldom be any need for an alternative distance measure e.g. to make the results more robust against the influence of outlying data points. Secondly it will be seen that our conclusions are essentially independent of the distance measure used, though of course the actual form of the improved position estimate proposed in Sect. 3 will have to be rederived if another metric is assumed.

It does not matter which numerical algorithm is used to minimize the chosen distance measure, provided that it does yield the correct extremum (up to numerical noise). However, the user should be warned that the discussion in

this paper does *not* apply if a technique is used which does not involve an actual minimization for each set of data but rather an interpolation between the results obtained for closely similar data sets (see e.g. Bishop & Roach 1992): if the form of the interpolating function is not prescribed by the physics of the problem at hand, then the accuracy of the result for a given data set will depend also on the form of the interpolating function and on the accuracy of the results between which the interpolation is done; this may lead to systematic errors of a kind which is not considered in this paper.

The problem we are dealing with can be formulated as follows. Let $\bar{f}(x_n)$, $n \in \{-q, \ldots, p\}$ be a set of *data values* representing the observed feature and let $g(x;\boldsymbol{\pi})$ be an analytical function, henceforth referred to as the *model*, depending on an array $\boldsymbol{\pi}$ of parameters. The model is chosen so that, with appropriate values of the parameters, its values in the data points x_n may closely approach the data values. The optimal parameter values are found by requiring that they minimize the quantity $X^2(\boldsymbol{\pi})$, given by

$$X^2(\boldsymbol{\pi}) = \sum_{n=-q}^{p} [g(x_n;\boldsymbol{\pi}) - \bar{f}(x_n)]^2. \qquad (1)$$

Obviously, if the data values can be regarded as having been *sampled* from a function with known analytical form $f(x;\boldsymbol{\kappa})$, evaluated with unknown values $\boldsymbol{\kappa}_0$ of its parameters, the minimization must yield the *exact* parameter values if the *correct* model is used, i.e. $g = f$. However, in practice this is seldom possible and the model will not quite match the data. This "mismatch" is the source of the systematic errors we aim to investigate.

Systematic errors in a position obtained by least-squares fitting are often referred to as *discretization* errors or *pixelation* errors; yet they are not inherent to the discrete nature of the data (except perhaps in the case of extreme undersampling) since an exact solution is obtained if the correct model is used. Therefore we think that the term *model-mismatch errors* is more appropriate in this case.

In the kind of applications we have in mind, there is usually at least a *minimal* mismatch due to the fact that the data result from *binning* and that, even if one knows the exact analytical form of the observed intensity distribution, the integration process of binning does not necessarily result in an expression in closed form to be used as the model, especially if the sensitivity of the detector is not uniform over a pixel (see e.g. Rutten et al. 1992; Jorden et al. 1993). As will be seen further on, larger errors may occur if also in other respects the model differs from the "original" $\bar{f}(x)$, and the worst cases are likely to arise if $\bar{f}(x)$ has no symmetry while (usually for want of precise information) one still uses a symmetrical function as a model.

In general it could be argued that the problem of finding the "position" can be formulated properly only if either the correct model is known or if the features in question are *intrinsically symmetric*, i.e. they result from the (possibly asymmetric) binning of a *symmetric* intensity distribution. The first case is obvious; in the second case the required position is unambiguously and naturally defined as the centre of symmetry of the observed intensity distribution, and this definition applies consistently to any model, provided only that it is symmetric as well. On the other hand, if a feature is intrinsically asymmetric and one cannot fit it with the correct model, one needs to introduce a parameter which characterizes its position; such a parameter can be defined in several ways which lead to different results (even without discretisation). Although one may be forced to adopt some such definition as "adequate" in a particular case, this is seldom quite satisfactory.

So in any centering applications one should attempt to avoid or discard intrinsically asymmetric features, or at least handle them with special care. When, for instance, one is analysing a comparison spectrum in which some important line is blended, one is not interested in the position of a blend as such, but in the position of its components, so the structure of the blend must be analysed in detail anyway. Besides, the position of the blend as a whole, by any definition, may depend critically on the relative strengths of the components as well as on the PSF (not to mention discretisation) so it is necessary to adopt a specific model to handle the blend. As another example, when cross-correlating two spectra, any asymmetry (apart from noise) in the correlation peak signifies some kind of mismatch between the spectra which should be investigated and minimized prior to the cross correlation; asymmetry in a cross-correlation peak may lead to serious systematic errors in the radial velocity (Verschueren 1991).

In practice one may be faced also with asymmetries of a more basic origin, such as optical aberrations or *asymmetrical* intra-pixel sensitivity variations (see e.g. Jorden et al. 1993), but in any case the effort should be to minimize the asymmetry, rather than to concentrate on improving the fitting strategy.

In Sect. 2 we make a few suggestions on the detection of systematic errors and on the way they could be corrected a posteriori (if a more correct re-analysis of the data is not feasible for some reason). In Sect. 3 we show that, for intrinsically symmetric features, model mismatch leads to an error in the position if the bin positions are not symmetrical with respect to the symmetry centre of the original feature; we estimate the resulting error and show how it can be reduced considerably. In Sect. 4 we discuss some numerical tests to verify the previous results.

A general treatment of data with intrinsic asymmetry cannot be given, as each case must be looked at individually and handled in view of both the nature of the asymmetry and the finality of the required results. For instance the way to handle line-blending in spectral classification

using low-resolution spectra was discussed by Hensberge et al. (1994), and in a forthcoming paper, De Cuyper & Hensberge (1994) will discuss the treatment of blends in wavelength calibration. However, in Sect. 5 we do consider the case of *weak* intrinsic asymmetry (i.e. where the asymmetry can be treated as a perturbation on the symmetrical case) as well as the influence of mere binning mismatch, in the case of arbitrary asymmetry.

Throughout this paper, except for a few brief remarks in Sect. 6, we shall assume that the data are noise-free. The influence of noise on the accuracy of centering by means of a least-squares fit, has been dealt with elsewhere (see e.g. Brown 1989; Verschueren 1991).

We shall also assume, without loss of generality, that the position variable x has been scaled so that bins have length 1 and bin positions are integer numbers, the n^{th} bin extending from $x = n - \frac{1}{2}$ to $x = n + \frac{1}{2}$ etc. Position- and error estimates will thus be expressed in pixel units.

2. Empirical detection of systematic errors in centering

A systematic effect cannot be traced empirically from a single measurement, or from a set of measurements on which the effect is identical: only a variation of the effect may betray its existence. Such a variation will be related to a change in the value of at least one parameter that characterizes the systematic effect. Systematic errors, whether on a sub-pixel scale or not, are thus most easily detected from exposures containing many features which differ only in a limited number of such parameters (ideally just one), or from series of exposures whereby minor differences between corresponding features can be modelled as being continuous in one such changing parameter.

As far as all features concerned have intrinsically the same intensity distribution, the only parameter controlling model-mismatch errors in centering is their actual sub-pixel location (also referred to as *pixel fraction*) as will be seen in the following sections; therefore, exposures have to be obtained in which the observed features exhibit a wide range in sub-pixel location. If the mismatch depends on the location of a feature on the detector (e.g. as a consequence of a variation of the PSF due to certain optical aberrations) then several such sets of data are required from different parts of the detector, so that one or two more parameters will be needed to characterize the resulting centering error. Model mismatch may also vary with the wavelength (e.g. due to an imperfect elimination of interference fringes on a CCD), with telescope position (e.g. through a combination of varying flexure and vignetting) etc., involving yet other parameters (besides the sub-pixel location) controlling possible systematic errors.

Often, dedicated test exposures will have to be taken to obtain all necessary information. For instance, in direct imaging studies, it will be useful to make different exposures of a test stellar field, with random or controlled sub-pixel shifts of the image on the CCD. On the other hand, certain kinds of observations may by themselves yield a sufficiently large number of measurements to detect systematic errors. As an example, one may think of all the different line positions (at sub-pixel level) on a single wavelength calibration exposure. Furthermore, the sub-pixel position of all lines may change collectively between successive wavelength calibration exposures (due to time variations of the refractive index of the air or due to different flexure effects in different telescope positions) thus vastly increasing the amount of data available for tests; moreover the information from these successive exposures can be used to detect model-mismatch on individual lines, e.g. due to an unresolved blend (Hensberge & Verschueren 1990). And to study centering errors which may occur in fitting the cross-order profile of echelle spectra (Verschueren & Hensberge 1990), a continuous range of mid-order sub-pixel positions is available because of the inclination of the echelle orders on the CCD.

In practice, then, systematic centering errors can be detected and studied by examining any quantity, deduced from the data, which is expected to present a random behaviour as a function of sub-pixel location.

Firstly, as an easy starting point, one should check whether the distribution of the *measured* sub-pixel locations themselves, obtained for many independent objects (be it stars, spectral lines, spectral orders, etc.) is random, since model mismatch will make it deviate strongly from randomness (Hensberge & Verschueren 1989).

Secondly, in the case that the respective positions of a set of objects (located either on a single frame or on successive frames) are expected to obey some physical or mathematical relation (as in the case of a dispersion relation, group proper motion, orbits), then residuals with respect to this relation can be checked against sub-pixel location (see e.g. Hensberge & Verschueren 1989, Verschueren 1991). As an example we note that Gulliver & Hill (1990), considering essentially only the rms of the residuals for several centering methods applied to wavelength calibration lamp spectra, concluded that there were no systematic differences between the several centering methods they tested; however, in their Table 1 they also quote the actual values of the residuals with two fit models (a parabola and a Gauss function), and if we plot these against sub-pixel location (Fig. 1) then the residuals for the parabola very clearly exhibit non-random behaviour, while the Gaussian data exhibit a much weaker dependence on sub-pixel location.

Thirdly, consider the case that there is no (or no precisely known) relation between the respective positions of the objects on a given frame and suppose furthermore that there are not enough frames available (or that the data do not cover the required range in values of the parameter which controls a systematic effect, as could happen with a very stable instrumental configuration) to detect

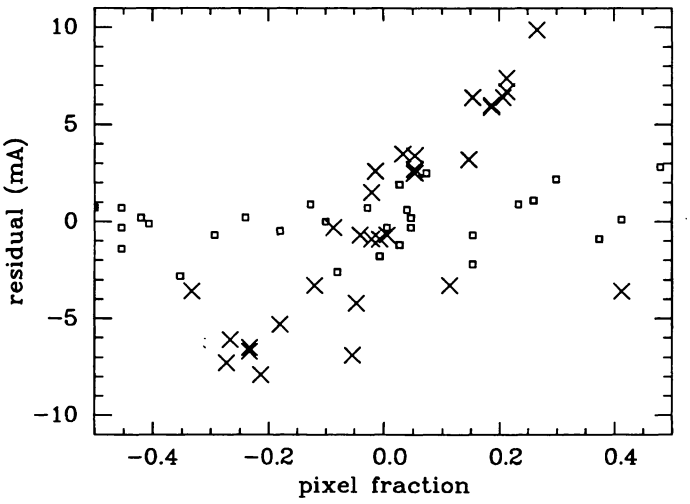

Fig. 1. The residuals (observed - calculated wavelengths) as a function of pixel fraction in a test of several centering algorithms by Gulliver & Hill (1990) using different models to fit the spectral line profiles (crosses: parabola fitted to 5 pixels, dots: Gauss function fitted to 7 pixels); the unit on the ordinate, 1 mÅ, corresponds to 0.028 pixels

non-random behaviour from the distribution of the positions. Then, one can still rely on a detailed study of the *relative* position of objects with respect to each other: e.g. if there is an effect related to the sub-pixel location, then distances which do not equal an integer number of pixels, will be different on different frames and the systematic effect will show up as a correlation between the difference in the respective distances measured on two frames, and the fractional part of these distances.

Once a systematic effect has been detected, and if the relation between the error and the controlling parameter(s) can be modelled, then this model can be used to largely remove the errors from the data without restarting the whole measuring process. Such improved positions are undoubtedly most directly needed when the positions of individual objects must serve as input for fitting some ad-hoc function wherein the parameters do not have a clear physical meaning. In that case, they can serve the purpose to derive a more suitable and more robust (i.e. with less degrees of freedom) fit function. These principles will be applied and discussed by Hensberge et al. (1995) in their analysis of wavelength calibration lamp spectra from the echelle spectrographs CASPEC and ECHELEC at ESO and long-slit exposures taken with CARELEC at OHP.

3. Model mismatch with intrinsically symmetric data

Considering a given feature, we define the origin $x = 0$ of the co-ordinate system as the centre of the highest pixel of the feature[1]; then, with the scaling we have adopted, it follows that the actual position, $x = d$, of the feature is limited to $-\frac{1}{2} \leq d \leq \frac{1}{2}$ (which implies that henceforth we may omit the distinction between the sub-pixel position and the actual position); furthermore, in that case, the sign of d is the sign of the position of the next highest pixel. It should be stressed, however, that in principle this information on the position can be inferred *only* if the data are intrinsically symmetric and noise-free; if, in practice, a measured position does not have these properties, then the measurement is suspect and the origin of the discrepancy should be investigated; in some cases a simple correction will be possible (e.g. if $d \cong \pm 0.5$, then even weak noise fluctuations may lead to the wrong choice for the origin but the (suspect) result of a first measurement may well indicate which other pixel centre should be used).

In most of the following discussion we may, for the sake of convenience, restrict the value of d to the interval $[0, \frac{1}{2}]$, if we define the positive sense of x to be from the highest to the next highest pixel. However, expressions to be applied in practice will be cast into a form which is valid for the whole interval $[-\frac{1}{2}, \frac{1}{2}]$, since the positive sense of the position co-ordinate will usually be fixed by other constraints and the necessity of inverting it for some of the features would imply an undesirable complication in the data analysis software.

Let $f(x)$ and $g(x)$ be symmetrical with respect to the origin, i.e.

$$f(-x) = f(x)\ , \quad f'(-x) = -f'(x)\ , \quad \text{etc..} \tag{2}$$

and let

$$\bar{f}(x) = f(x - d). \tag{3}$$

be the function[2], centred at $x = d$, from which the data points are sampled in the pixel positions $x_n = n$. Other parameters than the one characterizing the position do not concern us here: they are implied but need not be mentioned explicitly.

We now proceed to determine the position d by fitting the function g to the data points, i. e. by minimizing

$$X^2(\delta) = \sum_{n=-q}^{p} [g(n - \delta) - \bar{f}(n)]^2. \tag{4}$$

If the model cannot exactly match the data, the solution δ of

$$\sum_{n=-q}^{p} [g(n - \delta) - f(n - d)]g'(n - \delta) = 0. \tag{5}$$

will be only an approximation to d. The difference $\epsilon = \delta - d$ is the *model mismatch error.*

[1] if there happen to be two such pixels, either of these is suitable for a start.

[2] Generally f will be the result of a binning process, but that is immaterial for the discussion in this section.

If the data are intrinsically symmetric there are two exceptional cases in which the error may vanish even though the model cannot match the data exactly, notably when $d = 0$ or $d = \frac{1}{2}$. In fact, in that case it follows from (2) that $\delta = d$ provided the pixels used for the fit are chosen symmetrically with respect to this position (i. e. $p = q$ or $p = q + 1$ respectively). This suggests that for any other position a better approximation will be obtained by choosing

$$p = q \quad \text{if} \quad 0 \leq d < 0.25. \tag{6}$$

$$p = q + 1 \quad \text{if} \quad 0.25 < d \leq 0.5. \tag{7}$$

This may seem a rather trivial observation, but we find that in practice most people overlook it (e.g. Gulliver & Hill (1990) in their tests always used an odd number of fitpoints). Numerical tests by Verschueren (1991) indicated that the proper choice of the pixels used in a fit could reduce the average magnitude of the error on individual line positions by a factor $\frac{1}{2}$. This may be quite relevant if one is interested in fitting individual features (e.g. a cross-correlation peak). On the other hand, if the final result depends on many measured features, as in a wavelength calibration, it may be unimportant for the accuracy but it will undoubtedly decrease the rms of the residuals.

However, as we shall presently show, a careful analysis of the model mismatch error leads to a way to reduce it much more than by a proper choice of the number of pixels.

It follows from the symmetry of f and g that their k^{th} derivatives satisfy respectively

$$f^{(k)}(x) = (-1)^k f^{(k)}(-x) \; ; \quad g^{(k)}(x) = (-1)^k g^{(k)}(-x). \tag{8}$$

If we now consider a Taylor expansion of the l.h.s. of Eq. (5) in powers of d and δ (which have both the same order of magnitude), it is easily seen that, due to the properties (8), in the case (6) all even powers of d and δ will disappear. In the case (7) the same occurs provided one first substitutes

$$d = \frac{1}{2} - \hat{d} \quad \text{and} \quad \delta = \frac{1}{2} - \hat{\delta}. \tag{9}$$

and considers the expansion in powers of $\hat{d}$ and $\hat{\delta}$. Therefore we can expect to obtain a good approximation to the solution of Eq. (5) by considering only first order terms in d and δ (or in $\hat{d}$ and $\hat{\delta}$ in the case (7)). After some algebra one finds respectively

$$\delta^{(1)} = D_{\text{o}} d \quad , \quad \hat{\delta}^{(1)} = D_{\text{e}} \hat{d}. \tag{10}$$

where the superscript indicates that it is only a first-order approximation, while the subscript identifies the result obtained with an odd or an even number of data points respectively; with the notation n_{h} defined by

$$n_{\text{h}} = n - \frac{1}{2}. \tag{11}$$

the coefficients in Eq.(10) are given by

$$D_{\text{o}} = \frac{\sum_{n=-p}^{p} f'(n) g'(n)}{\sum_{n=-p}^{p} [g'^2(n) + (g(n) - f(n)) g''(n)]}. \tag{12}$$

$$D_{\text{e}} = \frac{\sum_{n=-p}^{p+1} f'(n_{\text{h}}) g'(n_{\text{h}})}{\sum_{n=-p}^{p+1} [g'^2(n_{\text{h}}) + (g(n_{\text{h}}) - f(n_{\text{h}})) g''(n_{\text{h}})]}. \tag{13}$$

In practice, because of the symmetry the summations may of course be reduced to half the number of terms, but for later convenience we have omitted this simplification in the equations above.

Notice that the coefficients D_{o} and D_{e} equal one if $g = f$, i.e. if the correct model is used. In general however the coefficients in (10) will differ from one and this difference could be considered as an estimate for the relative model-mismatch error of the fit. Furthermore it is easily seen that D_{o} and D_{e} must be very close to each other in magnitude; in fact, if Δx is the pixel size in physical units then in the limit for $\Delta x \to 0$, $p \to \infty$ with $p\Delta x = \text{const}$ one has $D_{\text{o}} = D_{\text{e}}$.

Suppose now that, with any given set of data, we do the least-squares fits corresponding to both of the possibilities (6) and (7); noting δ_{o}, δ_{e} for the respective results, it follows from Eqs. (9) and (10) that a first-order approximation for these is given by

$$\delta_{\text{o}}^{(1)} = D_{\text{o}} d \quad , \quad \delta_{\text{e}}^{(1)} = \frac{1}{2} - D_{\text{e}}(\frac{1}{2} - d). \tag{14}$$

and that a first order approximation for the respective errors $\epsilon_{o,e} = \delta_{o,e} - d$ is given by

$$\epsilon_{\text{o}}^{(1)} = (D_{\text{o}} - 1) d. \tag{15}$$

$$\epsilon_{\text{e}}^{(1)} = (D_{\text{e}} - 1)(d - \frac{1}{2}). \tag{16}$$

At this point it is appropriate to consider also the case of negative pixel fractions: with an odd number of fitpoints (centred around 0) $\delta_{\text{o}}^{(1)}$ and $\epsilon_{\text{o}}^{(1)}$ are given by exactly the same expressions as before, but an even number of fitpoints must now be centred around $x = -0.5$ (i.e. $q = p + 1$ in Eq. (7)) so that the expressions for $\delta_{\text{e}}^{(1)}$ and $\epsilon_{\text{e}}^{(1)}$ become, resp.

$$\delta_{\text{e}}^{(1)} = -\frac{1}{2} + D_{\text{e}}(\frac{1}{2} + d) \qquad (-0.5 \leq d < -0.25). \tag{17}$$

$$\epsilon_{\text{e}}^{(1)} = (D_{\text{e}} - 1)(d + \frac{1}{2}). \tag{18}$$

while the summation index in Eq. (13) must take the values $-(p+1), \ldots, p$.

Incidentally, we may now compare the behaviour of the model mismatch error in the case of an odd number of fitpoints, as predicted by the first-order approximation (15), to the data in Fig. 1: in fact the Gaussian residuals exhibit a very weak linear trend (implying that $|D_{\text{o}} - 1| << 1$, as will be confirmed in Sect. 4) while the parabolic residuals

follow a relatively strong slope for ($-0.25 \lesssim d \lesssim 0.25$); actually the magnitude of the parabolic residuals decreases again outside this range, but this particular behaviour will be explained in Sect. 4.

From Eqs. (15), (16) and (18) it follows that, up to first order in d, the errors will have opposite signs and roughly the same slope over the interval $[-0.5, 0.5]$, $\epsilon_e^{(1)}$ having a discontinuity at $d = 0$. This behaviour, already noted by Verschueren (1991) in the results of numerical experiments, suggests that a combination of the two measurements might lead to an estimate which is considerably more accurate than either of them separately, and moreover uniformly so over the whole interval.

In order to derive this combination we notice in particular that, if we had $D_o = D_e$ and $\delta_{o,e}^{(1)} = \delta_{o,e}$ then we could simply solve the equations (14) for the pixel fraction d. The form of this solution suggests the following definition for a combined estimator, which we shall denote by $\bar{\delta}$:

$$\bar{\delta} = \frac{\delta_o}{1 \pm 2(\delta_o - \delta_e)} = \frac{\delta_o}{1 + 2(|\delta_o| - |\delta_e|)}. \qquad (19)$$

where the $-$ sign applies in the case $-0.5 \leq d < 0$ and where the second form is based on the (safe) assumption that both position measurements δ_o and δ_e will have at least the correct sign. In the special cases $d = 0$ and $d = \pm 0.5$, where we know the fit must yield resp. $\delta_o = 0$ and $\delta_e = \pm 0.5$, this expression also yields the correct values. Furthermore, substituting (14) in (19) we obtain a first-order approximation for the error $\bar{\epsilon} = \bar{\delta} - d$ in terms of the coefficients D_o and D_e:

$$\bar{\epsilon}^{(1)} = \frac{(D_o - D_e)(1 - 2|d|)d}{D_e + 2|d|(D_o - D_e)}. \qquad (20)$$

So if one uses (19) as an estimate for the position, the magnitude of the maximal error (occurring at $d \cong 0.25$) will be proportional to the difference between D_o and D_e instead of the difference between these coefficients and one; therefore we expect that in general the combined estimator (19) will be considerably more accurate than either of the individual results δ_o, δ_e.

The application in practice of the improved position estimate (19) is straightforward in the case of cross-correlation peaks where, for other reasons, at some stage each peak has to be considered individually anyway, and a decision on the number of pixels to be used in a fit will be made in the course of an interactive session with some software package like e.g. CORSPEC (Verschueren 1991). In a wavelength calibration it may seem less so because in each frame a relatively large number of lines has to be located which requires some automatic procedure. Such procedures use either a fixed number of pixels to be specified by the user (based on the justifiable assumption that the PSF does not vary very much over the spectral image) or a variable one determined automatically on the basis of the levels of background and noise. Still, in any procedure it will require only a minor modification to have it repeat its normal operation with the usual number of pixels plus or minus one, and afterwards for each line to combine the results according to (19).

4. Numerical tests

In order to illustrate quantitatively the results from the previous section, we performed a number of tests, fitting several models to different sets of synthetic data, using a variety of choices for the set of fitpoints. We looked at the errors made by individual fits using an odd or an even number of fitpoints, and verified the improvement obtained by using the combined position estimate (19). The "data" we used are noise-free, binned (i.e. integrated over 1 pixel) Gaussians with a full width at half-maximum (FWHM) ranging from 2 to 32 pixels and a central position d in the range [0–0.5] pixel. The models are, resp., a 2^{nd} degree polynomial, a symmetric 4^{th} degree polynomial ($y = a_1 + a_3(x - a_2)^2 + a_4(x - a_2)^4$) and a Gaussian (+ constant to fit the background). The number of fitpoints used ranges from 5 (3 for the parabola) to a number equal to 2 (3 for FWHM=2) times the FWHM of the data.

We can be very brief about the Gaussian fit, where the only model mismatch is due to the binning of the data: fitting errors ϵ_o and ϵ_e turn out to be less than 10^{-5} pixel in absolute value, while typical errors $\bar{\epsilon}$ on the combined estimate are one order of magnitude smaller still. When binning of the data is the only model mismatch, it is clearly negligible.

It is appropriate here to remark that, when fitting a Gaussian (or any other simple peak function which vanishes asymptotically), convergence problems may arise if the model mismatch is not extremely small and if the fitpoints sample only the data near their centre; in order to avoid such problems it is best to *impose* the correct asymptotic behaviour to the model by fixing the background parameter in the fit.

The behaviour of the positional errors ϵ_o and ϵ_e as a function of the position d and of the number of fitpoints is shown in Fig. 2 for the case of a 2^{nd} degree polynomial fit. As predicted by Eqs. (15, 16), ϵ_o and ϵ_e are opposite in sign and they generally increase in magnitude when d, resp. $0.5 - d$ increases. It is interesting to note that the curves corresponding to 3, 4 or 5 fitpoints do not exhibit this monotonic behaviour over the whole of the interval; reconsidering Fig. 1 one sees that there as well the magnitude of the parabolic residuals decreases again for pixel fractions approaching ± 0.5; the same occurs with the symmetric 4^{th} degree polynomial and 4, 5 or 6 fitpoints. In fact with the number of fitpoints equal to the minimum number required for the model (3 for the parabola, 4 for the symmetric 4^{th} degree polynomial), the fit even gives the *correct* position for $d = 0.5$ as well as for $d = 0$. The reason for this behaviour is the fact that the model is able to pass *exactly* through symmetrically positioned data points

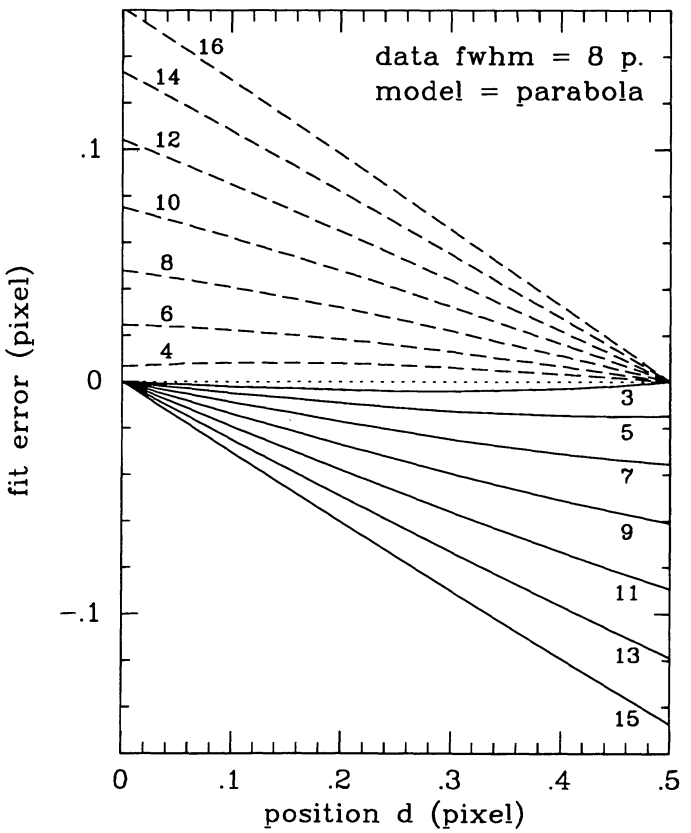

Fig. 2. Fit errors ϵ_o and ϵ_e as a function of centre position d for odd and even numbers of fitpoints as indicated; the data were binned Gaussians of FWHM = 8 px, centered at d, and the fit model used was a 2^{nd} degree polynomial

(2 or 3 resp.) and therefore necessarily will have the correct centre of symmetry. This effect also influences fits for a number of fitpoints *close to* the minimum number, and it is responsible for the bending of the curves towards a lower error in the region of d=0 for even fitting and d=0.5 for odd fitting.

Figure 2 also shows that the mismatch errors increase, the more fitpoints are used, evidently as the result of a poorer match of the model to the data. In fact, if one tolerates deviations of, say, up to 10% between the model values and the (Gaussian) data, then it is easily seen that the use of a parabola and a symmetric 4^{th} degree polynomial should be restricted to an interval of 0.8×FWHM and 1.1×FWHM respectively; if the data had been obtained from a Lorentzian curve, these limits would be even more stringent. So whenever it is necessary to extend the range of fitpoints as far as possible (e.g. to reduce the effect of noise) one must choose a model which, both at the centre and asymptotically, behaves in a similar way as the data.

In Fig. 3 are shown, for each model, the individual fitting errors ϵ_o and ϵ_e as well as the error $\bar{\epsilon}$ on the combined estimate of the centre position, as a function of the number of fitpoints divided by the FWHM (expressed in

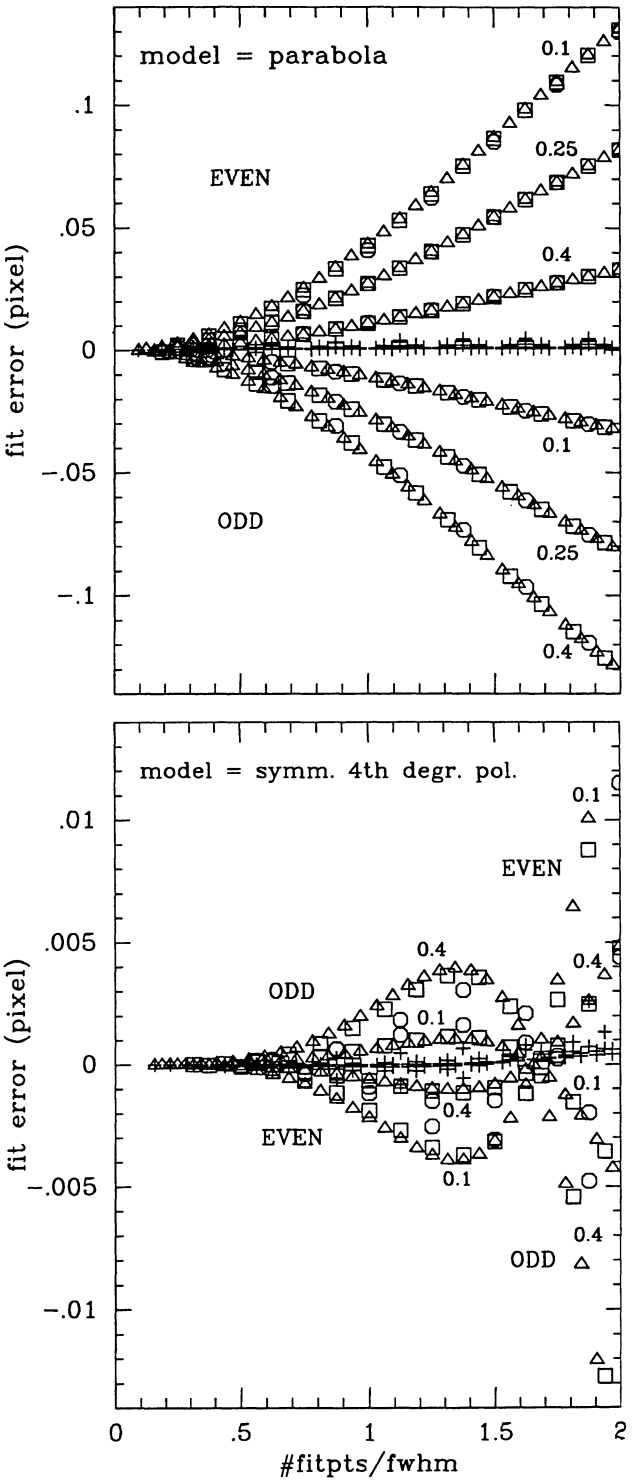

Fig. 3. Fit errors ϵ_o and ϵ_e as a function of the number of fitpoints in units of the FWHM of the data, for 2 models; different open symbols are used for values of the FWHM of the data = 8 (○), 16 (□), 32 (△) px; each position d of the data gives rise to two branches, corresponding respectively to even and odd numbers of fitpoints, as indicated. The error $\bar{\epsilon}$ on the combined estimate (19) of the central position is plotted with the same symbol (+) for all cases

pixel units) of the data; different symbols are for different FWHM of the data, and the different branches represent different central positions d, each of them giving rise to an odd and an even branch as indicated. It is clear that all results, viewed as a function of the number of fitpoints, scale with the FWHM of the data (except for a number of fitpoints close to the minimum number for the given model, for the reason explained above). Therefore, in choosing the optimal number of fitpoints to be used in practice, the relevant parameter characterizing the importance of model mismatch is the number of fitpoints in units of the FWHM.

The number of fitpoints in these figures has been extended beyond the limits discussed above, in order to point out a spurious decrease of the fit errors in Fig. 3b (symmetric 4$^{\text{th}}$ degree polynomial) around $\cong 1.7\times$FWHM. and to warn against the possible interpretation of such behaviour as indicative of an "optimal" range for the fit: in fact the figure shows the instability of the situation (i.e. a slight change in the FWHM with the same number of fitpoints could lead to very large errors) and a direct comparison of the data with the model values resulting from the least squares fit shows that the mismatch in such cases is very large indeed so that the seeming correctness of the position must be considered as quite accidental. Of course the enormous mismatch here is due to the fact that the model is unable to exhibit the right asymptotical behaviour and the match at the centre becomes worse, the more the model is forced to adapt to points representing the "background" level.

Comparing both models, one sees that the errors made by odd or even fitting are much larger for the parabola than for the symmetric 4$^{\text{th}}$ degree polynomial. This is of course not surprising since the latter has somewhat more freedom to match the data values. And although it is not discernible on the scale of these figures, the same is true for the errors on the combined estimate (see also Fig. 4).

However, the most important conclusion is that our proposed estimate (19), combining the odd and nearest even fit, results in a final error $\bar{\epsilon}$ on the centre position of less than 0.01 pixel in all cases (and mostly even below 0.001 pixel). It was also checked that the first order approximation $\bar{\epsilon}^{(1)}$ (20) for the error $\bar{\epsilon}$ is, on average, indeed a reasonably good estimate of the latter. However, the correlation between $\bar{\epsilon}^{(1)}$ and $\bar{\epsilon}$ is *not* sufficiently strict to warrant the use of $\bar{\epsilon}^{(1)}$ to decide which model or how many fitpoints would yield the smallest final error $\bar{\epsilon}$.

In Fig. 4, we compare the errors $\bar{\epsilon}$ on the combined estimates $\bar{\delta}$ of the centre position, for the 2nd degree and the symmetric 4th degree polynomial fits. As pointed out before, the latter model yields considerably smaller errors, except in some of those cases (marked with circles) where neither of the models should actually be used because the range of fitpoints extends too far beyond the appropriate range discussed above. This difference in performance is

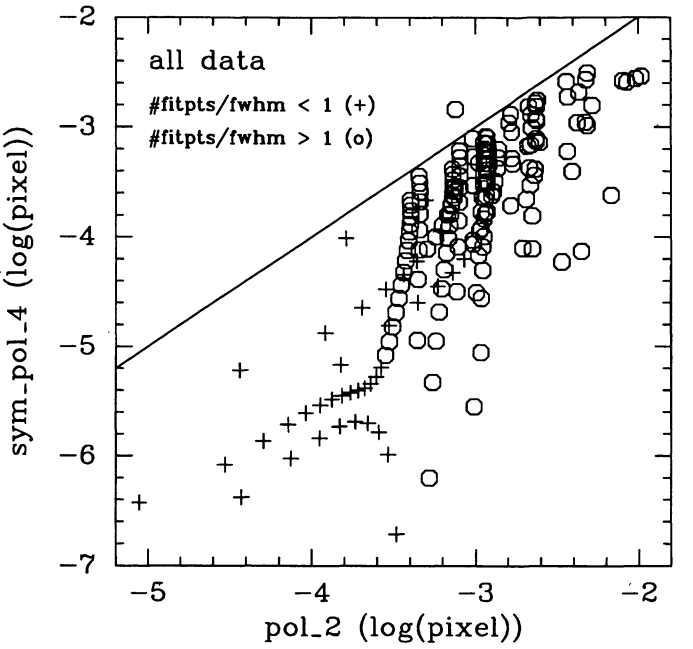

Fig. 4. Error $\bar{\epsilon}$ on the combined estimate (19) of the central position of the data for the symmetric 4th degree polynomial model, compared to that for the 2nd degree polynomial; all data except for FWHM=2 px (see text) are included

relevant e.g. for the position of cross-correlation peaks, the shape of which usually cannot be represented by a known analytical function, and which are often fitted with a parabola (e.g. Tonry & Davis 1979).

Finally, it is worth considering separately the case of data with a very small FWHM, say of the order of 2 pixels, which is typical for e.g. the lines in a wavelength calibration spectrum. Some packages determine line positions by fitting a parabola to either 3 (max. error $\cong 0.05$ px at $d \cong 0.3$), 4 (max. error $\cong 0.10$ px at d=0) or 5 (max. error $\cong 0.17$ px at d=0.5) pixels. The use of our estimate (19) will reduce these errors to less than 0.02 px in these cases; note that in high resolution spectroscopy, systematic errors of the order of 0.1 pixel are often unacceptably large, and their reduction to 0.02 px may be barely sufficient. Fitting a symmetric 4$^{\text{th}}$ degree polynomial cannot be expected here to yield reliable results at all, since it requires the use of a number of fitpoints in the "instability region" in Fig. 3. Generally speaking, one should always try to use a model representing the data as close as possible; for the lines in a comparison spectrum, a Gaussian is often the best solution; it can be fitted to 5, 6 or 7 points, and if necessary the combined expression (19) may be used to reduce the remaining (small) errors.

5. Some special cases with asymmetry

5.1. *Weakly asymmetric data*

Suppose the intrinsic asymmetry of the data is weak (e.g. slightly deformed line profiles, the cross-correlation peak of two spectra which have not been quite perfectly rectified,...); then it may still make sense to define the position of a feature as the position of the maximum of the function from which it was binned, and so locate it by fitting a model to the data and finding the maximum of the model. Then of course the question arises whether it is still possible to improve the measurement by combining the results obtained with an even and an odd number of fitpoints, as in the case of intrinsic symmetry.

We use the same notations as in Sect. 3, assuming that the maximum of the data and of the model are characterized uniquely by the parameters d and δ resp., but we drop the assumption that f and g are symmetrical functions. We may again derive a first-order estimate for the result δ of a least-squares fit, but of course, in this case the second order terms will not vanish anymore so that a first-order approximation is likely to be less reliable. However, the coefficients of d^2, $d\delta$ and δ^2 differ from zero *only* as a consequence of asymmetry, so they will be small in the present case. The first order approximation can be written as:

$$\delta_o^{(1)} = D_o d + R_o \quad , \quad \delta_e^{(1)} = \frac{1}{2} - D_e(\frac{1}{2} - d) + R_e. \quad (21)$$

where D_o and D_e are still given by Eqs. (12, 13), while

$$R_o = -\frac{\sum_{n=-p}^{p}(g(n) - f(n))g'(n)}{\sum_{n=-p}^{p}[g'^2(n) + (g(n) - f(n))g''(n)]}. \quad (22)$$

$$R_e = -\frac{\sum_{n=-p}^{p+1}(g(n_h) - f(n_h))g'(n_h)}{\sum_{n=-p}^{p+1}[g'^2(n_h) + (g(n_h) - f(n_h))g''(n_h)]} \quad (23)$$

The argument which led us to conclude that $D_o \cong D_e$, applies to R_o and R_e as well. Replacing R_o and R_e by their average $\bar{R}$ and repeating the operations which led to Eq. (19), we obtain:

$$\bar{\delta} = \frac{\delta_o - \bar{R}}{1 + 2(\delta_o - \delta_e)}. \quad (24)$$

This result is not of much practical use in itself, since it still contains the quantity $\bar{R}$ for which, at best, one could try and calculate a first- order estimate from Eqs. (22, 23). However, if the asymmetry is such that the contribution of R to $\bar{\delta}$ is important indeed, then the validity of the first-order approximation which suggests the form (24) is likely to be doubtful so that Eq. (24) simply should not be used at all.

Actually the point of Eq. (24) is mainly to show in what way some weak asymmetry in the data will affect the result of (perhaps inadvertently) applying the position estimator (19) to these data and to provide a test for its validity in that case: if the quantity $\bar{R}$ (as estimated from Eqs. (22, 23) turns out to be well within the error bounds one has accepted for the problem at hand, then it may be neglected in Eq. (24) and the weak asymmetry may be said to have no effect on the centering.

Furthermore, in marginal cases the expression for R_o and R_e might suggest an "improved" choice for the model, i.e. a choice which reduces the magnitude of these quantities. If a symmetrical model is used, in the denominator of Eqs. (22) and (23) the contributions $g(n)g'(n)$ and $g(n_h)g'(n_h)$ resp. will cancel and the magnitude of R_o and R_e will depend entirely on the asymmetry of the data. However, if the model is allowed to have some asymmetry as well, its values in the fitpoints may approach the data values more closely, with the effect of reducing the magnitude of R_o and R_e below their values for a symmetrical model. In other words, with weakly asymmetric data it might be justified to allow the model some freedom to become asymmetric (constrained of course by any knowledge one has on the nature of the data); however, when doing so one should always check carefully that the result remains meaningful.

5.2. *Binning errors with intrinsically asymmetric data*

Let $f(x)$ be a peaked function reaching its maximum at $x = 0$ and tending to zero at some distance from the origin; let $\bar{f}(n)$ be its binned value on the n^{th} pixel; then

$$\bar{f}(n) = \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} f(x)\mathrm{d}x = \sum_{l=0}^{\infty} \frac{1}{(2l+1)!}(\frac{1}{2})^{2l} f^{(2l)}(n). \quad (25)$$

In order to see the effect of binning on the position of $\bar{f}$, as determined from a least squares fit, we use $f(x)$ itself as the fit model. The position δ which minimizes $X^2(\delta) = \sum_{n=-p}^{p}[f(n-\delta) - \bar{f}(n)]^2$ is a solution of

$$\sum_{n=-p}^{p}[f(n-\delta) - \bar{f}(n)]f'(n-\delta) = 0. \quad (26)$$

Since the feature was located at the origin by definition, the solution of (26) represents the error due to binning; it can be estimated by expanding (26) in powers of δ, considering

$$\Delta_n = \bar{f}(n) - f(n). \quad (27)$$

i.e. the differences between the original values $f(n)$ and the binned values $\bar{f}(n)$, as quantities of the order of δ:

$$\sum_{n=-p}^{p}\left\{\sum_{k=1}^{\infty}\left[\frac{(-\delta)^k}{k!}f^{(k)}(n) - \Delta(n)\right] \times \sum_{l=0}^{\infty}\frac{(-\delta)^l}{l!}f^{(l+1)}(n)\right\} = 0 \quad (28)$$

In a first order approximation this leads to

$$\sum_{n=-p}^{p}[\delta f'(n) - \Delta(n)]f'(n) = 0. \quad (29)$$

so that

$$\delta = \frac{\sum_{n=-p}^{p} \Delta(n) f'(n)}{\sum_{n=-p}^{p} (f'(n))^2}. \quad (30)$$

and since, from (25), $\Delta(n) \cong \frac{1}{24} f''(n)$:

$$\delta \cong \frac{1}{24} \frac{\sum_{n=-p}^{p} f'(n) f''(n)}{\sum_{n=-p}^{p} (f'(n))^2}. \quad (31)$$

Tests in a number of realistic cases indicate that the order of magnitude of this error is at most a few times 10^{-3} pixels, owing to the fact that the numerator in Eq. (31) would vanish if the function f were symmetrical. However, one should bear in mind that, for a given set of data, this is only a lower bound for the actual error which will be made, and that this lower bound could only be reached if one could use as a fit model the original function itself from which the data values were obtained by binning.

6. Mismatch errors and noise

Naturally, the question arises as to the importance of model mismatch errors - and the relevance of a correction thereof - in the presence of noise. As far as individual positions are concerned, the preceeding results are certainly irrelevant if the noise-induced standard error on a position measurement is, say, several times the expected model mismatch error. On the other hand, if a sufficiently large number of data is combined (as e.g. in the case of wavelength calibration) systematic errors may yet show up even if noise-induced errors are considerably larger than the systematic ones (cf. e.g. the Gaussian residuals in Fig. 1), so in that case the relevance of our results will depend on the number of data available and on the purpose for which they are used.

Specializing now to the case of an individual position measurement, in order to answer the above question one needs an estimate for the noise-induced error and for the correction obtained by applying Eq. (19). Concerning the latter, one can take the difference $|\delta_o - \delta_e|$ as a measure for the magnitude of the mismatch correction, at least in those cases where application of Eq. (19) is valid (see further). As shown in Sect. 4, this quantity depends on the number of fitpoints scaled to the FWHM of the feature, on the fit model used and (weakly) on the sub-pixel location of the feature. On the other hand, Brown (1989) derived a general expression for the photon-noise induced centering error for any observed feature, as a function of the morphology of the feature and of the actual flux level; for a single Gaussian-shaped feature, this error depends essentially on its width, its strength, and on the signal to (photon) noise ratio (Verschueren 1991). It is important however to realise that this noise-induced error estimate was derived under the assumption that no systematic model-mismatch is present and that the full data feature is considered, and that therefore it represents only a lower bound; extensive Monte Carlo tests on synthetic features (Verschueren 1991) show that, in the presence of model mismatch, this lower bound is only attained for a limited range in fitpoints which decreases strongly as the mismatch increases.

Whenever an estimate of both the model-mismatch error and the noice-induced error are available, one can judge whether it is sufficient to take the fit result obtained with an odd or with an even number of fitpoints, or whether it is worth applying Eq. (19) to obtain a better sub-pixel precision. In the latter case, it is important to realise that Eq. (19) can only be applied when both δ_o and δ_e have values within the interval $[-0.5, 0.5]$, that they have the same sign, and that their sign is consistent with the choice of the fitpoints for the even fit, all of which is not necessarily true in the presence of noise (even very weak noise, if $d \cong 0.0$ px or $d \cong \pm 0.5$ px). In those cases where the above conditions are not fulfilled but the expected noise-induced error is nevertheless small enough to aim at sub-pixel precision, one should choose the odd fit result when both are close to 0.0 px and the even fit result when both are close to $\pm$ 0.5 px.

7. Concluding remarks

We have pointed out how systematic errors on the sub-pixel level can be detected and, in some cases, corrected empirically. In particular for the case that 1-dimensional centering is performed by the least-squares fit of a model profile to an image feature, we have discussed in detail the kind of errors that may result from *model mismatch* and we have shown that they may be considerably reduced if the intensity distribution of the feature to be localized, is intrinsically symmetric.

Although our discussion was limited to the 1-D case, it may be noted that to some extent it is also relevant for 2-D centering problems as e.g. in CCD-astrometry. The detection and empirical correction of systematic errors in 2-D centering are based on the same principles as in the 1-D case. If positions are determined by fitting a model to the data, the resulting systematic errors can be analyzed in the same way as in Sect 3. However, such an analysis is likely to become fairly complicated unless the features to be localized have circular symmetry, so in practice an empirically derived "ad hoc" correction may often be preferable.

On the other hand, our results are not relevant for CCD-photometry where centering is only part of the problem and where it is *essential* anyway that the correct model (or at least the best possible approximation to it) be used in the fit of stellar images (see e.g. Stetson 1990).

Acknowledgements. This research work was carried out in the framework of the project 'Service Centres and Research Networks', initiated and financed by the Belgian Federal Scientific Services (DWTC/SSTC). The authors are indebted to H.

Hensberge for suggesting the problem and for several stimulating discussions.

References

Bishop C.M., Roach C.M. 1992, Rev. Sci. Instrum. 63, 4450
Brown T.M. 1989, in: CCD's in Astronomy, ed. G.H. Jacoby, PASPC 8, 335
De Cuyper J.P., Hensberge H. 1994, in preparation
Dick J., Jenkins C., Ziabicki J. 1989, PASP 101, 684
Gulliver A.F., Hill G. 1990, PASP 102, 1200
Hensberge H., Verschueren W. 1989, The Messenger 58, 51
Hensberge H., Verschueren W. 1990, in: Errors, Bias and Uncertainties in Astronomy, eds. C. Jaschek, F. Murtagh (Cambridge Univ. Press) 335
Hensberge H., Verschueren W., De Cuyper J.-P. 1995, in preparation
Hensberge H., Verschueren W., Rogl J. 1994, in: The MK Process at 50 years, eds. C.J. Corbally, R.O. Gray and R.F. Garrison, PASPC 60, 358
Jorden P., Deltorn J.-M., Oates P. 1993, Gemini 41, 1
Rutten R., Dhillon V., Horne K. 1992, Gemini 38, 22
Stetson P.B. 1990, The techniques of least-squares and stellar photometry with CCDs, DAO preprint, May 1990
Tonry J., Davis M. 1979, AJ 84, 1511
Verschueren W. 1991, High Precision Radial Velocity Determination from CCD Echelle Spectra of Early Type Stars in the Young Cluster NGC 2244, Ph.D. Thesis, Free Univ. Brussels
Verschueren W., Hensberge H. 1990, A&A 240, 216