Acquisition, Archiving, and Analysis (3A) Software Package

User's Manual

Ilya Ilyin

Department of Geosciences and Astronomy, University of Oulu,
P.O.Box 333, FIN-90571 Oulu, Finland
Ilya.Ilyin@oulu.fi

May 1, 1997

Contents

Chapter 1

Data reduction

The data reduction of the échelle spectra consists of the following steps:

- Image processing
  - Removal of cosmic spikes.
  - Removal of the bias offset and trimming the image.
  - Correction for the non-linearity effects of the CCD.
  - Correction for the flat field.
  - Removal of scattered light.
  - Finding spectral orders and extraction of the orders.
- Processing the spectra
  - Correction for interference and vignetting.
  - Wavelength calibration of the spectra.
  - Normalization of the spectra to the continuum.
  - Fine tuning of the wavelength scale of the spectra.
  - Wavelength transformations
    * From the laboratory wavelength scale to heliocentric wavelengths.
    * From heliocentric wavelengths into the stellar rest frame.
    * From wavelengths to radial velocities.
  - Measurements of radial velocities, equivalent widths of stellar lines, creation of maps of line variability and so on.

1.1 Removal of cosmic spikes

The échelle CCD image obtained after observation of the stellar spectrum consists of different structures such as spectral orders, which include continuum curves superimposed with absorption and/or emission lines, interorder space or background and comic spikes. To detect cosmic spikes, we use a digital filter which removes all the structures from the image except the cosmic spikes. In other words, we have to apply a high-pass filter to remove all the low spatial frequencies from the image. After it is done, we can easily distinguish the cosmic spikes from the rest of the image by using a simple thresholding algorithm which will detect all the spikes above a specified level.

In presence of photon noise the question is how faint cosmic spikes we can detect and remove, and, can we clearly detect the the bottom parts (tails) of the cosmic spikes. Basically, we can smooth the image and then take the difference of the original and filtered images. But then we will find that, at the top of the bright orders, the noise due to photon statistics is higher and we can not distinguish, at least, the bottom of the spikes from the noise. The way to avoid this problem is to use the statistical property of this kind of the noise. The smoothed image can be considered as an estimate of the mean level of the original image in each pixel. If we take the ratio between the original and smoothed images we will get an estimate of the relative error, which is decreasing as the level of the signal grows. The following equation describes this process:

$$\text{ratio} = 1 + \frac{\text{noise}}{\text{bias} + \text{signal}} + \frac{\text{spike}}{\text{bias} + \text{signal}}$$

The image we are working with is a biased image which still contains the bias offset. In this way one can avoid division by zero at the very low accumulation levels between the orders or within the overscan bias strip. A special consideration should be taken into account about the bias offset in the presence of spatial noise which can not be completely removed by the filter applied. Such a problem occurs in

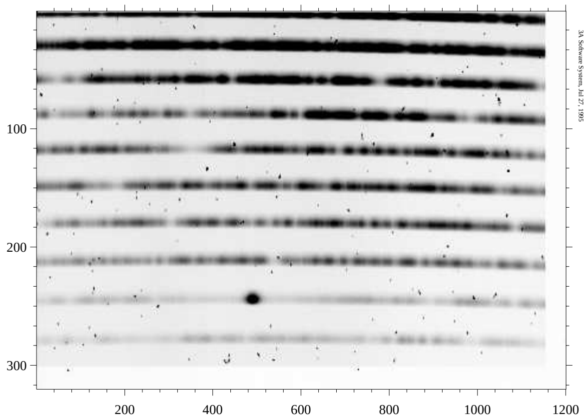

Figure 1.1: The original image of II Peg obtained with the SOFIN spectrograph with the medium resolution camera in the Ca II H spectral region. Ca II H 3933Å in emission is visible in the second order from the bottom.

case of close and narrow absorbtion lines.

The best way to filter the image is to use a median filter which filters across the rows of the CCD image. It means that the height of the filter is one column and the width should be adjusted to the features of the given image. A width of about 11-15 pixels is enough in most of the cases.

The procedure of detecting and removing the cosmic spikes consists of the following steps:

1. Find the best median filter length. For that, display the original image, point the horizontal marker line onto some cosmic spike, and display the cross-cut. Apply a median filter of such a width, that the smoothed curve is not affected by the spike. If there are some very broad cosmic spikes, then the tendency to increase the filter length as far as it does not affect the spectral features. If the length cannot be increased anymore, then detection and removal should be done in two stages by using a broader window after removal of all sharp spikes or by means of manual correction of the image data.

2. Process the original image(s) by a median filter of the selected length. That can be done from the database process menu (Fig 1.4):

   - Mark the images or put the table bar on the single image to be processed.
   - Select the `Filtration` command in the `Process` menu and choose the `Median filter` submenu.
   - Specify the filter parameters, choose ratio of original and filtered images as the result, and

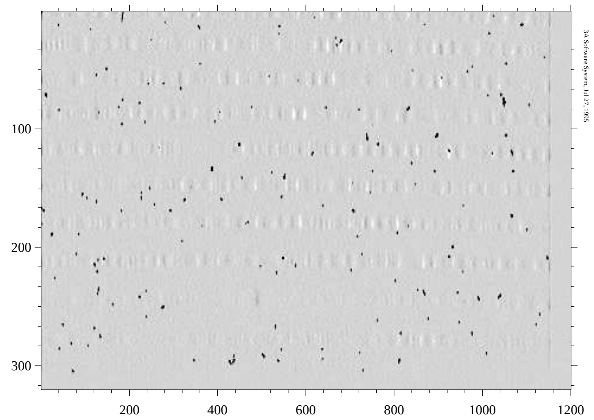

Figure 1.2: After processing by a median filter with window length of 15 pixels. The ratio between the original and the filtered images is shown.

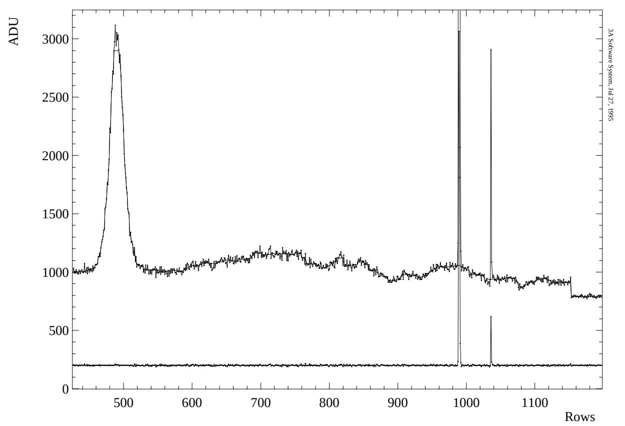

Figure 1.3: The plot shows a column taken from the original image near the emission line together with the cosmic spikes. A median filter of 15 pixels window length goes through the data and the ratio between the original and the filtered data is shown at the bottom. The ratio was enhanced with a scale factor.

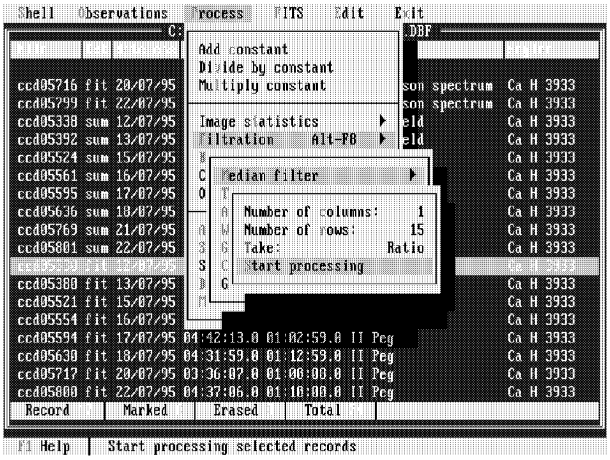

Figure 1.4: Two-dimensional median filter menu window.

start processing.

3. Make the cosmic spike mask image using the ratio of the original and the filtered images. The histogram of a region of the image (for instance, 100×100 pixels) with some number of spikes included is used in order to find a threshold level which could separate the spike intensities from the main part of the histogram. The histogram should be calculated and displayed with the appropriate resolution which allows to see and to set the separation level interactively. The most deviation points of data could significantly reduce the resolution. Cutting off the deviating tails of the histogram solves the problem. That can be done by changing the cutoff levels of the minimal and the maximal points. Here, the cutoff of minimal points is irrelevant, it could be 1-3%. The cutoff of the maximal points, which are the spikes, depending from their number, but the value 0.05% looks suitable.

   After the histogram was displayed, choose the threshold to separate the spikes by moving the graphical marker box in the histogram. The lower left end of the box is irrelevant, but the position of the right side of the box defines the proper level. In order to accept the data boundaries, press `Enter`, to disregard use `Alt-X`. On exit from the histogram, the imager program redisplays the image according to the newly selected interval: for the cosmic mask making procedure it could not be needed. Press `Esc` to stop displaying.

   The last step is to detect the most deviating points of the data according to the selected threshold. Before that, one should specify the direction of the detection: the program could detect the points which are deviating either below, or above, or in both directions regarding to the selected interval. For the ordinary

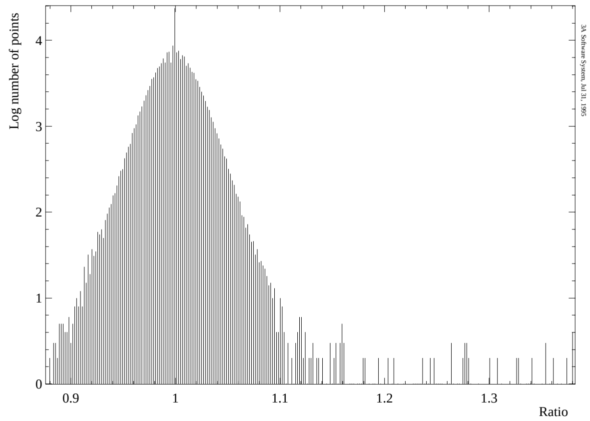

Figure 1.5: Histogram of the ratio of the original and the filtered images. A threshold at level 1.11 separates cosmic spikes from all other points of the data distribution.

   cosmic spikes, the second option is the actual. The option can be changed in the threshold mode menu command. Run the thresholding command in order to create the cosmic ray mask. The new image is written on the disk, appended to the database of images, and displayed on the screen to check the result.

   If the level was selected too low, then a noisy random pattern could be seen in the image which systematically follows the interorder space. On the other hand, if the level is too high, then only the uppermost parts of the spikes were detected and marked in the mask image as the points to be corrected. In both cases, a second trial is necessary: exit from the mask image display by `Alt-X` and display the histogram again by `Alt-R`. Change the level and repeat thresholding. The result should look better.

4. The command to remove cosmic spikes can be found in the `Process` menu of the database of images. The command has two arguments which have to be specified correctly: they define the area around the spike where its replacement is to be found. The median is used as the replacement of the spike, which is calculated in the specified window. The larger the window the more distinct points of the image could be involved (most likely, the neighbouring spectral orders), and the median value cannot be trusted. The smaller the window the less points are available; the program, furthermore, checks that the points involved into median calculations are not marked as well and extends the window, if necessary, to the specified size. This could lead to a comparatively large distance between the spike and the points used to calculate its replacement, even if the specified window size is small.

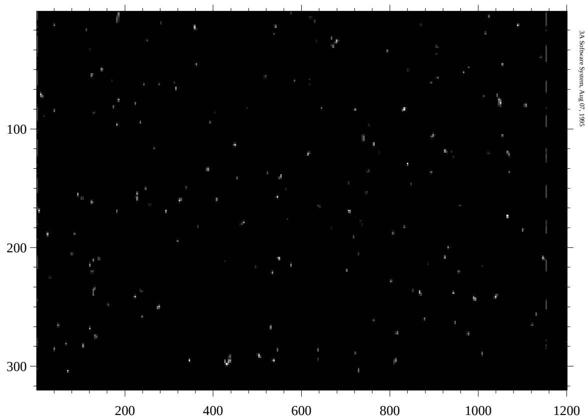

Figure 1.6: The image of the cosmic spikes mask with the features marked to be removed from the original image.

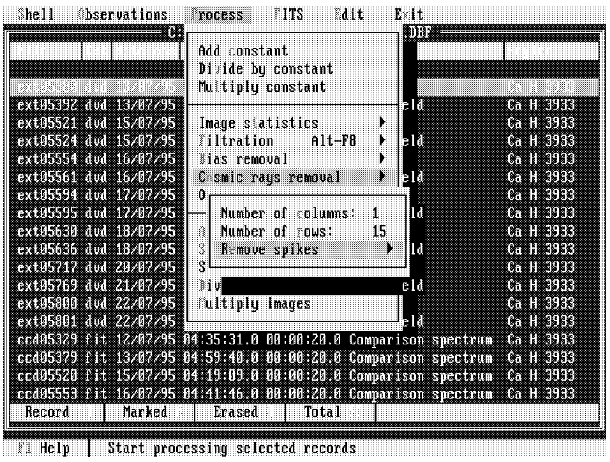

Figure 1.7: Menu window of cosmic spikes removal

The most reasonable way to select the removal window size is to say, that the width should be about the median filter length (used for the filtering of the original image to obtain the mask), and the height should be one, especially in the case of the échelle images where a rather strong cross-order gradient occurs for the short optical camera.

The program can be run in batch mode, in case the cosmic masks are prepared for a number of images, which can be marked so that the program could process one image after the other.

The output file, co-named as the original and with the extension `*.clr` is created. The image is appended to the database of objects.

5. In some cases the described semi-automatic cleaning procedure could not remove all the spikes completely. In another case, instead of filtering, thresholding, and removing, it could be much faster to remove 2–3 spikes interactively by using one of the two data correction procedures:

   - Correction of the deviating data points manually, as it is described in the imager program. The spikes should be displayed within a cross-cut of the image, given by the marker box, and replaced by interpolaton.
   - Correction by editing the image data in a window. Bad point then can be replaced by a median taken on the window size specified, as it is similar to the spikes removal procedure above.

   Free the disk space by deleting unnecessary images and their records from the database; they have the extensions:

   `*.dvd` which are the normalized images,

   `*.cr` which are the cosmic ray masks.

   The images which should remain in the database are the cleaned images with the extension `*.clr`.

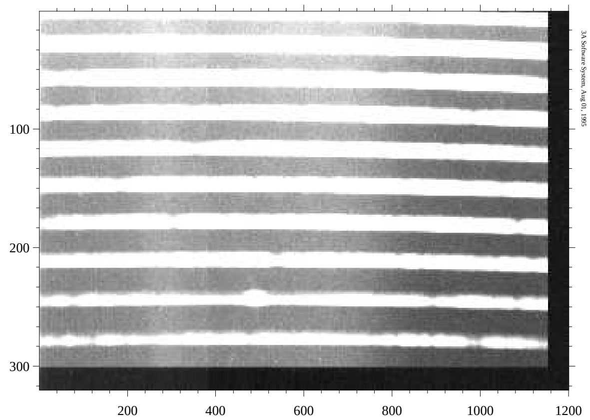

Figure 1.8: The same image after removing the cosmic spikes.

1.2 Removal of the bias offset and trimming the image

The analog-to-digital converter (ADC) of the CCD electronics transforms the signal into an unsigned integer number, which is in the range 0-65535 ADU (analog-to-digital units). To avoid negative numbers, which can occur for low signal levels due to noise, a constant bias offset is introduced into the value of each pixel by means of the CCD hardware setup. Usually, the value of the bias offset is about 500-1000 ADU and can easily be estimated by the level of a bias frame. But the bias offset may be changing with time, due to a drift of the CCD and/or the electronics' temperature, and also the value can change during readout of the CCD due to a small variance of the offset voltages. It means that removal of the bias offset must be done for each image obtained from the CCD separately, and instead of a constant level subtraction a bias structure should be removed. For each CCD image we can easily obtain overscan bias pixels by reading out some more pixels than the actual number of CCD columns and rows. The amount of additional columns and rows can be defined by expanding the readout region of the CCD. Usually, we obtain an extra 20-30 bias columns (direction of serial transfer) and 30-40 bias rows (direction of parallel transfer).

To remove the bias offset structure, we will extract an average column at the bottom of the image or an average row at the rightmost part of the image, then smooth it and subtract it from the whole image. Smoothing of the average vector (a one-dimensional set of data) is necessary to remove as much as possible random noise in the data of the vector, otherwise we will add this noise into the whole image during subtraction (note, that this noise would be correlated in the result, because the same column/row is subtracted from each original column/row). The origin of the random white noise in the overscan bias strips is in the CCD hardware, and it is known as readout noise. The readout noise is produced mostly by the reset noise of the output transistor of the CCD.

Due to some features of the CCD structure and the readout procedure, in the case of different CCD systems, the CCD frame may start and end with a very bright or dark column or row. This means in practice, that we should remove from the CCD frame a few first and last rows and a few first columns. Also we should remove the overscan bias strips in columns and rows after subtraction of the bias offset is finished. Omitting this step can produce severe boundary problems in the following reduction procedures. We call this operation trimming of the image. Trimming can be done in two steps: during extraction of the bias strip and, after bias subtraction, by extracting a subset of the image.

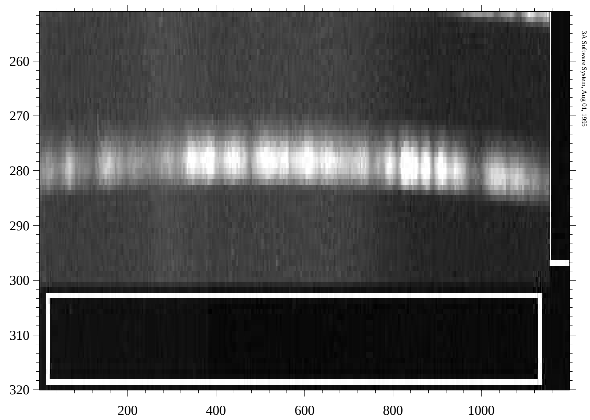

Figure 1.9: Enlarged area of the bottom part of a CCD image with the marker box located on the overscan by columns bias strip.

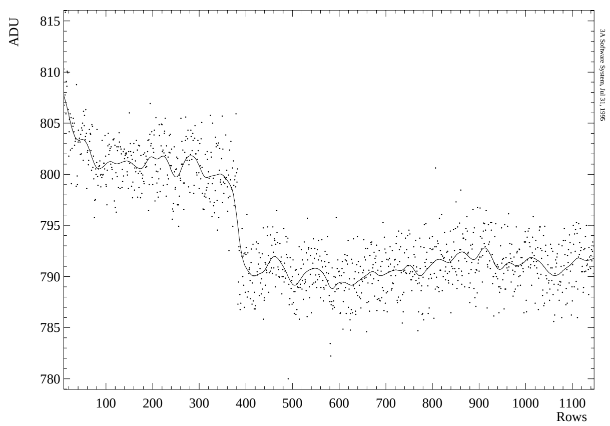

Figure 1.10: An average column taken from the previous image. A jump down in the bias level is seen. The smoothed curve is subtracted from the image column by column.

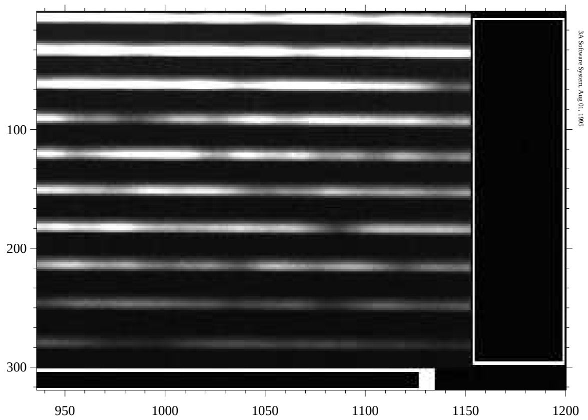

Figure 1.11: Enlarged area of the rightmost part of the CCD image with the marker box located on the overscan by rows bias strip.

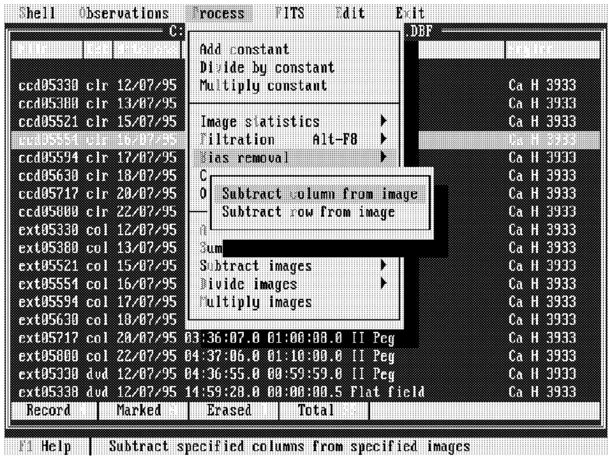

Figure 1.12: Menu window of the bias removal commands.

At first, we have to define the area, which will be used as the overscan bias strip. Usually, it is a good idea to use the overscan by columns for the stellar images, when we can mostly remove any variations in the bias across the whole image, but to use the overscan by rows for the flat field frames, where, due to the high accumulated signals, the bottom overscan strip would be affected by the effect of poor serial transfer efficiency. This case can be easily checked in two ways: either, point the marker box onto the overscan area and enhance the image by calculating and applying the histogram; then, some "leaking" of light into the overscan strip can be clearly seen. Or, take a cross-cut and enlarge the overscan area; then, a decreasing trend in the bias should be visible.

Two different options are used during bias subtraction: subtract the overscan vector as a column or as a row.

- Overscan bias by **columns**: bottom part of the image. The marker box should be pointed to the bottom overscan area and moved away from the beginning of the image, so that a few first rows will be omitted as well as some last rows of the image, including the vertical overscan strip and a few last rows of the exposed area of the CCD. The subtract bias column option should be selected in the bias menu window. The subtraction of the bias vector will be done from the whole image column by column, and the width, or number of rows, of the new image will be equal to the length of bias vector. The new unbiased image will still include areas which should be removed: the first few columns and the bottom overscan by columns strip. To do that, extract a subset of the image.
- Overscan bias by **rows**: rightmost part of the image. The marker box should be pointed to the right overscan area and lowered away from the very first few columns, and the height should be adjusted to avoid the bottom part of the image, including the horizontal overscan area and the last few columns of the exposed area of the CCD. The subtract bias row option should be selected in the bias menu window. The subtraction of the bias vector will be done from the whole image row by row, and the height, or number of columns, of the new image will be equal to the length of bias vector. The new unbiased image will still include areas which should be removed: the first few rows and the rightmost overscan by rows strip. To do that, extract a subset of the image.

The following is a summary of the steps:

1. Point the marker box onto the selected overscan area.
2. Put the marker borders into the list of regions: press `Alt-L`, create a new (`F5`) or edit (`F4`) an existing region. The extraction method must be either `Avrcols` or `Avrrows`. Highlight the region created by `Ins`. Exit by `Esc`.
3. Extract the region by `F8` or, if there are many images to process, from the `FITS` menu of the database program.
4. Transfer into the database of spectra and smooth the extracted vectors. Gaussian, median or trimmed mean filters are suitable for that. A filter length of about 15-25 is enough to get a suitably smoothed curve. If there are cosmic spikes, it is better to edit them manually by using `Ctrl-Enter` and `Alt-I`. After that, the data can be smoothed.
5. Highlight only the vectors to be used for bias subtracting.

6. Transfer back into the database of images and mark the images to be processed.
7. Go to the `Process/Bias removal` menu and select one of the two commands to subtract bias vectors, as a column or as a row. New images with extension `*.col` or `*.row` will be created and appended to the database.
8. To extract a subset of the unbiased images, display one of them. Select a suitable area of the image to extract and save it into the list of regions. It is convenient to change the file template prefix to `ext`, so that the extracted image file name will be e.g. `ext005302.col`. The extraction method must be `subimage`. Highlight this record and unmark all other lines in the table.
9. Extract the region by `F8` or, if there are many images to process, from the `FITS` menu of the database program.

1.3 Correction for the flat field

The total noise of the signal registered by the CCD can be expressed in the following form:

$$\delta_{total}^2 = \delta_{stat}^2 + \delta_{spn}^2 + (\sigma_{dark}/C)^2$$

where δ_{total} is the total relative noise of the registered signal C (in e^-), δ_{stat} is the relative error of the photon noise, which is equal to $1/\sqrt{C}$, δ_{spn} is the relative high frequency pixel-to-adjacent-pixel spatial noise due to variations of the pixels' sensitivities, σ_{dark} is the noise due to the dark current accumulated during the exposure time, including the readout noise of the CCD. For weak signals, the dark noise component becomes dominating. For strong signals, the high frequency spatial noise dominates over all the others. The photometric quality of the registered signal mostly depends on how accurate the spatial noise can be removed during the flat field correction.

The following spatial structures are affecting the object image which can be corrected by using the flat field image:

- High frequency spatial pixel-to-adjacent-pixel fixed pattern noise.
- Dirt stains and particles on the camera entrance window and the CCD chip.
- Optical interference on the CCD coating layers.
- Product of the vignetting function of the optical camera and the blaze function of the échelle.

In order to correct for these structures, we have to make a flat field exposure by using a special lamp situated in the spectrograph. The lamp has no spectral lines in its spectrum, but only a smooth continuum. During the flat field lamp exposure, the light passes through the spectrograph exactly in the same way as during stellar exposures. In order to achieve maximum accuracy of the flat field correction, we have to accumulate a much stronger signal in the flat field than in the object image. Because of the limited capacity of the CCD pixels, we have to obtain a number of flat fields, which depends on the maximal signal of the object image. For instance, in case of weak accumulated signals in the object image the number of the flat fields can be limited to 5, for strong signals it could be 50. Thus, we have to make a series of exposures with subsequent adding them together. Of course, this time consuming procedure cannot be done during observations, but it should be done during the following day.

In our recently developed technique of flat fielding, we are using merging of the spectral orders to obtain the mas-

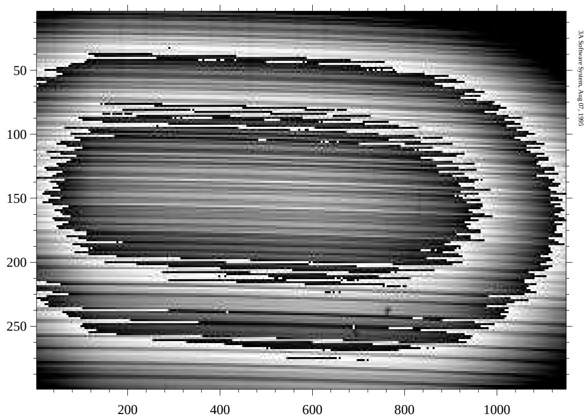

Figure 1.13: A typical flat field image obtained in the merged spectral orders mode in the 6427 & 7516Å spectral region with the 2nd optical camera of the SOFIN spectrograph. The sum of 31 individual flat fields yields a total signal of 1 200 000 ADU in the centre of the image. Wrap of the palette colours was used to enhance details on the image, where some dirt stains on the CCD window are visible at the bottom, between rows 600-800. Atmospheric oxygen absorption lines are faintly visible at the top of the image.

ter flat field. This simply means, that the slit height is increased to the maximum during flat field exposures, so that the spectral orders in the image overlap. There are several reasons to do that:

- To avoid the problem how to get the échelle and the cross-dispersion prism *exactly* into the same position as for the object image in the night.
- To get more light in the flat field at the wings of the spectral orders of the object image.
- To get a smoother structure of the flat field across the dispersion to avoid low frequency spatial contributions.
- To get more signal from the flat field lamp (which is red) in blue spectral regions where also the CCD is less sensitive.

To get an accumulation level which is strong enough for a reasonable exposure time, it is necessary, also, to increase or decrease the entrance slit width.

The disadvantages of this method are:

- The optical interference in some red spectral regions on the CCD coating layers has a different effect compared to the object exposure. This happens because of different spectral contents of the light per fixed area on the CCD, mostly, due to the mixture of the light from consecutive spectral orders. As a result, the amplitude, phase and periods of interference on the merged flat field are different from those on the object image. To get the same pattern of the interference, we have to make a flat field exposure with exactly the same slit setup as for the object image. Such an exposure can be done immediately after the object exposure, and we do not have to care about how strong a signal can be obtained as long as it is enough to extract the interference pattern (see below).
- The high frequency pixel-to-pixel spatial structure is wavelength dependent. Fortunately, at least for the EEV CCDs we are using, this function is very smooth between several spectral orders.

A basic approach to flat field correction is to extract spatial frequencies from the flat field image which are present in the object image and then divide the object image by the correction image. Extraction of the spatial frequencies means that we have to avoid, or filter out, the extra spatial structures which are present in the flat field image:

- The structure of the échelle orders in cross-dispersion direction.
- The product of the vignetting function of the optical camera and the blaze function of the échelle in dispersion direction.
- Absorption lines in the flat field. In some spectral regions from 7000Å to 9000Å we can see very strong bands due to molecular oxygen present in the air in the optical path of the spectrograph. In case of the merged flat fields, when the slit height has been increased, these lines are expanded over a few orders.
- The optical interference on the CCD coating layers which is different from the interference present in the object frame.

To remove these structures, we can use a lowpass filter which works across the rows in dispersion direction column by column over the whole unbiased flat field image, so that the structures present in the object image remain unchanged. The amplitude of the spatial noise is linear, i.e. depends on the amplitude of the signal (if the CCD response is linear, of course), and that means that we have the case of multiplicative spatial noise. To get the surface of the relative noise, which does not depend on the illumination level, we have to divide the original flat field by the filtered image. Then, the unbiased object image should be divided by the normalized flat field image.

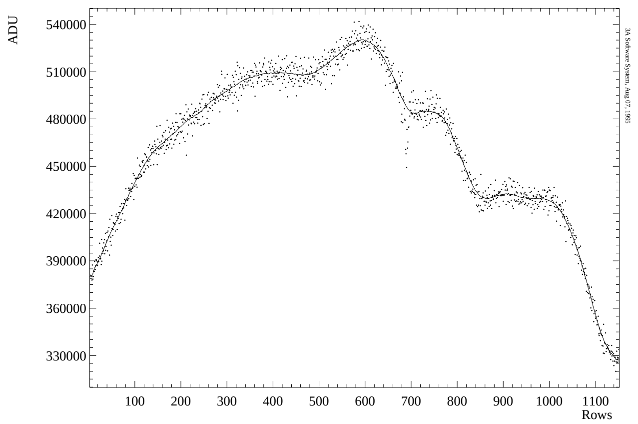

Figure 1.14: One column taken from the flat field image showing three components of the multiplicative noise: low frequency noise as variations of the illumination level across the image, medium frequency noise due to dirt stains, and high frequency noise as pixel-to-pixel fixed pattern. A trimmed mean filter with parameters 51/11+11 was used to eliminate the low frequency noise.

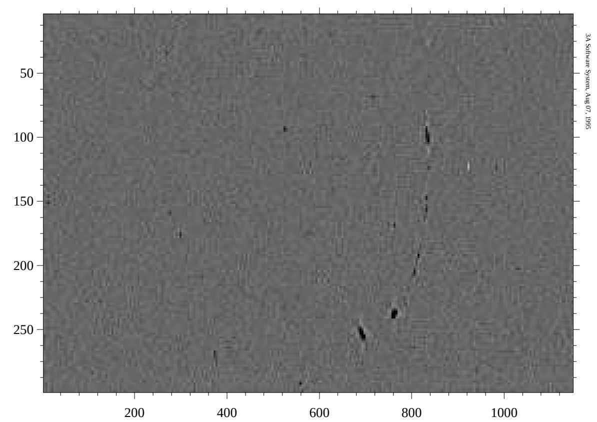

Figure 1.15: A normalized flat field containing only high frequency spatial pixel-to-pixel noise and dark spikes due to dirt stains/particles on the CCD window.

In order to remove atmospheric absorption lines from the master flat field, the normalized image can be filtered by a lowpass filter which works across the columns in cross-dispersion direction row by row over the whole image. Then, the normalized source image has to be divided once more by the new filtered image.

Optical interference patterns in the object image can be extracted from a single flat field image obtained with exactly the same conditions as the object image by the following steps:

1. Divide the unbiased flat field image by the normalized master flat field (the same way as the object image) to remove all the structures already corrected in the object frame.
2. Apply a lowpass filter working across the rows in dispersion direction in order to extract only the interference pattern.
3. Divide the object image by the normalized smoothed surface which consists of only the interference pattern.

It is obvious, that the vignetting function cannot be removed in this procedure, because it was removed as the low frequency spatial noise from the master flat field. Furthermore, the function goes along the spectral orders, which are in general tilted in the CCD image. But we are working along the columns because we are mostly concerned with the high frequency noise, which contributes most to the total noise. The vignetting function can be easily excluded after the extraction of orders in the stage of continuum fitting. Another possibility is to extract spectral orders from the ordinary flat field image corrected for the master flat field. Sufficiently smoothed and normalized to the maxima, the extracted flat field orders can be used to straightening the continuum of the object orders.

After division by the master flat field, it is worthwhile to check the value of the high frequency spatial noise reduction to be sure that all corrections were done properly. To do that, we have to extract the same *single* column from the original object image (before flat field correction), from the original unbiased flat field image and from corrected object image. We need to find in the object image, if possible, a flat column without steep gradients and spectral lines to get enough pixels to estimate the value of the noise. Then, for the known gain factor of the CCD (g), the accumulation level in ADU in the object column (C_*) and in flat field ($C_{\rm ff}$), we can estimate the lower limit of the relative error of the data in the column of the corrected image:

$$\delta^2_{\rm ratio} = \delta^2_* + \delta^2_{\rm ff} = \frac{1}{C_* \times g} + \frac{1}{C_{\rm ff} \times g}$$

It is assumed here, that the fixed pattern has been removed, but some amount of the random statistical noise from the flat field (the second term in the equation) is introduced into the corrected object image. Then, we have to compare this value to the measurements of the noise in the column of the image after correction. To measure the noise, we can smooth the column by a gaussian filter, take the ratio and use the statistical window to get the value of the relative error. If there are some weak spectral lines in the column, two consecutive columns of about the same level could be extracted and the ratio taken. The relative error then should be divided by $\sqrt{2}$.

For instance, the data shown in Figure 1.16 give a relative noise level in the object column before flat field correction of 1.37% and after the correction of 0.93% . The accumulated counts in the object column give a relative error of 23 000 ADU $\times$ 0.51 e^-/ADU = 0.92% and in the flat field 1 040 000 ADU $\times$ 0.51 e^-/ADU = 0.14%. Then, the sum of squares of these errors gives a relative error of 0.93%, as it was measured. Thus, the signal to noise ratio *per pixel* was improved from 73 to 108. During spectral order extraction, when 10-15 pixels are summed, the signal to noise ratio will be increased again by a factor of 3-3.5 for this particular exposure.

The following is a summary of the steps:

1. Process the unbiased summed flat field to remove low frequency structures. The trimmed mean filter is the most suitable for this purpose. The parameters of the filter depend on the structure of the flat field. To process the image column by column separately, the height of the filter should be set to 1. The width could be 31 and the number of points to reject must be equal for the lower and upper parts of the window, say, 11 and 11. The ratio should be taken while filtering.
2. If absorption lines are present in the flat field, the normalized image can be filtered in perpendicular direction to remove slow trends produced by the lines. The ratio should be taken while filtering.
3. Divide the unbiased object image by the normalized flat field.
4. Check the noise reduction.

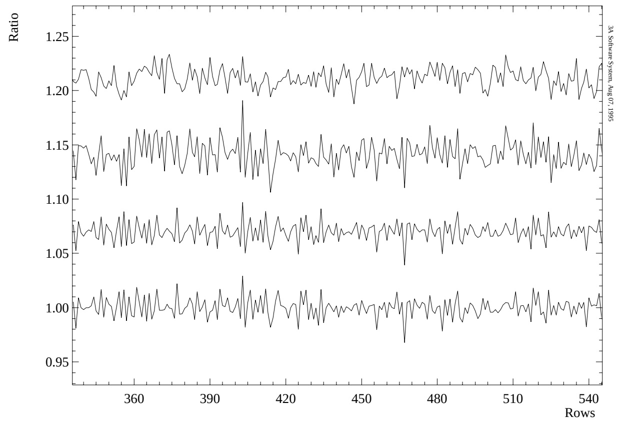

Figure 1.16: High frequency spatial noise of the CCD. The lower two panels are fragments of the same column taken from two merged flat fields obtained at different times. Statistical noise is negligible compared to the high frequency noise due to the large amount of the accumulated signal, and it is clearly seen, that the pattern is repeating from image to image. The third panel from the bottom shows the same column taken from the object image, filtered and normalized to the smoothed curve. The same pattern is visible with a larger statistical noise because of less accumulated signal. The upper panel shows the same column after division by the flat field column. The fixed pattern has been removed, only the signal + statistical noise remain.

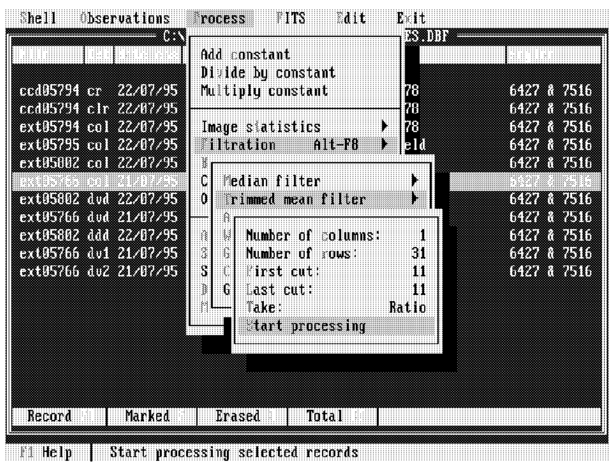

Figure 1.17: The trimmed mean filter menu in the database of images.

1.4 Removal of the scattered light

Scattered light in the échelle spectrograph originates from different sources:

- Scattering of the light which is out of the optical beam from the frames of the optics and the internal parts of the spectrograph.
- Scattering of the light from very bright emission lines or from very bright orders, so-called interorder scattering. In this case, the light is also scattered from the optical surfaces of the spectrograph, as well as on the CCD window.

The level of the scattered light increases in the blue spectral regions in case of observations of an object with strong red emission, when out of the beam red light diffuses into the optics of the spectrograph. Protection screens, baffles and optical filters limit this effect to a minimum, but still the scattered light, which has amplitudes of 1% to 10% of the level of the signal in the neighbouring orders, should be removed during image processing. If the scattered light on the échelle image is not removed or not completely removed, then it will cause reduction of the line depths in the extracted spectral orders (veiling).

To subtract scattered light, we have to extract a smooth surface by using information from the interorder space. In general, this surface turns out to be *not* describable by simple functions like polynomials. Therefore, we use a digital filter to remove all the spatial frequencies produced by the spectral orders except the interorder space.

A gaussian filter working in rejection mode across the columns can give a good estimate of the scattered light. The length of the filter should be selected as the maximal distance between the orders. Several iterations must be performed in order to reach the interorder level. After each iteration, points *above* the filtered curve will be replaced by the corresponding points taken from the filtered curve. The smaller the filter length, the more details in the structure can be extracted. The more iterations, the smoother the resulting surface. The introduction of a binning factor in the direction of the rows has the following reasons:

- Saving of computation time.
- The scattered light surface is generally a smooth surface without many small details. Instead of using each row to get an estimate of the scattered light in that row, we can use the average information from a few consecutive rows to get less noise from the source data and thus a more reliable estimate.

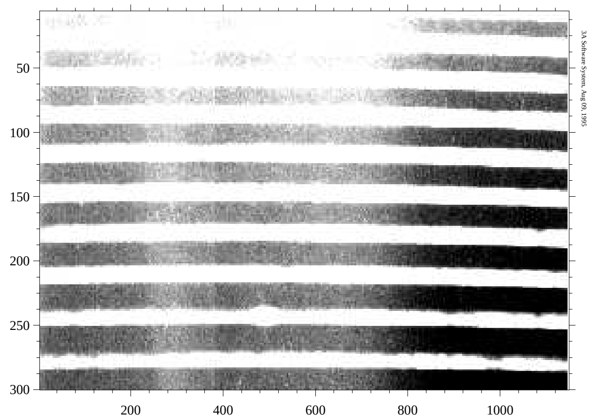

Figure 1.18: An image of II Peg taken with the 2nd camera in the region of Ca II H 3933Å. The scattered light is seen to be strongest in the blue. The interorder space has been enhanced to make the irregular structure of the scattered light better visible.

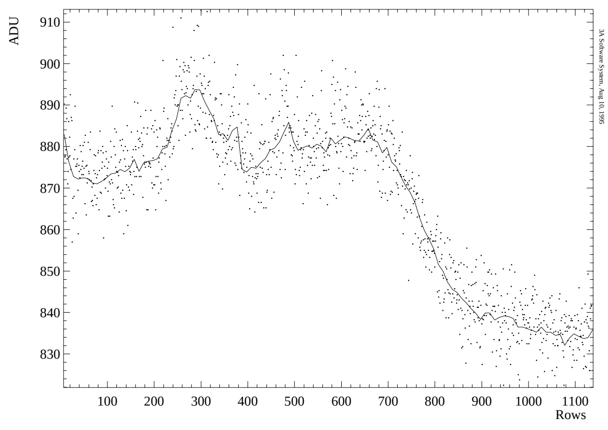

Figure 1.19: Average column (4 columns starting at 227) taken between the orders of the previous image to show the amplitude and the scale of variations of the scattered light.The solid line is the same average column taken from the final scattered light image.

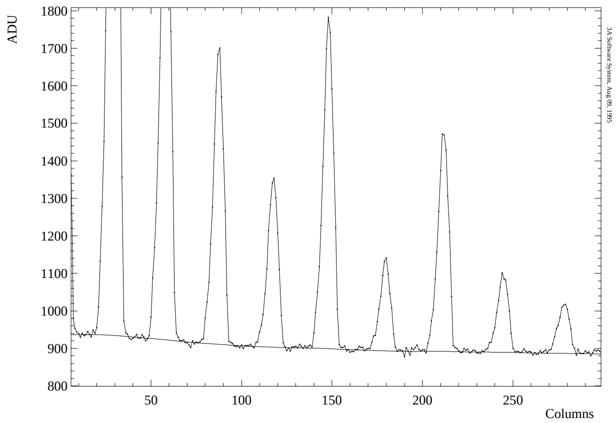

Figure 1.20: Cross-dispersion cut to adjust the gaussian filter parameters in order to fit the curve to the interorder points. The data represent an average of 10 rows taken from the central part of the image. A Gaussian filter with the parameters length= 22, number of iterations= 11 and *above* rejection was used.

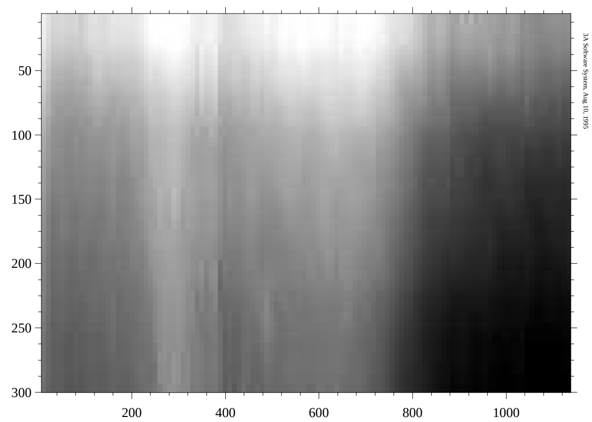

Figure 1.21: Surface of the extracted scattered light.A binning factor in the direction of the rows of 10 was chosen.

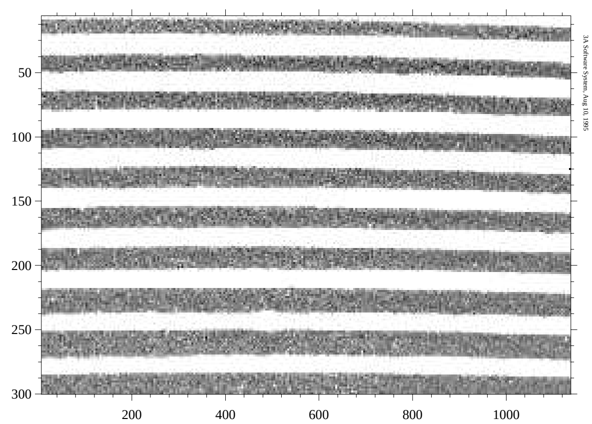

Figure 1.22: The same image as at the beginning but after subtraction of the scattered light surface.

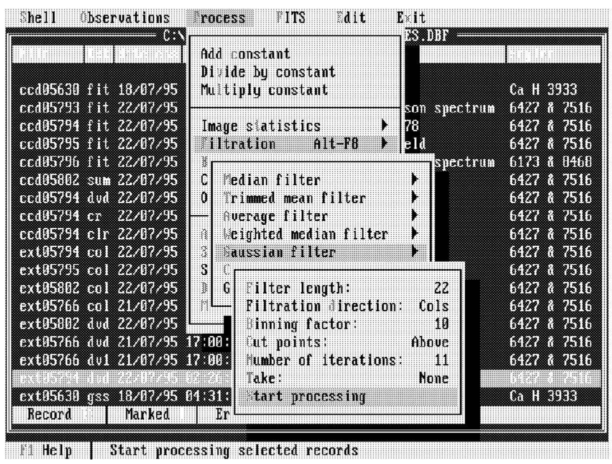

Figure 1.23: Gaussian filter setup menu commands in the database of images.

The value of the binning factor depends on the structure of the scattered light, but usually a value of at least 10 rows is convenient. A binned output image will be produced with an accordingly reduced number of rows. Subtraction of two images with different binning factors (step sizes or dispersions) by using the `ProcessSubtract` menu command will preserve the *minimal* step size of the two in the resulting image. It is important not to use the subtraction mode of the gaussian filter if the chosen binning factor is different from 1. In this case, the difference image would have a step size equal to the binning factor selected. If no binning was chosen, then it is a good idea to smooth the image by using, say, an average filter with a window of about 15×15 pixels.

1.5 Extraction of the spectral orders

Extraction of the spectral orders from the échelle image consists of the following steps:

- Find each order position in the two-dimensional image plane.
- Identify the échelle order numbers on the image.
- Find the width of each order.
- Extract all the orders to get spectra of the object.
- Assign orders from the object image to the comparison spectrum image (optionally, to the flat field image) and extract comparison spectra.

In order to find spectral orders on the image we have to specify two parameters, which depend on the order structure and the noise level of the image: the length of a gaussian filter and a binning factor. The program works across the image row by row with the specified binning step size and calculates the average row in each step. The average row is filtered using a gaussian window of the specified length. top of the orders from zero background parts. Pixels which are above of the smoothed curve are considered as points belonging to a detected spectral order. A fit of a gaussian to the upper part of the order is used to find a reliable position of the order in the given row. So far, the program knows the centre of the orders in each binned row and the number of orders found. A two-dimensional fit is used to link X and Y-coordinates of the orders together. In dispersion direction we are using a polynomial of second degree and in cross-dispersion direction a polynomial of 5th degree. Finally, we thus obtain a fit which excludes all the random errors in the coordinates of the orders.

In practice, we have sometimes a much more complicated situation with particular échelle images.

- Some of the orders could be too weak to be detected by the algorithm described. The program identifies missing orders by interpolation/extrapolation; to do this it analyzes the coefficients of the 2nd order polynomials as functions of relative order number.
- Some artifacts on the image could cause a detection of an artificial order. In case of very weak images, the gaussian filter curve is quite low near the background where a higher noise can trigger a false order detection. The same could happen in case of the not completely removed cosmic spikes in the interorder space. In such cases, the program tries to exclude not reliable orders with very low accuracy of the polynomial fit along the X-coordinate by performing a series of tests.

Both effects can produce an error, or gap, in the relative counter of the orders and create a "squeezing" or "expanding" mesh of the orders displayed on the image. If the length of the gaussian filter is chosen correctly, changing of the binning factor and a new trial could help to avoid the problem.

The next step is to find the widths of the all orders found. To find the width, we calculate an average profile of the each order. The program goes along the polynomial curve at the top of each order row by row and sums up the rebinned cross-dispersion profile. The full width is defined as the distance from the upper to the lower wings of the average profile where the intensity falls down to 1% from its maximum.

The final step is extraction of the orders into one-dimensional spectra. All the orders found are integrated within their borders as described in the following.

Let i and k be indices (sequential number) of rows and columns on the image, respectively. Then, the shape of the order can be described as:

$$y_i = a + bx_i + cx_i^2$$

where y_i is the position of the order centre across the columns in world coordinates for a given row with coordinate x_i. The world coordinates are:

$$x_i = X_0 + i \cdot X_1, \ \ y_k = Y_0 + k \cdot Y_1$$

where X_0 and X_1 are the reference pixel value `CRVAL2` and increment value along second axis (CCD rows) `CDELT2` of the FITS image, respectively. Y_0 and Y_1 are the corresponding values `CRVAL1` and `CDELT1` for the first axis (CCD columns) of the FITS image, respectively. Hereafter, the reference pixel value is referring to the first pixel on the axis (denoted by `CRPIX=1`).

Let w_1 be the width of the upper wing of the order (as measured from the center of the order) and w_2 be that of the lower wing; then, the order boundaries in index scale are:

$$r_1 = \frac{y_i - w_1 - Y_0}{Y_1}, \ \ r_2 = \frac{y_i + w_2 - Y_0}{Y_1}$$

Integration of the order along the row i across the columns starting from r_1 to r_2 can be described as:

$$S_i = I_{k_1,i} \cdot (1 - f_1) + \sum_{k=k_1+1}^{k_2-1} I_{k,i} + I_{k_2,i} \cdot f_2$$

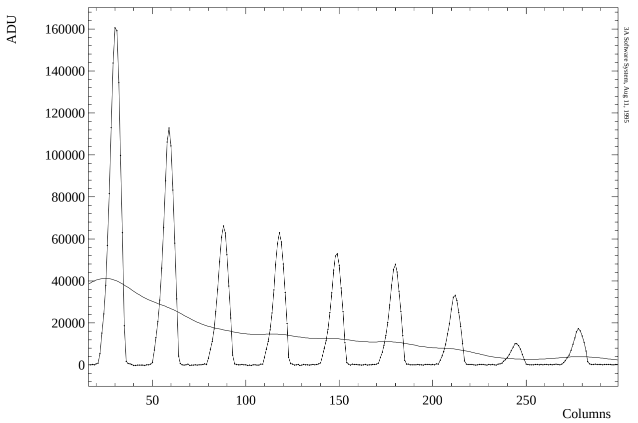

Figure 1.24: A gaussian filter with a window length of 33 pixels, applied to the sum of 10 cross-order rows in order to check separation of the upper parts of the spectral orders from the background area.

Here,

$$k_1 = \mathrm{int}(r_1) \text{ and } k_2 = \mathrm{int}(r_2)$$

are the integer parts of the indices of the first and the last *complete* columns, and

$$f_1 = \mathrm{frac}(r_1) \text{ and } f_2 = \mathrm{frac}(r_2)$$

are the fractional parts of the columns which are *partially inside* the order boundaries.

A summary of the orders extraction procedure:

- Display the image where the orders should be extracted and use the respective commands in the `Process` menu.
- Display a cross-dispersion cut with the number of rows set equal to the binning factor, which could be, as a first guess, around 10-20 rows for most of the cases.
- Select the `Find orders` command, setup the binning factor and start processing. The orders found will be displayed by curves along the top of the orders. If the picture is wrong, change the binning factor. Check the upper and the lower parts of the image by enlarging the areas and enhancing the colours by using the histogram facility to make the curves of the orders better visible. Information about orders' position is saved into the binary file with the same name as the image but with the extension `*.spp` (we will call it *SPP* file). Spectral orders found can be redisplayed by using the hotkey `F7`, if the file containing the order parameters is present in current directory.

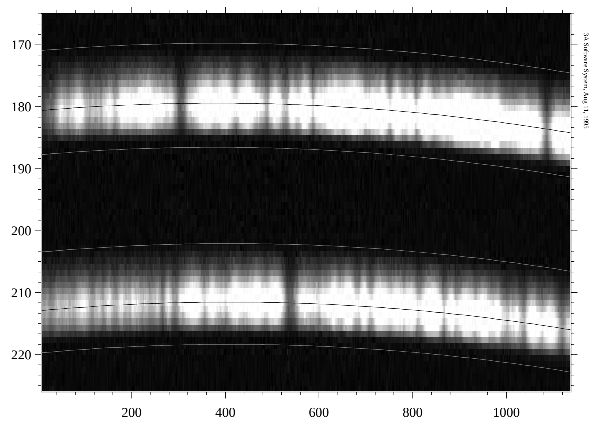

Figure 1.25: Enlarged area of the image with white dashed lines showing the borders of the orders. A central solid black line indicates the tops of the orders.

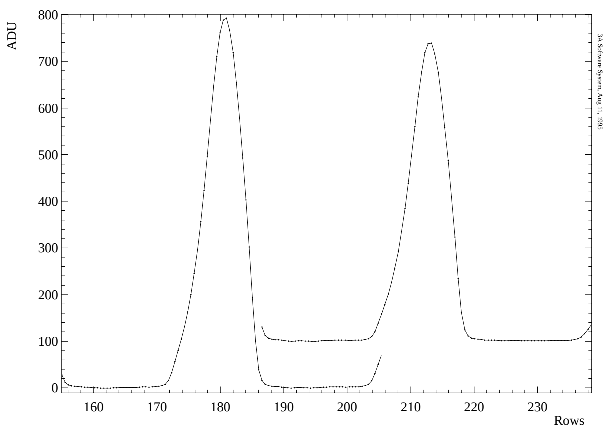

Figure 1.26: A few shifted average profiles taken from the image. Borders of the orders are not symmetrical with respect to the orders' maxima because of bad seeing/guiding during observations.

- Before the next step, we have to identify the échelle spectral order numbers we want to extract. So, we need to know the absolute échelle order number of the uppermost order found in the image and the direction in which the numbers are increasing. It is very important to specify the correct number, because these numbers will be used later for the two-dimensional dispersion curve procedure where the optical properties of the échelle are used. Furthermore, it is convenient to have the same enumeration of the same orders when observing in slightly different spectral regions (say, shifted in cross-dispersion direction by a few orders), i.e. to use the same absolute order number instead of different relative numbers for the same spectral order.

  In order to get the absolute échelle spectral order number, we can use the model of the focal plane of the spectrograph (called from the database of spectral regions) together with an identification of a few spectral lines on the comparison spectrum image or even using some features of the stellar image.

- Select the `Average orders` command and specify the absolute number of the uppermost échelle order found in the image. To give the program the information about how the order number is changing downwards, we have to give a *sign* to that order:

  Positive order number, if the order numbers are increasing downwards: this is the case for the $2nd$ and $3rd$ optical cameras of the SOFIN spectrograph. It means that the blue spectral orders are at the bottom of the échelle image.

  Negative order number, if the order numbers are decreasing downwards: this is the case for the $1st$ optical camera of the SOFIN spectrograph. It means that the red spectral orders are at the bottom of the échelle image.

  When processing the image, each average order is written into a separate FITS file and appended to the database of spectra. The file name of an average order is e.g. `37w05830.dat`, where the first two digits are the absolute order number, the letter means "*w*eight" function and the rest is the image file sequential number. The file name or its parts and extension are not changeable.

  When all the orders are extracted, the program displays the borders of the orders. Information about the orders' widths is also written into the *SPP* file.

  The following lines appear in the FITS header to be used later by the two-dimensional dispersion curve facility:

```
HISTORY  The order was extracted from CCD04701.FIT                      \\
ORDER   =                   33  / Spectral order number            \\
SPC0    =     70.805024526841  / Spectral order parabola:         \\
SPC1    =    -0.0214467894556  / Ycol=c0+Xrow*(c1+Xrow*c2)        \\
SPC2    = 2.0730447666694e-05                                     \\
SPFW    =     5.2206666255819  / Full width of the order in pixels  \\
```

  Changing these values (by editing the FITS header) does not have any effect, because the programs take the order parameters from the binary *SPP* file.

- Extract the orders by using the last command. The file name of the spectra is e.g. `37x05830.dat`, where the first two digits are the order number, the letter means "e*x*tracted" and the rest is the image file sequential number. The file name or its parts and extension are not changeable.

- Finally, assign the order structure to the comparison spectrum image. That can be done from the database of images by using the `Process/Assign orders` command as follows:

  1. Select the object image with the orders found and press `Enter`. To discard and exit, use `Esc`.
  2. Select the corresponding comparison spectrum image and press `Enter`. To discard and exit use `Esc`.

  The program creates a copy of the *SPP* file for the new image. The name of this *SPP* file is the same as the name of the image assigned. So, orders cannot be assigned to an image with the same name but a different extension.

  If the initial coordinates of the two images are different, then the parameters of the spectral orders are corrected for the offset. The initial coordinates of the object image `CRVAL1` and `CRVAL2` are always shifted compared to the comparison image after the trimming procedure, when a few first unwanted columns and rows were removed. If the second image boundaries are a subset of those of the first one, then outer orders will disappear.

- Display comparison spectrum image, display the orders and extract the comparison spectra. The initial order number will be taken from the *SPP* file belonging to the image.

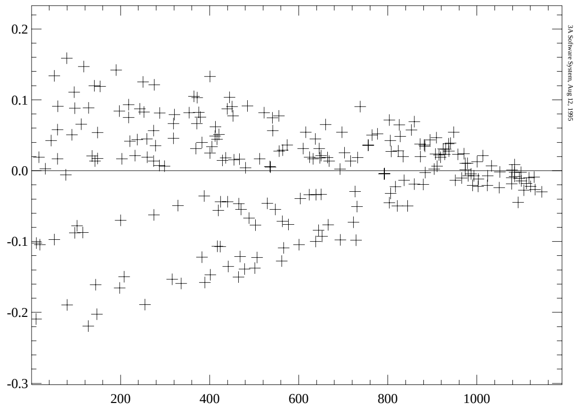

Figure 1.27: Residuals of the two-dimensional dispersion curve fit of 316 spectral lines taken from 10 different spectral orders across the comparison spectrum image obtained with the 3rd optical camera of the SOFIN spectrograph. The polynomial fit of second order was obtained with the assumption that the blaze line of the échelle is aligned to the CCD rows. Spectral lines from the reddest orders are located in the upper part of the picture and the bluest spectral lines are in the lower part.

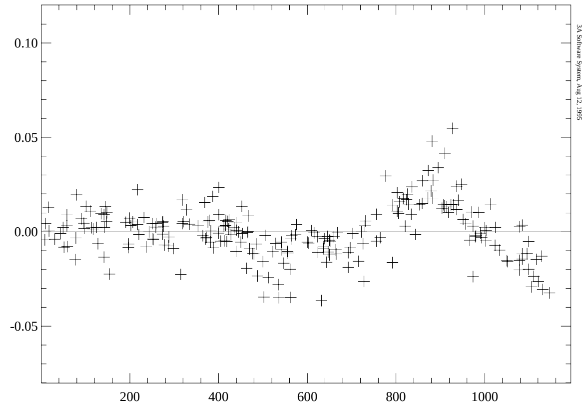

Figure 1.28: Residuals of the same lines after the distortion was excluded. The tilt of the blaze was found to be -899″ and the rate of change of the tilt, or distortion, across the image to be +1.73″/pixel. The root mean square deviation was reduced from $73m$Å to $15m$Å.

1.6 Wavelength calibration

- Locate, identify and fit spectral lines in some spectral order of the comparison spectrum.
- Make a two-dimensional dispersion curve fit based on one order to get *approximate* dispersion curves for the rest of the orders.
- Locate, pickup and fit spectral lines in several spectral orders across the image.
- Make a two-dimensional dispersion curve fit to get the *accurate* dispersion curve for each spectral order.

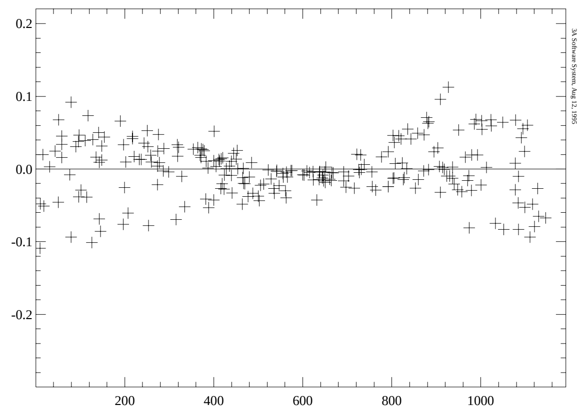

Figure 1.29: The behaviour of the residuals of a two-dimensional dispersion curve fit with different parameters. The distortion was set to zero, whereas the tilt of the blaze is set equal to the optimal value -899″.

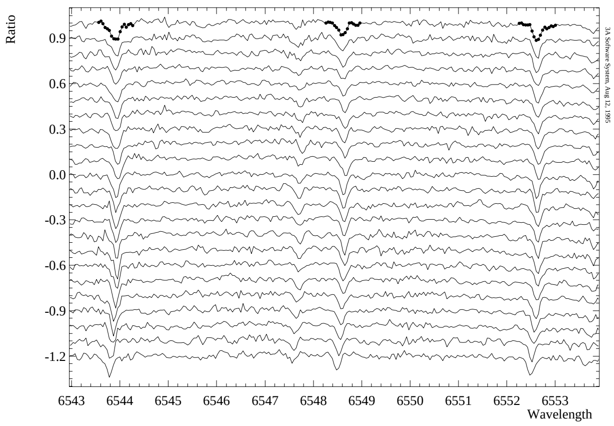

Figure 1.30: Fragment of the spectra with water vapour lines around Hα. A series of 23 observations of BW Vul were taken during two consecutive nights with the 2nd optical camera of the SOFIN spectrograph. The shift of the lines due to bending of the spectrograph at different zenith distances is clearly visible. Only the points marked around three strong water lines in the first spectrum (which was used as the reference spectrum) at the top of the plot were used during cross-correlation with all other spectra.

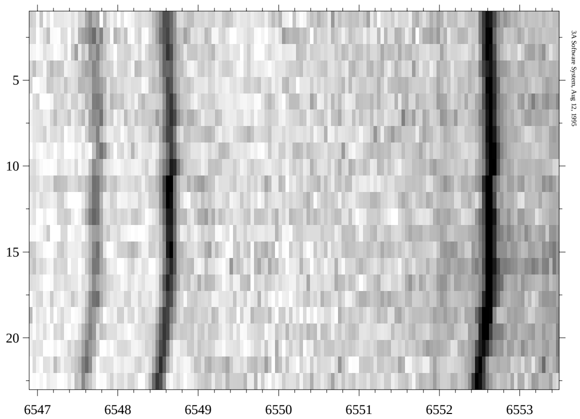

Figure 1.31: The same spectra stacked into an image. The vertical axis gives indices of the spectra.

1.7 Fine tuning of the wavelength scale

- Removal of the random relative shifts of the spectra in the reference order.
- Introducing the absolute offset into the spectra in the reference order.
- Applying the shift values to the spectra for all other spectral orders.

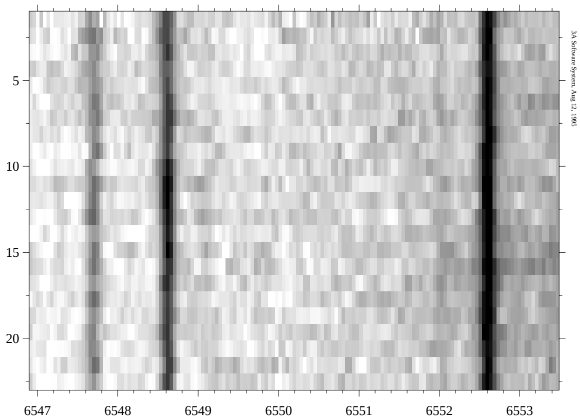

Figure 1.32: After correction for the random shift of the spectra. Still the whole picture is shifted in wavelengths by the value of the offset of the first reference spectrum.