Observations and data reduction

3.1-Observation modes

Two main modes of observations will be available:

1)Normal integration, with sky subtraction using the sky IFU, or with separate integration on the sky.
2)Nod-and-shuffle observations.

The first mode is convenient to take advantage of the full spatial coverage of the IFU, which is important for extended objects. It will be usually possible to place the sky IFU, situated at 5' from the main IFU, in a region free of object emission. The variability of the sky emission is taken into account, since the sky is observed simultaneously with the object. The idea, then, is to compute the median of the 25 sky spectra and subtract it from each 1300 object spectra. Differences in transmission from individual fibers can be compensated, using relative transmission factors obtained from flat field measurements. The accuracy of the sky subtraction is expected to be of the close to 1%, based on the experience of other groups. Our experience with the prototype is that the relative transmission of the fibers is very stable, even with strong curving of the fiber bundle.

When a mosaic of many exposures of an extended object is performed, it may be interesting to take one or two exposures of the sky with the main IFU. In this way direct subtraction can be performed, that takes into account in a simple way of differences in fiber transmission and wavelength shifts.

When very accurate sky subtraction is required, like for instance for faint objects, the nod-and-shuffle mode is probably better. In this mode, a mask is used in front of half of the lens array, and only a square region of 26x25 array elements will be used. Therefore on the CCD, groups of 25x3 lines corresponding to the observed region are alternated with groups of the same number of lines without signal. The telescope is switched between the object and a sky position, every 60 seconds, and at the same time, the charges are shuffled 75 lines on the CCD. The sky integration starts over the charges that where previously in the masked regions. After 60 seconds the charges are moved back to their previous positions and the telescope returns to the object. After 30 minutes of observations, the CCD contains 15 minutes of object integration and 15 minutes of sky observation that can be directly subtracted. The advantage is that for each spectrum, the sky is obtained with the same fiber and same pixels. This mode is useful when the sky is stable (no clouds).

3.2- Data Reduction

We intend to provide the means for doing quick look at the telescope and the data reduction software to reduce the data during the observations or in the next day.

The only step that is specific to SOAR-IFU reduction is the extraction of the fiber spectra from the initial CCD image. It is expected that everything else could be done using standard packages such as IRAF. As for looking at the final 3D data, there are also existing packages that can do this. We intend to develop a task similar to 'ldisplay' in the 'gmisc' package as a quick look tool.

In (3.2.1) we summarize the main steps of the data reduction procedure and in (3.2.2) we discussion of the problem of the superposition of spectra. This step of data reduction is described in more detail, as it poses new challenges.

3.2.1. Main steps
Considering that the acquisition system will provide the raw data in FITS.MEF, the steps of the reduction procedure can be summarized as follows:

a. The images are first processed for bias and dark subtraction; flat-field and fiber throughput correction; cosmic ray removal; sky subtraction (in the case of nod-and-shuffle).

b. Extracting the signal from each fiber, by using the multiple-Gaussian fitting code. It is possible that this step takes a relatively long time to be performed, and a simpler extraction could be provided for a quick view of the results.

c. Extraction of the spectra in the two dimensional images and their scattered light subtraction and extraction to one dimensional spectra in multispec format.

d. Wavelength calibration (using 2D spectrum calibration). In the case of sky subtraction using the sky IFU, it may be preferable to perform sky subtraction after this step, since small wavelength shifts may appear between different regions of the slit.

e. Writing reduced data on data cube format.
From the data cube, it will be straightforward to produce maps of line intensities, for a number of selected lines. For those who are interested in velocity maps (rotation curves of galaxies, kinematics of HII regions, etc), one possibility is to rely on the wavelength calibration already contained in the data cube. More sophisticated reduction, like cross-correlation with template spectra, will be tedious.

3.2.2. Deconvolution of overlapping spectra

In this sub-section we discuss step (b) above in more detail, describing the problem of the superposition of spectra. To solve this problem a solution that uses a code to fit multiple Gaussians is suggested.

We have developed a code based on the ROOT routines and applied it to theoretical curves in order to deconvolve multiple Gaussians. The procedure tries to fit simultaneously the observed profile of the whole set of fibers.

In Figure 2.1 the cut of a flat-field image obtained with the IFU-prototype is shown. This will illustrate how many pixels are contaminated by the signal of the other fibers.
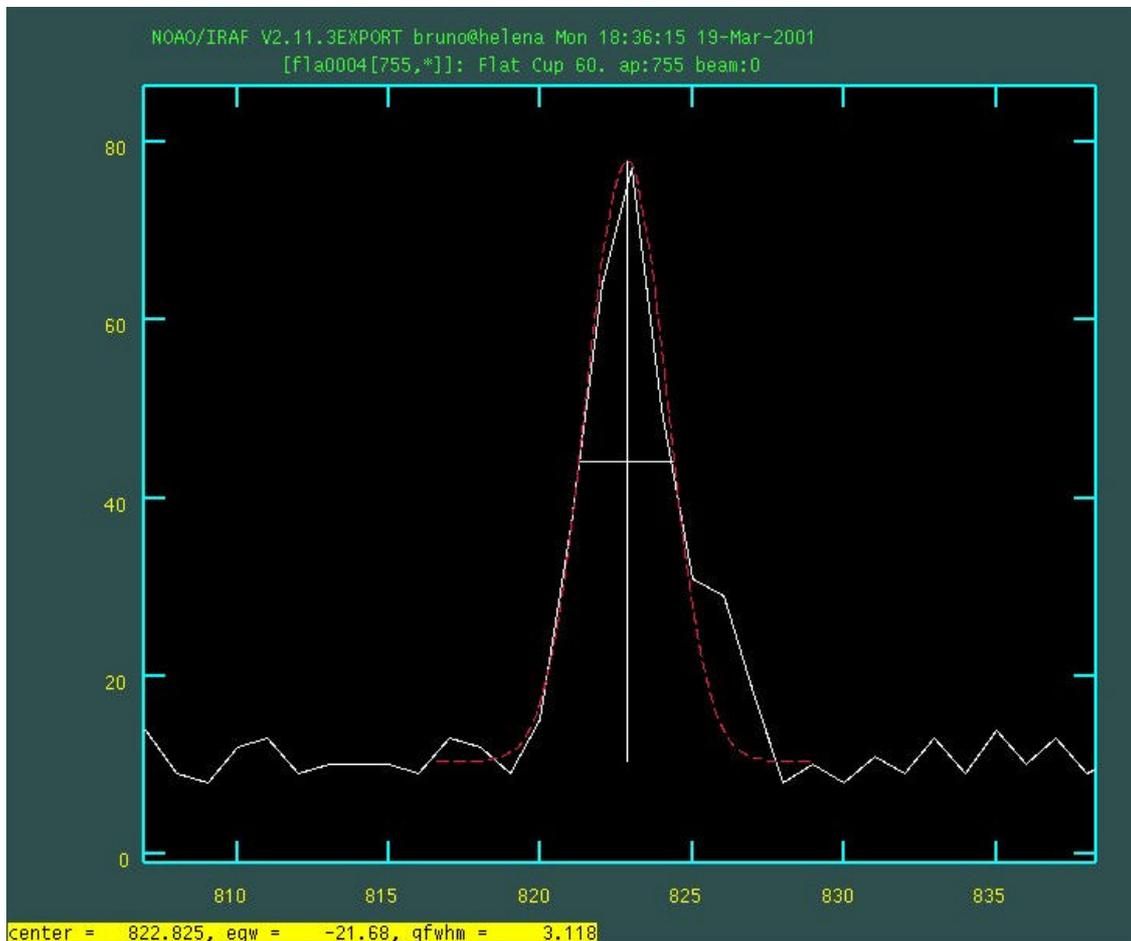

Figure 3.1. *The cut of a flat-field image obtained with the IFU-prototype at the OPD. A mask was use in order to assure that the signal detected by one fiber is not contaminated by the signal of the neighboring fibers. As the mask was not roughly located, a small contamination is detected in the right side of the profile. The FWHM is 3.12 pixels.*

The first step in the calculations is to determine the center and the FWHM of each fiber profile. In principle, it is expected that the signal of the fiber could be represented by a Gaussian with constant width, which is centered at a fixed position. When these parameters are known, the next step is to fit the real data (representative calibration data). In this case, only the amplitudes of the observed curves are fitted. The final step is to extract the observed signal of each fiber, removing the contamination from the other fibers. A numerical mask should be applied in the case of non-Gaussian profile, in order to correct for the spurious contributions. This could occur in the case of contamination due to the cylindrical curvature of the micro-lens, for example.

If the signal profile of the fibers does not vary appreciably by changing the instrument configuration (changes of gratings , grating angle and camera angle), once the main parameters have been determined, the observed data will be fitted in a fast way. Eventually, a number of files containing the parameters for different configurations could be necessary.

The first tests of the fit of multiple Gaussians are available in www.crux.fsc.ufsc.br/~kanaan/ifu, where the problem of dealing with the superposition of the spectrum of one fiber with the spectra of its neighbors is described. Some simulations are given to illustrate the solution to this superposition problem, which will be worst as the ratio fwhm/fiber-separation becomes larger.

The simplest case considers one single cut across the spatial dimension of the IFU-gram. The light from only a single fiber should pass through the spectrograph in order to have the signal intensity and shape (width and center of the Gaussian) measured without contamination from the other fibers. In principle, this procedure should be repeated as many times as the number of pixels in the dispersion direction; however, in practice, the parameters will probably vary linearly in the dispersion direction, and 3 to 5 measurements at different wavelengths will be sufficient to describe the width and center as a function of wavelength. Note that if we use a mask in the form of a slit illuminating a row of the lens array, we obtain 26 "individual" spectra at a time.

 The individual spectra will give us the width and center of the profile of each one of the fibers;. The second step is to fit the intensity of each fiber profile in the overlapping spectrum (calibration data) and then to apply the previously measured widths and centers. The more complex fit (intensity, width and center) is done first to the calibration data, which has more photons counting and do not suffer from superposition. In this way, the science data have to be fitted for a single parameter (times the number of fibers in one spatial cut). The fit to the science data will give us the intensity of each fiber at every wavelength, by repeating this procedure several times in the dispersion direction.

Probably the profile of the fibers on the CCD will not be a Gaussian. In order to approximate the profile into an analytic model (a Gaussian for instance) we will need to add a numerical model, by following, for example, the Stetson's model for stars in 'daophot' IRAF procedure.

In the next example the fit parameters obtained with 30 Gaussians are presented. There is no initial value as these Gaussians were not mathematically created but were the result of some real work at the AAO SPIRAL spectrograph.

It may be that we do not need to separate the fibers using the masks. The fit is almost perfect except for the two edges. This implies we must allow for some extra pixels at the CCD edges.

Figure 3.2. *The fit to the flat-field real data obtained during tests of the IFU-prototype at OPD. Blue represents the data and red represents the fit.*

3. 3 Real data from the IFU-prototype

The IFU-prototype has been installed at the OPD 1.6m telescope to test the spectrograph by using the 600 l/mm grating. Several objects were observed, and flat-field and arc comparison data were obtained. The next figures show some of the data, obtained in the February run. These results give us an idea about the alignment of the fibers, the rough profile of the signal, and the degree of the spectra overlapping in the CCD. The images are available in the IFU-SOAR home-page (www.lna.br/~bruno/IFU)
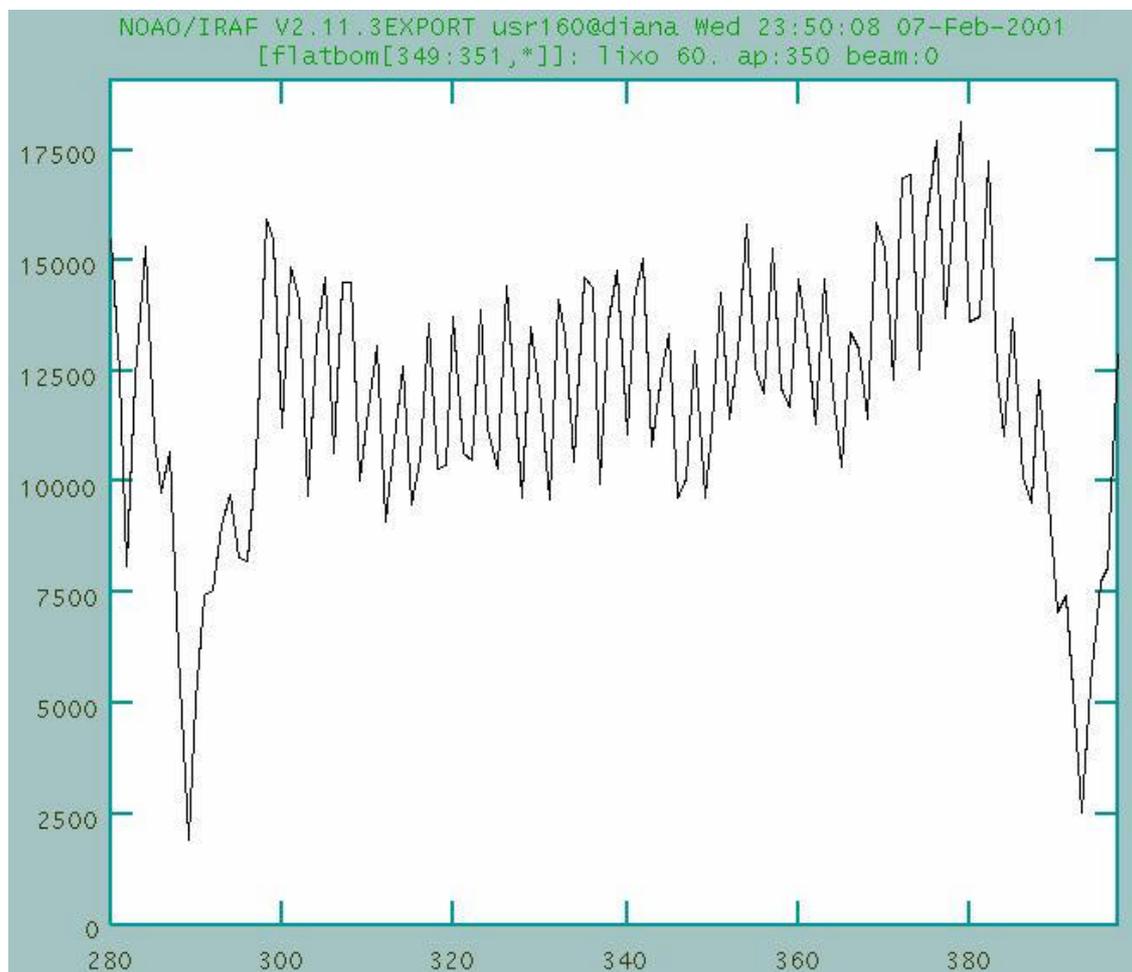

Figure 3. *Cut of a flat-field image obtained with the IFU-prototype, in the direction perpendicular to dispersion, illustrating the overlap of the signal of neighboring fibers.*
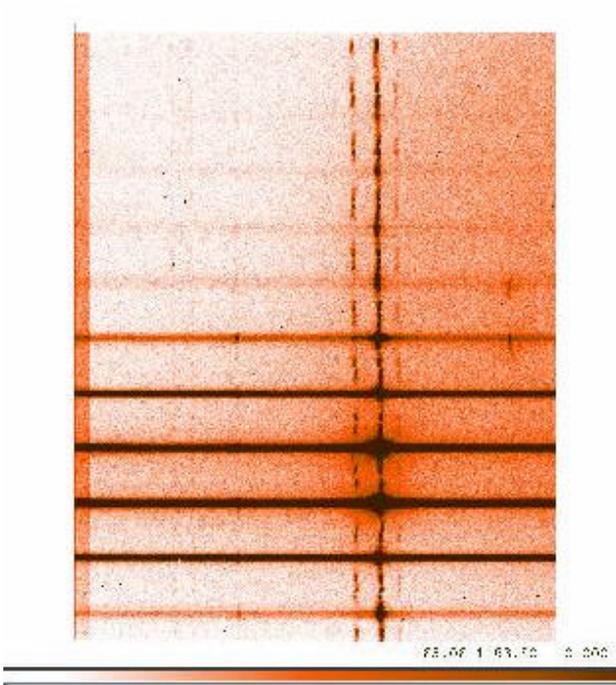
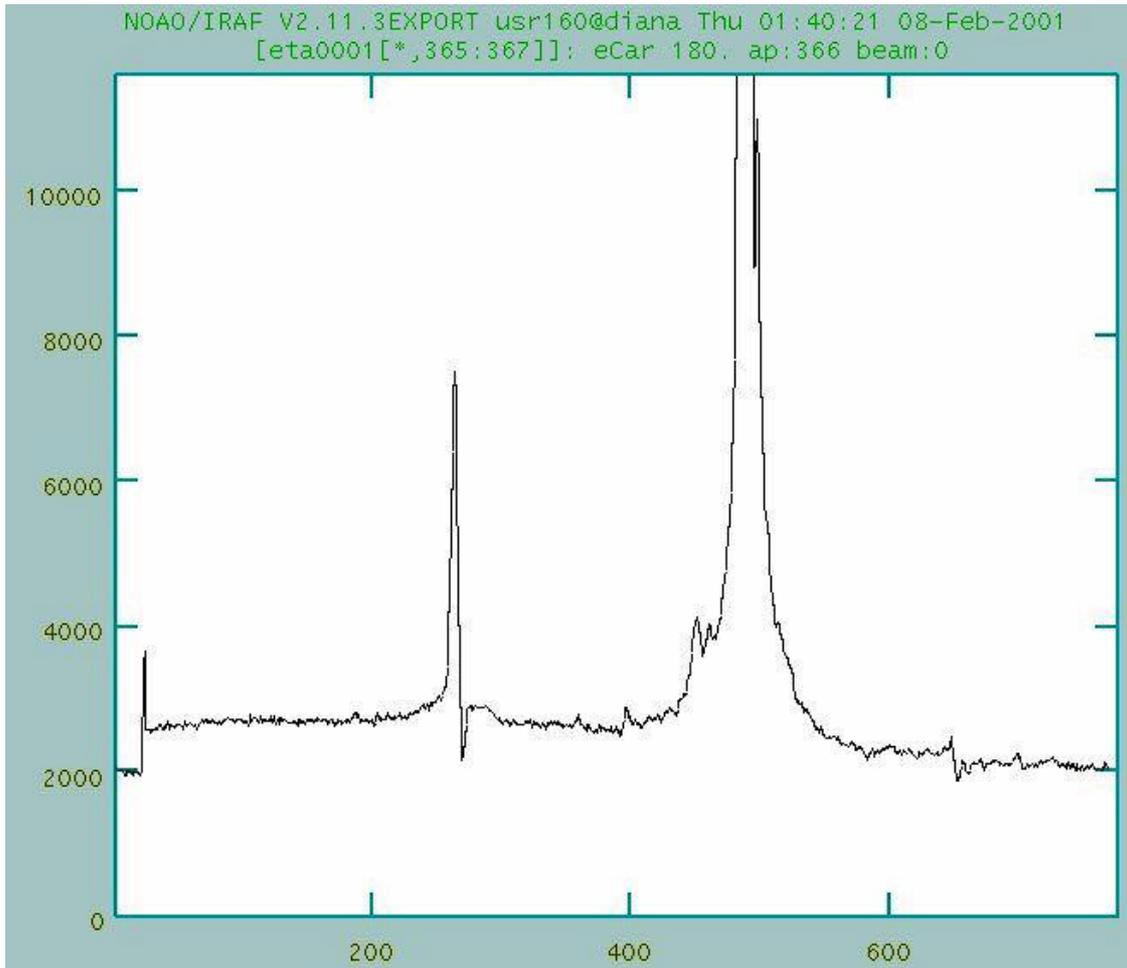Figure 3.4. The IFU image for the star AG Car

Figure 3.5. A cut of the image shown in Fig. 4 (AG Car spectrum)

3.4. Sharing the Gemini codes

We expect to share the codes to be used in the data reduction of the Gemini IFU spectrograph. We hope to share at least the common tasks that could be applied to the SOAR-IFU. We present here some information that we received from people we have contacted:

(Brian Miller - 4 Dec 2000)
The Gemini group developing these codes is in the process of writing IRAF scripts to reduce the GMOS data. These codes will be designed for multi-object rather than IFU spectroscopy. B. Miller suggests that a few modifications of the basic IRAF packages for spectra can do simple extractions. They intend to use 'ldisplay' in the 'gmisc' package as a quick look tool. Eventually they would have the option of doing "optimal" extraction that deconvolves overlapping spectra. This effort will be shared with the CIRPASS group. All Gemini data will be in multi-extension FITS files so the data will be

packaged a little differently than normal 'multispec'. Also, the IRAF group is supposed to be working on a new spectral format that will handle IFU data better.

(Frank Valdes - 2 Nov 2000)
Tools might be provided by IRAF to support the higher level tools being explored by the UK and Gemini people. According F. Valdes, extracting the spectra to some data format along with calibration would be related in some way with the IRAF 'apextract' tools. There are some issues about how to specify the initial data information and how to handle cases where the spectra overlap (merging of fiber profiles). Visualization and analysis is the type of thing that others are exploring as well as the discussion of intermediate and end point data formats. The IRAF project will be considering new data formats for the basic 1D-extracted spectra. This might then be turned into data cubes or other formats either as disk files or internal on-the-fly conversions.

The following has been adapted from the discussion list at the home-page http://www.astro.keele.ac.uk. Different ideas are presented.

The common standard for the data cube file will be a multi-extension FITS file. It is a single 3-D image with the axes in the order x/ra (pixels), y/dec (rows), lambda (planes) and the basic FITS keywords - CRVAL1, CD1_1, CRPIX1 etc. The logical extension is to have corresponding 3-D variance and data quality images in extra FITS extensions.

Some people do not think general software should constrain the order of the axes. There should be tools to rotate (tumble?) a data cube to any orientation to optimize the most common access. The software will recognize the axes by CTYPE keywords. However if someone insists the most obvious choice is the first two axes being spatial.

What they expect will happen with IRAF IFU packages is that extraction from the raw data will be a collection of 1D spectra with a position coordinate attached as well as some information about the size and shape of the aperture on the sky. Wavelength and flux calibration will proceed in 1D. Finally there will be a task that will build a data cube from the 1D spectra using the header information.

General software, as written by software developers such as Starlink and IRAF, should be able to find the primary data cube either in the primary HDU or as an extension. Users may provide only a simple primary HDU while more sophisticated systems may provide more complex FITS structures. But an image extension and a primary HDU are basically the same thing. In a pinch it is always possible to extract an image extension into a simple FITS image file. The IRAF syntax for specifying an extension would be nice to use: image[extno] or image[extname] or image [extname,extver] where extno is the extension position in the file (0 for primary), and extname/extver are the extension name and version following the image extension convention. In IRAF it is an error not to specify an extension if extensions exist but other software could adopt a default that no extension means the primary array.