

Gemini-South IFU Data Reduction User Manual

Elaine M. Snyder
Dept. of Physics & Astronomy
University of North Carolina at Chapel Hill
`emsnyder@live.unc.edu`

February 22, 2016

Contents

1	Introduction	4
1.1	What is an IFU and why would we want to use one?	4
1.2	Observing setups	4
1.3	Observing patterns	5
1.4	Files you'll need for reduction	6
1.5	How the data is Stored	6
1.5.1	The raw fits files	6
1.5.2	What is an MDF file?	6
2	Setting Up Your Workspace	9
3	Data Reduction Guide	11
3.1	Entering and Exiting the Pipeline	11
3.2	Organizing and Identifying Your Data	11
3.3	Reduction of the Flats	11
3.4	Reduction of the Arcs	13
3.5	Identifying and Removing Bad Pixels from the Flats and Arcs	14
3.6	Extracting the Arcs	14
3.7	Wavelength Calibration of the Arcs	14
3.7.1	Find the Wavelength Calibration	14
3.7.2	Apply the Calibration to the Arcs	15
4	Derivation of Kinematics	16
4.1	Deriving Velocity Dispersions	16
4.2	Deriving Rotation Curves	16
5	Timeline of Pipeline Updates	17

List of Tables

1.1 RESOLVE's Observation Setups	5
--	---

List of Figures

1.1	IFU Fibers and Data Format	7
1.2	Example of DS9 cube dialog box	8
1.3	Example of DS9 cube dialog box	8
3.1	Overscan Subtraction Example	12
3.2	Examples of the Fiber ID Step	13
3.3	Example Image of a Bad Fiber	14

Chapter 1

Introduction

Greetings, Gemini Multi-Object Spectrograph (GMOS) integral field unit (IFU) data reduction pipeline user! Before beginning your data reduction journey, it will be important to understand the structure of an IFU and the way the Gemini IFU data is stored in the fits files you'll be reducing. If you are familiar with this already, feel free to skip ahead. Most of this information can be found online on <https://www.gemini.edu>, but I've compiled it all here for easy reference and so that I may reference it in the context of the RESOLVE¹ survey's instrument setups.

1.1 What is an IFU and why would we want to use one?

The basic idea of any spectrograph is that incoming light from an object goes through a slit, is dispersed by a grating/prism/grism, and then focused onto your detector to create an image of your spectrum. Most of the time the slit is long and narrow ($\sim 0.5 - 2''$ typically), and most astronomers would align it along a galaxy's major axis. This image will thus have spectral information in the x direction and spatial information in the y direction. The GMOS IFU is different in that instead of having one slit aligned along a galaxy's major axis, you will have multiple fibers that cover the entire galaxy. These fibers act as "mini slits", and each one will be dispersed and refocused onto the detector. This means you receive spectral information from all parts of the galaxy, not just where a single slit falls, and therefore are getting x, y, and wavelength at once. Thus, the main benefit of using an IFU is that it enables the 3D spatial analysis of the kinematics, metallicities, or whatever else you may want to derive from your spectra.

In the case of the GMOS IFU setups for RESOLVE observations, you will have either 750 or 1500 hexagonal fibers (depending on which setup was used) that cover an on-sky area of $5'' \times 3.5''$ or $7'' \times 5''$. See the top portion of [Figure 1.1](#), which shows the positions of the fibers over an example galaxy. As [Figure 1.1](#) shows, the GMOS IFU is designed in such a way that not all of the fibers are centered on the object – either 250 or 500 of the fibers are positioned on a region of sky away from the target (but not too far away), which enables the sky level to be later subtracted during data reduction. The sky level can be affected by moon illumination, clouds, etc. so it is important to be able to subtract any extra sky light there may be. Note that sometimes in online documents, the fibers may be called "spaxels", which is short for "spectral pixels" – don't be confused by this hip IFU lingo.

1.2 Observing setups

As alluded to before, there are two main setups for RESOLVE's IFU observations. Similarly to our SOAR observations, there is a blue setup for absorption line galaxies and red setup for emission line galaxies. [Table 1.1](#) highlights the main differences between the two setups. To determine which galaxy is best for which setup, we ideally like to use the SOAR broad spectrum as a guide. Seeing strong emission in the broad spectrum will point to the emission line setup on Gemini being best, while no emission will point to the absorption line setup being preferred. However, since there is not always a SOAR (or even an SDSS) spectrum available, we often rely on their photometric colors as an indicator of whether emission or absorption is more likely.

¹<http://www.resolve.astro.unc.edu>

quantity	blue setup	red setup
wavelength range	4200 – 5600 Å	5500 – 6900 Å
used for	absorption line galaxies	emission line galaxies
central wavelengths	4850, 4900 Å	6300, 6400 Å
features of interest	H β , Mgb, Fe5270, Fe5335	H α , [NII], NaD
grating	B1200	B600
filter	none (open)	r-G0326
slits	1 slit	2 slits
field of view	3.5'' \times 5''	5'' \times 7''

Table 1.1: Comparison of RESOLVE’s red and blue observation setups for the Gemini-South IFU.

Using the GMOS IFU is ideal for RESOLVE’s purposes for two different reasons. In the red setup, our goal is to obtain velocity fields for RESOLVE’s smallest emission line galaxies. Thus, the spatial information we receive from the IFU spectra is ideal for this purpose. On the other hand, for the blue setup, our goal is to derive velocity dispersions. The small, absorption line galaxies we target with the IFU in the blue setup are often quite faint, so instead of focusing on the spatial information the IFU gives, we sum all the spaxels out to the effective radius of the galaxy, thereby summing the light enough to derive accurate dispersion measurements.

We estimate exposure times for both setups using the online Integration Time Calculators² (ITC) supplied by Gemini. We use an elliptical galaxy SED template for time estimates in the blue setup and adjust our exposure time calculations so that we may achieve a signal-to-noise ratio of 25 per Å when we bin all the fibers out to the effective radius of the galaxy. We use a model H α line with estimates of the line flux and continuum flux density for time estimates in the red setup, and adjust our exposure times to achieve a centroiding accuracy of $\sim 5.3 \text{ km s}^{-1}$. For both setups, we bin our spectra by 2 in the spectral direction to increase the signal-to-noise as well.

We also use the ITC to select central wavelength positions that ensure our features of interest do not fall in a chip gap (see bottom of [Figure 1.1](#)). We choose two different central wavelengths so that we can fill in the chip gaps as well when we sum the individual exposures – this is sometimes referred to as “spectral dithering”. The central wavelength positions selected for our setups are listed in [Table 1.1](#).

1.3 Observing patterns

In 2-slit (aka red setup) mode, we oftentimes will tile the IFU field of view in order to cover the entire extent of the galaxy. The 5'' \times 7'' field of view can either be tiled to 10'' \times 7'' for larger, rounder galaxies or to 5'' \times 14'' for longer, skinnier galaxies. This doesn’t make a large difference to the reduction routine, but there will be more science frames to reduce and more data cubes to stack at the end.

The basic observation sequence is different for each mode. Both will start with acquisition images to ensure the IFU field of view is centered on the galaxy. Then flats, arcs, and science frames are taken in each dither. For example, a flat, arc, science, and arc frame will be taken at one central wavelength, and then an arc, science, arc, flat frame will be taken at the other central wavelength, in that order. For reduction, we will want to flat field and wavelength transform the science images using the arcs and flats taken in the same dither. Once the data cubes are made, we can then sum the data cubes, both spatially and spectrally.

²<http://www.gemini.edu/node/10479>

1.4 Files you’ll need for reduction

The fits files you’ll want for your data reduction can all be found online in the Gemini Science Archive (<http://archive.gemini.edu>). You will have to create an account with your unique observing program ID and program key (usually sent in an email for each new semester). RESOLVE’s IDs are GS-2013B-Q-51 for the fall 2013 semester, GS-2014B-Q-13 and GS-2014B-Q-52 for the fall 2014 semester, and GS-2015B-C-1 (and carry-over GS-2014B-Q-13 band 1 time) for the fall 2015 semester (email Elaine or Sheila³, if you need the program keys for any of these programs). Once your account is created, you can look at each observing program and download the data you want.

As alluded to § 1.3, you’ll have flats, arcs, and science data for each galaxy in each wavelength dither. Normally, there are 2 flats, 2 or 4 arcs, and 2 or 4 science frames, depending on the estimated exposure times and/or the spatial offsets needed.

We also obtain a suite of baseline standards each semester – one standard for both the red and blue setups that will be used for each galaxy observation taken that semester. This includes a science observation of standard star along with flats, arcs, and twilight flats in both wavelength dithers. We use the twilight flats to create response functions for each fiber in the IFU, and use the standard star science observations to correctly flux calibrate our galaxy data.

Lastly, we’ll want a bias image in order to perform the bias subtraction. Each night after observing is complete, a standard set of bias frames are observed by the Gemini South operators. They are a little tricky to find in the online database, but you can go to the “View Associated Calibrations” tab to find all of the biases taken. Usually, searching on the observation date, the correct binning (CCDSUM 2 1), and the correct filter (see Table 1.1) will yield 5 bias frames to download and process with the Gemini IRAF routine, `gbias`. I have downloaded these already for the RESOLVE data, and have trimmed, overscan subtracted, and stacked them.

1.5 How the data is Stored

1.5.1 The raw fits files

1.5.2 What is an MDF file?

³[sheila\[at\]physics\[dot\]unc\[dot\]edu](mailto:sheila[at]physics[dot]unc[dot]edu)

GMOS IFU Example Data: NGC 221

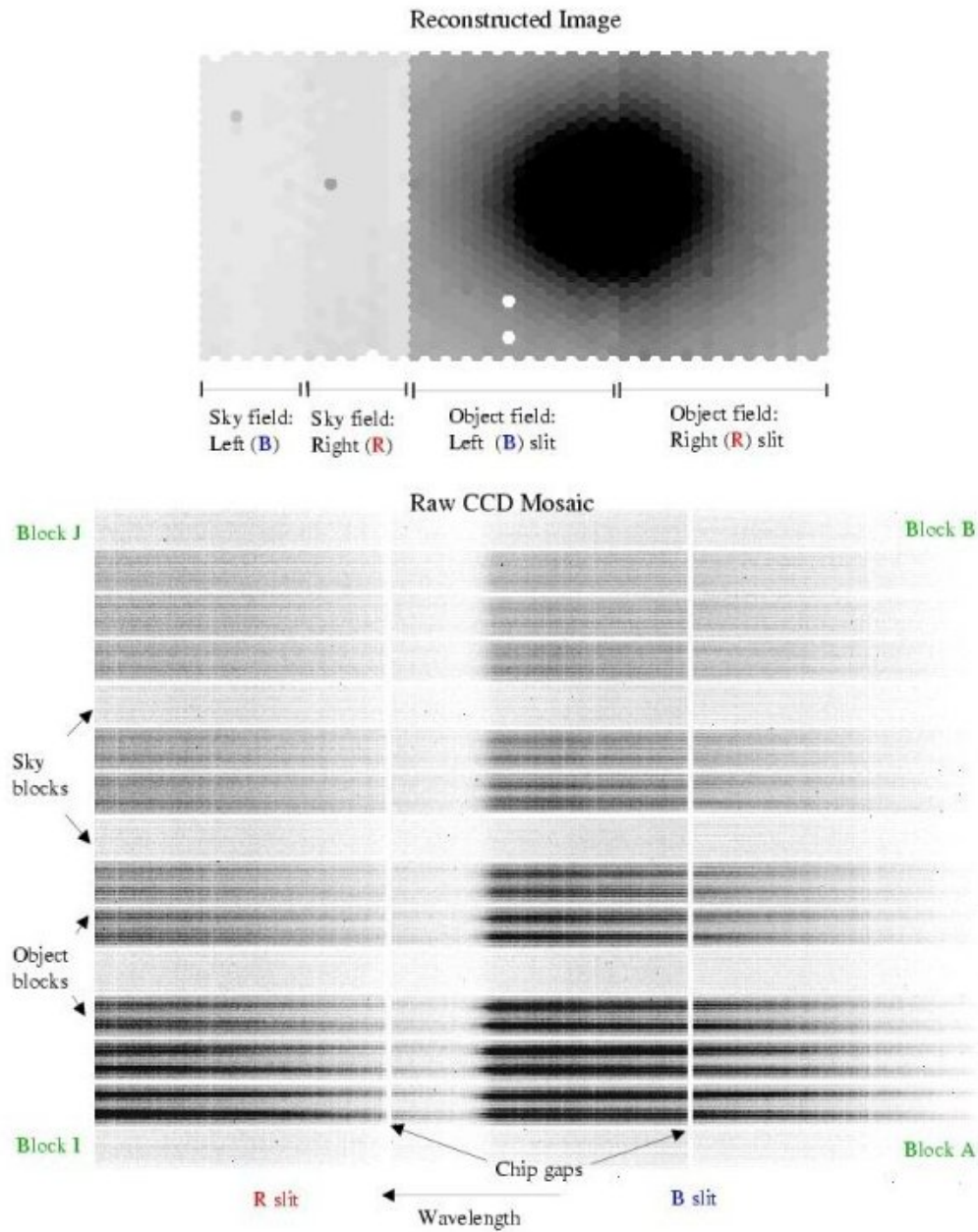


Figure 1.1: The top part of this image shows the IFU field of view over NGC221 (not a part of RESOLVE). The left portion of the image shows the sky spaxels, and the right portion shows the light of the galaxy. This is a 2D representation of the data we are getting – the third dimension is in the wavelength direction, and the darkness over where the galaxy is from summing the light inside each spaxel there. The bottom portion of this image shows the raw fits data from the above galaxies. Image from <https://www.gemini.edu/sciops/instruments/gmos/spectroscopy/integral-field-spectroscopy/field-slit-mapping>

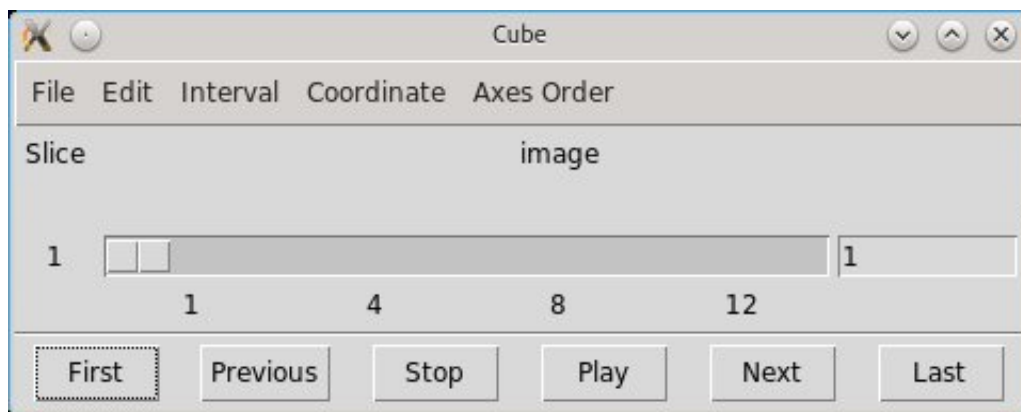


Figure 1.2



Figure 1.3

Setting Up Your Workspace

```
> pyraf
```

This is the EXPORT version of IRAF V2.16 supporting PC systems.

```
--> gemini
```

```

+----- Gemini IRAF Package -----+
|      Version 1.13.1, December 7, 2015      |
|      Requires IRAF v2.14.1 or greater        |
|      Tested with Ureka IRAF v2.16           |
|      Gemini Observatory, Hilo, Hawaii         |
|      Please use the help desk for submission of questions |
|      http://www.gemini.edu/sciops/helpdesk/helpdeskIndex.html |
+-----+

```

You may also need to download L.A.Cosmic, an algorithm that finds and removes cosmic rays. You can read more about it at <http://www.astro.yale.edu/dokkum/lacosmic/>. You can download the IRAF spectroscopic version from that website, and will need to add a line in your login.cl: “task lacos_spec = /path/to/where/you/put/it/lacos_spec.cl”. The pipeline will call this task during the cosmic ray removal step.

9

If you are new to data reduction in general, you will want to download DS9, an astronomical imaging application that you'll use to view our data as you progress through the pipeline. Download DS9 from this site: <http://ds9.si.edu/site/Home.html>.

Next, you'll want data to actually reduce! You should have flats, arcs, science frames, and a bias image before getting started. You should also have response flats from the baseline twilights and a flux calibration from the standard star data. See § 1.4 for more info on how to get your data and on each specific file you'll need. There should also be an observing log to inspect (RESOLVE data has `obslog.txt` or just `log.txt`) as well as an alternative line list file “`smalllinelist.dat`” that will be helpful during the wavelength calibration steps. Place all of these files in the same folder, since that's where the pipeline looks for them.

Lastly, you'll want to download the actual pipeline. You can access the code and this handy user guide at <https://github.com/emsnyder/geminiDRpipeline>. Always be on the lookout for updates in the future! You can place the code anywhere you like, but a nice folder structure would be to have individual folders for each galaxy with the pipeline code a level above. Now you should be all set, congrats!

Chapter 3

Data Reduction Guide

3.1 Entering and Exiting the Pipeline

Here are the basic start up commands:

1. cd to the directory where you put your data
2. enter the Ureka environment by typing `ur_setup` in the command line
3. enter PyRAF by typing `pyraf` in the command line
4. open a DS9 window by typing `!ds9 &` at the prompt
5. enter the pipeline by typing `execfile('/path/to/your/code/gemreductionpipeline.py')`

To end your session:

1. type `CTRL-C` to exit the pipeline, if not out already
2. enter `.exit` to exit PyRAF
3. use `ur_forget` to exit the Ureka environment

3.2 Organizing and Identifying Your Data

The first thing the pipeline does is look in your folder and identify which files are there. It will print out the files it found and whether there are arcs/flats/science/etc. It is good to check this against your observing log! If you find mistakes, please email Elaine, since the pipeline doesn't have a way to let you manually fix IDs yet. Every files is identified by looking in the image headers, so problems here most likely will be missing images. If these all look correct, type `1` to go on.

3.3 Reduction of the Flats

This very first step is often (in the author's experience) the most time-consuming and frustrating step of the entire reduction. We use the flats not only for flat fielding the science data, but also to identify the fibers as a reference for all the other files.

The package used here is a main part of the Gemini IRAF package and is called 'gfreduce'. It has many different options, but for the flat, we are just going to attach the MDF file, do a bias subtraction, overscan subtraction, and fiber ID so that we may extract all the spectra.

1. The MDF attachment is done without user input needed. A file is produced in your working directory with the prefix 'g'.
2. The bias and overscan subtraction is mostly done without user input needed. For the overscan subtraction, a PyRAF window will open with the overscan level being plotted. There is a also fit to the overscan level being plotting, and this normally looks fine as is. Press `q` inside the PyRAF window to move on. There is a plot for each amplifier in the CCDs, so you will see 12 individual plots. If the fit does not look fine, you can type `:order #` and then `f` inside the PyRAF window to change the order of the fit to a different number (with that number instead of `#` in the previous command). Higher orders typically do better at fitting any strange features. Unless something is very wrong, you probably won't need to worry about this. See Figure [Figure 3.1](#) for an example of what you should be seeing. This step produces a file with the prefix 'rg', where the 'g' comes from the previous step, i.e. the 'rg' is both overscan and bias subtracted and has the MDF attached.

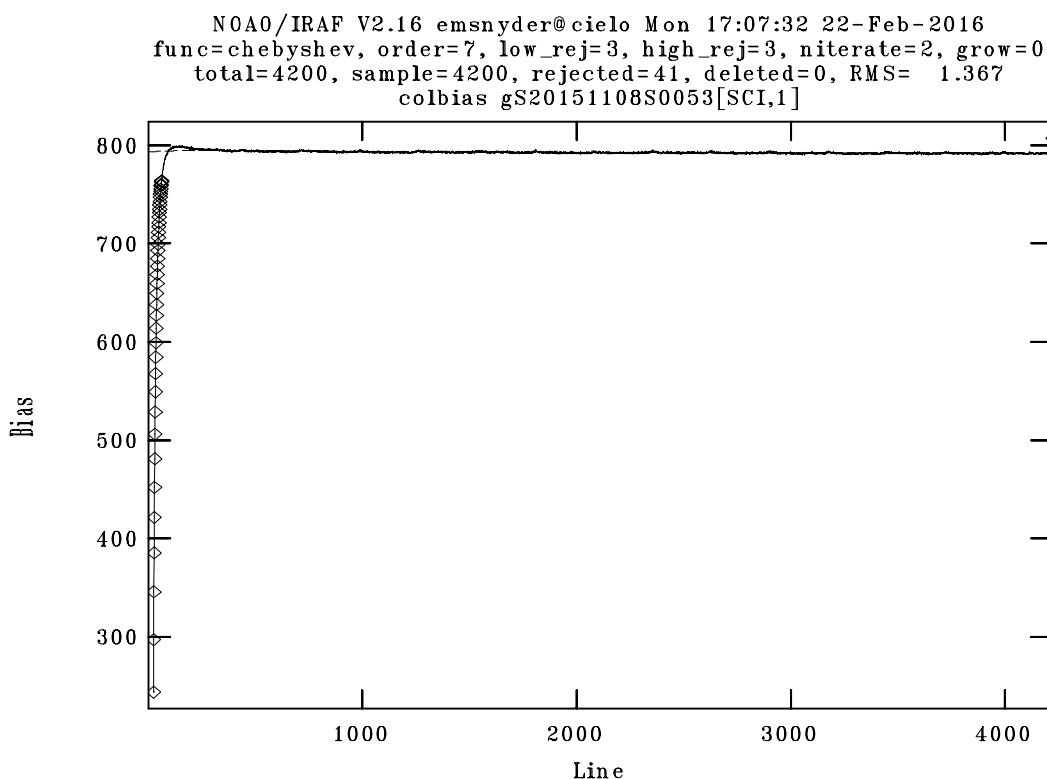


Figure 3.1: An example of the overscan subtraction window that PyRAF will open during the reduction of the flats step. The solid line represents the data, while the dashed line shows the fit to that data. The diamonds are data points that are being rejected from our fit.

3. Lastly, we must tackle the fiber identification. Your command window is going to ask you a few questions:

Extracting slit 1

Find apertures for ergS20151108S0053_1? ('yes'):

Here I recommend saying 'yes', as usually IRAF will be able to identify the fibers correctly for you.

Edit apertures for ergS20151108S0053_1? ('yes'):

Say 'yes' to this one as well, so we can inspect the fiber IDs and make sure everything looks correct. Figure 3.2 shows what this will look like – a complete mess! But we can zoom in by typing **w** in the PyRAF window, and then typing **e** at the bottom-left of the first 'fiber bundle' and again at its top-right. [Note: use **w** and **a** to un-zoom the window.] You should now have a screen that looks like Figure 3.2. I find that making this plot full-screen helps immensely. Each 'bundle' contains a set of 50 fibers, which you will see numbered at the top of your screen (though not very apparent in Figure 3.2). Your job is to make sure each fiber is identified correctly – this means no repeated numbers or missing fibers. Often I find that the low level of fiber 50 makes it get skipped in IRAF's auto-finding scheme, which makes the first fiber of the second 'bundle' start at 50 instead of 51. If this happens, you can use **d** in the plot window to delete the IDs for fibers 50–750, and then use **m** to re-mark the fibers. This is where it can get frustrating!

Figure 3.3 shows an example of a bad fiber not being marked in the identification process. This is exactly what we want to happen in this case. If you find another fiber that looks bad, you

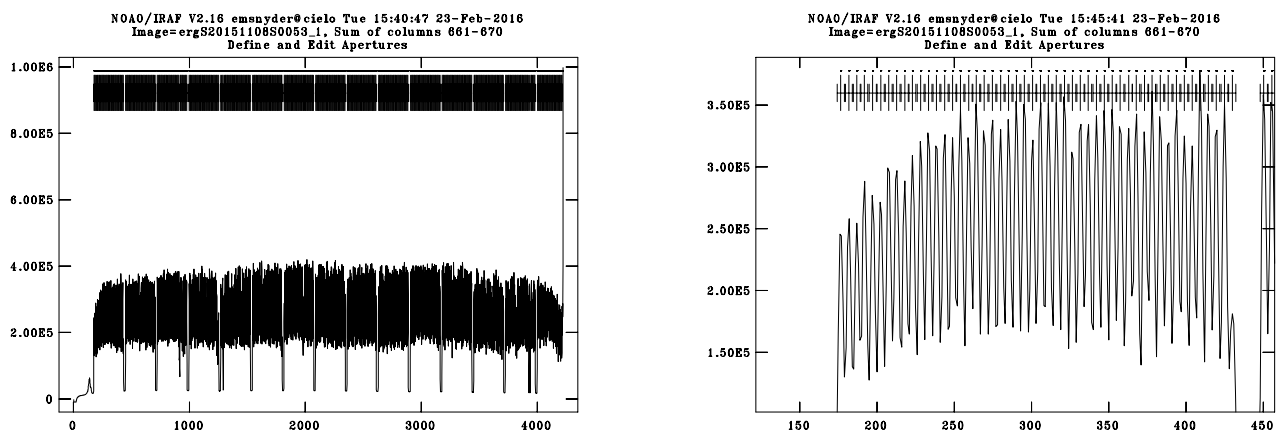


Figure 3.2: **(a)**. A zoomed-out example of the fiber ID window that PyRAF will open during the reduction of the flats step. **(b)**. An example of the fiber ID window that is zoomed-in on one set of fiber bundles.

will have to edit the MDF file to ‘turn off’ that fiber number. For the 2014-2015B slit-1 data, you will end with either fiber 742 or 743 (the rest up to 750 are bad).

If you are pleased with your IDs, type `q`. More questions will appear in the terminal, but it will want your answer inside the PyRAF graphics window.

Trace apertures for ergS20151108S0053_1?

yes

Fit traced positions for ergS20151108S0053_1 interactively?

NO (the capitals will tell IRAF no to all)

Write apertures for ergS20151108S0053_1 to database

yes

Extract aperture spectra for ergS20151108S0053_1?

yes

Review extracted spectra from ergS20151108S0053_1?

NO

If you’re in the blue setup (1-slit mode), this is the end of IDing for you! If you’re in the red setup (2-slit mode), you’ll have to repeat this process for fibers 751–1500 (usually fibers 1493 and above are missing for this group). Once complete, IRAF will extract all the spectra and create a file with the prefix ‘e’, so in full you should have an ‘ergS....fits’ file in your working directory. Since we have two different wavelength dithers, you’ll have to repeat this process for the second flat, but after that, we’ll use these files to identify the fibers for the rest of the bunch.

3.4 Reduction of the Arcs

The next step is to add the MDF and overscan subtract the arcs. A bias subtraction isn’t needed here since the arcs are read out “faster” than the bias frames. There is nothing interactive for this step.

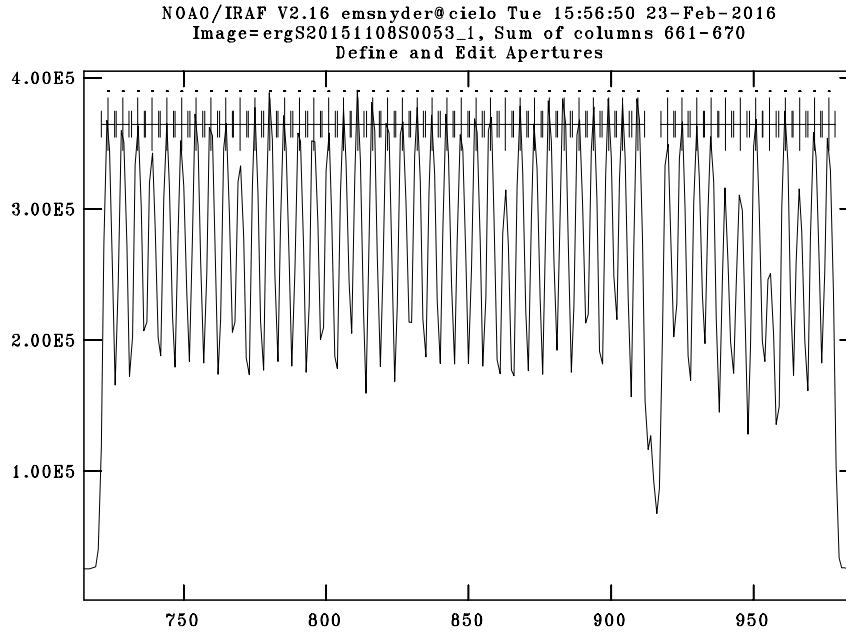


Figure 3.3: For this image, fiber 138 is bad. The MDF knows this though, and that number is skipped during the identification routine.

3.5 Identifying and Removing Bad Pixels from the Flats and Arcs

Before we get too into the reduction, we must look at our bias-subtracted images and find any hot pixels that could later affect the reduction. The pixels affected could be different for images taken at different exposure times, so we'll need to create a bad pixel map (BPM) for the flats, arc, and science frames separately (though the separate dithers can use the same BPM).

The pipeline will open a DS9 window with one of the flats first. It will then prompt you to look through all twelve extensions and identify any bad columns or other sections you see. If you don't see any bad pixels, you can enter 0 and move on to the arcs. If you do see some bad areas, enter 1, which will bring up another prompt for you to enter the extension number, and x and y values of the bad region. The extension number will be 1-12, and x and y can extend to the size of your image. To enter a bad rectangle, you can say x = 132:140 and y = 444:450. For a full column, you can enter *.

It is important to discern between cosmic rays and bad pixels too. I usually find that cosmic rays will be curved and erratic looking, while actual bad pixels will be in straight lines and will be much larger. See figure ## for examples. You will repeat this for an arc image, and later in the reduction will do this for the science frames as well.

3.6 Extracting the Arcs

This step will use the fiber IDs you created from the flat images to extract the spectra from the arcs. There is nothing interactive for this step.

3.7 Wavelength Calibration of the Arcs

This step is twofold: first we must find the wavelength solution for each arc lamp, and second, we must actually apply it to the arc as a check.

3.7.1 Find the Wavelength Calibration

Now we are going to

Image Data	Found	Fit	Pix Shift	User Shift	Z Shift	RMS
------------	-------	-----	-----------	------------	---------	-----

ergS20151108S0106_001 - Ap 440	10/10	10/10	-0.00557	0.00522	1.82E-6	0.0749
Fit dispersion function interactively? (no yes NO YES) ('no'):						
ergS20151108S0106_001 - Ap 441	10/10	10/10	-0.0793	0.081	1.36E-5	0.0731
Fit dispersion function interactively? (no yes NO YES) ('no'):						
ergS20151108S0106_001 - Ap 442	10/10	10/10	-0.0694	0.0711	1.12E-5	0.034
Fit dispersion function interactively? (no yes NO YES) ('no'):						
ergS20151108S0106_001 - Ap 443	10/10	10/10	-0.00833	0.00873	9.11E-7	0.0476
Fit dispersion function interactively? (no yes NO YES) ('no'):						
ergS20151108S0106_001 - Ap 444	10/10	10/10	0.0222	-0.0225	-4.0E-6	0.07
Fit dispersion function interactively? (no yes NO YES) ('no'):						
ergS20151108S0106_001 - Ap 445	10/10	10/10	0.0256	-0.0259	-4.7E-6	0.0896
Fit dispersion function interactively? (no yes NO YES) ('no'):						
ergS20151108S0106_001 - Ap 446	10/10	10/10	0.084	-0.0855	-1.4E-5	0.126
Fit dispersion function interactively? (no yes NO YES) ('no'): yes						

3.7.2 Apply the Calibration to the Arcs

Chapter 4

Derivation of Kinematics

4.1 Deriving Velocity Dispersions

4.2 Deriving Rotation Curves

Chapter 5

Timeline of Pipeline Updates