Building the Pipeline for Hubble Legacy Archive Grism data PHLAG

Martin Kümmel[1], R. Albrecht[1], R. Fosbury[1], W. Freudling[1], J. Haase[1], R.N. Hook[1], H. Kuntschner[1], M. Lombardi[1], A. Micol[2,1], M. Rosa[1], F. Stoehr[1], J.R. Walsh[1]
(1) Space Telescope - European Coordinating Facility,
Karl-Schwarzschild-Str. 2, D-85748 Garching b. München, Germany
(2) European Space Astronomy Centre,
P.O. Box - Apdo. de correos 50727, 28080 Madrid, Spain

Abstract

The Pipeline for Hubble Legacy Archive Grism data (**PHLAG**) is currently being developed as an end-to-end pipeline for the Hubble Legacy Archive (**HLA**). The inputs to **PHLAG** are slitless spectroscopic HST data with only the basic calibration applied; the outputs are fully calibrated, Virtually Observatory-compatible spectra, which will be made available through a static **HLA**-archive. We give an overview of the various aspects of **PHLAG**. The pipeline consists of several subcomponents – data preparation, data retrieval, image combination, object detection, spectral extraction using the aXe software, quality control – which is discussed in detail. As a pilot project, **PHLAG** is currently being applied to **NICMOS G141** grism data. Examples of **G141** spectra - data and metadata - reduced with **PHLAG** are shown.

Channel	Disperser	Wavelength Range [Å]	Resolution [Å/pixel]	FOV ["]
ACS/WFC	G800L	5500 – 10500	38.5	202x202
ACS/HRC	G800L	5500 – 10500	23.5	29x26
ACS/HRC	PR200L	1600 – 3900	20 [@2500Å]	29x26
ACS/SBC	PR130L	1250 – 1800	7 [@1500Å]	35x31
ACS/SBC	PR110L	1150 – 1800	10 [@1500Å]	35x31
NIC3	G141	11000 – 19000	80	51x51

Table 1: Some HST slitless modes under consideration for HLA activities

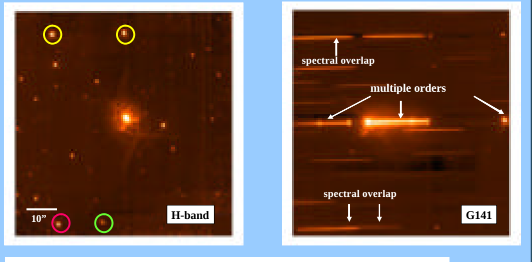

Figure 1: A basic dataset of slitless spectroscopic images (right) and their corresponding direct H-band (**F160W**) images (left)

1. The Hubble Legacy Archive

The current Hubble Space Telescope (**HST**) data archives at the **ST-ECF**, the Canadian Astronomy Data Centre (**CADC**) and the Space Telescope Science Intstitute (**STScI**) offer the images with calibrated pixel values for download. Further processing, such as the coadding of single images, must be done by the archive researcher individually.

After almost 17 years of **HST** observations, a large amount of expertise has been accumulated for all past and currently active **HST** instruments, and it is the right time now to start building a Hubble Legacy Archive (**HLA**) containing high level science data products for immediate scientific use [1].

While the **CADC** and the **STScI** concentrate on providing image products (e.g. stamp images and cutout services), the contribution of the ST-ECF to the HLA focuses on the area of slitless spectroscopy. Currently the archives deliver slitless data with only basic calibration applied (bias, dark), hence the improvements of an **HLA** with science ready spectra is large.

To generate high level science data products as input to the **HLA,** the Pipeline for Hubble Legacy Archive Grism (PHLAG) data was developed. Table 1 lists slitless **HST** modes which could be reduced for the **HLA**. In a pilot project, **PHLAG** is being developed and first applied to **NICMOS G141** slitless spectroscopic data.

2. Slitless spectroscopy

Slitless spectroscopic images show some special features and properties, which are depicted in Figure 1. The right and the left panels show a basic dataset consisting of slitless images and their corresponding direct images taken in the H-band (filter **F160W**).

Due to the low spectral resolution, there are multiple spectral orders of the same object visible in a single Field of View.

The absence of slits makes contamination, which is the mutual overlap of spectra, an ubiquitous phenomenon. There is contamination in the spatial direction, as can be seen in the dispersed image for the two objects in the green circle in Fig. 1, contamination in the dispersion direction with the object in the pink circle, and even contamination across different spectral orders, as it is the case for the two objects in the yellow circles.

The zero point of the wavelength calibration must be derived from the object positions on the direct image. Therefore a slitless spectroscopic dataset consists always of a set of direct images and dispersed images, which were taken at the same positions on the sky, and the processing of the direct images is an integral part of any slitless spectroscopy pipeline such as **PHLAG** [2].

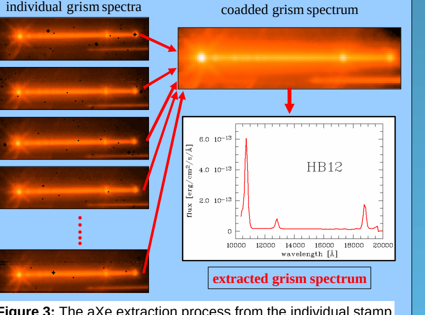

Figure 3: The aXe extraction process from the individual stamp images to extracted, deep spectra

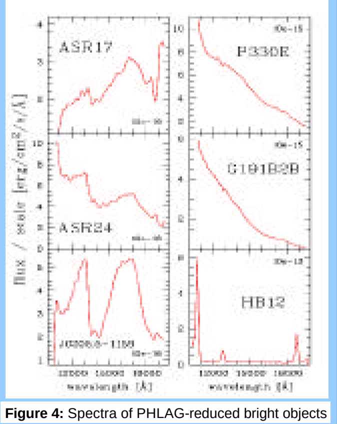

Figure 4: Spectra of PHLAG-reduced bright objects

3. PHLAG

PHLAG consists of a series of modules with each performing a certain reduction step on the data. The pipeline is implemented in Python, but utilizes existing software (**Pyraf** [3], **MultiDrizzle** [4], **SExtractor**) whenever possible. Important modules of **PHLAG** are:

- **data preparation:** In this step the data are prepared for the pipeline reduction. The direct images are grouped according to the filter. Every slitless image is paired with the direct image that was obtained with the smallest positional difference. The set of direct image - slitless image relationships created here is an essential input to the spectral extraction.
- **image combination:** To prepare for the object extraction, the filter images are combined to create a deep direct image. This is done using the **MultiDrizzle** software [4].
- **object detection:** The object detection software **SExtractor** is run on the combined direct image. Conservative parameter settings are used, and the prime aim is to detect all objects which also have detectable spectra in the slitless images.
- **spectral extraction:** The spectrum extraction package aXe [6,7] is used to extract the object spectra from the slitless images. As shown in Figure 3 for one spectrum, the flatfielded 2D spectral stamp images extracted from the individual grism images are coadded using the aXedrizzle technique [5]. Deep spectra are finally extracted from these combined 2D spectral images. For each spectrum, an estimate of the contamination caused by neighbouring objects is derived using the photometric information from the direct images.
- **metadata:** The spectra are post-processed and prepared for ingestion into the **HLA** archive. Metadata are collected (e.g. object positions, extraction geometry) or derived (e.g. signal-to-noise estimates for the spectra). To ensure compatibility with the Virtual Observatory (**VO**), we closely follow the rules and recommendations of the IVOA Spectral Data Model [8] in the selection of the metadata.
- **data ingestion:** This last module inserts the fully reduced, quality controlled and **VO**-ready spectra into the **HLA** archive.

```
# 1  LAMBDA
# 2  COUNT
# 3  ERROR
# 4  FLUX
# 5  FERROR
# 6  CONTAM
#Spectrum.DataModel = SPECTRUM 1.0
#Spectrum.Type = Slitless Spectrum
#Spectrum.CoordSys.SpaceFrame.Name = FK5
#Spectrum.CoordSys.TimeFrame.Name = TT
#Spectrum.CoordSys.TimeFrame.Zero = 2400000.5 ; as for all HST instruments
#Spectrum.Curation.Publisher = Space Telescope - European Coordinating Facility
#Spectrum.DataID.Title = HLA/ST-ECF NICMOS grism
#Spectrum.DataID.Creator = Space Telescope - European Coordinating Facility
#Spectrum.DataID.Version = 1.0
#Spectrum.DataID.Logo = http://www.stecf.org/img/stecflogocolor.gif
#Spectrum.TimeSI = T
.
.
.
#Spectrum.SpectralSI = 10-10 L
#Spectrum.FluxSI = 10+7ML-1T-3
#Spectrum.Char.FluxAxis.Name = FLUX
#Spectrum.Char.FluxAxis.Unit = erg/s/cm^2/Angstrom
#Spectrum.Char.FluxAxis.ucd = phot.fluDens ; em.wl;src.net
#Spectrum.Char.SpectralAxis.Name = WAVELENGTH
#Spectrum.Char.SpectralAxis.Unit = Angstrom
#Spectrum.Target.Name = HNG_J232614.77+581056.0_N8BR02T1Q ; the target name
#Spectrum.Target.pos.ra = 351.56154101 ; the target right ascension
#Spectrum.Target.pos.dec = 58.182230155 ; the target declination
#Spectrum.CoordSys.SpaceFrame.Equinox = 2000.0 ; coordinate equinox
#Spectrum.CustomParams.NumGrismImage = 15 ; Number of contributing grism images
#Spectrum.Length = 113
#Spectrum.DataID.Date = 2007-3-8T0:12:56 ; file creation date
 1.050321e+04  1.674639e+01  7.258485e-02  6.882959e-14  3.285606e-16  5.449870e-18
 1.058328e+04  3.281029e+01  6.593926e-02  8.340781e-14  2.364864e-16  9.115426e-18
.
.
.
 1.938325e+04  1.639907e+02  8.015385e-02  2.855464e-14  5.878995e-17  9.886373e-19
 1.946328e+04  9.394627e+01  6.736218e-02  2.870008e-14  6.097758e-17  1.775339e-18
```

Table 2: A spectrum in the internal ASCII format

4. Results:

Two complete test runs of **PHLAG** have been done in December 06 and March 07 on all available (~1000) **NICMOS G141** datasets. From the ~5000 grism images around 40,000 object spectra were extracted in one week of computing time. Figure 4 shows the spectra of bright objects as extracted by **PHLAG**.

Table 2 shows part of a spectrum in our internal ASCII format. The column definition is followed by the metadata and the data. The metadata resemble very closely the hierarchical structure defined in the IVOA Spectral Data Model, and a transformation from this internal to other formats for data distribution (`XML`, `VOTable` or `FITS`) is very easy. The column `CONTAM`, which contains an estimate of the contaminating flux from neighboring objects, was added to the "standard" columns in order to provide this additional information, which is essential for the interpretation of slitless spectra.

As of now, **PHLAG** is in the late phase of development, testing and optimization. After the next production run on all data, a first release of **NICMOS G141** spectra is expected in the second half of 2007. This released dataset is expected to become a formal part of the Hubble Legacy Archive (**HLA**) and to be accessible through both, custom interfaces and the **SSAP** protocol.

References

1. Walsh, J.R., Hook, R.N., 2006, ST-ECF Newsletter 40, p. 6
2. Kümmel, M., Larsen, S.S., & Walsh, J.R.2006, in "The 2005 Calibration Workshop", eds. Koekemoer, A., Goudfrooij, P., & Dressel, L., p. 85
3. PyRAF: http://www.stsci.edu/resources/softwarehardware/pyraf
4. Koekemoer, A.M., Fruchter, A.S., Hook, R.N., Hack, W. & Hanley, C. 2006, in "The 2005 Calibration Workshop", eds. Koekemoer, A., Goudfrooij, P., & Dressel, L., p. 423
5. Kümmel, Walsh, Larsen, Hook, 2005, ADASS XIV, eds. Shopbell, P., Britton, M. & Ebert, R., p. 138
6. aXe webpage: http://www.stecf.org/instruments/ACSgrism/aXe
7. Kümmel, M., Larsen, S.S., & Walsh, J.R.2005, ST-ECF Newsletter 38, p. 8
8. IVOA Spectral Data Model: http://www.ivoa.net/Documents/latest/SpectrumDM.html