Physics 346 – Spring 2021
Project # 7: Identifying AGN from emission-line spectroscopy

Emission lines from interstellar gas in a galaxy can provide a wealth of information about its chemical composition (e.g., "metallicity" as represented by the abundance of oxygen relative to hydrogen) and prevailing physical conditions (e.g., temperature and density). Emission lines can also reveal what physical processes are influencing those physical conditions — and in particular, whether atoms are being ionized by newly formed stars (i.e., in H II regions) or by accretion onto supermassive black holes (i.e., active galactic nuclei, or AGN). The basis for such a distinction is that AGN can produce arbitrarily high-energy photons capable of raising atoms to very high ionization states (i.e., stripping off many electrons), whereas young stars only produce photons up to energies corresponding to their surface temperatures. An important reference for the classification of optical emission-line spectra is the paper of Baldwin, Phillips, & Terlevich (1981, PASP, 93, 5), whose Figure 5 plotting the "BPT diagram" of [O III] λ5007/Hβ vs. [N II] λ6583/Hα shows how specific line ratios can be used to trace ionization state and identify AGN.

An additional tool for identifying AGN in emission-line spectra is the measurement of line *widths* (especially for the Balmer α or "Hα" line, which is typically the brightest emission line in a galaxy's optical wavelength spectrum). Broad ($\Delta v \geq 1000\,\mathrm{km\,s^{-1}}$) emission lines are produced when gas deep in a supermassive black hole's potential well is moving at very high velocities, and are not plausibly produced by other mechanisms. It can however be challenging to detect broad Hα emission, because emission from ionized gas in the immediate vicinity of a black hole may be "outshone" by emission from less rapidly moving gas on larger scales, and because it can be tricky to disentangle the emission from a galaxy's stars and from its gas.

For this project, you will be using both of the techniques described above to identify AGN within a sample of bright northern galaxies. For the first part of your analysis, you will use observations of the full sample of 418 galaxies to identify candidate AGN on the basis of emission line ratios. Starting from the tab-delimited file Project7_DataTable_1.tsv, which can be read into a spreadsheet or a programming environment of your choice, you should do the following:

1. Determine whether there are any galaxies you feel should be excluded from your analysis on the basis of limited or poor-quality data.

2. Plot the "BPT diagram" described above for all of the galaxies in the sample that you did not exclude.

3. On the basis of their location in the BPT diagram, identify galaxies whose gas is mainly being ionized by young stars, those whose gas is mainly being ionized by AGN, and those whose gas is being ionized by a combination of both.

4. Determine whether the location of a galaxy in the BPT diagram (and its status as star-formation-powered or AGN-powered) is correlated in any way with (a) the [O III] λ5007/[O I] λ6300 emission line ratio, which is a measure of ionization state; (b) the Hα/Hβ emission line ratio, which is a measure of dust content; (c) the Hα luminosity; (d) the Hα equivalent width (i.e., line-to-continuum ratio); and (e) the [N II] λ6583

velocity width. Please comment on the possible implications of any correlations you identify (or the lack thereof).

In addition, for the set of ten spectra of emission-line galaxies *and* the ten spectra of associated "partner" absorption-line galaxies that will be provided to you in ASCII format (which can be read into a spreadsheet or a programming environment of your choice), you should do the following:

1. For each pair of spectra, determine the appropriate wavelength shift and intensity scaling needed to subtract the (stellar-only) absorption-line partner spectrum from the emission-line spectrum, in order to produce a star-free version of the latter after subtraction.

2. After obtaining a star-free emission-line spectrum, isolate the Hα and [N II] portion of the spectrum, and decompose the line profile into a combination of narrow and/or broad velocity components. Please discuss how you determine whether a broad velocity component is present, and any uncertainties that attach to your determination of its width.

3. For any galaxies where you conclude a broad Hα velocity is present, determine whether there is any correlation between the width and the luminosity *of the broad component.* To allow you to convert flux to luminosity, the distances to the galaxies in your sample are listed below. Please comment on the possible implications of such a correlation (or the lack thereof).

Galaxy	Distance
1	71.1 Mpc
2	75.1 Mpc
3	22.0 Mpc
4	18.2 Mpc
5	40.9 Mpc
6	24.6 Mpc
7	16.6 Mpc
8	14.6 Mpc
9	32.8 Mpc
10	16.2 Mpc

• **Multiwavelength astronomy:** What data from other wavelengths would be scientifically interesting to add to the data you have analyzed here? How might light from other wavelengths help the analysis of your data? What multiwavelength analyses would you like to undertake?

Physics 346 – Spring 2021
Project # 8: Galaxy metallicities from emission-line spectroscopy

Emission lines from interstellar gas in a galaxy can provide a wealth of information about its prevailing physical conditions (e.g., temperature and density) and its chemical composition, which is typically parameterized using a "metallicity" that is defined in terms of the abundance of oxygen relative to hydrogen. (At $z = 0$, oxygen has the highest cosmic abundance of any element in the periodic table other than hydrogen or helium.) Studies of galaxy metallicities often focus on two questions: (i) which emission lines (and which calibration of them) should be used to estimate metallicity, and (ii) what factors account for a galaxy's observed metallicity. The most obvious answer to (i) would involve a set of oxygen and hydrogen lines, with the oxygen lines tracing multiple ionization states to ensure that none of a galaxy's oxygen is "missed," and indeed, one of the most widely used indicators is the ratio of line fluxes

$$R_{23} \equiv \frac{[\mathrm{O\,II}]\,\lambda 3727 + [\mathrm{O\,III}]\,\lambda 4959 + [\mathrm{O\,III}]\,\lambda 5007}{\mathrm{H}\beta} \tag{1}$$

which has been calibrated, e.g., by McGaugh (1991, ApJ, 380, 140). However, since R_{23} requires the measurement of multiple emission lines that may not all be accessible for a galaxy at a particular redshift, alternatives have also been proposed, like the ratio of line fluxes

$$N2 \equiv \frac{[\mathrm{N\,II}]\,\lambda 6583}{\mathrm{H}\alpha} \tag{2}$$

first calibrated by Pettini & Pagel (2004, MNRAS, 348, L59). To address (ii), it is often appealing to measure the metallicities of large numbers of galaxies and explore whether relationships between metallicity and other properties (e.g., stellar mass or star formation rate) suggest what factors determine a galaxy's chemical enrichment.

For this project, you will be using observations from the Sloan Digital Sky Survey (SDSS) — as extracted by a team from Johns Hopkins University and the Max Planck Institute for Astrophysics — to consider both of the above questions. The data release at https://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/ includes

- a FITS table containing "raw" spectral line fluxes, gal_line_dr7_v5_2.fit
- a FITS table containing "derived" total stellar masses, totlgm_dr7_v5_2.fit
- a FITS table containing "derived" total star formation rates, gal_totsfr_dr7_v5_2.fits

(you will need to click down 1–2 levels to access the files and the descriptions of their contents; make sure to compare the information in the FITS file headers with the file descriptions in case there are differences). Using these resources, you should do the following:

1. Identify an unbiased working sample of ~ 1000 galaxies that you will use for your analysis, and extract the corresponding subsets of the FITS tables described above. (This step is recommended so that you do not need to deal with the very large full datasets every time you work on this project.)

2. For each galaxy in your working sample with measurements of the necessary spectral lines, calculate the values of the R_{23} and $N2$ metallicity indicators.

3. Use the R_{23} and $N2$ values you have calculated, in combination with calibration relations from the literature, to make two separate measurements of 12 + log(O/H) — the standard expression for metallicity in the extragalactic literature.

4. Generate two plots of 12 + log(O/H) vs. stellar mass for your entire working sample (one for the R_{23}-derived metallicities and one for the $N2$-derived metallicities). Each plot should include a curve that represents the median metallicity as a function of stellar mass. Please comment on the possible implications of any relations you can identify (or the lack thereof), and on any notable differences between the two plots.

5. Create three subsets of your working sample that are binned by star formation rate, setting the boundaries between your bins so that each subset has a roughly equal number of galaxies. Now generate two new plots of 12 + log(O/H) vs. stellar mass for each of those three subsets, again showing median metallicity as a function of stellar mass. Please comment on the possible implications of any notable differences among these *six* plots.

6. **Multiwavelength astronomy:** What data from other wavelengths would be scientifically interesting to add to the data you have analyzed here? How might light from other wavelengths help the analysis of your data? What multiwavelength analyses would you like to undertake?

Physics 346 – Spring 2021
Project # 9: SALT+RSS Multi-Object Spectroscopy of Emission Line Galaxies

Congratulations! Your proposal for multi-object spectroscopy of emission-line galaxies at $z < 0.5$ using the Robert Stobie Spectrograph (RSS) on the Southern African Large Telescope (SALT), in which Rutgers has a 10% share, was accepted. Last night, the Support Astronomer took a one-hour block from this program and sent an email informing you that it met the requested data quality in terms of clear skies and seeing better than $2''$. A downloaded package of the dataset, including raw and pipeline-processed versions of each exposure along with brief documentation of the observations and a few other relevant files can be found in Prof. Gawiser's /project9 directory on astrolab.

The data reduction you are asked to undertake below will be challenging but ultimately rewarding. Most of the later steps can be done following the longslit spectroscopy approach of Workshop 2, but some basic .fits image read-in and processing in either Python or IDL will be required to set that up via the earlier steps. Those who are comfortable with Python may want to explore the astropy/specutils package; similar packages exist in IDL. However, the pre-RUPhAst steps below should not require such packages.

1. Use tar to unzip the data package into raw, product, and doc directories. Figure out what the difference is between raw and product files; you may find it easier to display them using the ds9 software available on astrolab than with RUPhAst. It should also be possible to process the multi-extension fits (MEF) format product files into updated .fits files that RUPhAst can read using the squeeze_extensions.py script.

2. Identify the flat-field exposures and combine them in a robust manner to make a single flat-field image.

3. Use the combined flat to flat-field the two main science exposures. This corrects for variations in the CCD pixel response and the overall telescope illumination pattern. One way to do this is to use the pipeline in RUPhAst.

4. Combine the two main science exposures in a way that rejects cosmic rays. The result should look much better!

5. Identify the slitlet whose central spectrum has the brightest continuum and emission lines.

6. Use the Neon arc lamp spectra that are part of the data package for wavelength calibration. You can use the included list of night sky lines (sky_lines.dat) to cross-check if the wavelength calibration makes sense.

7. Extract a one-dimensional spectrum. Be careful to adjust your sky background estimation regions to lie within the slitlet in order to properly subtract the many night sky lines.

8. Identify the [N II] $\lambda6584$ and Hα emission lines in this spectrum and obtain their flux ratio.

9. Use the [N II]/Hα indicator introduced by Denicolo et al. (2002, MNRAS 330, 69, PDF included in /project9) to measure the metallicity of this galaxy. Metallicity is typically defined as $12 + \log(O/H)$, with solar metallicity being 8.7 in these units.

10. Repeat this process for as many of the other slitlets as you can. Note that the horizontal shifts between slitlets seen in the through-the-slitmask image will let you get an initial wavelength solution via shifting by the same amount versus the original choice of slitlet. Do all of these galaxies have similar metallicity? Can you identify any other emission lines? (see galaxy_lines.dat) How many galaxies can you obtain redshifts for, and what are they?

11. Time to panic: did we ever do flux calibration? No standard star data were even taken! What about correcting for Milky Way dust or dust in the galaxies? How could our measured [N II]/Hα ratios be meaningful without any of those steps?

12. **Multiwavelength astronomy:** What data from other wavelengths would be scientifically interesting to add to the data you have analyzed here? How might light from other wavelengths help the analysis of your data? What multiwavelength analyses would you like to undertake?

Physics 346 – Spring 2021

Project # 10: CO($J = 1 \to 0$) imaging of the active galaxy NGC 1068

Galaxies are defined as "active" if their central, supermassive black holes are accreting material from their surroundings at large rates, and are thereby converting potential energy into radiation. The resulting active galactic nuclei (AGN) can radiate strongly at all wavelengths across the electromagnetic spectrum, from radio to X-ray and beyond, and are thought to have significant effects on their host galaxies — for example, driving some fraction of their interstellar gas into intergalactic space. One of the best-studied prototypes among active galaxies is NGC 1068, a relatively nearby ($D \approx 14.4$ Mpc) spiral galaxy whose nucleus has a "Seyfert 2" spectrum in the optical. (A Seyfert 2 shows emission lines from highly ionized species, but does not show the *broad* emission lines from gas moving at very high velocities near the central black hole, which would define a Seyfert 1.) In this project, you will investigate observations of the CO($J = 1 \to 0$) rotational emission line (rest frequency $\nu_0 = 115.271$ GHz) in NGC 1068 that were obtained with the Owens Valley Radio Observatory (OVRO) millimeter array in California.[1] Using the fully processed data cube in the file `ngc1068co.fits`, which has two spatial dimensions and one frequency dimension and can be found in Prof. Gawiser's /project10 directory on astrolab, along with the DS9 tool and (as a starting point) the IDL code in `sectionC_2020_example.pro`, you should do the following:

1. Determine the dimensions of the data cube, i.e., how many spatial pixels and spectral channels it has, and how large (in arcsec and MHz) those pixels and channels are. Note that to sum emission over any set of pixels in an interferometric map, you will need to add up those pixels' values *and* then divide that sum by the number of pixels per (synthesized) beam, which is
$$\frac{\pi}{4 \ln 2} \frac{a \times b}{\Delta x \times \Delta y} \tag{1}$$

2. Determine the systemic redshift for NGC 1068, and the rest-frame velocity ranges on either side of that systemic redshift for which CO($J = 1 \to 0$) emission is detected.

3. Run the example IDL code in the file `sectionC_2020_example.pro`. You can begin by typing "idl" in a new terminal window, and then at the IDL prompt, cutting and pasting the following three commands:
```
addpath = expand_path('+/usr/local/src/idl')
!path = addpath + ':'  + !path
delvar, addpath
```
(these commands allow IDL to find and use a number of prewritten programs in other directories). You can now type in IDL:
```
.compile sectionC_2020_example.pro
test1
```
to compile the `sectionC_2020_example.pro` file and then execute the "test1" program within it. If you open up `sectionC_2020_example.pro` in a text editor like gedit, you can see the commands that it uses to

[1] The OVRO millimeter array later became part of the Combined Array for Research in Millimeter Astronomy (CARMA; http://www.mmarray.org/).

- read in the data cube;
- create and plot to the screen a 2D array containing only channel 50;
- create and plot to a *file* a 2D array containing the sum of channels 50–60 (this file is named chans50to60.ps; it can be viewed from a different terminal window using the command "gs chans50to60.ps"); and
- calculate the total of all pixel values in a specified sub-region of the summed image.

You should discuss with your lab partners and your instructor(s) what the different lines in this program do. (For the remaining steps of this analysis, you are welcome to adapt this IDL program to serve your needs, or to develop an alternative approach using, e.g., Python.)

4. Assuming that spiral galaxies rotate about their centers in the sense that makes their spiral arms "trailing" rather than "leading," determine which part(s) of the galaxy is/are closest to us along the line of sight, farthest from us along the line of sight, moving towards us most rapidly, and/or moving away from us most rapidly.

5. Determine whether there is any *continuum* emission in NGC 1068. In making this assessment, you should keep in mind that to detect faint continuum emission, you may need to average many channels together so that you can see it emerge above a lower ($\propto n_{\rm chan}^{-1/2}$) averaged noise level.

6. If you determine that there is continuum emission in NGC 1068, please calculate the total continuum flux density (in units of Jy) and comment on any implications of its strength and spatial distribution. In addition, you should construct a continuum image that contains no spectral line emission, and a spectral line data cube that contains no continuum emission. The latter, continuum-free spectral line cube would then be the basis for the subsequent steps of your analysis.

7. Determine the total CO($1 \rightarrow 0$) line flux (in units of $\mathrm{Jy\,km\,s^{-1}}$) for NGC 1068, and the total mass of its molecular gas. For the latter conversion, you may use the following relationship from Sanders et al. (1991, ApJ, 370, 158):

$$\frac{M_{\rm H_2}}{M_\odot} = 1.18 \times 10^4 \left(\frac{F_{\rm CO(1-0)}}{\mathrm{Jy\,km\,s^{-1}}}\right)\left(\frac{D}{\mathrm{Mpc}}\right)^2 \quad (2)$$

   Please comment on any implications of this value for the overall gas mass fraction ($f_{\rm gas}$) in NGC 1068.

8. Calculate and plot the circular rotation curve ($v_{\rm circ}(R)$), the enclosed mass ($M(<R)$), and the enclosed gas mass fraction ($f_{\rm gas}(<R)$) in NGC 1068. Please comment on any associated constraints you can place on the mass of the galaxy's central, supermassive black hole, and any systematic uncertainties that apply to this mass estimate.

9. **Multiwavelength astronomy:** What data from other wavelengths would be scientifically interesting to add to the data you have analyzed here? How might light from

other wavelengths help the analysis of your data? What multiwavelength analyses would you like to undertake?

Physics 346 – Spring 2021
Project # 11: Multi-line CO imaging of two ULIRGs

For a galaxy outside the Milky Way, we can use observations of the relatively rare CO molecule (typically 10^{-4} times as abundant as H_2) to estimate not only the total mass of the galaxy's molecular clouds, but also the total *gravitational* (a.k.a. "dynamical") mass that includes gas, stars, and dark matter. To achieve the latter goal, it is necessary to obtain a spatially resolved rotation curve $v_{\rm circ}(R)$, where "spatially resolved" means that one has information on the gas motions in at least two independent pixels. For galaxies that are relatively compact, a spatially resolved rotation curve can only be obtained by mapping with an interferometer — specifically, an array equipped with millimeter-wavelength receivers, since the rest frequencies of the CO($J = 1 \to 0$) and CO($J = 2 \to 1$) lines are $\nu_0 = 115.271$ GHz and 230.538 GHz, respectively.

For this project, you are being provided with four pre-calibrated sets of visibilities obtained with the IRAM Plateau de Bure Interferometer (PdBI), a facility in the French Alps that later became part of the Northern Extended Millimeter Array (NOEMA). The targets of these observations are two ultraluminous infrared galaxies (ULIRGs), whose substantial dust luminosities (by definition, $> 10^{12}\,L_\odot$, i.e., over 100 times that of the Milky Way) are due to bursts of star formation triggered by the merger of two gas-rich progenitor galaxies. For each system, you will have data for both the CO($J = 1 \to 0$) and CO($J = 2 \to 1$) lines; files are named `iras15250co10.uvf`, `iras15250co21.uvf`, `iras17208co10.uvf`, and `iras17208co21.uvf`, and can be copied out of Prof. Gawiser's /project11 directory on the astrolab server. Using these data and the Difmap software[1] installed on the astrolab server, you should first work though the brief tutorial on Difmap and then complete the requested data analysis.

Tutorial: To become familiar with Difmap and its capabilities, you are encouraged to do the following:

1. To make it possible to run Difmap on the astrolab server, you will need to use `gedit` or another editor to add several lines of text at the end of the `.bashrc` file in your home area (note the period at the start of "`.bashrc`"!):
   ```
   PGPLOT_DIR=/usr/local/pgplot/
   export PGPLOT_DIR
   PGPLOT_DEV=/xwin
   export PGPLOT_DEV
   ```
   Once you have added these commands, every new terminal window you open will be able to run Difmap.

2. Load in a data file with the `observe` command, and create a pseudo-continuum dataset that includes the first ten frequency channels with the command `select RR,1,10`.

3. Using natural weighting (`uvweight 0,-2`), create a map that has 256×256 pixels, each of size $0.125'' \times 0.125''$ (you may find it helpful to invoke the command `mapunit arcsec` before specifying the map and pixel size; note that Difmap by default will

[1] A set of instructions for the basic setup and use of Difmap is appended below.

only *show* the inner quarter of any map). Make a note of (a) the estimated beam size, (b) the noise in the map (for which you will need to use the command `print rms(map)`, rather than the "estimated noise"), and (c) the location and flux density of the map peak (for which you can use the commands `print peak(x)`, `print peak(y)`, and `print peak(flux,max)`). (Also note that the pixel size should be no larger than a third of the synthesized beam; if that is not the case, you should regenerate the map with a smaller pixel size.)

4. Repeat the above exercise for a pseudo-continuum dataset that covers the *last* ten frequency channels, and for a pseudo-continuum dataset that covers the *middle* ten frequency channels. As a check of understanding: how do the three beams compare to each other, and how would you account for any differences? How do the three map peaks compare to each other, and how would you account for any differences?

5. For the pseudo-continuum dataset near the center of the observed band, draw a clean box (interactively, using the left mouse button after invoking the `mapplot` command) and use the `clean` command to deconvolve the dirty map. When you have cleaned deeply enough (you may need to use more than one clean box), the source should have effectively disappeared from the residual map, and the positive and negative peaks in the residual map should be roughly comparable. As a check of understanding: is the rms in the residual map what you would expect given the rms levels you measured at the frequency extremes? What is the total cleaned flux density that Difmap reports (in units of Jy), and what does this actually correspond to in flux units of $\mathrm{Jy\,km\,s^{-1}}$?

6. Using the `uvplot` command, produce a plot of the *uv* sampling of the dataset. As a check of understanding: from looking at the central gap in the *uv* plane, what limits can you set on the diameter of the antennas in the PdBI? What can you say about the array's maximum projected baseline?

Data Analysis:

1. Determine the key properties of the spectral axis of each dataset, i.e., how many frequency channels it has, what the central frequency $\leftrightarrow$ redshift of the dataset is, and how wide each channel is in frequency $\leftrightarrow$ velocity.

2. Determine the range of velocities relative to the systemic redshift over which each galaxy shows CO emission.

3. Determine whether each of the four datasets shows evidence of continuum emission, and if so, how to adjust estimates of "spectral line" flux (see below) to correct for the contribution of such continuum emission.

4. Determine the total molecular gas mass of each galaxy. This calculation can leverage the scaling of $M_{\mathrm{H_2}}$ from CO($J = 1 \to 0$) spectral line flux $F_{\mathrm{CO(1-0)}}$

$$\frac{M_{\mathrm{H_2}}}{M_\odot} = 1.180 \times 10^4 \left(\frac{D}{\mathrm{Mpc}}\right)^2 \left(\frac{X}{3 \times 10^{20}}\right) \left(\frac{F_{\mathrm{CO(1-0)}}}{\mathrm{Jy\,km\,s^{-1}}}\right) \tag{1}$$

which assumes a conversion factor

$$X \equiv N_{\mathrm{H_2}}/I_{\mathrm{CO}(1-0)} = 3 \times 10^{20}\,\mathrm{cm^{-2}}\,(\mathrm{K\,km\,s^{-1}})^{-1} \tag{2}$$

like that measured in the Milky Way. You may use the expression $D \approx cz/H_0$ to approximate distance, where the Hubble parameter $H_0 \approx 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

5. Estimate the flux ratio of the CO($J = 2 \rightarrow 1$) and CO($J = 1 \rightarrow 0$) lines, and comment on whether this ratio changes significantly (a) as a function of velocity within a single galaxy, or (b) between your two targets.

6. Determine the total dynamical mass of each galaxy, for which you can use the formula for $M(< R)$ stated in a previous lab exercise. Note that a generous definition of a "rotation curve" may be possible and necessary here: as long as you can identify a separation on the sky ΔR between the locations where emission peaks at two velocities separated by Δv, you can estimate $v_{\mathrm{circ}} \approx 0.5 \times \Delta v$ and $R \approx 0.5 \times \Delta R$ for the purposes of this estimate.

7. Calculate the total molecular gas mass fraction $f_{\mathrm{gas}} \equiv M_{\mathrm{H_2}}/M_{\mathrm{dyn}}$ for each of your targets, and comment on any implications of these values.

8. **Multiwavelength astronomy:** What data from other wavelengths would be scientifically interesting to add to the data you have analyzed here? How might light from other wavelengths help the analysis of your data? What multiwavelength analyses would you like to undertake?

Physics 346 – Spring 2021
Project # 12: Measuring the H I Content of Early-Type Galaxies

The vast majority of galaxies can be classified according to their morphologies as *early-type* (ellipticals and lenticulars) or *late-type* (spirals and irregulars). Astronomers have determined that there are other properties of galaxies that correlate with these morphological classifications: early-type galaxies tend to be found in richer environments (i.e., they have more neighbors), have older stellar populations, and have less interstellar gas than late-type galaxies. Studying these trends (and the exceptions to them) has the potential to shed light on important aspects of the formation and evolution of galaxies.

For this project, you are being provided with observations of multiple early-type galaxies obtained with the 100 m diameter Effelsberg Radio Telescope (operated by the Max Planck Institute for Radio Astronomy in Bonn, Germany) that can be used to constrain their masses of neutral atomic gas (i.e., H I). These targets were chosen because of their relative isolation, which reduces the likelihood that their H I properties will have been strongly influenced by environmental effects. The observations of your target galaxies (along with two galaxies observed as spectral line calibrators, and a number of other systems) were taken on three different days, and can be found in three data files that you should copy out of Prof. Baker's home area on the astrolab server: `Effelsberg20010721.eff`, `Effelsberg20010722.eff`, and `Effelsberg20020610.eff`. Using these data and the CLASS software installed on the astrolab server, you should do complete the tutorial and then move on to the requested data analysis.

Tutorial: To become familiar with CLASS and its capabilities, you are encouraged to do the following:

1. To make it possible to run CLASS on the astrolab server, you will need to use `gedit` or another editor to add several lines of text at the end of the `.bashrc` file in your home area (note the period at the start of "`.bashrc`"!):
```
export GAG_ROOT_DIR=/root/gildas-exe-jul14c
export GAG_EXEC_SYSTEM=x86_64-redhat5.4-gfortran
source $GAG_ROOT_DIR/etc/bash_profile
alias rm='rm -i'
alias cp='cp -i'
alias mv='mv -i'
```
   Once you have added these commands, every new terminal window you open will be able to run programs from the GILDAS software package (including CLASS). The last three lines will help protect against files being inadvertantly deleted or overwritten.

2. Launch the CLASS software by typing "class" in a terminal window. The terminal window should acquire a `LAS90>` prompt, and a new graphics window should open up.

3. Load in and inspect the contents of your data file by typing, e.g.,
```
LAS90>file in Effelsberg????????.eff
LAS90>find
LAS90>list
```
   where "????????" is replaced by the digits appropriate for one of your data files. You

will see a long list of entries corresponding to the sequence of observations taken at the Effelsberg telescope on a given day. Each *scan* was taken for a particular astronomical target, and yielded four different spectra taken by four independent electronic *backends*. It is possible to use the `find` command to filter this list according to source, backend, and scan number — for example,
`LAS90>find /source IC5258 /telescope MPI-100M-AK1`
would select only observations of the galaxy IC 5258 with the first backend unit, and `/scan` could be used in the same command line to select a particular scan number or range.

4. Follow this series of steps to analyze observations of your spectral line calibrator (here, assumed to be Sextans B):

`LAS90>find /source sextansb`	Selects all observations of Sextans B.
`LAS90>average /2010`	Averages all selected observations (using "2010" algorithm).
`LAS90>smooth`	Smooths spectrum by a factor of 2.
`LAS90>set plot histogram`	Sets plot format.
`LAS90>set unit v`	Sets x-axis to velocity.
`LAS90>plot`	Plots mean spectrum (y-axis units are K).
`LAS90>set window`	Interactively defines the velocity range where H I is seen.
`LAS90>base 3 /plo`	Calculates and plots a third-order polynomial baseline.
`LAS90>plot`	Replots spectrum with baseline subtracted.
`LAS90>minimize`	Derives a Gaussian fit to line peak.
`LAS90>display`	Prints out the results of the Gaussian fit.
`LAS90>visualize`	Overplots the Gaussian fit on the spectrum.
`LAS90>header`	Prints out basic information about the spectrum.
`LAS90>greg sextansb.dat /formatted`	Writes out spectrum in tabular form.
`LAS90>hardcopy sextansb.ps`	Writes out plot of spectrum.

Note that several of these commands can be executed in different ways, e.g., you can smooth by a factor other than 2, specify a baseline fit of order other than 3, and specify a velocity range to exclude from your fit in a non-interactive way a la `set window 1000 1500`. For details on the syntax of a particular command, just type `help` and the name of the command within CLASS. Note also that you can execute Linux commands from within CLASS by prefacing them with a $ sign, and you can run a script of CLASS commands by prefacing the name of the script with an @ sign. If you would like to *force* CLASS to plot only a specific velocity range, you can specify

`LAS90>set mode x # #`	Limits velocity range to # through # in $\mathrm{km\,s^{-1}}$.
`LAS90>plot`	Replots spectrum with new velocity range.

Data Analysis:

1. For each of the three data files, determine which scan numbers correspond to observations of your assigned four targets — IC 1156, IC 5258, NGC 6172, and UGC 10115.

2. For each of the three data files, determine which scan numbers correspond to observations of one of the *spectral line calibrators* for the Effelsberg telescope. These are

two dwarf galaxies (compact enough to fit well within the Effelsberg telescope's beam): Holmberg I (sometimes abbreviated "Ho I") with a total H I flux of $38.6\,\mathrm{Jy\,km\,s^{-1}}$, and Sextans B (sometimes abbreviated "Sex B") with a total H I flux of $106.2\,\mathrm{Jy\,km\,s^{-1}}$.

3. For each of the three data files, determine the appropriate conversion factor between $\mathrm{K\,km\,s^{-1}}$ and $\mathrm{Jy\,km\,s^{-1}}$ using the observations of your spectral line calibrator, and the uncertainty in this factor. (Note: all the information you need is in your data or in these instructions!)

4. For each of your target galaxies and each of the three data files, produce an average, baseline-subtracted spectrum. (You may wish to explore use of the `drop` command in CLASS to exclude particular scan/backend combinations from the average if this helps eliminate radio frequency interference (RFI). RFI features typically manifest as very narrow spikes, either either positive or negative depending on whether a signal was detected at the on-source or off-source position within a given scan.) With the other members of your group, you will need to come up with strategies for deciding (a) what type(s) of baseline to subtract, (b) what type(s) of RFI to exclude, and (c) whether you have detected any H I line emission.

5. For each of your target galaxies, produce a single baseline-subtracted spectrum made from the *combination* of the spectra from the three individual datasets. When producing this spectrum, you may need to take into account differences in the amount of time for which a given target galaxy was observed on a given day, and in the conversion factor between $\mathrm{K\,km\,s^{-1}}$ and $\mathrm{Jy\,km\,s^{-1}}$ that you calculated for a given day. (Note: the combination of spectra from different days will need to be done outside CLASS, so writing out the three different spectra for a given galaxy will be a ncessary first step.)

6. For any galaxy in which you have detected H I line emission, calculate the total H I line flux (and its uncertainty) in units of $\mathrm{Jy\,km\,s^{-1}}$. For any galaxy in which you have not detected H I line emission, calculate a 3σ upper limit on its H I line flux in units of $\mathrm{Jy\,km\,s^{-1}}$. (A 3σ limit is defined in terms of the 1σ RMS noise in a spectrum; you can assume a fiducial velocity width of $100\,\mathrm{km\,s^{-1}}$ to calculate it.)

7. For each of your target galaxies, determine the distance using the formula $D \approx v/H_0$, where D is the distance in Mpc ($1\,\mathrm{Mpc} = 3.09 \times 10^{24}\,\mathrm{cm}$), v is the recessional velocity in $\mathrm{km\,s^{-1}}$, and the Hubble constant $H_0 \approx 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. (This approximation for D is reasonable when we are studying nearby galaxies.) Use this distance to calculate the neutral atomic hydrogen mass and its uncertainty (or an upper limit on the H I mass if you have not actually detected any H I emission) for each of your target galaxies, using the formula

$$\frac{M_{\mathrm{H\,I}}}{M_\odot} = 2.36 \times 10^5 \left(\frac{D}{\mathrm{Mpc}}\right)^2 \left(\frac{F_{\mathrm{H\,I}}}{\mathrm{Jy\,km\,s^{-1}}}\right) \tag{1}$$

In presenting your results, you should discuss any additional, specific, *systematic* uncertainties in these values that go beyond merely random uncertainties. (Note: the concepts of "observer error" and "human error" are not specific and are never invoked in scientific papers!)

8. Considering the distances and morphologies of your target galaxies (see the NASA Extragalactic Database at `http://ned.ipac.caltech.edu/`), consider whether you can draw any conclusions about the occurrence of H I in early-type galaxies from these data alone. If these data are not sufficient, what additional types of galaxies would you suggest observing in order to allow a more conclusive result?

9. **Multiwavelength astronomy:** What data from other wavelengths would be scientifically interesting to add to the data you have analyzed here? How might light from other wavelengths help the analysis of your data? What multiwavelength analyses would you like to undertake?