Lab 1
Unix Basics, Astronomical Images

Boxes contain questions that you are expected to answer (in the box). You will also be asked to put computer output into special directories, so it can be graded.

Name: ______________________________
Partner(s): ______________________________

These days, astronomical tools are often computational tools, and they often run under some flavor of the unix operating system. This lab will introduce you to unix commands and syntax that you will need to get things done, and to **ds9**, an image-visualization tool that is standard in astronomy. Most astronomical measurements involve counting things (though this is not always obvious). This lab will also illustrate that: "The uncertainty in any counted number of random events, as an estimate of the true average number, is *the square root of the counted number*." (Taylor, p. 48) We will repeat this mantra *ad nauseum*. Get used to it.

Log in to your computer. (Take turns at the keyboard. Watching is not as good as doing.)

login name student
password physics

You will get a blue screen with some icons. *Right click* on the blank background and select "Open Terminal" to bring up a command window.

Unix is largely command-line oriented (though you can usually cut and paste with the mouse, and many applications are GUIs.)

Where to get help
To get the manual for any unix command whose name you know, type **man command_name**.

The default web browser is Firefox. Start it by double-clicking its icon on the left side of the screen. For a general unix tutorial, go to http://www.ee.surrey.ac.uk/Teaching/Unix (also listed on the course web site).

Type **pwd**. What is the result? /home/student

Since you have just logged in, this is your *home directory*.

Type **ls**. Type **man ls** to see what you just did. Describe the function of the ls command in a few words.
Lists the files contained in the current directory

Try **ls –l** (this adds an option to the ls command). Type **man ls** to get the manual page for the ls command.

In a few words, describe what **ls –l** does. (Don't quote the manual.)
Lists the files as ls does, but in a long format, printing more information

In column 1 of the output, what does it mean if there is a leading letter "d"?
The listed file is a directory.
Write a command that will list the files in the current directory in a long format, sorted in order of increasing time since the last change.
ls -lt

The directory structure is central to unix, and it is important to learn to get around within it easily. Read the **unix tutorial** for a description of the possibilities. Some things are obvious, and work a lot like Windows folders. **mkdir** makes a new directory. **rmdir** removes one (but directories must be empty before you can remove them). **cd** changes directories, either to a child or parent directory of the current one, or to any place, with the appropriate syntax. A full file pathname contains the string of directories needed to find a file from some standard starting point, usually the "root" directory, denoted "/". Thus, a file in the "images" directory below your home directory is described as "/home/student/images/lab003.fit". There are a few non-intuitive conventions: "**.**" means the current directory. "**..**" means the parent directory of the current one. Thus, to move up a level in the directory tree, type "**cd ..**". To move to a child directory called "whatsis", type "**cd whatsis**". **cd** followed by nothing takes you to your home directory. And so on. RTM (*Read The Manual*).

Make a directory called results_yourname (eg results_JohnTravolta) as a subdirectory of your home directory. (You will put stuff in here to be graded.) What is the full pathname of this directory? /home/student/results_TimBrown

The **cat** command writes a file's contents to the screen, or with *redirection* (RTM), to another file.
From your home directory, type **cat hobbit.txt**
How many lines of output did you get? 10

Who first wrote these words? (Google it if you must. We will search the web for information again and again.)
J. R. R. Tolkien
Type **man wc** (word count). Write down a wc command below that will print the number of lines in the file hobbit.txt. How many does it find? wc -l hobbit 10 lines

The standard unix text editor is **vi**, and we will have to use it (or an equivalent such as **emacs**) a lot. Like most editors, it has many complex functions; getting used to them takes practice. The course web site has pointers to a vi manual, and to a 1-page cheat sheet that explains the most useful commands. RTM.

~~Using vi, open the file hobbit.txt for editing. (If you are fluent with a different unix editor, use it, but come talk to me.)~~

Do the following:
Insert a blank line at the beginning of the file, and write your first name into it.
Insert a blank line at the end of the file, and write your last name there.
Using a global character substitution, change every occurrence of the word "hobbit" to "halfling". Write the command that you used below:
(Hint: it will start with the character ":") :1,$ s/hobbit/halfling/g :%s/hobbit/halfling/g

Write the edited text out to a file named "halfling.txt".
Use the **mv** (move) command to move the output file into the directory "results_YourName" that you created earlier. Make sure it is there, and that you did not destroy "hobbit.txt". Use **ls –l** to count the bytes in halfling.txt. How many bytes? 254
(This will vary with your name)

ds9 Fundamentals

The program **ds9** allows you to display astronomical images, measure many of their properties, and much more. Unfortunately, the manual for ds9 is particularly terse and unhelpful. The only way to understand the program's capabilities is to try things, or to consult a guru (perhaps your own lab partner). Do the following few first steps.

From your home directory, **cd images** to get into the "images" directory.
Typing **ds9 lab000.fit** will open a window and display the FITS data file lab000.fit. About how many stars do you see? About a dozen

Left click on "scale", on "more…" if necessary to make "zscale" visible, and on "zscale". About how many stars now? A few hundred

Right click inside the image window, and drag the cursor around. What is the effect of moving it left-right? Changes zero point brightness.

What is the effect of moving the (right-clicked) cursor up-down? Changes contrast

Astronomical images ordinarily come with "metadata" = (data about the data). In images in FITS format, the metadata appear in a header that can be displayed separately from the data.
In the ds9 window, click on "File/Display Fits Header…" Answer the following questions, based on the header.

What are the x- and y- dimensions of the image (in pixels)? X = 1016 Y = 1024

What was the image exposure time (in s)? 240.

What filter was used? (We will go into what this means later) R

What was the name of the object being observed? SN 2006 EU

Google the object name you found above, and find an image of it on the web. Match this image to your ds9 display of lab000.fit.
Adjust the scaling so you get a recognizable picture of the object (which is a supernova) and its host galaxy.
(Try clicking "scale", then "min/max" and "square root". Then play with right-click-and-drag-the-cursor to get a good-looking image.)
What are the x,y coordinates of the supernova? 521, 528

What is the signal in the central pixel of the supernova? (units of counts) 2031

In the central pixel of the host galaxy? 3949

Save an image of your screen (under the "File" menu of ds9). Save it as a jpeg. Make the filename yourname.objectname.jpg
(eg JohnWayne.SN2000ab/jpg). Copy this file into the results_YourName folder.

Star Counts

Star counts (the number of stars found per square degree, down to some limit of faintness) are a useful way of understanding the structure of the Milky Way galaxy, in which we live. They reveal mostly large-scale structure, however: over small angular separations, the number density of stars seldom changes by very much.

Use **ds9** to look at the image **lab001.fit**
Notice that the image shape is different from **lab000.fit**: this image was taken with a different telescope and CCD camera.
There are many other differences, best seen by displaying the FITS header. Do this.

What are the dimensions of the **lab001.fit** image (in pixels)? 1536 (x) × 1024 (y)

What filter was used? Green

What was the exposure time? 30. s

Now open a new unix terminal window, **cd catalogs**, and use vi (or another editor) to look at **lab001.cat**. This file was written by a program **SExtractor** that analyzes an image to locate star-like objects, and reports their positions and fluxes (in this context, the flux is the total number of counts found within the fuzzy boundaries of each object). Normally this program reports a lot of additional information for each star, but these details are not needed yet. Notice there are *header* lines (3 in this case), one describing each column of the output. After this we find one line per object detected, each line containing all of the information about that object.

Choose 4 stars from the table, with flux values ranging from about 1000 to a few times 100,000. Find each star in the ds9 image, and In a table below, enter x, y, flux, and (measured with ds9) the value of the central pixel for each star.

241.8	649.7	2005	91
554.7	637.5	24468	723
1036.3	560.4	61875	1695
542.7	235.3	178952	4585

Give a (rough) relation between the central pixel value and the flux. (For now do this by eye. We will do better later.) Why are the flux and the central intensity different from each other?

flux ≃ 35 × (central pixel)

The number of stars (down to whatever limit SExtractor recognizes) may be obtained from the number of lines in each image's corresponding .cat file. Use the **wc** command to count lines in each of the files **lab001.cat** through **lab009.cat**. Below, make a table with filenames on different rows, showing the number of counted stars for each. (3 lines are header, not stars)

file	N		$\sqrt{N}$	N_{Top_half}	$N/2 - N_{Top\_half}$	$\sqrt{N/2}$
001	218	±	15	115	-6.0	10
002	204	±	14	104	-2.0	10
003	223	±	15	107	4.5	11
004	244	±	16	112	10.0	11
005	278	±	17	142	-3.0	12
006	267	±	16	147	-13.5	12
007	249	±	16	126	-1.5	11
008	290	±	17	155	-10.0	12
009	309	±	18	155	-0.5	12

Taylor (p. 48) says: "The uncertainty in any counted number of random events, as an estimate of the true average number, is *the square root of the counted number*." Enter this uncertainty for each image in the table above. Use Taylor's (chap 2) rules about how to express uncertainties.

We can test to see if the square-root-of-the-number-counted rule is followed by these star counts as follows: Suppose we count N stars in the observed field. If the true local density of stars down to some faintness limit is N stars per (area of our field of view), then in half of the area, we expect to count N/2 +/- sqrt(N/2).

This trick of subdividing the data to establish consistency has broad utility, and researchers in all fields of science use it a lot.

To make a pretty good count of stars in only one-half of the field of view for each file, examine the corresponding catalog file, eg **lab001.cat**. Notice that the stars are listed (almost) in order of decreasing y-coordinate. (*There are some obvious exceptions to this rule near the top and bottom of the file, and in the main body of the file you will also find occasional items that are out of sequence. Ignore these for now.*) Because of this ordering, we can estimate the number of stars in the top half (with larger y-coordinates) by counting the number of lines that lie above the first line having ($y < N_y/2$) where $N_y = 1024$ = the y dimension of the detector, measured in pixels.

Describe how to do this line count within the vi editor (or other editor of your choice). Find the last line with $y \geq 512$. Find the number of this line with ctrl-G. Subtract 3 (to account for header lines).

For each of the catalog files **lab001.cat** to **lab009.cat**, compute the difference between the expected and actual number of stars in the top half of the image, ie, ($N/2 - N_{top\_half}$) and enter both this value and sqrt(N/2) in your table on the previous page.

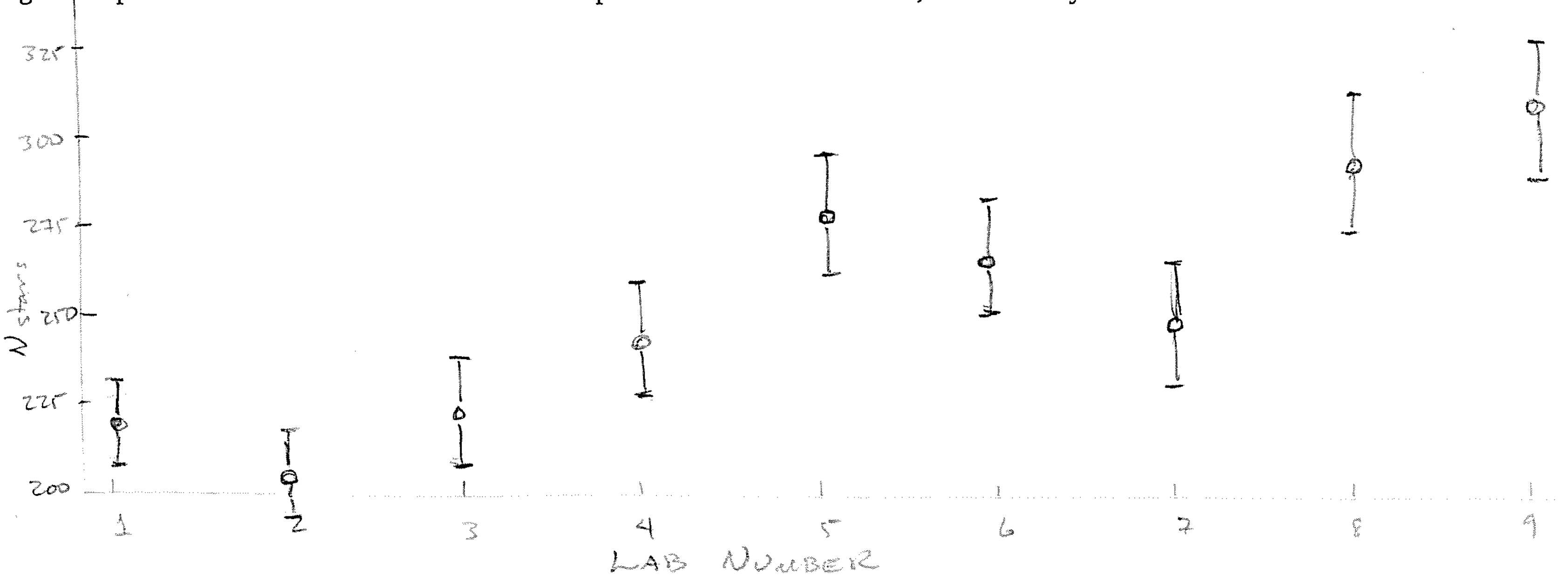
Are the differences you compute generally consistent with the square-root law? Are they *all* consistent? If not, is this a problem? Explain. The differences are broadly consistent with the expected $\sqrt{N/2}$ errors. If anything, the differences $N/2 - N_{TopHalf}$ are a little smaller than expected. As we will see, one normally expects that about 1/3 of the data points fall more than 1 expected error from the best value; here we see only 1/9. With

The out-of-sequence stars (ones that do not go in strict order of decreasing y-coordinate) are a source of error in the above measurement. In your judgment, is the error that they cause (a) Significant, (b) Not very important, or (c) Exactly zero? Justify your answer.

a sample this small, however, anything can happen.

(b) Not very important. The typical effect of sorting the lines into order would change the estimates of $N_{TopHalf}$ by ± a few, a smaller effect than the expected $\sqrt{N/2}$ errors.

Below, draw a plot that shows the star counts as a function of the image number 1-9. (This actually makes sense, because the images were obtained equally spaced along an arc of a great circle in the sky.) Label your axes, and choose the plot ranges to make the plot as legible as possible. Indicate the uncertainties with plotted error bars. For now, do all this by hand.



Do you think that the star density measured in the area around **lab009.fit** is larger than in **lab001.fit,** or is the difference in measured value a statistical fluke? Justify your answer. The star count in lab009 exceeds that in lab001 by 91 ± 35 stars (Taylor's provisional rule 2.18). The difference is about 2.6 times the expected error, which is quite unlikely to happen by accident. Thus, the star density in lab009 probably really is bigger than in lab001.

Do you think that the star density may **not** be increasing monotonically from **lab001.fit** to **lab009.fit** (yes/no)? Justify your answer.

Yes, it may not be. The average increase from lab001 to lab009 is clear, but the data suggest a small deviation from monotonicity near points 6-7. Evidence for the reality of this feature is not compelling, but cannot be ignored.

Do you think that **lab010.fit** might be the next image in the series **lab001.fit** to **lab009.fit**, or is it something different? Justify your answer.

Something different – i.e., drawn from a different distribution. lab010 has 477 stars, which differs from the avg of lab001-lab009 (~254) by 5.9 times the expected error in the difference.