ASTRONOMY 535

OBSERVATIONAL TECHNIQUES

CLASS NOTES

Spring 2003

Instructor: Rene Walterbos

Lecture notes from Jon Holtzman, with some modifications/additions by RW

Entire file in one Postscript file.

1 Class introduction

- Class content
  - how to optimally collect astronomical data and extract information, in principle and **in practice**
  - philosophy: need to understand how instruments work to understand information, accuracy, and to get most from the data. Consequently, we will discuss instruments. For us, the term instrument includes the telescope, camera, and the detector. Principally we will discuss imaging instruments and spectrographs. We may briefly discuss some other types of instruments (imaging spectroscopy, interferometry). We will cover instrumentation/methodology for optical/IR observations; radio and high energy covered in a separate class offered occasionally.
  - NOTE: hardly anything (telescopes, instruments, software) "just works" despite what other people say.
  - Not a class on how to use data reduction packages, although you will need to use these. It will be your responsibility to learn these by trying them, reading documentation, and talking with other graduate students.
  - NOTE: to get most from data (and often, to stand apart from the "crowd", you need to develop refined or new techniques of data analysis. Beware of canned procedures. Also, ALL instruments generally present slightly new problems which demand different techniques.
  - practical emphasis. Work with real data. Use NMSU telescopes/data when possible.
  - Note changing face of astronomy, e.g. automated 1m, HST, SDSS
- Work load
  - Roughly biweekly problems/tasks, including perhaps two projects, which might involve data taking, data reduction, and/or a simple optics/instrumentation project. Note graduate/undergrad distinction (if there are any undergrads in class).
  - Experience all of: observing, reduction, analysis, computing
  - Two quizes, one final,

  - Texts: Schroeder, Howell, Kitchin, Bevington, Numerical Recipes
  - Notes on Web

- NMSU telescopes:
  - ARC 3.5m: consortium with U Wash, U Chicago, Princeton, JHU, NMSU, U. Colorado. NMSU gets 15%.
  - NMSU 1m: status
  - NMSU Tortugas mountain observatory : 24" with CCD

- trip to APO?

- Instruments:
  - 3.5m:
    - DIS - double imaging spectrograph
    - GRIM - IR imager / spectrograph
    - SPICAM - drift scan camera
    - echelle spectrograph
      under development:
    - new IR instrument(s)
    - Wide field imager? Or multi-color imager?
  - 1m: 512x512 CCD camera, 1024x1024 CCD camera, new 2K x 2K camera under construction, under collaboration with Los Alamos National Laboratory.

- APO home page - http://www.apo.nmsu.edu

- NMSU astronomy home page: http://astro.nmsu.edu

- Homework
  - Look at the APO Web site. Look at observing schedule and some observing logs. Check out the 10 micron all-sky camera images. Look through the Instruments pages, and at anything else which might interest you!
  - Look at the NOAO Web site (http://www.noao.edu)
  - Look at the STScI Web site (http://www.stsci.edu)

  - Look at the SDSS Web site (http://www.sdss.org)

- Get experience of students: observing, data reduction, analysis, computing
  - HOMEWORK: assessment problems.

2 Properties of light and error analysis

Properties of light

- astronomy - learn about objects outside of our atmosphere. need information from them to do so: light!

- We can describe the amount of light an object emits or that we receive by three fundamental quantities: luminosity, L, intensity or surface brightness, I, and flux F. Each of these can be defined monochromatically, e.g. I_λ, or I_ν, or in integrated form where e.g.

$$F = \int F_\lambda d\lambda = \int F_\nu d\nu$$

  The monochromatic intensity may be referred to as specific intensity, while monochromatic flux may be called flux density, especially in radio astronomy.

- The intensity is defined as the amount of energy transmitted through a unit surface element per unit time per unit frequency (or wavelength) into a unit solid angle in the direction (θ, ϕ), where θ is the angle away from the normal to the surface element, and ϕ the azimuthal angle. The flux is the amount of energy passing through a unit surface element in all directions, defined by

$$F_\nu = \int I_\nu cos(\theta) d\Omega$$

  where $d\Omega$ is the solid angle element, and the integration is over the entire solid angle.

  The luminosity is the *intrinsic* energy emitted by the source per second. Thus we can also consider it as the power. For an isotropically emitting source,

$$L = 4\pi d^2 F$$

  where $d =$ distance to source.

- What do we measure for sources? For unresolved sources, we measure the flux, usually over some bandpass or wavelength interval, for resolved sources we can directly measure their intensity distribution on the sky. Only if we know the distance to a source, can we calculate the luminosity of that source.

- Questions: how do the three quantities, L, I, and F depend on distance to the source? What are the units of these three quantities? (Note: there are many variations and you should make sure you understand what is adopted by someone presenting data). There is also no concensus between disciplines in the terminology, so while astronomers are reasonably consistent in the use of these terms, you may run into things like radiant flux, irradiance, radiance, spectral intensity, and others, which are all related but not exactly equal to intensity or flux.

  To what quantity is apparent magnitude for a star related? To what quantity the absolute magnitude?

- Light, by quantum mechanics, is photons, has characteristics of both waves and particles. Wavelength/frequency corresponds to energy:

$$E = h\nu = \frac{hc}{\lambda}$$

- electromagnetic spectrum: gamma rays - X rays - UV - optical - IR - mm - radio

- from astronomical sources, much astronomy is done by measuring $flux = ergs/cm^2/sec$ Other general branch is study of spatial distribution, i.e. astrometry, morphology.

- for most sources, flux is a function of wavelength, so we actually are often interested in estimates of the monochromatic flux, or flux density = flux / unit wavelength. F_ν : flux per unit frequency. F_λ : flux per unit wavelength.

$$\int F_\nu d\nu = -\int F_\lambda d\lambda$$

$$F_\nu = -F_\lambda d\lambda/d\nu$$

$$= F_\lambda \frac{c}{\nu^2} = F_\lambda \frac{\lambda^2}{c}$$

- Terminology of flux measurements:

  - photometry (broad-band flux measurement),
  - spectroscopy (relative measurement of fluxes at different wavelengths),
  - spectrophotometry (absolute measurement of fluxes at different wavelengths).

- Generally, we measure flux, not flux density, though for narrow band and spectroscopy they're almost equivalent. Also, with most detectors, we measure photon flux (photon counting devices), rather than energy flux (bolometers). The photon flux is given by

$$\text{photonflux} = \text{flux/energyperphoton}$$
$$= \int F_\lambda \frac{\lambda}{hc} d\lambda$$

Magnitudes and the count equation

In astronomy, however, magnitudes are often used instead of fluxes:

$$m = -2.5 \log F + const$$

The flux density for astronomical objects may be specified by a monochromatic magnitude:

$$F_\lambda = K 10^{-0.4 m(\lambda)}$$

where the coefficient of proportionality, K, depends on the definition of photometric system. In the STMAG system, $K = 3.6E - 9 ergs/cm^2/s/\AA$, which is the flux of Vega at 5500Å. Alternatively, we can write

$$m_\lambda = -2.5 \log F_\lambda - 21.1$$

In the ABNU system, we have

$$F_\nu = 3.6 \times 10^{-20} erg/cm^2/s/Hz 10^{-0.4 m_\nu}$$

or

$$m_\nu = -2.5 \log F_\nu - 48.6$$

Related, but slightly different, are integrated magnitude systems; the STMAG and ABMAG integrated systems are defined relative to sources of constant F_λ and F_ν systems, respectively.

$$m_{ST} = -2.5 \log \frac{\int F_\lambda d\lambda}{\int 3.63 \times 10^{-9} d\lambda}$$

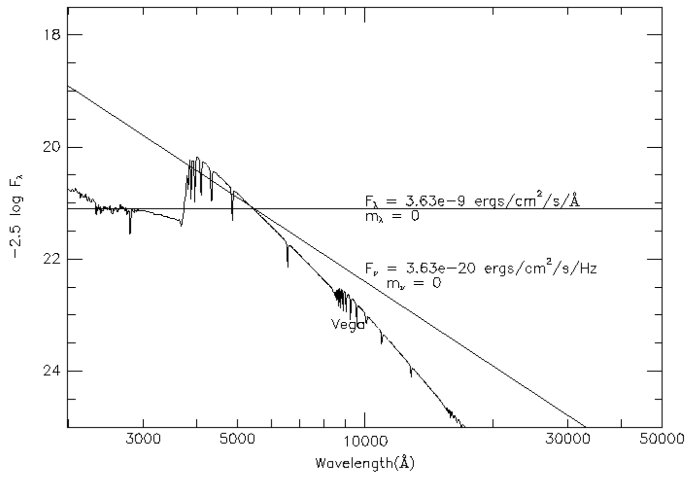


Figure 1: Differences between different magnitude systems

$$m_{AB} = -2.5 \log \frac{\int F_\nu d\nu}{\int 3.63 \times 10^{-20} d\nu}$$

The standard UBVRI broadband photometric system, as well as many other magnitude systems, however, are not defined for a constant F_λ or F_ν spectrum; rather, they are defined relative to the spectrum of an A0V star. Most systems are defined (or at least were originally) to have the magnitude of Vega be zero in all bandpasses.

For the broadband UBVRI system, we have

$$m_{UBVRI} \approx -2.5 \log \frac{\int_{UBVRI} F_\lambda(object) d\lambda}{\int_{UBVRI} F_\lambda(Vega) d\lambda}$$

Figure 1 shows a plot to demonstrate the difference between the different systems.

Given a magnitude of an object in the desired band and absolute spectrophotometry of Vega (Hayes 1985, Calibration of Fundamental Stellar Quantities, IAU Sym 111), one can estimate the flux. Alternatively, if one is given just the V magnitude of an object, a spectral energy distribution (SED), and the filter profiles (Bessell 1990, PASP 102, 1181), one can compute the integral of the SED over the V bandpass, determine the scaling by comparing with the integral of the Vega spectrum over the same bandpass, then use the normalized SED to compute the flux in any desired bandpass. Some possibly useful references for SEDs are: Bruzual, Persson, Gunn, & Stryker; Hunter, Christian, & Jacoby; Kurucz).

The photon flux from the source is not exactly what you measure. To get the number of photons that you count in an observation, you need to take into account the area of your photon collector (telescope), photon losses and gains from the Earth's atmosphere, and the efficiency of your collection/detection apparatus. Generally, the astronomical signal can be written

$$S = Tt \int \frac{F_\lambda}{\frac{hc}{\lambda}} q_\lambda a_\lambda d\lambda$$

where T is the telescope collecting area, t is the integration time, a_λ is the atmospheric transmission (next lecture) and q_λ is the system efficiency.

Usually, however, one doesn't use this information to go backward from S to F_λ because it is very difficult to measure all of the terms precisely. Instead, most observations are performed differentially to a set of *standard* stars, which themselves have been (usually painfully) calibrated. We'll discuss this process in more depth later.

The count equation is used, however, very commonly, for computing the approximate number of photons you will receive from a given source in a given amount of time for a given observational setup. This number is critical to know in order to estimate your expected errors and exposure times in observing proposals, observing runs, etc. Understanding errors in absolutely critical in all sciences, and maybe even more so in astronomy, where objects are faint, photons are scarce, and errors are not at all insignificant.

Errors in photon rates

For a given rate of emitted photons, there's a probability function which gives the number of photons we detect, even assuming 100% detection efficiency, because of *statistical* uncertainties. In addition, there may also be *instrumental* uncertainties. Consequently, we now turn to the concepts of probability distributions, with particular interest in the distribution which applies to the detection of photons.

Distributions and characteristics thereof

- concept of a distribution : define $p(x)dx$ as probability of event occuring in $x + dx$:
$$\int p(x)dx = 1$$

Some definitions relating to values which characterize a distribution:

$$\text{mean} \equiv \mu = \int xp(x)dx$$

$$\text{variance} \equiv \sigma^2 = \int (x-\mu)^2 p(x)dx$$

$$\text{standarddeviation} \equiv \sigma = \sqrt{variance}$$

median : mid-point value.

$$\frac{\int_{-\infty}^{x_{median}} p(x)dx}{\int_{-\infty}^{\infty} p(x)dx} = \frac{1}{2}$$

mode : most probable value

Note that the geometric interpretation of above quantities depends on the nature of the distribution; although we all carry around the picture of the mean and the variance for a Gaussian distribution, these pictures are not applicable to other distributions.

Also, note that there is a difference between the *sample* mean, variance, etc. and the *population* quantities. The latter apply to the true distribution, while the former are estimates of the latter from some finite sample of the population. The sample quantities are derived from:

$$\text{samplemean} : \bar{x} \equiv \frac{\sum x_i}{N}$$

$$\text{samplevariance} \equiv \frac{\sum (x_i - \bar{x})^2}{N-1}$$

$$= \frac{\sum x_i^2 - (\sum x_i)^2/N}{N-1}$$

The sample mean and variance approach true mean and variance as N approaches infinity.

Now we consider what distribution is appropriate for the detection of photons. The photon distribution comes from the *binomial* distribution, which gives the probability of observing the number, x, of some possible event, given a total number of events n, and the probability of observing the particular event (among all other possibilities) in any single event, p:

$$P(x,n,p) = \frac{n!p^x(1-p)^{n-x}}{x!(n-x)!}$$

For the binomial distribution, one can derive:

$$\text{mean} \equiv \int xp(x)dx = np$$

$$\text{variance} \equiv \sigma^2 \equiv \int (x-\mu)^2 p(x) dx = np(1-p)$$

In the case of detecting photons, n is the total number of photons emitted, and p is the probability of detecting a single photon out out of the total emitted. We don't know either of these numbers! However, we do know that $p << 1$ and we know, or at least we can estimate, the mean number detected:

$$\mu = np$$

.

In this limit, the binomial distribution asymtotically approaches the *Poisson* distribution:

$$P(x,\mu) = \frac{\mu^x \exp^{-\mu}}{x!}$$

From the expressions for the binomial distribution in this limit, the mean of the distribution is μ, and the variance is

$$\text{variance} = \sum_x [(x-\mu)^2 p(x,\mu)]$$

$$= np = \mu$$

This is an *important result*.

Note that the Poisson distribution is generally the appropriate distribution not only for counting photons, but for *any* sort of counting experiment.

Figure 2 shows some plots of the Poisson distribution for $\mu = 2, \mu = 10$, and $\mu = 10000$.

Note, for large μ, the Poisson distribution is well-approximated around the peak by a *Gaussian*, or *normal* distribution:

$$P(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

This is important because it allows us to use "simple" least squares techniques to fit observational data, because these generally assume normally distributed data. However, beware that in the tails of the distribution, the Poisson distribution can differ significantly from a Gaussian distribution.

The normal distribution is also important because many physical variables seem to be distributed accordingly. This may not be an accident because of the central limit theorem: if a quantity depends on a number of independent random variables

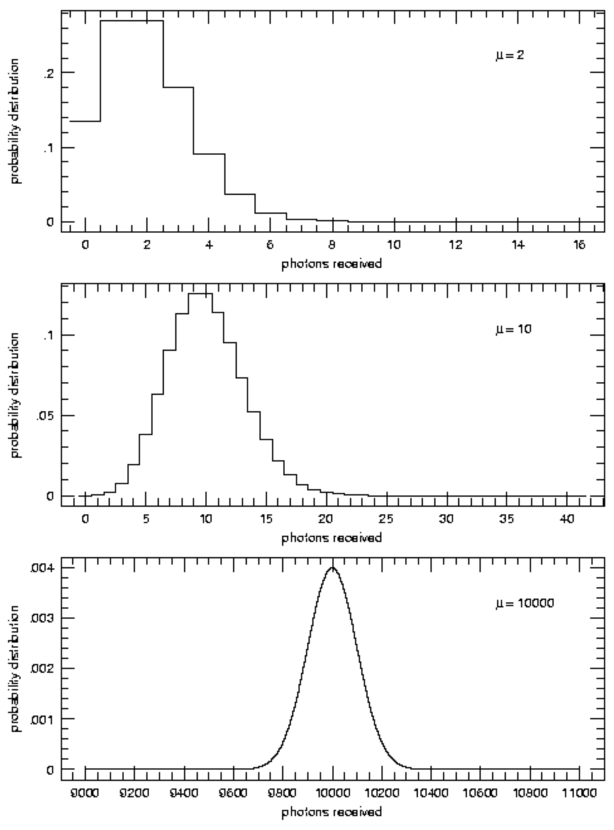


Figure 2: The Poisson distribution for three values of μ

.

with ANY distribution, the quantity itself will be distributed normally. (see statistics texts).

Importance of error distribution analysis

You need to understand the expected errors in your observations in order to:

- predict the amount of observing time you'll need to get errors as small as you need them to do your science,

- answer the question: is observed data consistent with expected errors? If the answer is no, they you know you've either learned some astrophysics or you don't understand something about your observations. This is especially important in astronomy where objects are faint and many projects are pushing down into the noise as far as possible. Really we can usually only answer this probabilistically. Generally, tests compute the probability that the observations are consistent with the expected distribution (the null hypothesis). You can then look to see if this probability is low, and if so, you can reject the null hypothesis.

Confidence levels

For example, say we want to know whether some single point is consistent with expectations, e.g., we see a bright point in multiple measurements of a star, and want to know whether the star flared. Say we have a time sequence with known mean and variance, and we obtain a new point, and want to know whether it is consistent with known distribution.

In the limit where the observed distribution is approximately Gaussian, we can calculate the *error function*, which is the integral of a gaussian. From this, we can determine the probability of obtaining some point given the known distribution.

Some simple guidelines to keep in mind (the actual situation often requires more sophisticated statistical tests) follow. First, for Gaussian distributions, you can calculate that 68% of the points should fall within plus or minus one sigma from the mean, and 95.3% between plus or minus two sigma from the mean. Thus, if you have a time line of photon fluxes for a star, with N observed points, and a photon noise σ on each measurement, you can test whether the number of points deviating more than 2σ from the mean is much larger than expected. To decide whether any single point is really significantly different, you would want to use a 5σ rather than a 2σ criterion. The 5σ has much higher level of significance of course. On the other hand, if you have far more points in the range $2 - -4\sigma$ brighter or fainter than you would expect, you may also have a significant detection of intensity variations (provided you really understand your uncertainties on the measurements!).

Noise equation: how do we predict expected errors?

Given an estimate the number of photons expected from an object in an observation, we then can calulate the *signal-to-noise* ratio:

$$\frac{S}{N} = \frac{S}{\sqrt{\sigma^2}}$$

which gives the inverse of the predicted fractional error (N/S).

Consider an object with photon flux (per unit area and time), S', leading to a signal, $S = S'Tt$ where T is the telescope area and t is the exposure time. In the simplest case, the only noise source is Poisson statistics from the source, in which case:

$$\sigma^2 = S = S'Tt$$

$$\frac{S}{N} = \sqrt{S'Tt}$$

In other words, the S/N increases as the square root of the object brightness, telescope area, efficiency, or exposure time.

A more realistic case includes the noise contributed from Poisson statistics of background light (next lecture), B', which has units of flux per area on the sky (also usually given in magnitudes).

$$B' = \int \frac{B_\lambda}{\frac{hc}{\lambda}} q_\lambda d\lambda$$

The amount of background in our measurement depends on how we choose to make the measurement (how much sky area we observe). Say we just use an aperture with area, A.

The total number of photons observed, F, is

$$F = (S' + B'A)Tt$$

The variance of the total flux, from Poisson statistics, is:

$$\sigma_F^2 = (S'Tt + B'ATt)$$

since the variances add for object+background. To get the desired signal, we will need to measure separately the total signal and the background signal to estimate:

$$S \equiv S'Tt = F - B$$

where B is some estimate we have obtained of the background level multiplied by the area A ($B = B'ATt$). The noise in the measurement is

$$\sigma_S^2 \approx \sigma_F^2 = (S' + B'A)Tt$$

where the approximation is accurate if the background is determined to high accuracy. This leads to a common form of the *noise equation*:

$$\frac{S}{N} = \frac{S'}{\sqrt{S' + B'A}}\sqrt{T}\sqrt{t}$$

In the *signal-limited* case, $S' >> B'A$, we get

$$\frac{S}{N} = \sqrt{S'tT}.$$

In the *background limited* case, $B'A >> S'$, and

$$\frac{S}{N} = \frac{S'}{\sqrt{B'A}}\sqrt{tT}.$$

S/N drops faster when you're background limited. This illustrates the importance of dark-sky sites, and also the importance of good image quality.

Consider two telescopes of collecting area, T_1 and T_2. If we observe for the same exposure time on each and want to know how much fainter we can see with the larger telescope at a given S/N, we find:

$$S_2 = \frac{T_1}{T_2} S_1$$

for the signal-limited case, but

$$S_2 = \sqrt{\frac{T_1}{T_2}} S_1$$

for the background-limited case.

Instrumental noise

In addition to the errors from Poisson statistics (statistical noise), there may be additional terms from instrumental errors. A prime example of this using CCD detectors is readout noise, which is additive noise (with zero mean) which comes from the detector, and is independent of signal level. For a detector whose readout noise is characterized by σ_{rn},

$$\frac{S}{N} = \frac{S'Tt}{\sqrt{S'Tt + B'ATt + \sigma_{rn}^2}}$$

if a measurement is made in a single pixel. If an object is spread over N_{pix} pixels, then

$$\frac{S}{N} = \frac{S'Tt}{\sqrt{S'Tt + B'ATt + N_{pix}\sigma_{rn}^2}}$$

For large σ_{rn}, the behavior is the same as the background limited case. This makes it clear that if you have readout noise, image quality (and/or proper optics to keep an object from covering too many pixels) is important for maximizing S/N. It is also clear that it is critical to have minimum read-noise for low background applications (e.g., spectroscopy).

There are other possible additional terms in the noise equation, arising from things like dark current, digitization noise, errors in sky determination, errors from photometric technique, etc.; we'll discuss some of these later on.

Error propagation

Consider what happens if you have several known quantities with known error distributions and you combine these into some new quantity: we wish to know what the error is in the new quantity.

$$x = f(u, v, ...)$$

The question is what is σ_x if we know σ_u, σ_v, etc.?

As long as errors are small:

$$x_i - < x > \sim (u_i - < u >)\left(\frac{\partial x}{\partial u}\right) + (v_i - < v >)\left(\frac{\partial x}{\partial v}\right) + ...$$

$$\sigma_x^2 = lim(N - > \infty)\frac{1}{N}\sum(x_i - < x >)^2$$

$$= \sigma_u^2\left(\frac{\partial x}{\partial u}\right)^2 + \sigma_v^2\left(\frac{\partial x}{\partial v}\right)^2 + 2\sigma_{uv}^2\frac{\partial x}{\partial u}\frac{\partial x}{\partial v}$$

The last term is the *covariance*, which relates to whether errors are /it correlated.

$$\sigma_{uv}^2 = lim(n - > \infty)\frac{1}{N}\sum(u_i - < u >)(v_i - < v >)$$

If errors are uncorrelated, then $\sigma_{uv} = 0$ because there's equal chance of getting opposite signs on v_i for any given u_i.

Examples:

- addition/subtraction:

$$x = u + v$$

$$\sigma_x^2 = \sigma_u^2 + \sigma_v^2 + 2\sigma_{uv}^2$$

- multiplication/division:

$$x = uv$$

$$\sigma_x^2 = v^2\sigma_u^2 + u^2\sigma_v^2 + 2uv\sigma_{uv}^2$$

- logs:

$$x = \ln u$$
$$\sigma_x^2 = \sigma_u^2/u^2$$

Distribution of resultant errors

When propagating errors, even though you can calculate the variances in the new variables, the distribution of errors in the new variables is not, in general, the same as the distribution of errors in the original variables, e.g. if errors in individual variables are normally distributed, errors in output variable is not necessarily.

When two variables are added, however, the output is normally distributed.

Determining sample parameters: averaging measurements

We've covered errors in single measurements. Next we turn to averaging measurements. say we have multiple observations and want the best estimate of the mean and variance of the population, e.g. multiple measurements of stellar brightness. Here we'll define the best estimate of the mean as the value which maximizes the likelihood that our estimate equals the true parent population mean.

For equal errors, this estimate just gives our normal expression for the sample mean:

$$\mu = \frac{\sum x_i}{N}$$

Using error propagation, the estimate of the error in the sample mean is given by:

$$\sigma_\mu^2 = \sum \frac{\sigma_i^2}{N^2} = \frac{\sigma^2}{N}$$

But what if errors on each observation aren't equal, say for example we have observations made with several different exposure times? Then we determine the sample mean using a:

$$\text{weightedmean} = \frac{\sum \frac{x_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}}$$

and the estimated error in this value is given by:

$$\sigma_\mu^2 = \sum \frac{\frac{\sigma_i^2}{\sigma_i^4}}{(\sum \frac{1}{\sigma_i^2})^2} = \frac{1}{\sum \frac{1}{\sigma_i^2}}$$

where the σ_i's are individual weights.

This is fine, but how do we go about choosing the weights, σ_i? We know we can *estimate* σ using Poisson statistics for a given count rate, but remember that this is a sample variance (based on a single observation!) not the true variance. This can lead to biases.

Consider observations of a star made on three nights, with measurements of 40, 50, and 60 photons. It's clear that the mean observation is 50 photons. However, beware of the being trapped by your error *estimates*. From each observation alone, you would estimate errors of $\sqrt{40}$, $\sqrt{50}$, and $\sqrt{60}$. If you plug these error estimates into a computation of the weighted mean, you'll get a mean rate of 48.64!

Using the individual estimates of the variances, we'll bias values to lower rates, since these will have estimated higher S/N.

Note that it's pretty obvious from this example that you should just weight all observations equally. However, note that this certainly isn't always the right thing to do. For example, consider the situation in which you have three exposures of different exposure times. Here you clearly want to give the longer exposures higher weight. In this case, you again don't want to use the individual error estimates or you'll introduce a bias. There is a simple solution here also: just weight the observations by the exposure time. This works fine for Poisson errors (variances proportional to count rate), but not if there are instrumental errors as well which don't scale with exposure time. For example, the presence of read noise can have this effect. With read noise, the longer exposures should be weighted even higher than expected for the exposure time ratios. The only way to properly average measurements in this case is to estimate a sample mean, then use this value scaled to the appropriate exposure times as the input for the Poisson errors.

Can you split exposures?

When observing, one must consider the question of how long individual exposures should be. There are several reasons why one might imagine that it is nicer to have a sequence of shorter exposures rather than one single longer exposure (e.g., tracking, monitoring of photometric conditions, cosmic ray rejection), so we need to consider whether doing this results in poorer S/N.

Consider the object with photon flux S', background surface brightness B', and detector with readout noise σ_{rn}. A single short exposure of time t has a variance:

$$\sigma_S^2 = S'Tt + B'ATt + N_{pix}\sigma_{rn}^2$$

If N exposures are summed, the resulting variance will be

$$\sigma_N^2 = N\sigma_S^2$$

If a single long exposure of length Nt is taken, we get

$$\sigma_L^2 = S'TNt + B'ATNt + N_{pix}\sigma_{rn}^2.$$

The ratio of the noises, or the inverse ratio of the S/N (since the total signal measured is the same in both cases), is

$$\frac{\sigma_N}{\sigma_L} = \sqrt{\frac{S'TNt + B'ATNt + NN_{pix}\sigma_{rn}^2}{S'TNt + B'ATNt + N_{pix}\sigma_{rn}^2}}$$

In the signal- or background-limited regimes, exposures can be added with no loss of S/N. However, if readout noise is significant, then splitting exposures leads to reduced S/N.

Random errors vs systematic errors

So far, we've been discussing *random* errors. There is an additional, usually more troublesome, type of errors known as *systematic* errors. These don't occur randomly but rather are correlated with some, possibly unknown, variable relating to your observations.

EXAMPLE : flat fielding

EXAMPLE : WFPC2 CTE

Note also that in some cases, systematic errors can masquerade as random errors in your test observations, but actually be systematic in your science observations

EXAMPLE: flat fielding, subpixel QE variations.

Note that error analysis from expected random errors may be the only clue you get to discovering systematic errors. To discover systematic errors, plot residuals vs. everything!

3 Intermezzo on Fourier Transforms, Convolution, and Sampling

Fourier Transforms

- Useful reference: R. Bracewell, The Fourier Transform and its Applications. An excellent monograph on this topic. Also includes discussion of the Fast Fourier Transform algorithm, which enabled rapid execution of FTs on the computer.

- The concept of a Fourier Transform (FT) of a function goes back to the idea of trying to represent a function by a sum of sines and or cosines of various frequencies. By having such a sum, one then also knows which frequencies the function contains, so the FT is related to the "power spectrum" of a function. The frequencies here are the inverse of the function variable, e.g. a function of time has frequencies in units of s^{-1}, a function of position has spatial frequencies in inverse units of length. We talk of "spatial frequencies" in that case.

- The significance of FTs is that many FT-relations exist between quantities connected with astronomical observing and data analysis, especially in interferometry, but also in optics of single telescopes, and the theory behind the "seeing" for the Earth's atmosphere. In addition, concepts such as FTs, convolutions, (auto)correlations, power spectra, etc play a role in signal and image analysis.

- The general definition of the FT of a function $f(x)$ is:

$$F(s) = \int_{-\infty}^{+\infty} f(x) e^{-i2\pi xs} dx$$

.

Note here the link with the cos and sin series, since

$$e^{i\theta} = cos\theta + i\ sin\theta$$

The inverse transform is given by:

$$f(x) = \int_{-\infty}^{+\infty} F(s) e^{+i2\pi xs} dx$$

.

If you combine these two relations, you express a function f(x) as a double integral of itself. Is this true for all functions? The mathematical conditions are often not met for real signals, but as long as the integral of $|f(x)|$ over the entire range of $f(x)$ is finite, and any discontinuities in $f(x)$ are finite.

Convolution, Cross-Correlation, Auto-Correlation

- A convolution of one function with another involves a smoothing or averaging of that function, generally resulting in loss of spatial or time resolution. Many processes in obtaining data involve convolutions, for example the blurring of images caused by the atmosphere is a convolution of the original image with a "point spread function" that broadens the images. In data analysis, simple actions such as taking running means of data involve a convolution, and the effects of such actions can be easily understood given the connection between convolutions and FTs that we will discuss later on.
- Mathematically, a convolution is defined as follows:

$$f(x) * g(x) = \int f(u) g(x-u) du$$

where the integration is over the entire range in u.

Convolving thus involves sliding the mirror image of one function along the x-axis, taking the product of that mirrored function and the other function at each position and integrating the area under this product (Draw this process for some simple functions, and visualize what happens!). By replacing g(x) with a simple running mean (e.g. a sequence such as (0.2, 0.2, 0.2, 0.2, 0.2) it is easy to see that this running mean is a convolution of f(x) with this sequence of numbers (in this case replacing each value of f(x) with the mean value of 5 points centered on f(x)). Note: the numbers here are normalized such that their sum is 1 (so the integral $\int_{-\infty}^{+\infty} g(u)du$ is 1. This preserves the overall original scale of values of the function f(x), for example in the case of imaging data, the intensities would be conserved. Since all data is digital, it doesn't hurt to replace "functions" by arrays of numbers when thinking about taking convolutions; FTs can also be obtained digitally, through the Fast Fourier Transform algorithm.

- In most cases the convolution functions may be symmetric, in which case it makes no difference whether the function is mirrored or not; however, it is important to recognize that in real life this mirroring does take place in actual situations involving a convolution. This reversal of g(u) to g(x-u) also implies that convolution is *commutative*; $f(x) * g(x) = g(x) * f(x)$. Convolution is also *associative*, i.e. $f * (g * h) = (f * g) * h$ and distributive over addition, $f * (g + h) = f * g + f * h$.

- If we do not reverse g(u) to g(x-u), we obtain the cross-correlation of the two functions:

$$f \star g = \int f(u)g(u - x)du$$

  This process is not commutative. Cross-correlating signals allows us to determine how similar the signals are. Likewise, cross-correlation can be used to find sources in data; for example, by cross-correlating the data with an idealized point source function one is in principle able to identify locations and amplitudes of sources as peaks in the cross-correlation map.

- If we take the cross-correlation of a function with itself, we obtain the *auto correlation*. Like convolution, and cross-correlation, auto correlation involves a loss of information, i.e. it is not possible always to uniquely determine a function if only its auto-correlation is known (the "phase" information is lost). Some instruments only measure the auto-correlation of a signal.

- An important process in astronomy involves the attempts to *deconvolve* signals. This involves trying to obtain higher resolution ("super resolution") than the images or time series provide, by reversing the convolution process. This only works under certain special circumstances; thinking in terms of FTs, one can only obtain higher frequency information on a function (which is what deconvolution is trying to do) if additional information or constraints on the signal are provided.

- The *Energy Spectrum* of a function is given by the square of its FT, $|F(s)|^2$. Just as in the case of an auto-correlation, the spectrum contains no phase information, so there is a loss of information compared to the original function.

Some Special Functions

The relationship between FTs and convolutions is best made clear with the help of some special functions.

- The *rectangle* function, $II(x)$ is defined such that

$$rect(x) = II(x) = 0 \ for \ |x| > 1/2, \ and \ 1 \ for \ |x| < 1/2$$

  Question: what would $4II(x/3)$ look like? How about $II(x-8)$?

  This function when multiplied with another function provides an easy way to cut out sections of other functions.

  Question: what would $II(x) * II(x)$ look like?

- The *sinc* function

$$sinc(x) = \frac{sin\pi x}{\pi x}$$

  This is the filtering or interpolating function; this looks almost, but not quite like a sinusoid of gradually decreasing amplitude. (Plot it!). Its properties include:

  $sinc\ 0 = 1$

  $sinc\ n = 0$ where n is a non-zero integer

  $\int_{-\infty}^{+\infty} sinc(x)dx = 1$

  The *sinc* function is very important. It is the FT of the II function, thus sinc(x) has all frequencies present with equal strength, up to a cutoff frequency. This function, in a convolution, therefore performs perfect low-pass filtering (as we will see when discussing the various theorems). It is also an important interpolation function. If the rectangle function is seen as a single narrow slit (where the *width* of the slit corresponds to the width of the rectangle function), $sinc^2$ describes the Fraunhofer diffraction pattern of a monochromatic light beam falling through the slit. This illustrates in a simple way the Fourier relation between "apertures" and the resulting "images" produced by telescopes.

The two-d analog of the sinc (hence the FT of a square rectangle function in x,y plane) is a Bessel function of the first order. In optics you will learn more about the corresponding analogs of a round 2-d rectangle function (which can be seen as a round telescope mirror).

- *Impulse symbol* or δ-function. Defined by

$$\delta(x) = 0 \ for \ x \neq 0, \int_{-\infty}^{+\infty} \delta(x) dx = 1$$

Note: this function is only defined in an integral, and you should *not* interpret this as $\delta(0) = 1$!

The role of $\delta(x)$ is quite important; any signal narrower than the resolving power of your instrument or telescope is essentially seen as an impulse function, in a mathematical sense. The function also plays an important role in sampling. The FT of $\delta(x)$ is a function that is 1 everywhere (an infinitely sharp function has all frequencies present at equal strength).

Question: what is $\int_{-\infty}^{+\infty} \delta(x-a) f(x) dx$?

Question: what is $\delta(x) * f(x)$?

- *Sampling or replicating* symbol, $III(x)$. This is an infinite sequence of unit impulses, or delta functions:

$$III(x) = \sum_{n=-\infty}^{+\infty} \delta(x-n),$$

and

$$III(x) f(x) = \sum_{n=-\infty}^{+\infty} f(n) \delta(x-n)$$

.

Thus, multiplication with $III(x)$ is equivalent to sampling of the function f(x) at regular intervals. How do we control the interval spacing?

This function is relevant, since in practise we always sample a signal and are never able to measure it completely. Likewise, any reproduction of a singal or data in the computer involves a digitized, sampled representation of the actual data, so all data is sampled. Under convolution, $III(x)$ has a replicating property, reproducing multiple, possibly overlapping, copies of f(x). Try to sketch this yourself. Overlapping will occur if $f(x)$ extends beyond $|x| > 1$. What happens if we convolve $f(x)$ with $III(x/a)$, for different values of a?

The FT of $III(x)$ is $III(s)$!

Basic Theorems and some other Facts

- We will not prove any of the theorems. If you are curious, Bracewell does this in his book.

- The FT of a Gaussian function is another Gaussian. A narrow Gaussian in one domain will have broad FT, and vice versa. Likewise, any "narrow" function will have a "broad" FT. Taking this in the limit, one can approximate a the δ function as an ever narrowing Gaussian of unit integral, and realize that the FT will approach an infinitely broad Gaussian, i.e. a function that is one everywhere, as we stated earlier.

- We can generalize the issue of the relation between the width of one function and that of its FT in the *Similarity Theorem*:

  If $f(x)$ has FT $F(s)$ then $f(ax)$ has FT $\frac{1}{|a|}F(\frac{s}{a})$.

- *Addition Theorem*:

  If $f(x)$ and $g(x)$ have FTs $F(s)$ and $G(s)$, respectively, then $f(x) + g(x)$ has FT $F(s) + G(s)$.

- *Shift Theorem*:

  If $f(x)$ has FT $F(s)$, then $f(x - a)$ has FT $e^{-2\pi ias}F(s)$.

  This implies that a linear shift in one domain corresponds to a phase shift in the FT domain. This makes sense; you are not really changing the frequency content of a function by shifting it, so the amplitude of the FT should not change, only its phase. This theorem finds many examples in optics and radio interferometry; consider a parallel light beam falling on an aperture. To shift the diffracted beam through a small angle, one changes the angle of incidence (i.e. one changes the phase of the illumination across the aperture).

- *Convolution Theorem*:

  If $f(x)$ has FT $F(s)$ and $g(x)$ has FT $G(s)$, then $f(x) * g(x)$ has FT $F(s)G(s)$.

  So, convolution in one domain corresponds to multiplication in the Fourier domain. This theorem makes it often quite easy to visualize the effects of convolutions or FTs, and also to obtain an impression of the FT of complex functions. Often a function can be seen as one or more products or convolutions of simple functions, and successive application of this theorem then makes it easy to visualize what happens in the Fourier domain. Example: an interferometer consisting of multiple dishes can be seen as a single telescope in which many "holes" have been cut, or as a replication of many single telescopes. The impact on the subsequent telescope beam (or PSF), which is related to the FT of the aperture plane, can be seen by applying this theorem.

- Properties related to the convolution theorem:
  - The area under a convolution is equal to the product of the areas under the individual functions:

    $$\int f(x) * g(x) dx = [\int f(x) dx][\int g(x) dx].$$

  - The abscissas of the centers of gravity add under convolution:

    $$< x >_{f*g} = < x >_f + < x >_g$$

    where $< x >_f = \frac{\int x f(x) dx}{\int f(x) dx}$.
  - The second moments add if $< x >_f$ or $< x >_g = 0$. In general:

    $$< x^2 >_{f*g} = < x^2 >_f + < x^2 >_g + 2 < x >_f < x >_g$$

    .

    This has the important consequence that under convolutions, typically the variances add.

- *Rayleighs Theorem*: The integral of the squared modulus of a function is equal to the integral of the squared modulus of its spectrum:

  $$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \int_{-\infty}^{+\infty} |F(s)|^2 ds$$

  .

  Essentially, this is a statement of energy conservation.

- *Autocorrelation Theorem*: If $f(x)$ has FT $F(s)$, then its autocorrelation function $\int_{-\infty}^{+\infty} f^*(u) f(u+x) du$ has FT, $|F(s)|^2$. This is a special case of the convolution theorem. So, the autocorrelation of a signal is the FT of its power spectrum. Some instruments work by recording the autocorrelation of a signal rather than the signal itself (in radio astronomy). By taking the FT of that, one has the power spectrum of the signal. Note that phase information is lost in the auto-correlation or power spectrum, so this is not as informative as recording the actual signal.

- Finally, an important property: The definite integral of a function from $-\infty$ to $+\infty$ is equal to the value of its FT transform at the origin:

$$\int_{-\infty}^{+\infty} f(x)dx = F(0)$$

.

A consequence of this is that in the case of an interferometer that does not include the "zero spacing", hence the 0-meter baseline, the net total recorded flux will be zero. In radio interferometry one speaks of the "missing short baseline problem."

Sampling

- Since all digital analysis of signals requires representation of data over finite sections of data (for example pixels on the sky, or time segments in a time series), the question is what the influence of this "sampling" of the signal is on the final analysis. In addition, the issue comes up of how frequent the signal needs to be sampled to avoid loss of information, or, if loss is unavoidable, what the nature of that loss of information is.

- Under a certain condition, it is possible to recover the intermediate values (hence the actual signal) with full accuracy. The condition is that the function or signal should be *band-limited* which means that its FT is nonzero over a finite range of the transform variable, not over an infinite range. In the case of a band-limited signal, the it sampling theorem states the minimum sampling rate required to fully recover the signal:

A function whose FT is zero for $|s| > s_c$ is fully specified by values spaced at equal intervals not exceeding $1/over2s_c^{-1}$ save for any harmonic term with zeros at the sampling points.

Examples: the *sinc* function is fully specified by the sequence $(x, f(x))$: $....(-2,0),(-1,0),(0,1),(1,0),(2,0),....$ even though all except one sample are 0! On the other hand, the sequence $....(-2,0),(-1,0),(0,0),(1,0),(2,0),...$ does not fully specify the function $sin(\pi x)$, because of the exception regarding harmonic terms stated above.

- In practise, a function or signal may not strictly be bandlimited. However, there is usually a range in s beyond which $F(s)$ becomes negligibly small; in that case we can characterize the error or loss of information resulting from the imperfect sampling. In this case, the measuring device may *impose* the cutoff frequency, so that the FT becomes $II(\frac{s}{2s_c})G(s)$. Here $G(s)$ is the FT of the actual signal and s_c is the cutoff frequency imposed by the measuring device (e.g. the pixel size on a CCD restricts the rate of sampling of the point spread function in imaging).

- Since the FT of $II(\frac{s}{2s_c})G(s)$ is $sinc 2s_c x * g(x)$, the effect imposed by the sampling is to convolve the original signal with a sinc function.

- How we recover the values of the original signal at intermediate, non-sampled points is the subject of *interpolation*. Interpolation is extremely common in astronomical data analysis. For example, one often has to re-grid an image to align it with another image. This re-gridding could involve a change in pixel size, image rotation, and image translation. It is important to recognize that interpolation is rarely perfect, i.e. usually it implies further smoothing (that is, loss of resolution) of the sampled signal. For a band-limited signal or function, convolution with $2s_c sinc(2s_c x)$ will yield values for all x which are exact if the sampling theorem was met.

- If the function is not band-limited, (e.g. Gaussian) then the *sinc* function is not necessarily the best interpolation function; a Gaussian would be better. See .e.g Press etal's monograph on Numerical Recipes for discussion of various interpolation methods.

- In practise, any interpolation will be done by using only a finite range of points in x, so even using a *sinc* interpolation function in practise will not produce the ideal result unless you take this *sinc* along over a very wide range in x. It is of course possible to estimate the consequences of this. Only including the *sinc* over a limited range in x implies that you are multiplying the *sinc* with a rectangle function, so using FT theorems, one may visualize what the consequences can be.

- *Undersampling*. What happens if you undersample a signal or function? See the figure I handed out. (Figures 10.3 and 10.6 in Bracewell). Essentially, undersampling a function causes overlap in the frequency domain, so the higher frequencies are messed up; this is called *aliasing*. Figure 10.6 demonstrates the effect of sampling a function at too low a rate to detect a spike. In astronomy, the situation in e.g. imaging is slight different, since the pixels have finite extent, and adjoin smoothly. The consequence of this is that even if we do undersample (by having pixels too large for the point spread function), we do not fully miss stars since the pixels adjoin; there is flux conservation, but what happens is that the stellar profiles are broader than they could be, and stars may overlap. We are still undersampling the signal though.

4 Effects of the earth's atmosphere

The earth's atmosphere has several different effects: it emits light, it absorbs light, it shifts the apparent direction of incoming light, and it degrades the coherence of incoming light, leading to degradation of image quality when collecting light over a large aperture.

4.1 Night sky emission

We first consider emission from the sky, which comes from several sources:

- air glow
- zodiacal light
- aurorae
- sunlight
- moonlight
- light pollution
- telescope and dome!
- unresolved stars and galaxies

Most of these source are emission line sources, not contiunuum sources, as shown in Figure 3.

How bright are these sources?

- air glow : strong in lines
- zodiacal light : $m_V \sim 22.2 - 23.5$ depending on ecliptic latitude (and to a lesser degree on ecliptic longitude)
- sunlight : small away from twilight

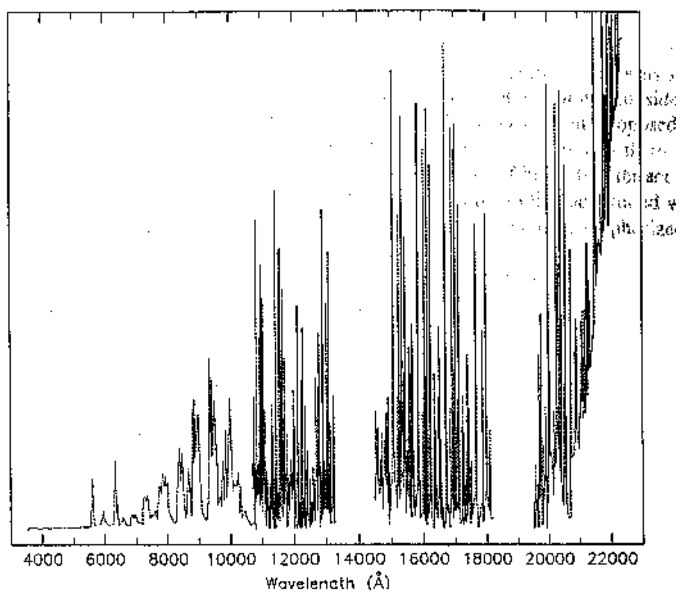


Fig 3. The night sky from 4000 A to 2.2 μm (courtesy D. Crampton), showing the strong night sky emission. The rise in the integrated night sky background is actually less frightening: The sky $(I\text{-}H)_{AB}$ ~ 3.3, comparable to the rise between B and I, where the sky $(B\text{-}I)_{AB}$ ~ 2.5.

Figure 3: Night sky emission, from article by Lilly

- moonlight : variable, can be very bright! ∼10 times brighter at full moon.

lunar age (days)	U	B	V	R	I (mag/square arcsec)
0	22.0	22.7	21.8	20.9	19.9
3	21.5	22.4	21.7	20.8	19.9
7	19.9	21.6	21.4	20.6	19.7
10	18.5	20.7	20.7	20.3	19.5
14	17.0	19.5	20.0	19.9	19.2

- light pollution : strong in distinct lines

Optical:

For broadband work, for example in the V band, $m_{sky} \sim 22$ mag/arcsec at good site so we switch over to background limited around m=22 for good image quality, switch over around m=20 for poorer image quality. Consequently, image quality matters for faint objects!!

One can choose narrower filters to improve the situation, but still have to then pay the price of longer exposure times.

Optical spectroscopy: sky emission generally not much of a problem except around lines.

IR: For broadband work in H band, $m \sim 13.5$ mag/arcsec; in K band, $m \sim 12.5$ mag/arcsec. So for all except bright objects, we're background limited. For infrared spectra, it's hard to estimate S/N : depends on where your feature is located.

Sky brightness from most sources *varies* with time and position in the sky in an *irregular* fashion. Consqeuently, it's essentially impossible to estimate the sky *a priori*: sky must be determined from your observations, and if your observations don't distinguish object from sky, you'd better measure sky close by in location and in time: especially critical in the IR.

4.2 Transmission of atmosphere

Earth's atmosphere doesn't transmit 100% of light. Various things contribute to the absorption of light:

- scattering
- aerosols
- ozone
- H_2O

- O_2
- CO_2
- N_2O
- CH_4

All are functions of wavelength, time to some extent, and position in sky.
Airmass and zenith distance dependence

Definition of airmass: path length that light takes through atmosphere relative to length at zenith: $X \equiv 1$ vertically (at $z = 0$). Given the zenith distance, z, which can be computed from:

$$\sec z = (\sin\phi \sin\delta + \cos\phi \cos\delta \cos h)^{-1}$$

where ϕ is the latitude, δ is the declination, and h is the hour angle (= local sidereal time− α, where α is the right ascension), we have

$$X \sim \sec z$$

which is exactly true in the case of a plane parallel atmosphere. Since the earth's atmosphere is not a plane, the plane parallel approximation breaks down for larger airmasses. For $X \gtrsim 2$, a more precise formula needed:

$$X = \sec z - 0.0018167(\sec z - 1) - 0.002875(\sec z - 1)^2 - 0.0008083(\sec z - 1)^3$$

Consider thin sheet of atmosphere, with incident flux F, and outcoming flux $F + dF$. Let the thin sheet have opacity $\kappa = N\sigma$, where N is the number density of absorbers/scatterers, and σ is the cross-section/absorber-scatterer.

$$dF = -\kappa F dx$$

$$F = F_0 e^{-\int \kappa dx} \equiv F_0 e^{-\tau}$$

where τ is the optical depth of atmosphere.

$$\tau(X) \sim \tau(z = 0)X$$

$$F = F_0 e^{-\tau_0 X}$$

$$m = m_0 + 1.086\tau_0 X$$

We can define the *extinction coefficient* k_λ:

$$m = m_0 + k_\lambda X$$

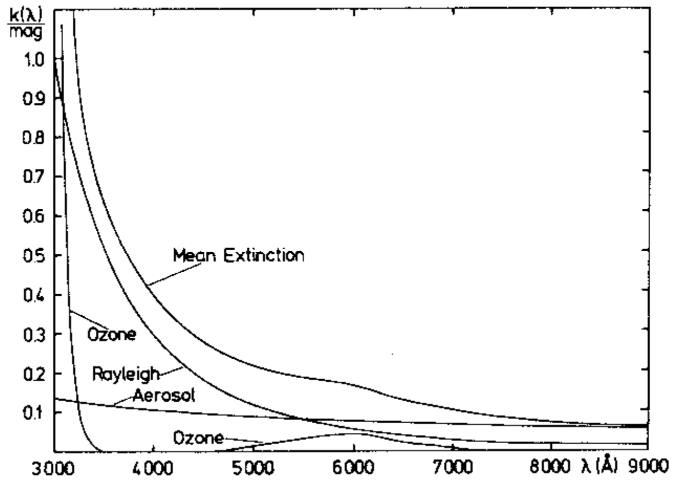


Fig. 1. Mean vertical extinction at Flagstaff, Arizona, in May-June 1976. The assumed ozone and Rayleigh contributions are shown separately

Figure 4: Extinction sources in the optical (Tug, Lockwood, and White,).

$$k_\lambda \equiv -1.086\tau(z=0)$$

so the extinction can be specified by a set of extinction coefficients. Of course, use of the scaling of τ or k with airmass assumes *photometric weather*!!

Sources of extinction

In the optical part of the spectrum, extinction is a roughly smooth function of wavelength and arises from a combination of ozone, Rayleigh scattering, and aerosols, as shown in Figure 4.

The optical extinction can vary from night to night or season to season, as shown in Figure 5. Because of this variation, you must determine the amount of extinction on each night separately if you want accuracy better than a few percent. Generally, the *shape* of the extinction curve as a function of wavelength probably varies less than the amplitude at any given wavelength. Because of this, one commonly uses *mean extinction coefficients* when doing spectroscopy where one often only cares about relative fluxes.

In the infrared, the extinction does not vary so smoothly with wavelength because of the effect of molecular absorption. In fact, significant absorption bands define the so-called infrared windows (JHKLM), as shown in the near IR in Figure 6. At longer wavelengths, the broad absoprtion band behavior continues, as shown in Figure 7. In this figure, $transmission = f(b_\lambda l)$ where l is path length (units of

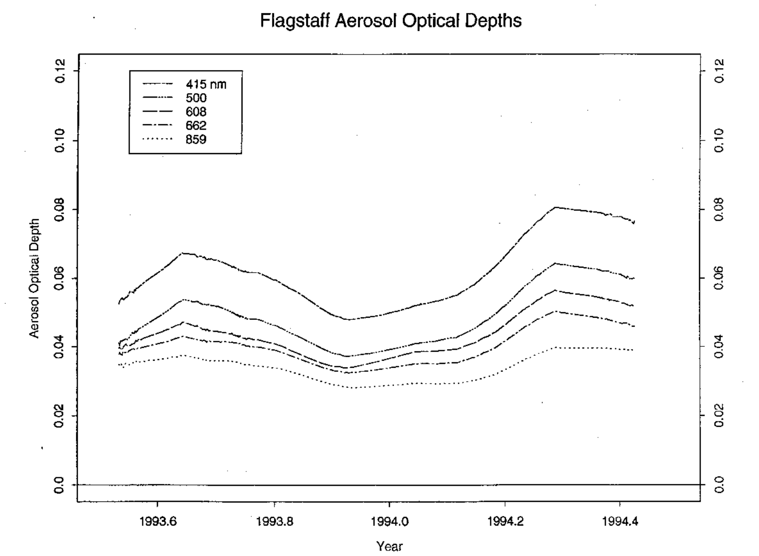


Figure 5: Variation in aerosol optical depth with time (Lockwood)

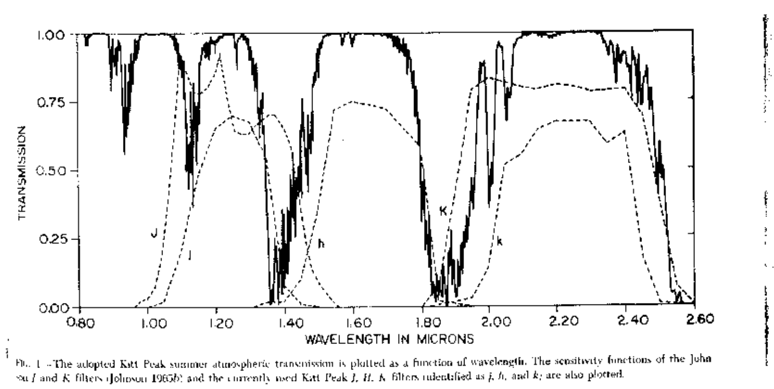
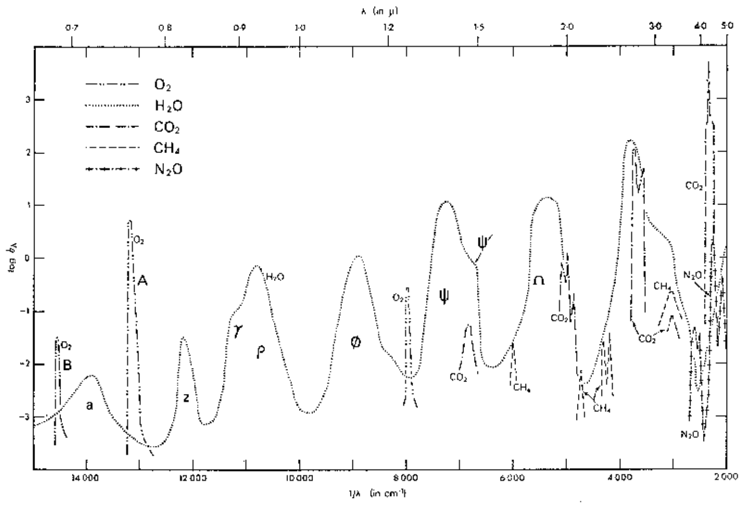


Figure 6: Atmospheric transmission in the near-IR (Manduca & Bell ...).



Infrared band absorption of atmospheric gases

Figure 7: Atmospheric transmission in the IR (Allen, Astrophysical Quantities).

airmass):

$b_\lambda l$	f
-3	1
-2	0.97
-1	0.83
0	0.5
1	0.111
2	0.000

The L band is at 3.5μ, M band at 5μ.

4.3 Atmospheric refraction

The direction of light as it passes through the atmosphere is also changed because of refraction since the index of refraction changes through the atmosphere.

Let z be the true zenith distance, α be the observed zenith distance, α_n be the observed zenith distance at layer n in the atmosphere, μ be the index of refraction at the surface, and μ_n be the index of refraction at layer n. At the top of the atmosphere:

$$\frac{\sin z}{\sin \alpha_N} = \frac{\mu_N}{1}$$

At each infinitessimal layer:

$$\frac{\sin\alpha_n}{\sin\alpha_{n-1}} = \frac{\mu_{n-1}}{\mu_n}$$

as so on for each layer down to the lowest layer:

$$\frac{\sin\alpha_1}{\sin\alpha} = \frac{\mu}{\mu_1}$$

Multiply these to get:

$$\sin z = \mu \sin\alpha$$

from which we can see that the refraction depends only the index of refraction near the earth's surface.

We define astronomical refraction, r, to be:

$$\sin(\alpha + r) = \mu \sin\alpha$$

so r gives the error in the measured zenith distance.

In cases where r is small (pretty much always):

$$\sin\alpha + r\cos\alpha = \mu\sin\alpha$$

$$r = (\mu - 1)\tan\alpha$$

$$\equiv R\tan\alpha$$

where R is the "constant of refraction".

A typical value of the index of refraction is $\mu \sim 1.00029$, which gives R = 60 arcsec (red light).

The direction of refraction is that a star apparently moves towards the zenith. Consequently in most cases, star moves in both RA and DEC:

$$r_\alpha = r\sin q$$

$$r_\delta = r\cos q$$

where q is the *parallactic* angle, the angle between N and the zenith:

$$\sin q = \cos\phi\frac{\sin h}{\sin\alpha}$$

Note that the expression for r is only accurate for small zenith distances ($z < 45$). At larger z, can't use plane parallel approximation. Observers have empirically found:

$$r = A\tan\alpha + B\tan^3\alpha$$

$$A = (\mu - 1) + B$$

$$B \sim -0.07"$$

but these vary with time, so for precise measurements, you'd have to determine A and B on your specific night of observations.

Of course, the index of refraction varies with wavelength, so consequently does the astronomical refraction r;

λl	R
3000	63.4
4000	61.4
5000	60.6
6000	60.2
7000	59.9
10000	59.6
40000	59.3

This gives rise to the phenomenon of *atmospheric dispersion*, or *differential refraction*. Because of the variation of index of refraction with wavelenth, every object actually appears as a little spectrum with the blue end towards the zenith. The spread in object position is proportional to $\tan\alpha$.

Note the importance of this effect for spectroscopy, and the consequent importance of the relation between a slit orientation and the parallactic angle.

4.4 Seeing: theory and practice

References: Coulson, ARAA 23,19; Beckers, ARAA 31, 13; Schroeder 16.II.

The Earth's atmosphere is turbulent and variations in the index of refraction cause the plane wavefront from distant objects to be distorted, as shown in Figure 8.

This distortion introduces amplitude variations, positional shifts and also image degradation.

This causes two astronomical effects:

- scintillation, which is amplitude variations, which typically vary over scales of cm: generally very small for larger apertures, and
- seeing: positional changes and image quality changes. The effect of seeing depends on aperture size: for small apertures, one sees a diffraction pattern moving around, while for large apertures, one sees a set of diffraction patterns (speckles) moving around on scale of ~ 1 arcsec. These observations imply:

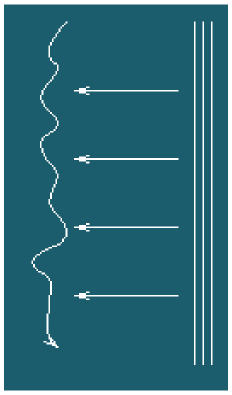


Figure 8: The effect of the Earth's atmosphere on a plane wavefront.

- local wavefront curvatures flat on scales of small apertures, and
- instantaneous slopes vary by $\sim$ an arcsec.

The time variation scales are several milliseconds and up.

The effect of seeing can be derived from theories of atmospheric turbulence, worked out originally by Kolmogorov, Tatarski, Fried. Here, I'll quote some pertinent results, without derivation.

A turbulent field can be described statistically by a *structure function*:

$$D_N(x) ==< |N(r+x) - N(r)|^2 >$$

where x is separation of points, N is any variable (e.g. tempereature, index of refraction, etc), r is position.

Kolmogorov turbulence gives:

$$D_n(x) = C_n^2 x^2/3$$

where C_n is the refractive index structure constant. From this, one can derive the phase structure function at the telescope aperture:

$$D_\phi(x) = 6.88 \frac{x}{r_o}^{5/3}$$

where the coherence length r_0 (also known as the Fried parameter) is:

$$r_0 = 0.185 \lambda^{6/5} \cos^{3/5} z [\int (C_n^2 dh)]^{-3/5}$$

where z is zenith angle, λ is wavelength. Using optics theory, one can convert D_ϕ into an image shape.

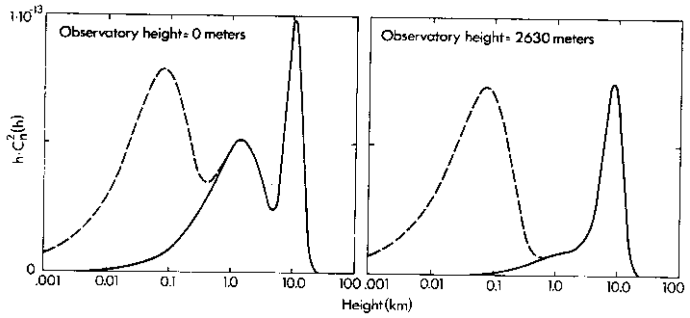


Figure 2 Average C_n^2 profile with local height h_L (in km). (*Left*) Profile for a sea level site. (*Right*) Profile for a 2630 meter high mountain site. The solid curve follows the expression given by Valley (1980) for height h above sea level: $C_n^2 = [2.05 \times 10^{-23} \cdot h^{10} \cdot \exp(-h) + 0.93 \times 10^{-16} \cdot \exp(-h/1.5)]\ m^{-2/3}$. It ignores near ground, local seeing. It is scaled to give 0.5 arcsec seeing at $=0.55\ \mu$m at sea level. The dashed line corresponds to $C_n^2 = (2.17 \times 10^{-15} + 5 \times 10^{-17} \cdot h_L^{-2/3}) \cdot \exp(-h_L/0.08)$ which also results in 0.5 arcsec seeing by itself. It approximates this local nighttime seeing. For the sea level site the resulting seeing is 0.76 arcsec; for the mountain site 0.63 arcsec. For daytime condition the local seeing will be worse. The $h \cdot C_n^2$ vs log h presentation was chosen to better visualize the contributions of the different heights to r_o.

Figure 9: From Beckers, ARAA 1993.

Physically, r_0 is (roughly) inversely proportional to the image size from seeing:

$$d \sim \lambda/r_0$$

as compared with the image size from diffraction-limited images:

$$d \sim \lambda/D.$$

Seeing dominates when $r_0 < D$; a larger r_0 means better seeing.

Seeing is more important than diffraction at shorter wavelengths, diffraction more important at longer wavelengths; effect of diffraction and seeing cross over in the IR (∼5 microns for 4m); the crossover falls at a shorter wavelength for smaller telescope or better seeing.

The meat of r_0 is in $\int (C_n^2 dh)$; as you might expect, this varies from site to site and also in time. At most sites, there seems to be three regimes of "surface layer" (wind-surface interactions and manmade seeing), "planetary boundary layer" (influenced by diurnal heating), and "free atmosphere" (10 km is tropopause: high wind shears), as seen in Figure 9.

A typical site probably has $r_0 \sim 10$ cm at 5000Å.

We also have to consider the coherence of the same turbulence pattern over the sky: coherence angle call the *isoplanatic angle*, and region over which the turbulence pattern is the same is called the *isoplanatic patch*.

$$\theta \sim 0.314 r_0/H$$

where H is the average distance of the seeing layer:

$$H = \sec z[\int (C_n^2 h^{5/3} dh) / \int (C_n^2 dh)]^{3/5}$$

For $r_0 = 10$ cm, $H \sim 5000$ m , $\theta \sim 1.3$ arcsec.

In the infrared $r_0 \sim 70$ cm, $H \sim 5000$ m, $\theta \sim 9$ arcsec.

Note however, that the "isoplanatic patch for image motion" (not wavefront) is $\sim 0.3 D/H$. For $D = 4$ m, $H \sim 5000$ km, $\theta_{kin} = 50$ arcsec.

4.5 What does seeing cause the image to look like?

The "quality" of an image can be described in many different ways. The overall shape of the distribution of light from a point source is specified by the *point spread function.* Diffraction gives a basic limit to the quality of the PSF, but any aberrations or image motion add to structure/broadening of the PSF.

Another way of describing the quality of an image is to specify it's *modulation transfer function* (MTF). The MTF and PSF are a Fourier transform pair. Turbulence theory gives:

$$MTF(\nu) = \exp[-3.44(\lambda\nu/r_o)^{5/3}]$$

where ν is the spatial frequency.

Note that a gaussian goes as ν^2, so this is close to a gaussian. The shape of seeing-limited images is roughly Gaussian in core but has more extended wings, as shown in Figure 10.

Empirically, seeing may be fit with a gaussian or, better, a Moffat function as shown in Figure 11.

$$I = p_1(1 + (x - p_2)^2/p_4^2 + (y - p_3)^2/p_5^2)^{-p_6}$$

Another way of characterizing the PSF is by giving the *encircled energy* as a function of radius, or at some specified radius.

A final way of characterizing the image quality, more commonly used in adaptive optics applications, is the Strehl ratio. The Strehl ratio is the ratio between the peak amplitude of the PSF and the peak amplitude expected in the presence of diffraction only.

Measuring the Seeing

POINT-SPREAD FUNCTION 701

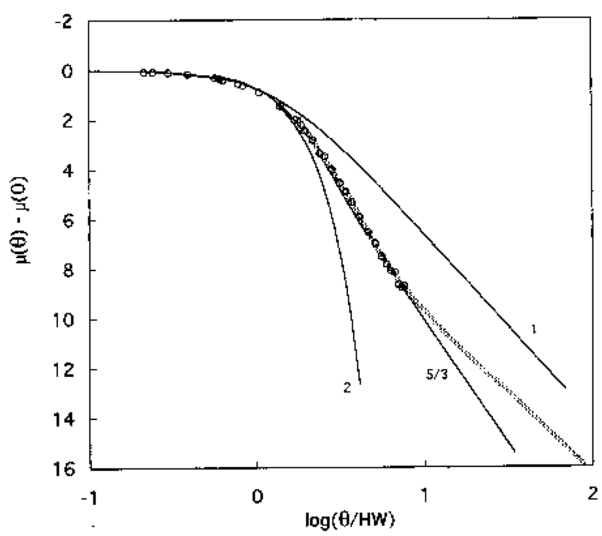


FIG. 2—The observed PSF (open circles, stippled line) is compared to theoretical models for MTFs of different indices n.

Figure 10: From Racine,

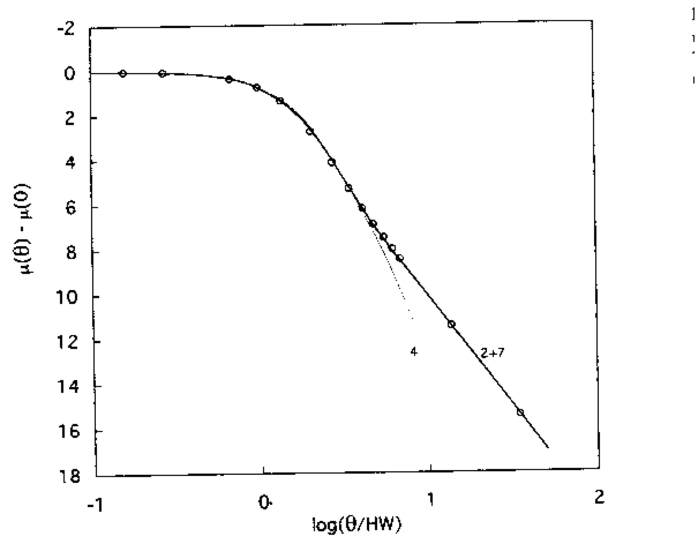


FIG. 3—The Kolmogorov PSF (open circles) fitted by a Moffat function with β=4 (thin line) and by the sum of two Moffat functions (β_1=7 and β_2=2). All functions have the same HW as the PSF.

Figure 11: From Racine,

Typically seeing is characterized by the FWHM. Usually the implication is that the form is Gaussian with sigma = FWHM /2.354 . Recall that FWHM doesn't fully specify a PSF, and one should always consider how applicable the quantity is.

Options: measure FWHM directly; fit a Gaussian or other function (we'll talk later about fitting).

However, beware of effects of sampling of the PSF: you're really getting the PSF integrated over pixels, not the PSF!

5 Astronomical optics

Because astronomoical sources are faint, we need to collect light. We use telescopes/cameras to make images of astronomical sources. Example: a 20th magnitude star gives $\sim 0.01 photons/s/cm^2$ at 5000 A through a 1000 A filter! However, using a 4m telescope gives 1200 photons/s.

5.1 Single surface optics and definitions

We will define an optical system as a system which collects light; usually, the system will also make images. This requires the bending of light rays, which is accomplished using lenses (refraction) and/or mirrors (reflection).

The operation of optical systems is given by Snell's law of refraction:

$$n \sin i = n' \sin i'$$

where n are the indices of refraction, i are the angles of incidence, relative to the normal to the surface. For refection:

$$i' = -i$$

An optical element takes a source at s and makes an image at s'. The source can be *real* or *virtual*. A real image exists at some point in space; a virtual image is formed where light rays apparently emanate from or converge to, but at a location where no light actually appears. For example, in a Cassegrain telescope, the image formed by the primary is virtual, because the secondary intercepts the light and redirects it before light gets to the focus of the primary.

The image will not necessarily be a perfect image: all rays regardless of height at the surface, y, may NOT cross at the same point. This is the subject of aberrations, which we will get into in a while. Obviously, the degree of aberration will depend on how much the different rays differ in y, as well as depend on the shape of the surface. We define *paraxial* and *marginal* rays, as rays near the center of the aperture and those on the edge of the aperture. We define the *chief ray* as the ray that passes through the center of the aperture. To define nominal quantities, however, we consider the *paraxial* regime. In this regime, all angles are small, aberrations vanish, and a surface can be wholly specified by its radius of curvature R.

The *field angle* gives the angle formed between the chief ray from an object and the z-axis. Note that paraxial does not necessarily mean field angle=0. One can have an object at a field angle and still consider the paraxial approximation.

Note also that for the time being, we are ignoring *diffraction*. But we'll get back to that too. We are considering *geometric* optics, which is what you get from diffraction as wavelength tends to 0. For nonzero wavelength, geometric optics applies as scales $x \gtrsim \lambda$.

We can derive the basic relation between object and image location as a function of a surface where the index of refraction changes (Schroeder, chapter 2).

$$\frac{n'}{s'} - \frac{n}{s} = \frac{(n' - n)}{R}$$

The points at s and s' are called *conjugate*. If either s or s' is at infinity (true for astronomical sources for s), the other distance is defined as the *focal length*, f, of the optical element. For $s = \text{inf}$, $f = s'$.

We can define the quantity on the right side of the equation, which depends only the the surface parameters (not the image or object locations), as the *power*, P, of the surface:

$$P \equiv (n' - n)/R = n'/f' = n/f$$

We can make a similar derivation for the case of reflection; in fact, we can treat reflection by considering refraction with $n' = -n$.

$$\frac{n'}{s'} + \frac{n'}{s} = \frac{(n' + n')}{R}$$

$$\frac{1}{s'} + \frac{1}{s} = \frac{2}{R}$$

This shows that the focal length for a mirror is given by $R/2$.

We define the *focal ratio* to be the focal length divided by the aperture diamter. The focal ratio is also called the *F-number* and is denoted by the abbreviation $f/$. Note $f/10$ means a focal ratio of ten; f is not a variable in this! The focal ratio gives the beam "width"; systems with a small focal ratio have a short focal length compared with the diameter and hence the imcoming beam to the image is wide. Systems with small focal ratios are called "fast" systems; systems with large focal ratios are called "slow" systems.

The *magnification* of a system gives the ratio of the image height to the object height:

$$h'/h = (s' - r)/(s - R) = ns'/n's$$

The magnification is negative for this case, because object is flipped. The magnification also negative for reflection: $n' = -n$. Magnification not relevant for source at infinity, but still it is a very important quantity for multi-element systems.

We define the *scale* as the motion of image for given incident angle of parallel beam from infinity. From a consideration of the chief rays for objects on-axis and at field angle α, we get:

$$\tan\alpha \approx \alpha = \frac{x}{f}$$

or

$$\text{scale} \equiv \frac{\alpha}{x} = \frac{1}{f}$$

In other words, the scale, in units of angular motion per physical motion in the focal plane, is given by $1/f$.

5.2 Multi-surface systems

To combine surfaces, one just takes the image from the first surface as the source for the second surface, etc., for each surface. We can generally describe the basic parameters of multi-surface systems by equivalent single-surface parameters, e.g. you can define an effective focal length of a multi-surface system as the focal length of some equivalent single-surface system. The two systems (single and multi) are equivalent in the paraxial approximation ONLY.
Example: a lens (has two surfaces).

Consider a lens in air ($n \sim 1$). The first surface give

$$\frac{n}{s'_1} - \frac{1}{s_1} = \frac{(n-1)}{R_1} = P_1$$

The second surface gives:

$$\frac{1}{s'_2} - \frac{n}{s_2} = \frac{(1-n)}{R_2} = P_2$$

but we have $s_2 = s'_1 - d$ (remember we have to use the plane of the second surface to measure distances for the second surface).

After some algebra, we find the effective focal length (from center of lens):

$$P = \frac{1}{f'} = P_1 + P_2 - \frac{d}{n} P_1 P_2$$

$$P = \frac{(n-1)}{R_1} + \frac{(1-n)}{R_2} - \frac{d}{n}\frac{(n-1)(1-n)}{R_1 R_2}$$

From this, we derive the *thin lens* formula:

$$P = \frac{1}{f'} = \frac{(n-1)}{R_1} + \frac{(1-n)}{R_2} = (n-1)(\frac{1}{R_1} - \frac{1}{R_2})$$

$$\frac{1}{f'} = \frac{1}{f_1} + \frac{1}{f_2}$$

Example: a plane-parallel plate. Zero power, but moves image laterally: $\Delta = d[1 - (1/n)]$.
Example: two-mirror telescopes:

In astronomy, most telescopes are two-mirror telescopes of Newtonian, Cassegrain, or Gregorian design. The Cassegrain is the most common and is outlined in Figure 12.

First, accept some basic definitions:

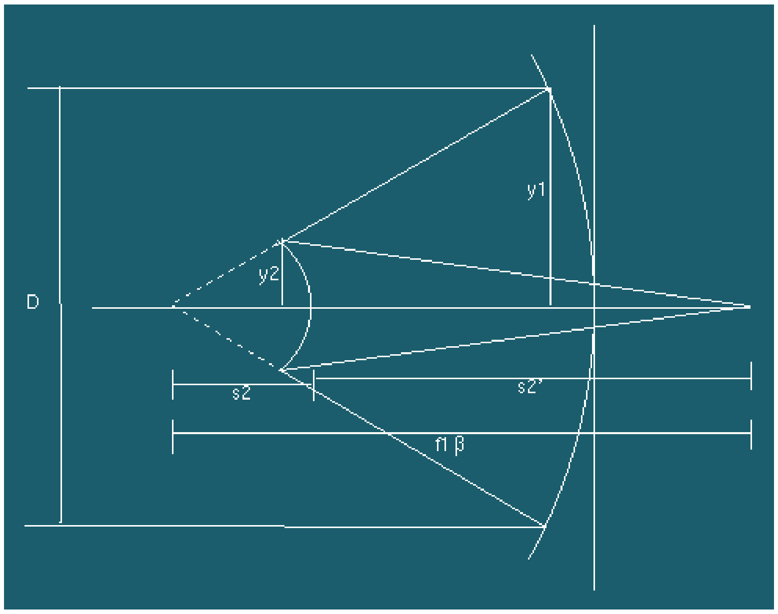


Figure 12: Schematic of a Cassegrain telescope.

- ratio of focal lengths, ρ:
$$\rho = R_2/R_1 = f_2/f_1$$
- magnification of the secondary, m:
$$m = -s_2'/s_2$$
- *back focal distance*, $f_1\beta$:
$$f_1\beta = D\eta$$
- primary focal ratio, F_1:
$$F_1 = f_1/D$$
- ratio of marginal ray heights, k:
$$k = y_2/y_1$$

From similar triangles,
$$k \equiv \frac{y_2}{y_1} = \frac{s_2}{f_1} = \frac{2s_2}{R_1}$$
$$s_2 = \frac{kR_1}{2}$$

For the secondary:

$$\frac{1}{s_2'} + \frac{1}{s_2} = \frac{2}{R_2}$$

$$1 - m = \frac{2s_2'}{R_2}$$

$$s_2' = \frac{R_2}{2}(1-m)$$

Dividing these two relations gives:

$$-m = \frac{R_2}{2}(1-m)\frac{2}{kR_1} = \frac{\rho(1-m)}{k}$$

So to summarize, we get:

$$\rho = \frac{mk}{(m-1)}$$

We also get

$$(1+\beta) = k(m+1)$$

We can derive the effective focal length and focal ratio from our thick lens formula:

$$f = f_1 m$$

$$F = F_1 m$$

So we find that the telescope basic parameters (paraxial) are determined by 3 of: $f, f_1, m, \rho, k, \beta$ (not m, ρ, k or m, β, k). Usually, f_1 is limited by technology. Then choose m to match desired scale. k is related to separation of mirrors, and is a compromise between making telescope shorter and blocking out more light vs. longer and blocking less light; in either case, have to keep focal plane behind primary!

Definitions for multi-surface system:

- aperture stop: determines the amount of light reaching an image (usually the primary mirror)
- field stop: determines the angular size of the field. This is usually the detector, but for a large enough detector, it could be the secondary.
- pupil: location where rays from all field angles fill the same aperture.
- entrance pupil: image of aperture stop as seen from source object (usually the primary).
- exit pupil: image of aperture stop formed by all subsequent optical elements.

In a two-mirror telescope, the location of the exit pupil is where the image of the primary is formed by the secondary. This can be calculated using $s = d$ as the object distance (where d is the separation of the mirrors), then with the reflection equation, we can solve for s' which gives the location of the exit pupil relative to the secondary mirror. If one defines the quantity δ, such that $f_1\delta$ is the distance between the exit pupil and the focal plane, then (algebra not shown):

$$\delta = \frac{m^2 k}{m + k - 1} = \frac{m^2(1+\beta)}{m^2+\beta}$$

The exit pupil is an important concept. When we discuss aberrations, it is the total wavefront error at the exit pupil which gives the system aberration.

One final thing to note is how we focus a Cassegrain telescope. Most instruments are placed at a fixed location, β, behind the primary. Focussing is usually then done by moving the secondary mirror. Clearly, if you move the secondary you change k. Since ρ is fixed by the mirror shapes, it's also clear that you change the magnification as you move the secondary; this is expected since you are changing the system focal length, $f = mf_1$. The amount of image motion for a given secondary motion is given by:

$$\frac{d\beta}{dk} = \frac{d}{dk} k(m+1) - 1$$

where

$$m(k) = \frac{\rho}{\rho - k}$$

so

$$\begin{aligned}
\frac{d\beta}{dk} &= \frac{d}{dk} k \frac{(2\rho - k)}{(\rho - k)} - 1 \\
&= \frac{2(\rho-k)^2 + 2\rho k - k^2}{(\rho-k)^2} \\
&= \frac{2(\rho-k)^2 + \rho^2 - (\rho-k)^2}{(\rho-k)^2} \\
&= 1 + \frac{\rho^2}{(\rho-k)^2} = 1 + m^2
\end{aligned}$$

where the motion of the secondary is $f_1 dk$ and the motion of the image plane is $f_1 d\beta$.

5.3 Aberrations

5.3.1 Surface requirements for unaberrated images

Next we consider non-paraxial rays. We first consider what surface is required to make an unaberrated image.

We can derive the surface using Fermat's principle. Fermat's principle states that light travels in the path such that infinitessimally small variations in the path doesn't change the travel time to first order: d(time)/d(length) is a minimum. For a single surface, this reduces to the statement that light travels the minimum path length. An alternate way of stating Fermat's principle is that the *optical path length* is unchanged to first order for a small change in path. The OPL is given by:

$$OPL = \int c dt = \int \frac{c}{v} v dt = \int n ds$$

Fermat's principle has a physical interpretation when one considers the wave nature of light. It is clear that around a stationary point of the optical path light, the maximum amount of light can be accumulated over different paths with a minimum of destructive interference. By the wave theory, light travels over all possible paths, but the light coming over the "wrong" paths destructively interferes, and only the light coming over the "right" path constructively interferes.

Fermat's principle can be used to derive the basic laws of reflection and refraction (Snell's law):

(do the derivation)

Now consider a perfect imaging system that takes all rays from an object and makes them all converge to an object. Since Fermat's principle says the only paths taken will be those for which the OPL is minimally changed for small changes in path, the only way a perfect image will be formed is when all optical path lengths along a surface between an image and object point are the same - otherwise the light doesn't get to this point!

Instead of using Fermat's principle, we could solve for the parameters of a perfect surface using analytic geometry, but this would require an inspired guess for the correct functional form of the surface.

We find that the perfect surface depends on the situation: whether the light comes from a source at finite or infinite distance, and whether the mirror is concave or convex. We consider the various cases now, quoting the results without actually doing the geometry.

Consider a concave mirror with one conjugate at infinity. Fermat's principle gives:

$$y^2 = 2Rz$$

where $R = 2f$, the radius of curvature at the mirror vertex. This equation is that of a parabola. Note, however, that a parabola makes a perfect image only for on axis images (field angle=0).

For a concave mirror with both conjugates finite, we get an ellipse. Again, this is perfect only for field angle = 0.

$$(z - a)^2/a^2 + y^2/b^2 = 1$$

$$y^2 - 2zb^2/a + z^2b^2/a^2 = 0$$

where

$$a = (s + s')/2.$$

$$b = \sqrt{(ss')}$$

$$R = ss'/(s + s') = 2b^2/a$$

For a convex mirror with both conjugates finite, we get a hyperbola:

$$(z - a)^2/a^2 - y^2/b^2 = 1$$

$$y^2 + 2zb^2/a - z^2b^2/a^2 = 0$$

where

$$a = (s + s')/2$$

$$b^2 = -ss'$$

(s is negative)

$$R = -2b^2/a$$

For a convex mirror with one conjugate at infinity, we get a parabola.

Note that in all cases we've considered a one-dimension surface. We can generalize to 2D surfaces by rotating around the z-axis; for the equations, simply replace y^2 with $(x^2 + y^2)$.

As you may recall from analytic geometry, all of these figures are *conic sections*, and it is possible to describe all of these figures with a single equation:

$$\rho^2 - 2Rz + (1 + K)z^2 = 0$$

where

$$\rho^2 = x^2 + y^2$$

and R is the radius of curvature at the mirror vertex, K is called the conic constant ($K = -e^2$, where e is the eccentricity for an ellipse, $e(b, a)$).

$K > 0$ gives a prolate ellipsoid
$K = 0$ gives a sphere
$-1 < K < 0$ gives a oblate ellipsoid
$K = -1$ gives a paraboloid
$K < -1$ gives a hyperboloid

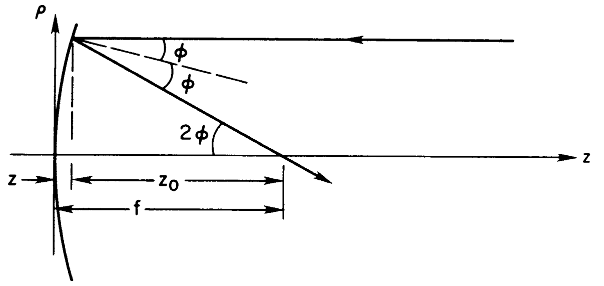


Fig. 4.1. Geometry of ray from distant object reflected from concave mirr

Figure 13: Spherical aberration diagram

5.3.2 Aberrations: general description and low-order abberrations

Now consider what happens for surfaces that are not perfect, e.g. for the cases considered above for field angle≠0 (since only a sphere is symmetric for all field angles), or for field angle 0 for a conic surface which doesn't give a perfect image?

You get *aberrations*; the light from all locations in aperture does **not** land at **any** common point.

One can consider aberrations in either of two ways:

1. aberrations arise from all rays not landing at a common point,
2. aberrations arise because wavefront deviates from a spherical wavefront.

These two descriptions are equivalent. For the former, one can talk about the *transverse* aberrations, which give the distance by which the rays miss the paraxial focus, or the *angular* aberration, which is the angle by which the rays deviate from the perfect ray which will hit paraxial focus. For the latter, one discusses the wavefront error, i.e., the deviation of the wavefront from a spherical wavefront as a function of location in the exit pupil.

First, consider where rays land as f(z). We can derive the effective focal length as f(z) for an arbitrary conic section for field angle=0:

in Figure 13.

$$z_0 = \rho/\tan(2\phi) = \rho(1-(\tan\phi)^2)/(2\tan\phi)$$
$$\tan\phi = dz/d\rho$$

from conic equation:

$$\rho^2 - 2Rz + (1+K)z^2 = 0$$

$$z = \frac{R}{(1+K)}\left[1 - \left(1 - \frac{\rho^2}{R^2}(1+K)\right)^{1/2}\right]$$

$$z \approx \frac{\rho^2}{2R} + (1+K)\frac{\rho^4}{8R^3} + (1+K)^2\frac{\rho^6}{16R^5} + \ldots$$

$$dz/d\rho = \rho/(R-(1-K)z)$$

$$z_0 = \frac{\rho}{2}\left[\frac{R-(1+K)z}{\rho} - \frac{\rho}{R-(1+K)z}\right]$$

$$\begin{aligned} f_e &= z + z_0 \\ &= \frac{R}{2} + \frac{(1-K)z}{2} - \frac{\rho^2}{2(R-(1+K)z} \\ &= \frac{R}{2} - (1+K)\frac{\rho^2}{4R} - (1+K)(3+K)\frac{\rho^4}{16R^3} - \ldots \end{aligned}$$

$$\Delta f = f_e - \frac{R}{2}$$

Note that f_e is independent of z only for $K = -1$, a parabola. Also note that Δf is symmetric with respect to ρ.

We define spherical aberration as the aberration resulting from $K \neq -1$. It is an aberration which is present on axis as seen in Figure 14.

Spherical aberration is symmetric in the pupil. There is no location in space where all rays focus at a point. In fact, the behavior as a function of focal position is not symmetric. One can define several criteria for where the "best focus" might be, leading to the terminology paraxial focus, marginal focus, diffraction focus, and the circle of least confusion.

The asymmetric nature of spherical aberration as a function of focal position distinguishes it from other aberrations and is a useful diagnostic for whether a system has this aberration. This is shown in in Figure 15.

We define *transverse spherical aberration* (TSA) as the image size at paraxial focus. This is not the location of the minimum image size.

in Figure 16.

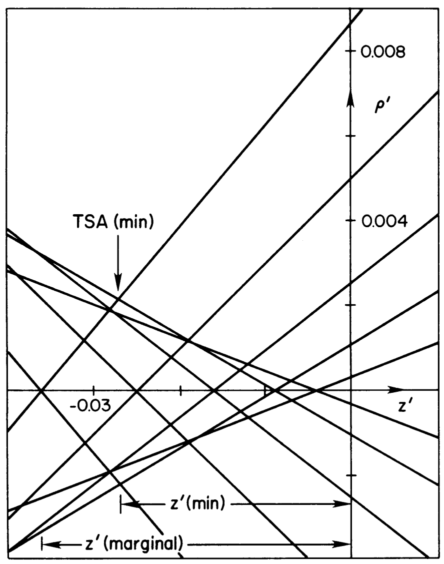


Fig. 4.5. Ray distribution near paraxial focus for image with spherical aberration. Paraxial focus is at (0, 0). See Eq. (4.2.8) for definition of parameters.

Figure 14: Spherical aberration

$$\frac{TSA}{\Delta f} = \frac{\rho}{(f - z(\rho))}$$

$$TSA = -(1+K)\frac{\rho^3}{2R^2} - 3(1+K)(3+K)\frac{\rho^5}{8R^4} + ...$$

We can also consider aberration as the difference between our wavefront and a spherical wavefront, which in this case is given by a parabolic surface.

in Figure 17.

$$\Delta z = z_{parabola} - z(K) = -\frac{\rho^4}{8R^3}(1+K) + \ldots$$

The difference in angle between the "perfect" ray from the parabola and the actual ray is called the *angular aberration*, in this case *angular spherical aberration*, or ASA.

in Figure 18.

$$ASA = 2(\phi_p - \phi) \approx \frac{d}{d\rho}(2\Delta z) \approx -(1+K)\frac{\rho^3}{R^3}$$

where $2\Delta z$ gives the optical path difference between the two rays.

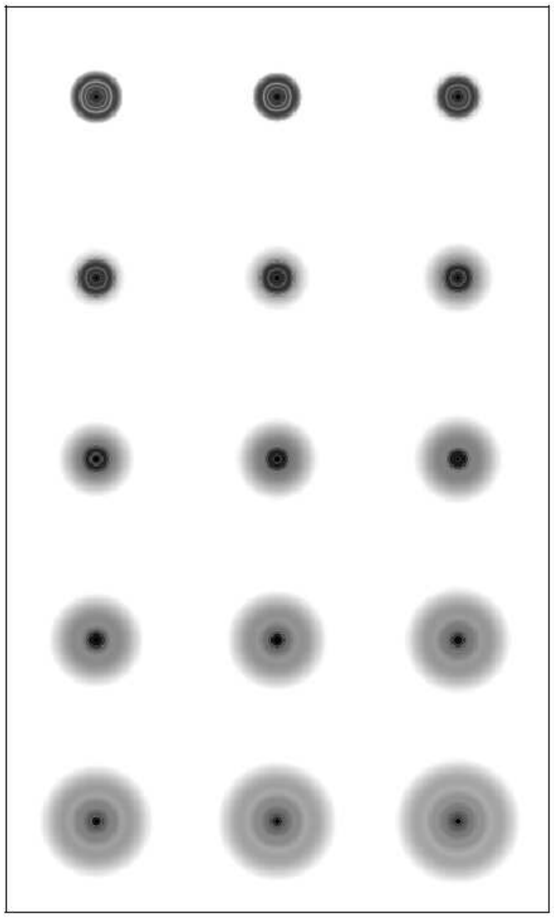


Figure 15: Spherical aberration image sequence as function of focal position

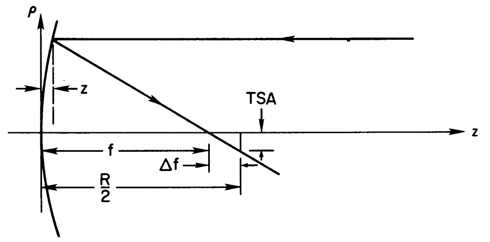


Fig. 4.2. Transverse spherical aberration (TSA) at paraxial focus. See Eqs. (4.1.6) and (4.2.1).

Figure 16: Spherical aberration diagram

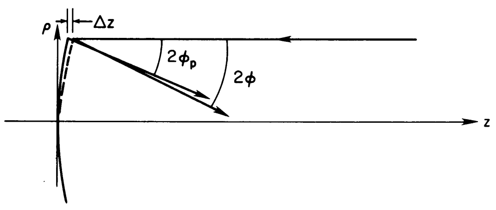


Fig. 4.3. Path difference between ray reflected from paraboloid (solid curve) and conic (dashed curve). Size of Δz, given in Eq. (4.2.2), is greatly exaggerated in the diagram.

Figure 17: Spherical aberration diagram

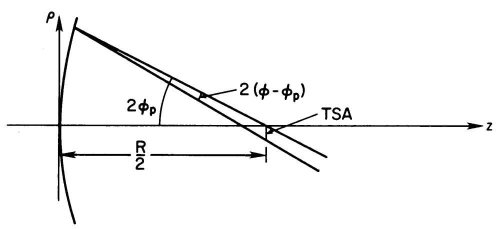


Fig. 4.4. Relation between TSA and angular difference between ray paths after reflection. See Eq. (4.2.3).

Figure 18: Spherical aberration diagram

This is simply related to the transverse aberration:

$$TSA = \frac{R}{2} ASA$$

This result can be generalized to any sort of aberration: the angular and transverse aberrations can be determined from the optical path difference between a given ray and that of a spherical wavefront. The relations are given by:

$$\text{angular aberration} = d/d\rho(2\Delta z)$$

$$\text{transverse aberration} = s' d/d\rho(2\Delta z)$$

If the aberrations are not symmetric in the pupil, then we could define angular and transverse x and y aberrations separation by taking derivatives with respect to x or y instead of ρ.

We can describe deviations from a spherical wavefront generally. Since all we care about are optical path *differences*, we write an expression for the optical path difference between an arbitrary ray and the chief ray, and in doing this, we can also include the possibility of an off-axis image, and get

$$OPD = OPL - OPL(chiefray)$$

$$OPD = A_0 y + A_1 y^2 + A_1' x^2 + A_2 y^3 + A_2' x^2 y + A_3 \rho^4$$

where we've kept terms only to fourth order and chosen our coordinate system such that the object lies in the y-z plane. The coefficients, A, depend on lots of things, such as (θ, K, n, R, s, s').

Analytically, people generally restrict themselves to talking about *third-order* aberrations, which are fourth-order (in powers of $x, y, \rho, \text{or} \theta$) in the optical path difference, because of the derivative we take to get transverse or angular aberrations. In the third-order limit, one finds that $A_2 = A_2'$, and $A_1 = -A_1'$. Working out the geometry, we find for a mirror that:

$$A_0 = 0$$

$$A_1 = \frac{n\theta^2}{R}$$

$$A_2 = -\frac{n\theta}{R^2}\left(\frac{m+1}{m-1}\right)$$

$$A_3 = \frac{n}{4R^3}\left[K + \left(\frac{m+1}{m-1}\right)^2\right]$$

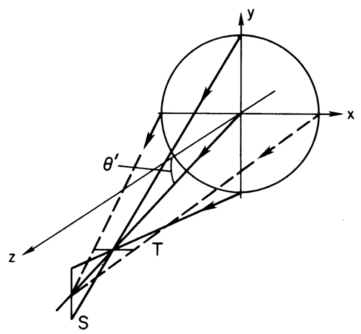


Fig. 5.2. Location and orientation of astigmatic line images. T and S denote tangential and sagittal images, respectively.

Figure 19: Rays in the presence of astigmatism.

From the general expression, we can derive the angular or the transverse aberrations in either the x or y direction. Considering the aberrations in the y direction, we find:

$$AA_y = 2A_1y + A_2(x^2 + 3y^2) + 4A_3y\rho^2$$
$$AA_x = 2A_1x + 2A_2xy + 4A_3x\rho^2$$

The first term is proportional to $\theta^2 y$ and is called *astigmatism*. The second term is proportional to $\theta(x^2 + 3y^2)$ and is called *coma*. The final term, proportional to $y\rho^2$ is *spherical aberration*, which we've already discussed (note for spherical, $AA_x = AA_y$ and in fact the AA in any direction is equal, hence the aberration is circularly symmetric).

Note that rays along the y-axis are called *tangential* rays, while rays along the x-axis are called *sagittal* rays.

For astigmatism, rays from opposite sides of the pupil focus in different locations relative to the paraxial rays. At the paraxial focus, we end up with a circular image. As you move away from this image location, you move towards the tangential focus in one direction and the sagittal focus in the other direction. At either of these locations, the astigmatic image looks like a line. Astigmatism goes as θ^2, and consequently looks the same for opposite field angles. Astigmatism is characterized in the image plane by the *transverse* or *angular* astigmatism (TAS or AAS), which refer to the height of the marginal rays at the paraxial focus. Astigmatism is symmetric around zero field angle.

Figure 19 shows the nature of astigmatism.

Figure 20 shows the behavior of astigmatism as one passes through paraxial focus.

For coma, rays from opposite sides of the pupil focus at the same location. However, the tangential rays focus at a different location than the sagittal rays, and neither of these focus at the paraxial focus. The net effect is to make an image that vaguely looks like a comet, hence the name coma. Coma goes as θ, so the direction of the comet flips sign for opposite field angles. Coma is characterized by either the *tangential* or *sagittal transverse/angular coma* (TTC, TSC, ATC, ASC) which describe the height/angle of either the tangential or sagittal marginal rays at the paraxial focus: $TTC = 3TSC$.

Figure 21 shows the nature of coma.

Figure 22 shows the behavior of coma as one passes through paraxial focus.

In fact, there are two more third-order aberrations which we haven't yet discussed: *distortion* and *field curvature*. Neither affects image quality, only location (unless you are forced to use a flat image plane!). Distortion goes as θ^3. The field curvatures can be derived from the aberration coefficients and the mirror parameters.

We can also determine the relevant coefficients for a surface with a displaced stop (Schroeder p 77), or for a surface with a decentered pupil (Schroeder p89-90); it's just more geometry and algebra. With all these realtions, we can determine the optical path differences for an entire system: for a multi-surface system, we just add the OPD's as we go from surface to surface. The final aberrations can be determined from the system OPD.

5.3.3 Aberration compensation and different telescope types

Using the techniques above, we can write expressions for the system aberrations as a function of the surface figures (and field angles). If we give ourselves the freedom to choose surface figures, we can eliminate one (or more) aberrations. For example, if we set the primary to be a parabola ($K_1 = -1$), we can use the aberration relations to determine K_2 such that spherical aberration is zero; this will give us perfect images on-axis. We find that:

$$K_2 = -\left(\frac{(m+1)}{(m-1)}\right)^2 + \frac{m^3}{k(m-1)^3)}(K_1+1)$$

satisfies this criterion. This type of telescope is called a *classical* telescope. Using the aberration relations, we can determine the amount of astigmatism and coma for such telescopes, and we find that coma gives significantly larger aberrations than astigmatism.

If we allow ourselves the freedom to choose both K_1 and K_2, we can eliminate both spherical aberration and coma. Designs of this sort are called *aplanatic*. The relevant

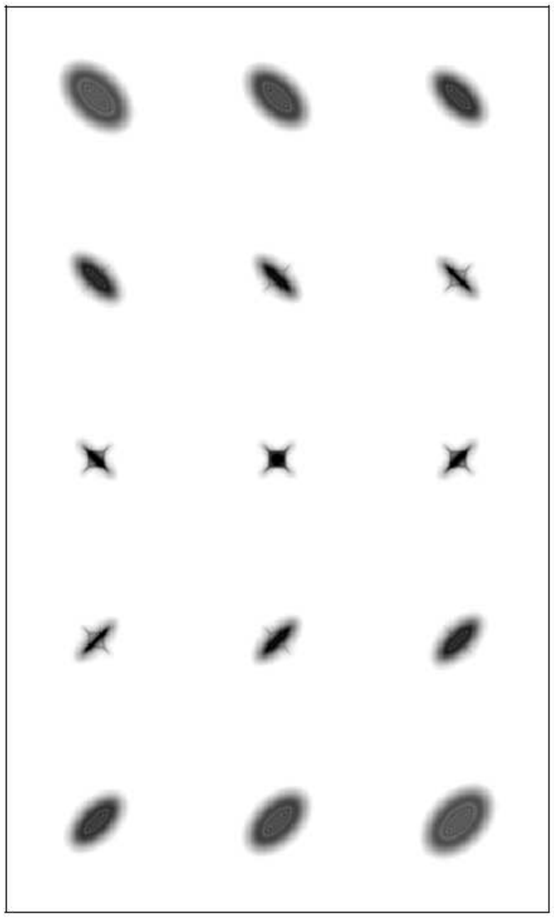


Figure 20: Astigmatism image sequence as function of focal position

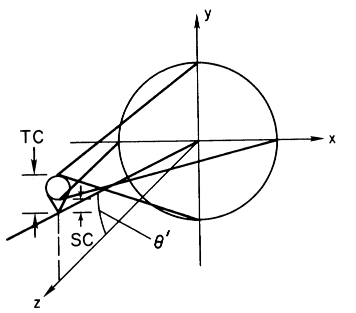


Fig. 5.5. Sketch of comatic image profile. TC and SC are tangential and sagittal coma, respectively.

Figure 21: Rays in the presence of coma.

expression, in terms of the magnification and back focal distance (we could use the relations discussed earlier to present these in terms of other paraxial parameters), is:

$$K_1 = -1 - \frac{2(1+\beta)}{m^2(m-\beta)}$$

We can only eliminate two aberrations with two mirrors, so even this telescope will be left with astigmatism.

There are two different classes of two-mirror telescopes: Cassegrain telescopes and Gregorian telescopes. For the classical telescope with a parabolic primary, the Cassegrain secondary is hyperbolic, whereas for a Gregorian it is ellipsoidal (because of the appropriate conic sections derived above for convex and concave mirrors with finite conjugates). For the aplanatic design, the Cassegrain telescope has two hyperbolic mirrors, while the Gregorian telescope has two ellipsoidal mirrors. An aplanatic Cassegrain telescope is called a *Ritchey-Chretien* telescope.

The following table gives some characteristics of "typical" telescopes. Aberrations are given at a field angle of 18 arc-mm in units of arc-seconds. Coma is given in terms of tangential coma.

Characteristics of Two-Mirror Telescopes

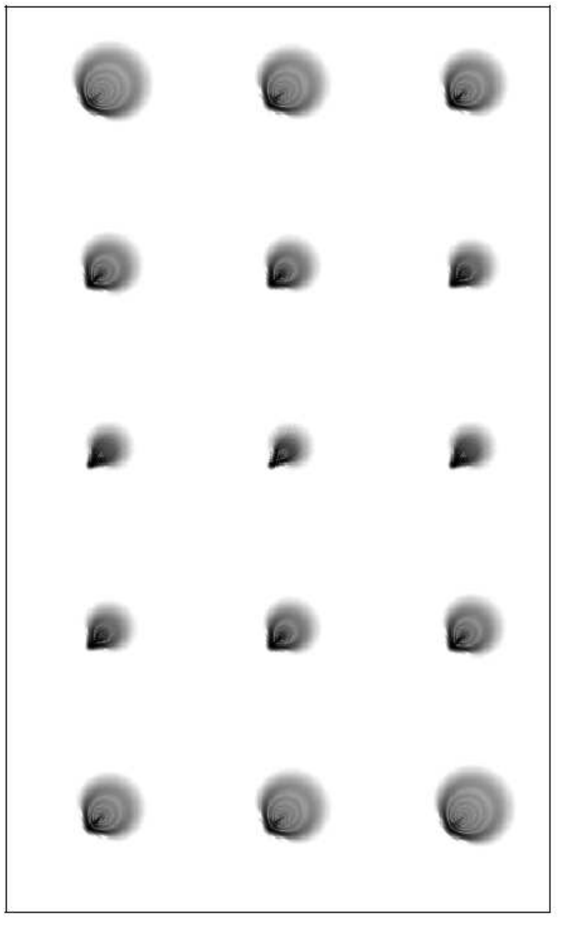


Figure 22: Coma image sequence as function of focal position

Parameter	CC	CG	RC	AG
m	4.00	-4.00	4.00	-4.00
k	0.25	-0.417	0.25	-0.417
1 - k	0.75	1.417	0.75	1.417
mk	1.000	1.667	1.000	1.667
ATC	2.03	2.03	0.00	0.00
AAS	0.92	0.92	1.03	0.80
ADI	0.079	0.061	0.075	0.056
$\kappa_m R_1$	7.25	-4.75	7.625	-5.175
$\kappa_p R_1$	4.00	-8.00	4.00	-8.00

The image quality is clearly better for the aplanatic designs than for the classical designs, as expected because coma dominates off-axis in the classical design. In the aplanatic design, the Gregorian is slightly better. However, when considerations other than just optical quality are considered, the Cassegrain usually is favored: for the same primary mirror, the Cassegrain is considerably shorter and thus it is less costly to build an enclosure and telescope structure. To keep the physical length the same, the Gregorian would have to have a faster primary mirror, which are more difficult (i.e. costly) to fabricate, and which will result in a greater sensitivity to alignment errors. Both types of telescopes have a *curved* focal plane.

5.4 Ray tracing

For a fully general calculation of image quality, one does not wish to be limited to third-order aberrations, nor does one often wish to work out all of the relations for the complex set of aberrations which result from all of the sources of aberration mentioned above. Real world situations also have to deal with *vignetting* in optical systems, in which certain rays may be blocked by something and never reach the image plane (e.g., in a two-mirror telescope, the central rays are blocked by the secondary).

Because of these and other considerations, analysis of optical systems is usually done using *ray tracing*, in which the parameters of an optical system are entered into a computer, and the computer calculates the expected images on the basis of geometric optics. Many programs exist with many features: one can produce *spot diagrams* which show the location of rays from across the aperture at an image plane (or any other location), plots of transverse aberrations, plots of optical path differences, etc., etc.

Demo ray trace program. Start with on-axis object, single mirror. Where is focus? What will image look like with spherical mirror? What do we need to do to make it perfect? How does it depend on aperture size? Now how do off-axis images look like? spot diagrams, through focus, ray fan, opd plots, etc. Now introduce second mirror. What determines where focus will be? Magnification? What shape to make

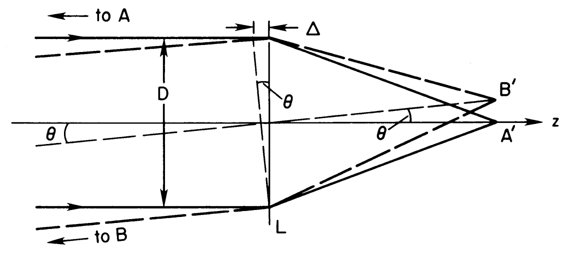


Fig. 3.11. Schematic of perfect optical system from which approximate diffraction limit is derived. See Eq. (3.6.1) and Section 3.VI.

Figure 23: The diffraction limit of a telescope

a perfect on-axis image? What do off-axis images look like? How do we make them better? Now how is performance? Real 3.5m and 1m prescriptions. Issue: guider.

5.5 Physical (diffraction) optics

Up until now, we have avoided considering the wave nature of light which introduces *diffraction* from interference of light coming from different parts of the aperture. Because of diffraction, images of a point source will be slightly blurred. From simple geometric arguments, we can estimate the size of the blur introduced from diffraction: as shown in Figure 23

We find that:

$$\theta \sim \frac{\lambda}{D}$$

Using this, we find that the diffraction blur is smaller than the blur introduced by seeing for $D \gtrsim 0.2$ meters at 5500 Å, even for the excellent seeing conditions of 0.5 arcsecond images. However, the study of diffraction has become important recently because of several reasons: 1) the existence of the Hubble Space Telescope, which is diffraction limited (no seeing), 2) the increasing use of infrared observations, where diffraction is more important than in the optical, and 3) the development of adaptive optics, which attempts to remove some of the distortions caused by seeing. Consequently, it's now worthwhile to understand some details about diffraction.

To work out in detail the shape of the images formed from diffraction involves understanding wave propagation. Basically, one integrates over all of the source points in the aperture (or exit pupil for an optical system), determining the contribution of each point at each place in the image plane. The contributions are all summed taking into account phase differences at each image point, which causes reinforcment at some points and cancellation at others. The expression which sums all of the individual source points is called the *diffraction integral*. When the details are worked out, one finds that the intensity in the image plane is related to the intensity and phase at the exit pupil. In fact the wavefront is described at any plane by the *optical transfer function*, which gives the intensity and phase of the wave at all locations in that plane. The OTF at the pupil plane and at the image plane are a Fourier transform pair. Consequently, we can determine the light distribution in the image plane by taking the Fourier transform of the pupil plane; the light distribution, or point spread function, is just the modulus-squared of the OTF at the image plane. Symbolically, we have

$$PSF = |FT(OTF(pupil))|^2$$

$$OTF(pupil) = P(x, y) \exp ik\phi(x, y)$$

$P(x, y)$ is the *pupil function*, which gives the transmission properties of the pupil, and usually consists of ones and zeros for locations where light is either transmitted or blocked (e.g., for a circular lens, the pupil function is unity within the radius of lens, and zero outside; for a typical telescope the pupil function includes obscuration by the secondary and secondary support structure). ϕ is the phase in the pupil. More relevantly, ϕ can be taken to be the optical path difference in the pupil with some fiducial phase, since only OPDs matter, not the absolute phase. Finally the wavenumber k is just $\frac{2\pi}{\lambda}$.

For the simple case of a plane wave with no phase errors, the diffraction integral can be solved analytically. The result for a circular aperture with a central obscuration, when the fractional radius of the obscuration is given by ϵ, the expression for the PSF is:

$$PSF \propto \left[\frac{2J_1(v)}{v} - \epsilon^2 \frac{2J_1(\epsilon v)}{\epsilon v}\right]^2$$

$$v = \frac{\pi r}{\lambda F}$$

where J_1 is a first order Bessel function, r is the distance in the image plane, λ is the wavelength, and F is the focal ratio ($F = f/D$).

This expression gives the so-called *Airy pattern* which has a central disk surrounded by concentric dark and bright rings. One finds that the radius of the first dark ring is at the physical distance $r = 1.22\lambda F$, or alternatively, the angular distance $\alpha = 1.22\lambda/D$. This gives the size of the *Airy disk*.

For more complex cases, the diffraction integral is solved numerically by doing a Fourier transform. The pupil function is often more complex than a simple circle, because there are often additional items which block light in the pupil, such as the support structures for the secondary mirror.

Figure 24

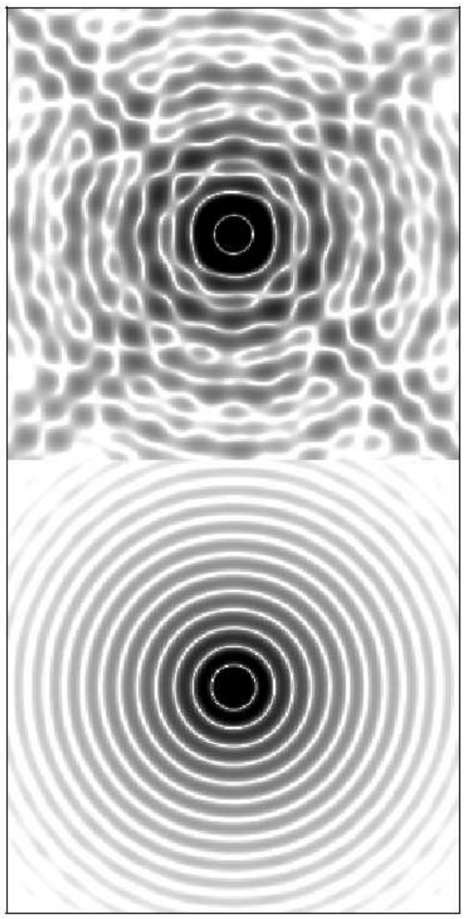


Figure 24: The Airy pattern, with and without obscurations

shows the Airy pattern, both without obscurations, and with a central obscuration and spiders in a setup typical of a telescope.

In addition, there may be phase errors in the exit pupil, because of the existence of any one of the sources of aberration discussed above. For general use, ϕ is often expressed as an series, where the expansion is over a set of orthogonal polynomials for the aperture which is being used. For circular apertures with (or without) a central obscuration (the case most often found in astronomy), the appropriate polynomials are called *Zernike* polynomials. The lowest order terms are just uniform slopes of phase across the pupil, called tilt, and simply correspond to motion in the image plane. The next terms correspond to the expressions for the OPD which we found above for focus, astigmatism, coma, and spherical aberration, generalized to allow any orientation of the phase errors in the pupil. Higher order terms correspond to higher order aberrations.

Figure 25

shows the form of some of the low order Zernike terms: the first corresponds to focus aberration, the next two to astigmatism, the next two to coma, the next two to trefoil aberration, and the last to spherical aberration.

A wonderful example of the application of all of this stuff was in the diagnosis of spherical aberration in the Hubble Space Telescope, which has been corrected in subsequent instruments in the telescope, which introduce spherical aberration of the opposite sign. To perform this correction, however, required and accurate understanding of the amplitude of the aberration. This was derived from analysis of on-orbit images, as shown in Figure 26.

Note that it is possible in some cases to try to recover the phase errors from analysis of images. This is called *phase retrieval*. There are several ways of trying to do this, some of which are complex, so we won't go into them, but it's good to know that it is possible. But an accurate amplitude of spherical aberration was derived from these images. This derived value was later found to correspond almost exactly to the error expected from an error which was made in the testing facility for the HST primary mirror, and the agreement of these two values allowed the construction of new corrective optics to proceed...

5.6 Sources of aberrations

So far, we have been discussing aberrations which arise from the optical design of a system when we have a limited number of elements. However, it is important to realize that aberrations can arise from other sources as well. These other sources can give additional third-order aberrations, as well as higher order aberrations. Some possible sources include:

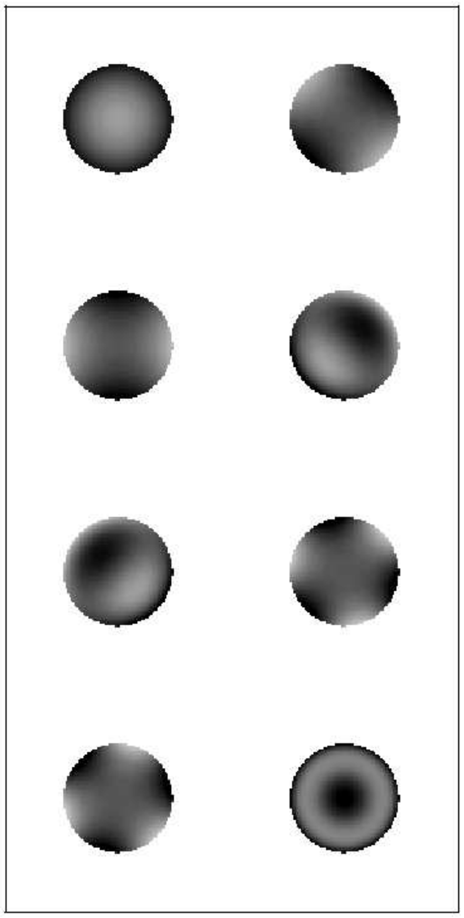


Figure 25: Images of Zernike terms, orders 4-11.

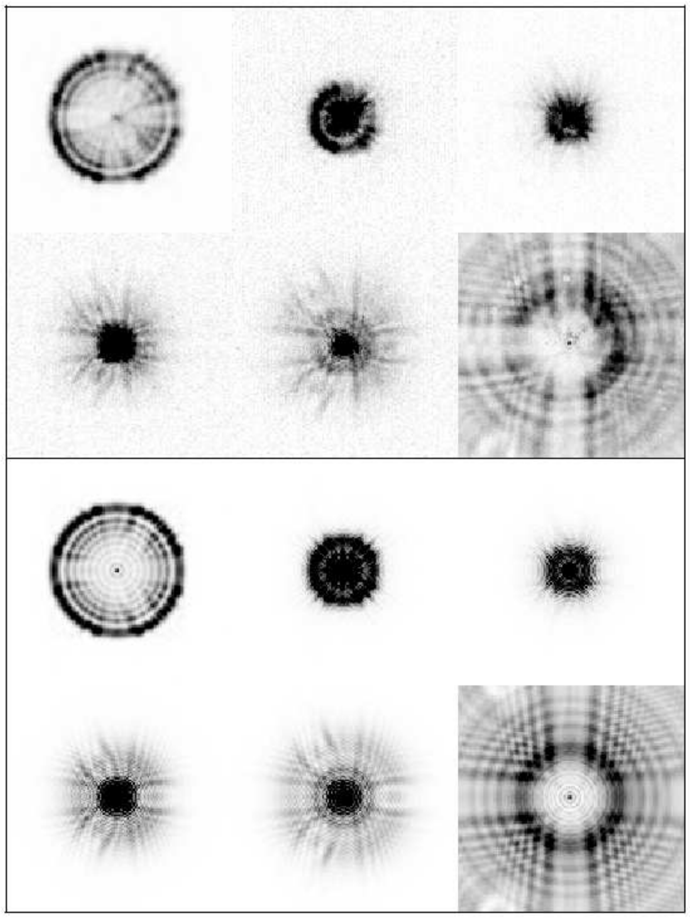


Figure 26: HST images and corresponding models

- misfigured or imperfectly figured optics : rarely is an element made exactly to specification!
- misalignments. If the mirrors in a multiple-element system are not perfectly aligned, aberrations will result. These can be derived (third-order) from the aberration expressions for decentered elements. For two mirror systems, one finds that decentering or tilting the secondary introduces a *constant* amount of coma over the field. Coma dominates astigmatism for a misaligned telescope.
- mechanical/support problems. When the mirrors are mounted in mirror cells the weight of the mirror is distributed over some support structures. Because the mirrors are not infinitely stiff, some distortion of the mirror shape will occur. Generally, such distortion will probably change as a function of which way the telescope is pointing. Separate from this, becuase the telescope structure itself is not perfectly stiff, one expects some flexure which gives a different secondary (mis)alignment as a function of where one is pointing. Finally, one might expect the spacing between the primary and secondary to vary with temperature, if the telescope structure is made of materials which have non-zero coefficients of expansion.
- chromatic aberration. Generally, we've only been discussing mirrors since this is what is used in telescopes. However, astronomers often put additional optics (e.g., cameras or spectrographs) behind telescopes which may use refractive elements rather than mirrors. There are aberration relations for refractive elements just as we've discussed, but these have additional dependences on the indices of refraction of the optical elements. For most refractive elements, the index of refraction varies with wavelength, so one will get wavelength-dependent aberrations, called chromatic aberrations. These can be minimized by good choices of materials or by using combinations of different materials for different elements; however, it is an additional source of aberration.
- seeing. The earth's atmosphere introduces optical path differences between the rays across the aperture of the telescope. This is generally the **dominant** source of image degradation from a ground-based telescope. Consequently, one builds telescopes in good sites, and as far as design and other sources of image degradation are concerned, one is generally only interested in getting these errors small when compared with the smallest expected seeing errors.

5.7 Adaptive Optics

The goal of *adaptive optics* is to partially or entirely remove the effects of atmospheric seeing. Note that these day, this is to be distinguished form *active optics*, which works at lower frequency, and whose main goal is to remove aberrations coming from the change in telescope configuration as the telescope moves (e.g., small changes in alignment from flexure or sag of the primary mirror surface as the telescope moves). Active optics generally works as frequencies less than (usually significantly) 1 Hz, whereas adaptive optics must work at 10 to 1000 Hz. At low frequencies, the active optics can be done with actuators on the primary and secondary mirrors themselves. At the high frequencies reqiured for adaptive optics, however, these large mirrors cannot respond fast enough, so one is required to form a pupil on a smaller mirror which can be rapidly adjusted; hence adaptive optics systems are really separate astronomical instruments.

Many adaptive optics systems under development: see http://www.cfht.hawaii.edu/Instruments/Imaging/AOB/other-aosystems.html

The basic idea of an adaptive optics system is to rapidly sense the wavefront errors and then to correct for them on timescales faster than those at which the atmosphere changes. Consequently, there are really three parts to an adaptive optics system:

1. a component which senses wavefront errors,
2. a control system which figures out how to correct these errors, and
3. an optical element which receives the signals from the control system and implements wavefront corrections.

There are several methods used for wavefront sensing. Two ones in fairly common use among today's adaptive optics system are Shack-Hartmann sensors and wavefront curvature sensing devices. In a Shack Hartman sensor, as shown in Figure 27,

an array of lenslets is put in a pupil plane and each lenslet images a small part of the pupil. Measuring image shifts between each of the images gives a measure of the local wavefront tilts. Wavefront curvature devices look at the intensity distribution in out-of focus images.

To correct wavefront errors, some sort of deformable mirror is used. These can be generically split into two categories: segmented and continuous faceplate mirrors, where the latter are more common. A deformable mirror is characterized by the number of adjustable elements: the more elements, the more correction can be done.

In general, it is almost impossible to achieve complete correction even for ideal performance, and one needs to consider the effectiveness of different adaptive optics systems. This effectiveness depends on the size of the aperture, the wavelength, the number of resolution elements on the deformable mirror, and the quality of the site.

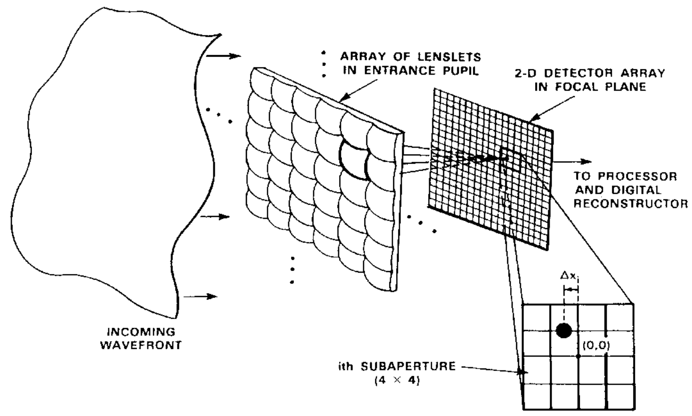


Figure 3 Principle of Hartmann-Shack wavefront sensor. The lenslets array produces an array of star images on the 2-dimensional detector array (generally a CCD or Intensified CCD). Tilt variations in the incoming, distorted wavefront result in position variations ($\Delta x, \Delta y$) of the star images on the detector. These are measured and fed to a digital processor which reconstructs the wavefront distortions. (From Murphy 1992; reprinted with permission of Lincoln Laboratory, MIT, Lexington, MA.)

Figure 27: Shack-Hartmann sensor

Clearly, more resolution elements are needed for larger apertures. Equivalently, the effectiveness of a system will decrease as the aperture in increased for a fixed number of resolution elements. One can consider the return as a function of zernike order corrected and aperture size. For large telescopes, you'll only get partial correction unless a very large number of resolution elements on the deformable mirror are available. The following table gives the mean square amplitude, Δ_j, for Kolmogorov turbulence after removal of the first j terms; the rms phase variation is just $\sqrt{\Delta_j}/2\pi$. For small apertures, you can make significant gains with removal of just low order terms, but for large apertures you need very high order terms.

Z_j	n	m	Expression	Description	Δ_j	$\Delta_j - \Delta_{j-1}$
Z1	0	0	1	constant	1.030 S	
Z2	1	1	$2r\cos\phi$	tilt	0.582 S	0.448 S
Z3	1	1	$2r\sin\phi$	tilt	0.134 S	0.448 S
Z4	2	1	$\sqrt{3(2r^2-1)}$	defocus	0.111 S	0.023 S
Z5	2	2	$\sqrt{6r^2\sin 2\phi}$	astigmatism	0.0880 S	0.023 S
Z6	2	2	$\sqrt{6r^2\cos 2\phi}$	astigmatism	0.0648 S	0.023 S
Z7	3	1	$\sqrt{8(3r^3-2r)\sin\phi}$	coma	0.0587 S	0.0062 5
Z8	3	1	$\sqrt{8(3r^3-2r)\cos\phi}$	coma	0.0525 S	0.0062 S
Z9	3	3	$\sqrt{8r^3\sin 3\phi}$	trifoil	0.0463 S	0.0062 S
Z10	3	3	$\sqrt{8r^3\cos 3\phi}$	trifoil	0.0401 S	0.0062 S
Z11	4	0	$\sqrt{5(6r^4-6r^2+1)}$	spherical	0.0377S	0.0024 S

r = distance from center circle; ϕ = azimuth angle; $S = (D/r_0)^{5/3}$.

Another important limitation is that one needs an object on which you can derive the wavefront. Measurements of wavefront are subject to noise just like any other photon detection so bright sources may be required. This is even more evident when one considers that you need a source which is within the same isoplanatic patch as your desired object, and when you recall that the wavefront changes on time scales of milliseconds. These requirements place limitations on the amount of sky over which it is possible to get good correction. It also places limitations on the sorts of detectors which are needed in the wavefront sensors (fast readout!).

band	λ	r_0	τ_0	τ_{det}	V_{lim}	θ_0	Coverage (%)
U	0.365	9.0	.009	.0027	7.4	1.2	1.8 E-5
B	0.44	11.4	.011	.0034	8.2	1.5	6.1 E-5
V	0.55	14.9	.015	.0045	9.0	1.9	2.6 E-4
R	0.70	20.0	.020	.0060	10.0	2.6	0.0013
I	0.90	27.0	.027	.0081	11.0	3.5	0.006
J	1.25	40	.040	.0120	12.2	5.1	0.046
H	1.62	55	.055	.0164	13.3	7.0	0.22
K	2.2	79	.079	.024	14.4	10.1	1.32
L	3.4	133	.133	.040	16.2	17.0	14.5
M	5.0	210	.21	.063	17.7	27.0	71
N	10	500	.50	.150	20.4	64	100

Conditions are: 0.75 arcsec seeing at 0.5 μ; $\tau_{det} \sim 0.3\ \tau_0 = 0.3r/V_{wind}$; $V_{wind} = 10$ mIsec; $H = 5000$; phnton detection efficiency (includes transmission and QE) = 20%; spectral bandwidth = 300 nm; $SNR = 100$ per Hartmann-Shack image; detector noise = $5e^-$.

The isoplanatic patch limitation is severe. In many cases, we might expect non-opticmal performance if the reference object is not as close as it should be ideally. In most cases, adaptive optics worts in the partially correcting regime. This typically gives PSFs will a sharp core, but still with extended wings.

The problem of sky coverage can be avoided if one uses so-called laser guide stars. The idea is to create a star by shining a laser up into the atmosphere. To date, two generic classes of lasers have been used, Rayleigh and sodium beacons. The Rayleigh beacons work by scattering off a layer roughly 30 km above the Earth's surface; the sodium beacons work by scattering off a layer roughly 90 km above the Earth's surface. Laser guide stars still have some limitations. For one, the path through the atmosphere which the laser traverses does not exactly correspond to the path that light from a star traverses, because the latter comes from an essentially infinite distance; this leads to the effect called focal anoisoplanatism. In addition, laser guide stars cannot generally be used to track image motion since the same telescope is being used to project and record the reference image; separate tip-tilt tracking is required. In principle, all of these problems can be solved by using multiple laser guide stars with natural stars for tip-tilt, but to date, I don't know of the implementation of any such system with multiple guide stars. Systems with single laser guide stars have certainly been tested and appear to work well; but remember, only over an isoplanatic patch.

Note that even with perfect correction, one is still limited by the isoplanatic patch size. As one moves further and further away from the reference object, the correction will gradually degrade. In principle, correction over a wider field of view is possible with *multiple* deformable mirrors and multiple reference objects, giving rise to the concept of *multi-conjugate* adaptive optics systems.

Science with adaptive optics. Typical AO PSFs. Morphology vs. photometry.

5.8 Telescope design considerations

Telescope mounts

We've talked about the optics that go into telescopes. However, it's clear that these optics need to be supported in some structure and kept in alignment with each other. The support structures needed are really an engineering issue (and a challenging one for large telescopes), and we won't disucss it here. In addition to supporting the optics, the structure also needs to be capable of tracking astronomical objects as they move across the sky because of the rotation of the earth.

There are two main different sorts of telescope mounts found in observatories: the *equatorial* mount and the *altitude-azimuth (alt-az)* mount. The equatorial mount is by far the most common for older telescopes, but the alt-az design is being used more frequently for newer, especially larger, telescopes. In the equatorial design, the telescope move along axes which are parallel and perpendicular to the polar axis, which is the direction parallel to the earth's rotation axis. In such a mount, tracking the earth's rotation only requires motion along one axis, the one perpendicular to the polar axis, and the tracking motion is at a uniform rate. In the alt-az mount, the telescope moves along axes which are perpendicular and parallel to the local vertical axis. With this mount, however, tracking of celestial objects requires motions of variable speed along both axes. An additional complication of an alt-az mount is the fact that, for a detector which is fixed to the back of the telescope, the image field rotates as the telescope tracks an object. Note, however, that the telescope pupil does not rotate with the object.

An equatorial mount is much easier to control for pointing and tracking. However, from an engineering point of view, it is much more demanding to construct, especially for large telescopes which have significant weight. The engineering complications generally result in a significantly larger cost than for an alt-az design. An alt-az telescope, however, has a significantly more complex control system, and must have an image rotator for the instruments.

Regardless of mount type, the mount is never built absolutely perfectly, i.e. with axes exactly perpendicular, exactly aligned as they should be, totally round surfaces, optics aligned with mechanics, etc. As a result, a telescope does not generally point perfectly. However, many effects of an imperfect telescope are quite repeatable, so they can be corrected for. This corrrection is done by something called a pointing model, which records the difference in true position from prediction position over the sky, and, once derived, the pointing model can be implemented to significantly improve pointing. A good telescope points to within a few arcseconds after implementation of a good pointing model.

Related to pointing is tracking performance. The issue here is how long the telescope can stay pointed at a given target. You can consider this question as how well the telescope can point over the area of the sky through which your object will drift. Since your required pointing stability should be significantly less than one arcsec, so that tracking does not degrade the image quality significantly, almost no telescopes have sufficiently good pointing to track to within an arcsecond for an arbitrarily long time. Most telescopes can track sucessfully for several minutes, but will give significant image degradation for exposures longer than this. Consequently, most telescopes/instruments are equipped with *guide cameras*, which are used to continually correct the pointing by observing an object somewhere in the field of view of the telescope (possibly the object you are interested in, but usually not, since that's where your detector is looking). These days, most guiders are *autoguiders*, meaning that they automatically find the position of the guide object, compute the pointing offsets needed to keep this object in one position, and send these offsets as commands to the telescope. The observer generally just has to choose a guide object for the autoguider to use, though they also may have to adjust the guide camera sensitivity or gain to insure that the guide star has a strong signal. These days, many autoguiders can automatically find guide stars from some on-line catalog (e.g., the

HST Guide Star Catalog, which catalogs stars down to V 14). However, if one is taking long exposures and knows that they'll need to use guide stars, make sure to find out whether such a facility is available ; if not, it may still be possible to find guide stars in advance of your observing run, e.g., from the sky survey. If so, you should seriously consider doing so, as it can take a frustratingly long amount of time to search for a guide star at the telescope in real time. Since telescope time is heavily oversubscribed at most facilities, you really want to maximize your efficiency, and doing so is a large part of what will make you a "expert" observer.

5.9 Large mirror types

Another real-world issue for large telescopes is the technology of how to build a large mirror which will not be so heavy that it will sag under its own weight. Additionally, since it has been recognized that good image quality requires that the mirrors be at the same temperature as the outside air, the mirror technology must be such that the mirror has a short thermal time constant, or, in other words, it must be able to change temperature to match the outside air fairly quickly. If necessary, one can consider thermally controlling the mirror, e.g., with heating or air conditioning.

In the large mirror regime, there are currently three leading technologies. The first is the construction of a single large mirror (monolithic) made from borosilicate glass, but having large hollowed out regions to keep the weight down. This *borosilicate honeycomb* design has been pioneered by Roger Angel at the Mirror Lab of the University of Arizona. This type of mirror has been successfully cast in a 3.5m size (used in the ARC 3.5m (APO), WIYN 3.5m (KPNO), and the Starfire Optical Range Telescope near Albuquerque), and in a 6.5m format for the MMT conversion (Mt. Hopkins, AZ) and the Magellan (Las Campanas Observatory, Chile) telescopes; in the works are an 8m format (x2) for the Large Binoculuar Telescope (Mt. Graham, AZ). The second design is also monolithic but has a mirror which is significantly thinner than the borosilicate mirror. These *thin mirrors* are being built primarily by two companies, Corning (USA) and Schott (Germany). They use materials with good thermal properties, ULE (Corning) and Zerodur (Schott). Thin mirrors are being used in ESO's 3.5m New Technology Telescope (La Silla, Chile), Japan's 8m Subaru telescope (Mauna Kea, Hawaii), the two 8m Gemini telescopes (Mauna Kea and Cerro Pachon, Chile), and ESO's Very Large Telescopes (4 8m's in La Silla). Finally, the third design make use of *segmented mirrors*, in which a large mirror is made by combining many small mirrors. This design is currently operational in the 10m Keck telescope (Mauna Kea), semi-operations for the 12m(?) Hobby-Eberly Telescope, and is planned for the 10m Gran Telescopio de las Canarias.

The borosilicate mirrors have the advantage that they are stiffer than the other designs, so the mirror support is less complicated. For thin mirrors, the support system must be activated to allow for changing shape as a function of telescope pointing. For segmented mirrors, each segment must be controlled to make sure the entire surface is smooth. The thick mirror is also less susceptible to wind shake, which can adversely affect image quality. The thin and segmented mirrors have the advantage of better thermal properties since they contain less total material.

The choice of a primary mirror technology can be complicated. In designing a large telescope, one generally first decides on an optical prescription which is chosen considering the main scientific goals for the project (e.g., large field, IR, good image quality, etc.). The primary mirror choice is made considering the choice of site (e.g, are there large temperature changes, lots of wind, etc.), availability, issues of engineering complexity, and, especially, cost (and politics). The choice of a mount and control system to use is basically a cost and operations issue.

5.10 Using telescopes

Generally it is usually fairly straightforward to use an astronomical telescope. Most of the time after arrival at an observatory can be spent checking the instrument and detector performance rather than checking the telescope performance. You should carefully consider, however, how to maximize your efficiency at the telescope; telescope time is expensive and hard to come by.

Before going to a telescope, you might consider the following checklist of things to do:

- Learn how the telescope is commanded. In particular, you may wish to find out whether scripts can be written (or are already available) to do routine motions on the sky (e.g. dither back and forth between positions, especially important for IR observing)

- Find out whether user object catalogs can be used, and if so, find out the format and prepare files. It can save a lot of time and headaches to have your coordinates preloaded in a file and save the agony of typing numbers in the middle of the night and getting them wrong.

- Find out the pointing accuracy of the telescope. You may not necessarily be able to count on the telescope pointing exactly at the coordinates that you tell it to go to. For fainter objects, it is highly recommended that you bring finding charts with you to the telescope. Note that it is now fairly easy to make finding charts from the Digitized Sky Survery (from Palomar and ESO plates). The program getimage is available for your use; this program can extract FITS images of arbitrary size from the digitized sky survey, which we have on CDROM. Any image processing package should be capable of reading these and producing hardcopy pictures. For checking a few images, you can reference the sky survey over the WWW at http://skyview.gsfc.nasa.gov.

- Find out the guiding performance of the telescope. For what exposure times is guiding required? If you will be taking exposures which need to be guided, how does one find guide stars? Can time be saved by finding guide stars in advance?

- Plan your observations. You generally want to observe objects at the lowest possible airmass. In combination with your scientific priorities, this will set the order in which you can plan to make observations. You can figure out transit times by considering the sidereal times for your night and the right ascension of your targets using

  $$HA = LST - \alpha$$

  You can get observing calendars using the program skycalendar and compute airmass tables, etc., for specific targets using the program skycalc (i'm sure lots of other programs are also available for these tasks).

  Remember to consider the calibration observations you will need to take (we'll talk more about this later). Also remember to plan for different possible conditions. For example, if you program requires photometric weather, what will you do if it's not photometric? what if the seeing is horrible, etc.?

One of the first things to be done just after dark is focussing the telescope. This generally involves taking images at a range of focus settings and comparing them to determine the best focus. One should be prepared with software for analyzing image quality to make this determination - also, software for looking at all images simultaneously. Generally, the focus position in encoded somehow so one gets a quantitative measure of the secondary location. One should be aware, however, of the possibility of slack in the gears controlling the focus mechanism, which can make the focus not repeat even when the readout position is the same; because of this, it is generally wise to always move to a focus position from one direction.

While focussing, one can generally also get an idea of the quality of the seeing of the night. Remember that seeing varies from frame to frame, and because of this, multiple exposures even at the same focus can look very different. To minimize seeing effects, one may wish to choose a focus star on which exposures of several seconds can be made: for a brighter star with very short exposures, seeing changes may confuse you. Clearly, however, one doesn't want to use a very faint star because one would like to get the focussing procedure over as quickly as possible so you can get on with your science. Remember, however, the signal-to-noise gains are substantial for a more concentrated image, so it will be worth your while to do a good job: if you rush it, you may regret it later when you have more time to notice how blurred your images are!

You also need to remember that the focus is likely to change throughout the night as the temperature changes. So continue to inpsect your images as you take them, and if the quality appears to be degrading, you should redo a focus run. Most likely, the telescope focus will consistently change in one direction (as it gets colder) and you may even be able to get a good estimate of how much it changes as a function of the temperature with experience. Which direction focus goes is a good thing to write down at a telescope, as you can save significant time during mid-night focus changes if you already know which direction you need to go. (but always beware of someone coming and rewiring the focus motor/control since the time of your last run!).

One may wish to quickly inspect an out-of-focus image for signs of large aberrations in the system. Almost certainly, nothing will be done about these immediately, but if the image quality is poor enough, it may be possible to have something (e.g., alignment) done the next day, so you still may possibily help your observing run, or certainly, you will help subsequent observers. At least, if something seems strange, you should let someone know so they can judge for themselves if there is really a problem.

Overall, this is a key point; you need to be vigilant to look for peculiarities in your data, and if you see something that hasn't previously been documented or that you don't understand, you need to ask someone about it rather than just assume that it is "normal"!

5.11 INSTRUMENTATION

5.11.1 Auxiliary Optics

Often, astronomers use additional optics between the telescope and their detector. We can loosely define a *camera* to consist of any additional optics between the telescope and the detector. This is the most general definition; sometimes, camera is just used to refer to optics which contribute to an *imaging* system, as opposed to a *spectrograph*. Here we consider some of the more common optics used in imaging cameras.

5.11.2 Field Flatteners

As we've discussed, all standard two-mirror telescopes have curved focal planes. It is possible to make a simple lens to correct the field curvature. One can derive that if a lens is placed in (or near) a focal plane, it does not contribute any significant aberrations (though if one were really working in a diffraction-limited system, one would have to check this carefully). Consequently, we can put a lens right near the telescope focal plane to flatten the field. For a field which curves towards the secondary mirror, one finds that the correct shape to flatten the field is just a plano-concave lens with the curved side towards the secondary.

5.11.3 Focal Reducers

A focal reducer is a camera which changes the focal ratio, or the scale, of the telescope. It is simply a reimaging system which demagnifies the telescope focal plane. In its simplest form, it consists of just two lenses: the collimator and the camera lens. The collimator lens has the same focal ratio as the telescope, and converts the telescope beam into a collimated beam. The camera lens then refocuses the light light with the desired focal ratio. The magnification of the system is given by:

$$m = \frac{f_{camera}}{f_{collimator}}$$

Consequently, the focal length of the entire system (telescope plus focal reducer) is:

$$\begin{aligned} f_{system} &= f_{telescope} m \\ &= f_t \frac{f_{cam}}{f_{col}} \\ &= F_t D_t \frac{F_{cam} D_{cam}}{F_{col} D_{col}} \\ &= D_{tel} F_{cam} \end{aligned}$$

where we have used the fact that $F_{col} = F_t$ (or else it won't collimate the light) and that the diameters of the beams on the camera and collimator are the same. Consequently, the scale in the image plane of the focal reducer is just the scale in the telescope focal plane multiplied by the ratio of the focal ratio of the camera to that of the telescope.

Note that with a focal reducer, one does not just get a new scale "for free". The focal reducing system may introduce additional aberrations giving reduced image quality. In addition, one **always** loses some light at each additional optical surface from reflection and/or scattering, so the more optics in a system, the lower the total throughput.

5.11.4 Pupil reimagers

Often, an additional lens, called a *field lens* is placed in the telescope focal plane. This does not affect the focal reduction but is used to reimage the telescope pupil somewhere in the focal reducer. One reason this may be done is to minimize the size that the collimator lens needs to be to get off-axis images. The size of the field lens itself depends on the desired size of the field that one wishes to reimage.

Another use of reimaging the pupil is when one is building a *coronagraph*, an imaging system designed to observe faint sources nearby to very bright ones. The problem in seeing the faint source is light from the bright one, both from scattered light and from diffraction. The solution is to put an occulting spot in the telescope focal plane which removes most of the light from the bright object. However, the diffraction structure is still a problem. It turns out you can remove this by reimaging the pupil after the occulting spot and putting a mask in around the edges which are the source of the diffraction (Lyot stop). The resulting image in the focal plane of the focal reducer is free of both bright source and diffraction structure. To do really well with a coronagraph, one also needs to minimize scattering on the optical surfaces, which requires very smooth lenses which are very clean.

Also, pupil reimagers are widely used in IR systems to reduce emission via cold pupil stops. The issue here is that the telescope itself contributes infrared emission which acts as additional background in your observations. There is little you can do about emission from the primary, since you need to see light from the primary to see your object! However, you can block out emission from regions of the pupils which are obscured already, for example, by the secondary and/or secondary support structures. To do this you put a mask in the pupil plane. Obviously, however, the mask needs to be colder than the telescope itself or else the mask would contribute the background.

5.11.5 Filters

Filters are used in imaging systems to restrict the observed wavelength range. Using multiple filters thus provides color information on the object being studied. Generally, filters are classified as broad band ($\gtrsim$ 1000Åwide) medium band (100$\lesssim$ 1000 Å), or narrow band (1 $\lesssim$ 100 Å). Broad band filters work by using colored glass, which has pigments which absorb certain wavelengths of light and let others pass. Bandpasses can be constructed by using multiple types of colored glass. These are generally the most inexpensive filters. Narrowband filters, on the other hand, work by the principle of *interference*, and are often called interference filters. They are made by using two partially reflecting plates separated by a distance d apart. The priciple is fairly simple:

When light from the different paths combines constructively, light is transmitted; when it combine destructively, it is not. Simple geometry gives:

$$m\lambda = 2nd \cos\theta$$

It is clear from this expression that the passband of the filter will depend on the

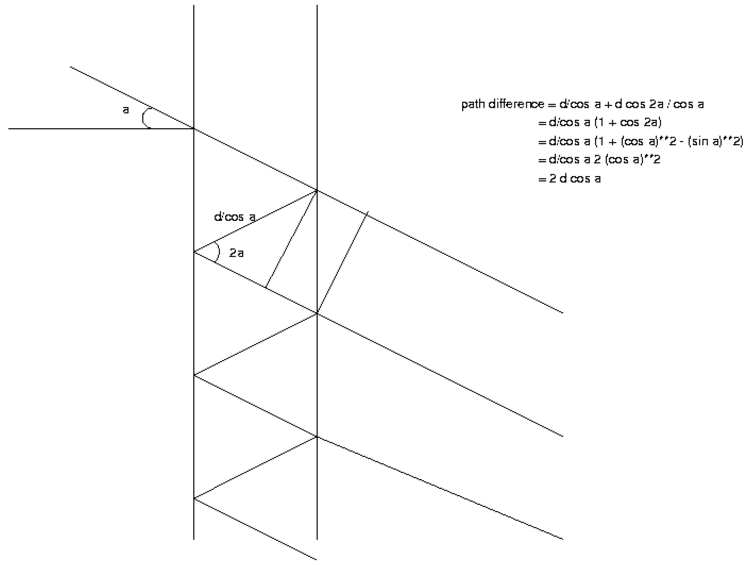


Figure 28: Schematic of an interference filter

angle of incidence. Consequently narrowband filters will have variable bandpasses across the field if they are located in a collimated beam; this can cause great difficulties in interpretation! If the filter is located in a focal plane or a converging beam, however, the mix of incident angles will broaden the filter bandpass. This can be a serious effect in a fast beam.

Since interference filters will pass light at integer multiples of the wavelength, the extra orders often must be blocked. This can be done fairly easily with colored glass.

The width of the bandpass of a narrowband filter is determined by the amount of reflection at each surface. Both the wavelength center and the width can be tuned by using multiple cavities and/or multiple reflecting layers, and most filters in use in astronomy are of this more complex type.

Note filters can introduce aberrations, dust spots, etc.

5.11.6 Fabry-Perot Interferometer

A *Fabry-Perot* system makes use of a tunable interference filter. The filter is tuned in wavelength by adjusting one of

- the spacing,
- the index of refraction (usually changed by chaning the pressure), or

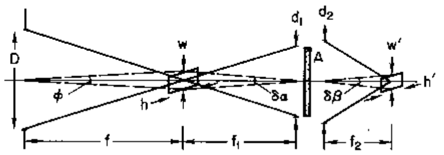


Fig. 12.4. Schematic layout of slit spectrometer with dispersing element *A*. See text, Section 12.II, for definitions of parameters.

Figure 29: Schematic of a basic spectrograph

- the tilt of the interference filter.

A tunable interference filter is called an *etalon*. Generally, etalons are made to provide very narrow bandpasses, on the order of 1Å.

A picture taken with a Fabry-Perot system covers multiple wavelengths because the etalon is located in the collimated beam between the two elements of the focal reducer. At each etalon setting, one observes an image which has rings of constant wavelength. By tuning the etalon to give different wavelengths at each location, one build up a "data cube", through which observations at a constant wavelength carve some surface. Consequently, to extract constant wavelength information from the Fabry-Perot takes some reasonably sophisticated reduction techniques. It is further complicated by the fact that to get accurate quantitative information, one requires that the atmospheric conditions be stable over the entire time when the data cube is being taken.

5.11.7 Spectrographs

A spectrograph is an instrument which separates different wavelengths of light so they can be measured independently. Most spectrographs work by using a *dispersive* element, which directs light of different wavelengths in different directions.

A conventional spectrograph has a collimator, a dispersive element, a camera to refocus the light, and a detector. The performance of a spectrograph is characterized by the *dispersion*, which gives the amount that different wavelengths are separated, and the *resolution*, which gives the smallest difference in wavelength that two different monochromatic sources can be separated. There are different sorts of dispersive elements with different characteristics; two common ones are prisms and diffraction gratings, with the latter the most commonly in use in astronomy.

The dispersion depends on the characteristic of the dispersing element. Various elements can be characterized by the angular dispersion, $d\theta/d\lambda$, or alternatively, the reciprocal angular dispersion, $d\lambda/d\theta$. In practice, we are often interested in the linear disperion, $dx/d\lambda = f_2 d\theta/d\lambda$ or the reciprocal linear dispersion, $d\lambda/dx = \frac{1}{f_2} d\lambda/d\theta$ where the latter is often referred to simply as the dispersion in astronomical contexts, and is usually specified in Å/mm or Å/pixel.

If the source being viewed is extended, it is clear that any light which comes from regions parallel to the dispersion direction will overlap in wavelength with other light, leading to a very confused image to interpret. For this reason, spectrographs are usually used with slits or apertures in the focal plane to restrict the incoming light. Note that one dimension of spatial information can be retained, leading to so-called *long-slit* spectroscopy. Also, if there is a single dominant point source in the image plane, or if they are spaced far enough (usually in combination with a low dispersion) that spectra will not overlap, spectroscopy can be done in *slitless* mode. However, note that in slitless mode, one can be significantly impacted by sky emission.

The resolution depends on the width of the slit or on the size of the image in slitless mode, because all a spectrograph does is create an image of the focal plane after dispersing the light. The "width" of a spectral line will be given by the width of the slit or the image, whichever is smaller. In reality, the spectral line width is a convolution of the slit/image profile with diffraction and the disperser response function. Note that throughput may also depend on the slit width, depending on the seeing.

Given a linear slit or image width, ω (or an angular width, $\phi = \omega/f$, where f is the focal length of the telescope) and height h (or $\phi' = h/f$), we get an image of the slit which has width, ω', and height, h', given by

$$h' = h\frac{f_2}{f_1}$$

$$\omega' = r\omega\frac{f_2}{f_1}$$

where we have allowed that the dispersing element might magnify/demagnify the image in the direction of dispersion by a factor r, which is called the anamorphic magnification.

Using this, we can derive the difference in wavelength between two monochromatic sources which are separable by the system.

$$\delta\lambda = \omega'\frac{d\lambda}{dx}$$

$$\delta\lambda = r\omega\frac{f_2}{f_1}\frac{d\lambda}{dx}$$

The bigger the slit, the lower the resolving power.

The resolution is usually characterized in dimensionless form by

$$R \equiv \frac{\lambda}{\delta\lambda} = \frac{\lambda f_1}{r\omega f_2(d\lambda/dx)}$$

Note that there is a maximum resolution allowed by diffraction. This resolution is given aproximately by noting that minimum angles which can be separated is given by approximately λ/d_2, where d_2 is the width of the beam at the camera lens, from which the minimum distance which can be separated is:

$$\omega_{min} = f_2\frac{\lambda}{d_2}$$

The slit width which corresponds to this limit is given by:

$$\omega' = r\omega\frac{f_2}{f_1} = f_2\frac{\lambda}{d_2}$$

or

$$\omega = \frac{f_1}{r}\frac{\lambda}{d_2}$$

and the maximum resolution is

$$R_{max} = \frac{d_2}{f_2(d\lambda/dx)} = d_2\frac{d\theta}{d\lambda}$$

Image slicers: preserving resolution and flux.
Fiber spectrographs: multiobject data.
Slitlets: multiobject data.

5.11.8 Dispersing elements

Prisms

Perhaps the simplest conceptual dispersing element is a prism, which disperses light because the index of refraction of many glasses is a function of wavelength. From Snell's law, one finds that:

$$\frac{d\theta}{d\lambda} = \frac{t}{d}\frac{dn}{d\lambda}$$

where t is the base length, and d is the beamwidth. Note that prisms do not have anamorphic magnification ($r = 1$). The limiting resolution of a prism, from above is:

$$R_{max} = \frac{d_2}{f_2(d\lambda/dx)} = d_2\frac{d\theta}{d\lambda}$$

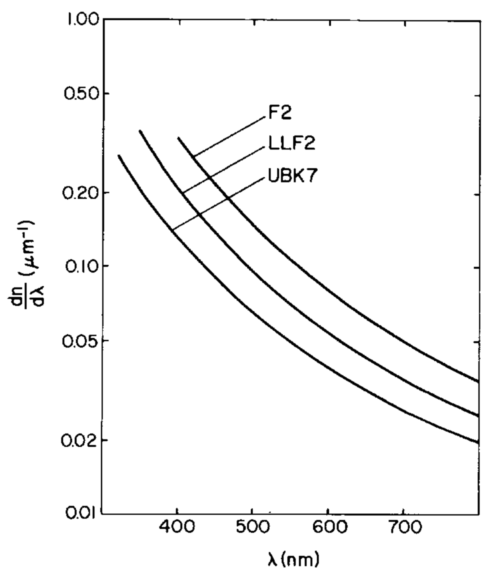


Fig. 13.1. Dispersion curves for three glasses from the Schott glass catalog.

Figure 30: $dn/d\lambda$ for typical glasses used in prisms

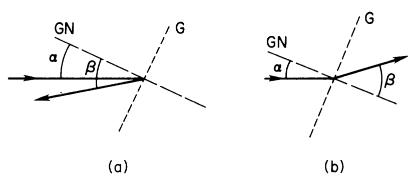


Fig. 13.2. Schematic showing angles of incidence α and diffraction β for (a) reflection grating and (b) transmission grating. See the discussion following Eq. (13.2.1) for the sign convention.

Figure 31: Schematic of a grating

$$R_{max} = t\frac{dn}{d\lambda}$$

One finds that $dn/d\lambda \propto \lambda^{-3}$ for many glasses.

So dispersion and resolution are a function of wavelength for a prism. In addition, the resolution offered by a prism is relatively low compared with other dispersive elements (e.g. gratings) of the same size. Typically, prisms have $R < 1000$. Consequently, prisms are rarely used as the primary dispersive element in astronomical spectrographs. They are occasionally used as cross-dispersing elements.

Gratings

Diffraction gratings work using the principle of multi-slit interference. A diffraction grating is just an optical element with multiple grooves, or slits. Diffraction gratings may be either transmissive or reflective. Bright regions are formed where light of a given wavelength from the different slits constructive interferes.

Figure 31 outlines the principle of a grating; light comes in at some incidence angle, and light comes out at a variety of different angles of diffraction. At a given angle of diffraction, light of some wavelengths constructively interferes, while light at another wavelength destructively interferes.

The location of bright images is given by the *grating equation*:

$$m\lambda = \sigma(sin\theta + sin\alpha)$$

for a reflection grating, where σ is the groove spacing, m is the order, and α and θ are the angles of incidence and diffraction as measured from the normal to the grating surface.

The dispersion of a grating can then be derived:

$$\frac{d\theta}{d\lambda} = \frac{m}{\sigma \cos\theta}$$

Fig. 13.3. Change in beamwidth due to anamorphic magnification of grating. See Eq. (13.2.3).

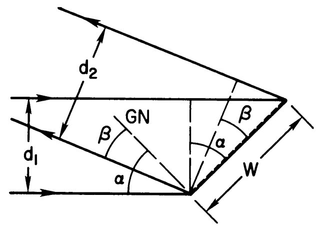


Figure 32: Schematic of anamorphic magnification in a grating

One can see that the dispersion is larger at higher order, and for a finer ruled grating. The equation can be rewritten as

$$\frac{d\theta}{d\lambda} = \frac{\sin\theta + \sin\alpha}{\lambda\cos\theta}$$

from which it can be seen that high dispersion can also be achieved by operating at large values of α and θ. This is the principle of an echelle grating, which has large σ, and operates at high m, α and θ, and gives high dispersion and resolution.

Typical gratings have groove densities between 300 and 1200 lines/mm. Echelle gratings have groove densities between 30 and 300 lines/mm.

One can derive the anamorphic magnification for a grating by looking at how θ changes as α changes at fixed λ. One finds that:

$$r = \frac{d\theta}{d\alpha} = \frac{\cos\alpha}{\cos\theta} = \frac{d_1}{d_2}$$

where the d's are the beam diameters. Note that higher resolution occurs when $r < 1$, or $\theta < \alpha$.

The limiting resolution can be derived:

$$R_{max} = \frac{d_2}{f_2(d\lambda/dx)} = d_2\frac{d\theta}{d\lambda}$$

$$R_{max} = \frac{d_2 m}{\sigma\cos\theta} = \frac{mW}{\sigma} = mN$$

where W is the width of the grating ($= d_2/\cos\theta$), and N is the total number of lines in the grating.

Note that light from different orders can fall at the same location, leading to great confusion! This occurs when

$$m\lambda' = (m+1)\lambda$$

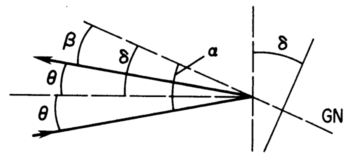


Fig. 13.4. Relation between blaze angle δ, grating normal GN, and angles of incidence and diffraction.

Figure 33: Schematic of how a blazed grating works

or

$$\lambda' - \lambda = \frac{\lambda}{m}$$

The order overlap can be avoided using either an *order-blocking filter* or by using a cross-disperser. The former is more common for small m, the latter for large m.

One can compare grating operating in low order, those operating in high order, and prisms, and one finds that higher resolution is available from grating, and that echelles offer higher resolution than typical low order gratings.

We can also discuss *grating efficiency*, the fraction of incident light which is directed into a given diffracted order. One finds that for a simple grating, less light is diffracted into higher orders. However, one can construct a grating which can maximize the light put into any desired order by *blazing* the grating, which involves tilting each facet of the grating by some blaze angle. The blaze angle is chosen to maximize the efficiency at some particular wavelength in some particular order; it is set so that the angle of diffraction for this order and wavelength is equal to the angle of reflection from the grating surface.

5.11.9 Operational items

Choice of slit width, grating (science, seeing).

How to put object in slit. Imaging the slit. Slit viewing cameras.

Calibration: wavelength, detector, etc.

Show examples of typical spectra: line lamps, flat fields, stellar spectra, galaxy spectra. Night sky emission.

Aberrations/distortions: line curvature, S-distortion, flexure.

5.11.10 Non-dispersive spectroscopy

It is also possible to use interference effects to measure spectral energy distributions instead of a dispersing element. The Fabry-Perot is an example of such a type of instrument, although it does not record all wavelengths simultaneously.

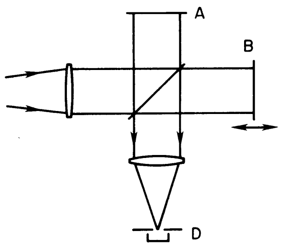


Fig. 13.15. Schematic of Fourier transform spectrometer. A, Fixed mirror; B, movable mirror; D, detector.

Figure 34: Schematic of a fts spectrograph

Another instrument which uses interference to infer spectroscopy information is the Fourier Transform Spectrometer (FTS), which is basically a scanning Michaelson interferometer. The light from the source is split into two parts using a beamsplitter. One part of light is reflected off a fixed flat mirror and the other is reflected off a mirror which can be moved laterally. The two images are combined to form fringes. The fringe pattern changes as the path length of the second beam is changed. The intensity modulation for a given wavelength (λ) or wavenumber ($k = 2\pi/\lambda$) is given by:

$$T(k, \Delta x) = \frac{T_{max}}{2}[a + \cos(2k\Delta x)]$$

and the flux after integrating over all wavelengths is:

$$F(\Delta x) = C \int I(k)T(k, \Delta x)dk = C \int I(k)cos(2k\Delta x)dk$$

where $I(k)$ is the input spectrum. Consequently it is possible to recover the input spectrum by taking the Fourier cosine transform of the recorded intensity. In practice, a discrete Fourier transform is used.

The FTS requires scanning in path spacing. But unlike the Fabry-Perot, it yields information on intensity at all wavelengths simultaneously.

6 Detectors

6.1 Basic Principles and Properties

To illustrate basic terminology used for detectors, we will consider 3 examples of detectors: photographic plate, photomultiplier, and array detector.

Detectors work because they are made of some material which interacts with photons. An incoming photon generates a chemical reaction (photographic) or photoelectron. Photoelectrons are counted, either

- one at a time, usually with amplification, or
- by collection, then subsequent measurement.

In a photomultiplier, some fraction of incident photons hit a photosensitive material and eject a photoelectron. This electron is amplified numerous times to create a large "swarm" of electrons which is detected as a pulse. Thus, photons are "counted" as they come in. Generally, simple photomultipliers do not retain any information about the location on the detector where the photon hits. There are some modern devices, however, called microchannel plates, which are essentially arrays of small photomultipliers where positional information can be obtained; one of the more common of these is called a MAMA, and exists in several instruments on the Hubble Space Telescope. Traditional photomultipliers were the workhorse of photometry from the 50's to the late 70's.

Array detectors are in more common usage today. In these devices, incoming photons create photoelectrons which are trapped in local potential wells. The amount of energy needed to eject a photoelectron depends on the type of material used. In the optical, silicon provides a good choice, but the excitation energy for silicon is too high to be used in the infrared. In the IR, various different substances are used, including HgCdTe, InSb, and PtSi. After a specified amount of time, the photoelectrons are "counted". The method by which this is done differs between different types of arrays. In CCDs, the charge is physically clocked down columns of the device, a single row at a time (a parallel transfer) then read out of serial register; CCDs are inherently asymmetric in rows and columns. In IR devices, each pixel is read individually, in sequence.

When discussing detectors, one should be familiar with some terminology:

- Quantum efficiency: fraction of photons detected. Of course, the quantum efficiency is usually a function of wavelength. Some typical values are:
  - photographic : $\sim 0.1\%$
  - photomultiplier : $\sim 10 - 20\%$
  - array detector : $\sim 20 - 90\%$
- Size and resolution elements.
  - photographic: large size (many inches), good resolution.
  - photomultiplier: several inches, no resolution (although there are arrays of photomultipliers, e.g. MAMA)

  - array detector: individual detectors started out small but are continually growing. The largest current CCDs are around 2.5 inches square, whereas the largest IR arrays are roughly an inch square. Larger effective sizes are now available (for CCDs) because some modern devices are made to be "buttable", i.e. several can be placed side by side with only very small gaps between them. Pixel sizes of array detectors are typically 15-25 microns.

- Full well: maximum number of photons possible to detect.
  - photographic: limited by number of grains
  - photomultiplier: unlimited
  - array detector: $\sim$100,000 photons.

- Readout noise. The process of counting electrons is never totally exact, so noise is introduced. Generally this is at a level of 1-100 electrons rms. If photon events are counted, rather than counting the number of photoelectrons, as is the case for a photomultiplier, there is no readout noise. In array detectors, readout noise is often important. Readout noise can actually come in two different forms: fixed pattern and random noise. The latter is what people usually refer to as readout noise. Fixed pattern noise is noise which is introduced in the readout process but which has the same spatial pattern over the detector from exposure to exposure.

- Dark current. Electrons in a given substance will be moving around at speeds which are correlated with the operating temperature of the device. If there is enough thermal motion, an electron can actually be liberated from the substance and then counted as a (spurious) photon detection. This is called dark current. Devices with lower photoelectric threshholds (used at longer wavelengths) are more susceptible to dark current, thus they must be operated at a colder temperature. CCDs are typically operated betwee -70 and -120 C, IR arrays are operated colder (LN_2 = 77K, Liquid He =4K).

- Linearity: relation between number of output electrons and input photons. In a fully linear detector, this relation is linear, with the slope of the relation given by the quantum effeciency. An alternate way to consider nonlinearity is to consider that the quantum efficiency is a function of the count level.
  - photographic: nonlinear (characteristic curve)
  - photoelectric: linear except for dead-time correction. A dead-time correction applies for bright sources; if two photons arrive essentially simultaneously, they will only be counted as a single photon.

  - array detector: CCDs *usually* linear up to 50-90% of full well. IR arrays are usually slightly nonlinear over their entire range, but repeatably so.

- Reciprocity: sensitivity changes as a function of photon incidence rate. Exists in photographic plates, *maybe* in some array detectors? Dead-time correction is a reciprocity failure.

- Defects. Most detectors are not perfect; there are often small regions which are unusable because of very low quantum efficiency, blocked columns, etc. Many devices are rejected for astronomical use because of too many defects.

- Readout speed: typically 25 microsec/pixel for array detectors. However, the chips can in many cases be read at a variety of speeds. In general, when chips are read faster, the readout noise increases.

- Digitization. In array detectors, after the charge is collected and read out, it is sent through a chain of electrons which digitizes the signal, often after amplifying it. The digitization is made by a device known as an A/D convertor; these work by comparing an input signal with a set of reference voltages which successively differ by factors of two. Thus an input signal is translated into a series of bits depending on whether the input voltage exceeds a series of reference voltages. Typical A/D correctors in use in astronomy consider 16-bits. The digital signal which comes out of the CCDs is variously referred to as counts, digital numbers (DN), or analog-to-digital units (ADU). The number of output counts is related to the number of input counts by a constant which depends on the amplification in the electronics. The amplification factor is known by most people as the gain, but astronomers define the gain of a device by the number of input electrons divided by the number of output counts (i.e., the inverse gain); the "astronomical" gain is usually specified in units of e^-/DN. Because the number which we receive from the electronics chain differs from the number of input electrons, the calculation of noise must take this into account. The photon counting noise (rms) is given by the square root of the number of detected photons. The number of detected photons is given by GC, where G is the (inverse) gain and C is the number of detected counts. Consequently, the noise *in electrons* is $\sqrt{GC}$, and in units of counts is given by $\sqrt{C/G}$. This is apart from readout noise; the latter is usually specified in units of electrons, giving a total noise in electrons of $\sqrt{GC + R^2}$, or, in units of counts, by $\sqrt{C/G + R^2/G^2}$.

  A/D convertors can only measure a positive incoming signal. At low light levels, the true input signal can be negative in the presence of readout noise. To avoid trucation of the negative signals, a constant voltage, called the bias, is added to the signal before it passes through the A/D. This bias must later be removed to preserve the correct count ratios between different sources.

A/D convertors can introduce small systematic errors in recorded count rates if the reference voltages are not carefully controlled.

- Dynamic range. A detector system can be characterized by its dynamic range, which is the ratio of the signals of the brightest and faintest sources which can be detected (with some definition of "detection"). At the bright end, the system is limited by either the full well of the detector (or the number of electrons at which the detector goes significantly non-linear), or alternatively by the limitation of the A/D convertor (e.g., if an A/D convertor has 16 bits, you can never see counts higher than $2^{16} - 1 = 65535$. At the faint end, the system is limited either by the A/D convertor (you can't detect less than one count), or by the readout noise (source buried by readout noise cannot be detected). The gain of a system is often set to maximize the dynamic range; if the readout noise is ~ 10 electrons, one can maximize dynamic range by digitizing the signal by several electrons/DN if the detector has sufficient full well.

- Saturation behavior. When full well is reached, the chip is said to be saturated. In some array detectors, especially CCDs, when a given pixel is saturated, additional charge often leaks into adjacent pixels, usually into adjacent rows rather than columns because of the construction of the CCD. Separately from detector saturation, one occasionally sees saturation of the readout electrons which can have the effect of a saturated pixel affecting the counts in subsequent pixels.

- Hysteresis generally refers to processes in which the prior exposure history of a detector affects subsequent exposures. A common example is *residual image*, in which an area of a detector which was subject to a very bright source in a previous exposure will continue to "glow" in subsequent exposures. Another form of hysteresis is so-called quantum efficiency hysteresis (QEH) in which the quantum efficiency of the chip changes as a function of previous exposure history. One manifestation of this effect has been used by astronomers, mostly with older CCDs; for some reason, when these CCDs were subjected to an extended influx of ultraviolet light, their optical quantum efficiency was found to be increased, and to stay increased at a stable level as long as the chip was kept cold. This led to the practive of "UV-flooding" CCDs as they were being cooled.

6.2 CCDs

By the nature of their operation, there are some additional effects which are peculiar to CCDs. The unique feature of CCDs is that they are used as shift registers to transfer the charge through the detectors themselves in the process of readout. The parallel transfer efficiency (transfer efficiency from one row to the next) must be extremely good in order not to lose any significant amount of charge over the large number of parallel transfers which must be performed (especially in larger devices). For example, a charge transfer efficiency (CTE) of 0.999 per transfer, which sounds good, will result in a loss of 64% of the signal over 1024 transfers, or 83% of the signal over 2048 transfers! For detectors of these sizes, CTE's of order 0.99999 or better are required. Fortunately, they are achievable, though not trivially so; many devices are rejected because of inadequate CTE.

Charge transfer problems can lead to a variety of effects which are often encountered by CCD users, especially those pushing for the most accurate photometry. An example of one such problem is known as *deferred charge*. This occurs because some CCDs transfer charge less efficiently at very low light levels. Essentially, this makes the detector non-linear at low light levels. Deferred charge can be corrected for if exposure levels are above the level where the nonlinearities are important. Alternatively, if low light levels are expected, detectors which exhibit this problem can be "pre-flashed" in which a background level of photons is put on the chip before the exposure is started; when one does this, however, one must pay the price of additional background noise.

The manufacture of CCDs, especially the low noise devices with high quantum efficiency which are needed by astronomers, is a complex procedure; only several manufacturers currently attempt this. The quality of a CCD is generally specified by its quantum efficiency (as a function of wavelength), its readout noise, and charge transfer efficiency (though there are other figures of merit as well).

CCDs can either be illuminated from the front side (where the electronics are implanted) or from the back side. To work efficiently when back-illuminated, the chips must be *thinned* by some process. Generally, thinned back-side illuminated chips have higher quantum efficiencies than front-side illuminated chips. However, the thinning process can be difficult, with a relatively high fraction of attempts at thinning ending in failure. Front-side illuminated chips not only have lower quantum efficiencies, but tend to have particularly poor blue response.

An additional way to improve blue (or any) response is to coat the chip either with some sort of anti-reflection coating to minimize reflective losses, or with some sort of lumogen which converts blue (or UV) photons to longer wavelengths where the chip is more sensitive.

The quantum efficiency is never totally uniform over the entire chip. Pixel-to-pixel variations in q.e. are typically a few percent. Over larger scales, q.e. variations can be larger; larger q.e. variations are often found in thinned chips because the thinning process may be non-uniform. The variations in q.e. across the chip, along with possible differences in illumination pattern across the chip, leads to the necessity of flat-fielding.

An additional consideration rarely considered is whether there are quantum efficiency variations *within* each pixel. In the limit where sources are very well sampled (i.e. cover many pixels), these are irrelevant, but they would lead to direct systematic photometric errors in the situation where sources are undersampled. Little is known about the possibility of such variations in different devices, but it is likely that they exist at some level.

6.3 IR detectors

Infrared arrays operate under totally different principles than CCDs. In general, there is substantially less experience with materials which are sensitive to infrared photons than there is with silicon. As a result, the infrared arrays cannot be used as readout registers in the way in which silicon arrays (CCDs) are. Instead, the IR arrays are generally coupled to a silicon array (a multiplexor). For some reason, they are generally not coupled to CCDs, however. Each pixel from the multiplexor array is read out individually, in sequence; charge is not transferred from one pixel to another. In IR arrays, each pixel can be thought of as a capacitor; as photons are detected, charge builds up on the capacitor. The amount of charge on the capacitor can be read out at any time, *without affecting the accumulated charge*. This leads to so-called nondestructive readouts for IR devices; a given pixel can be read out many times without removing the accumulated charge. The removal charge is done in a seperate reset operation.

Before charge accumulation begins, each pixel is reset to some initial value. Because of thermal noise, however, it is not possible to know precisely what this initial value is from one reset operation to the next. This would introduce a fundamental uncertainty in the total charge measured if one only read each pixel once at the end of the desired integration period. To avoid this, most astronomical IR detectors perform *doubly-correlated sampling*, in which the array is read shortly after reset (non-destructively) and then again at the end of a specified integration period. The difference between the two readouts give the desired counts per integration period. To lower the effect of readout noise, it is possible that during each of these readouts, the chip is actually read several times; averaging the successive differences reduces the effective readout noise.

PtSi, HgCd, InSb. Typical QE. Subpixel QE.

6.4 UV and other detectors

7 Data Reduction

7.1 Data reduction procedure

Generally, remove instrument signatures in reverse order from which they were introduced. Possible correction steps include:

1. A/D correction.
2. Bias level subtraction.
3. Bias structure subtraction (superbias).
4. Linearity correction.
5. Dark current subtraction (superdark).
6. Preflash subtraction, or deferred charge correction.
7. Shutter shading correction.
8. Flat field correction.
9. additonal complicated things: scattered light (has different flat field characteristics). Amplifier hysteresis correction (probably only good qualitatively!). Similarly, CTE correction. Hot pixel fixup. Residual image subtraction. Pixel area correction.
10. Fringing subtraction/sky subtraction.

Bias level subtraction

Use overscan/underscan if available. If not, you must use bias frames or dark frames of the same exposure time. Note IR detectors should be already bias-subtracted if they are using doubly-correlated reads, but often there is residual bias which is often exposure-time dependent, and if so, can be subtracted out with darks of the same exposure time (at the same time doing dark subtraction!). Warnings about overscan regions: amplifier hysteresis. Vertical and horizontal overscan. (vertical often not taken, rarely applied). Note possible drift of bias level with time (apparent in vertical overscan), so possibly fit. If fitting for varying bias, beware of bias jumps.

Bias structure subtraction

Readout electronics sometimes introduce fixed pattern noise, and in some cases, it is possible to subtract this from your data. However, you can *only* do this if the fixed pattern is *repeatable*. You must check; beware of a fixed pattern which changes with, e.g., telescope pointing. Often, this noise is variable in time or location, and consequently you only make things worse by attempting to correct for it.

Check bias frame histograms. Should be zero mean and Gaussian with width given by readout noise. This gives you first observed estimate of readout noise. If it's not zero-mean or has significantly larger width, look for fixed pattern. Construct superbias frame from lots of bias frames. Repeat several times and see if histogram is repeatable.

Required number of bias frames depends to some level on how bad the bias structure is. If it's really bad, i.e., much worse than the readout noise, you'll improve things with only a few bias frames. However, usually, the bias structure is subtle and less than the readout noise. In this case, you can often *add* noise by doing a superbias subtraction. In the absence of any bias structure, the superbias will have noise of $rn/\sqrt{N}$, and consequently, your superbias subtracted frame will have effective readout noise of $\sqrt{rn^2(1+1/N)}$. So probably want at least 10 frames, more if you really want to go faint and will be readout noise limited.

IR data should not required bias structure subtraction with doubly-correlated reads, but note that GRIM does.

Note that it is possible to do bias subration simultaneous with dark subtraction if you take calibration of the same length.

Linearity correction

For some detectors, especially in the infrared, you must make a correction for the nonlinearity of the devices. To do this, you must measure the output intensity at a range of known input intensities. To do this, you must be able to accurately control the input intensities. This is probably best done by taking data at different exposure times; since shutter timing is usually quite precise, the ratio of the exposure times give the ratio of the input intensities. However, for this to be true, you must use a light source which is stable, i.e., does not vary with time (at the one percent level); this is often hard to come by.

To make a linearity curve, you take a series of frames of increasing exposure level, usually by increasing the exposure time. To do linearity, you must know the relative levels of your different frames, which is the ratio of exposure times *if* the light source is stable. However, most light sources are not stable. To do a linearity correction, you need at least a light source which does not vary on short time scales. Then you can intersperse lamp monitoring exposures to try to calibrate out the light variations. Generally you can do this by interspersing some fiducial exposure between each of your successive light levels. For example, a linearity sequence might consist of an exposure sequence like: 1s, 2s, 1s, 4s, 1s, 8s, 1s, 16s, etc. The longer exposures are used to make the linearity curve, and the 1s exposures are there to test for source variability; if the latter occurs, the short exposures can be used to correct the other exposures for changing source intensity (as long as the changes are of sufficiently long timescale that they are well sampled by the short exposures). When you are done, make a plot of light level in each of the fiducial exposures. Look for any drifts in the lamp brightness, and if they exist, you will need to calibrate them out of your linearity exposures. Plot the linearity exposures as DN/s vs mean count level. You need to do this in regions which has a roughly constant sensitivity; don't average over the whole array; it is even possible to do it on a per pixel basis. Fit some sort of curve (low order polynomial) through the data. Then for every observation, use this curve to correct the observed count rate to the count rate you would have received at some fiducial level.

Note that the correction will be larger the further your program observation is from the fiducial light level. To minimize errors from errors in your linearity correction, take your program exposures as close as possible to a single light level; when exposure levels differ a lot, you will need to be trusting your linearity correction more.

Dark current subtraction

Detectors generate dark current from thermal excitation. Generally, the dark current in a device is different for each different pixel. Most pixels have very low dark current (because an operating temperature is chosen to bring dark current to an acceptably low level.) However, some pixels will have larger dark current (warm pixels), and some pixels may have very high dark current (hot pixels). The amount of dark current on a frame will obviously be larger for longer exposure times.

If the dark rate in a given pixel is constant in time, one can subtract the dark current by using dark frame, i.e. frames with the shutter closed but of finite exposure time (not bias frames), and making "superdark" frames by combining multiple dark frames (to beat down photon and readout noise). Sometimes the dark rate in a pixel does not appear to scale linearly with time, so sometimes you have to beware of using dark frames at some exposure time to correct object frames of a different exposure time; in the ideal world, you would take dark frames at each different exposure time you will be using for your object frames. However, dark frames are time-consuming to take. Also, you need to be careful about taking dark frames during the daytime, because sometimes instruments have small light leaks which make daytime darks useless. Cloudy nights are an ideal time to take dark frames!

For pixels with very low dark rates, which is often the majority of the pixels, you need to take very large numbers of dark frames to avoid being dominated by readout noise. Often, however, you only have time to take a relatively small number of darks, since they must be of long exposure teim. Conseqeuntly, you might wish to consider a strategy where you make a superdark frame, but set all pixels for which dark rate is too small to be accurately measured to be identically zero. This way you avoid just adding readout noise to these pixels, but you still use the measured dark rate to correct for warmer pixels.

Preflash subtraction, or deferred charge correction

For CCDs with charge transfer problems at low light levels (mostly older devices), nonlinearities can be avoided by preflashing each exposure with a small pedestal of light, although this is not ideal as it adds background noise to your object frames. If you add a preflash, it must be subtracted during data reduction; to do so, one would take a series of preflash-only frames, and combine them to make a preflash subtraction frame.

If your object frames have sufficient background by themselves so that all pixels are above the nonlinear low levels, it is possible to correct for deferred charge effects without the use of a preflash frame. Details of how to obtain the required calibration data and construct the relevant calibration frames are given in Gilliland in "Astronomical CCD Observing and Reduction Techniques" (PASP Conf. Series 23).

Shutter shading correction

Since shutters take a finite amount of time to open, the true exposure time will be a function of location on the frame. Generally, shutters move quite fast, so this is only a measureable effect for short exposure times, and even then, only for detectors which are physically large (i.e. have large shutters). However, this is becoming more common with very large format CCDs. If you care about photometry at the one percent level on short exposures with such a device (and you often do, because standard star exposures are often relatively short), you may need to consider a shutter shading correction. This can be derived by comparing N short exposures with exposure time t with a single long exposures of exposure time Nt; for example 20 1s exposures with 1 20s exposure.

For any frame, the observed number of counts, O, will be given by:

$$O = S(t + \Delta t(x, y))$$

where S is the true count rate, t is the commanded exposure time, and Δt gives the difference in exposure time at different positions on the detector. From this you would get:

$$NO_t = NS(t + \Delta t)$$

$$O_{Nt} = S(Nt + \Delta t)$$

If a sufficiently long exposure is taken in the latter case, then, to first order,

$$O_{Nt} \approx SNt$$

so

$$\frac{NO_t}{O_{Nt}} \approx \frac{(t + \Delta t)}{t}$$

so from a ratio of the two frames, one could compute Δt, and use the first equation above to correct every observed frame to the true count rate for the commanded exposure time:

$$S = \frac{Ot}{t + \Delta t}$$

Flat fielding

Flat fielding is the calibration step which corrects for varying throughput as a function of position in the field. This arises from variations in pixel-to-pixel sensitivity, illumination patterns, and filter throughput variations. All of these source cause multiplicative field-dependent errors.

In principle, these are easy to correct for; all one needs to do is to observe a source which is spatially flat, and any deviations from a flat response give the field-dependent response variations. One then corrects all program observations by dividing by this flat-field frame. However, for accurate correction, one needs to observe a perfectly flat source, which can be difficult. Often, a white spot on the inside of the dome is used which is way out of focus, but, at some level this provides a different illumination than sky because it is a near-field target. Also,one needs to consider the color dependence of the flat field response; you need separate flats for every filter. For broad filters, you also need to consider the color of the light source and how it compares with the color of your objects.

Individual pixel response requires high S/N; generally, you should aim for accuracy of a small fraction of a percent. However, be aware that computed S/N doesn't give true accuracy for large spatial scales, because the intrinsic limitation is the uniformity and color of light source. The main problem in flat fielding is systematic errors, not random errors.

Generally, flats are constructed using one of three techniques, or some combination of these three: dome flats, twilight flats, and dark sky flats.

Dome flats are good for measuring the pixel-to-pixel variation if color match is good or if your array doesn't have strong color dependence; often, however, color match is poor. Because of illumination problems, however, dome flats are usually limited to a few percent accuracy for large scales.

Twilight flats can be good for pixel-to-pixel if you don't have too many filters and if they're not too narrow and if it doesn't take too long to read array. They are generally much better than dome flats for large scale structure. They can have some color mismatch problems. The biggest problem with twilight flats is that twilight is short, so it can be difficult to get data in all filters; certainly, one usually needs to use both evening and morning twilight. A lesser problem is that one will occaisionally see stars in the twilight flats; this requires that you take several frames in each filter, moving the telescope between frames, and then reject stars in the process of combining the multiple frames.

Alternatively you can use your dark sky exposures to make the flats. However, generally, the level level is low (in the optical) so you can't build up high S/N. Consquently, when using these "superflats", one determines the pixel-to-pixel response with dome flats, and then you get the large scale variations with "superflats". Combine many frames from the night rejecting stars and fit surface to get large scale variation. The great advantage of this technique is that you match the color of the night sky perfectly.

What are the effects of flat fielding errors? For no backgrounds, the fractional error in your measurment is the fractional error in your flat. However, in the real world, we need to subtract off background, and the presence of flat-fielding errors make you estimate the sky wrong. You observe S + B. If you subtract xB, this gives S + (1-x)B, which differs from S by:

$$error = \frac{S + (1-x)B}{S} = 1 + (1-x) * \frac{B}{S}$$

so, for example, if $S = B$ and $x = 0.99$ (one percent error in sky), then you make a one percent error. However, if $B = 100S$, then you make a factor of two error!

Thus flat-fielding becomes critically important if you are measuring sources which are fainter than the sky level. This occurs in the optical as you go very faint, and almost *always* occurs in the IR, because the background is so great.

Note that the same problem occurs if the background is really non-uniform. This can be the case in the IR because of thermal contributions from the telescope/dome itself. This is especially troublesome because the background non-uniformity may very well change with time and telescope pointing.

What can you do about this? If all your frames are subject to the same flat-fielding or additive background errors, then you can beat the problem by observing the sky *on the same pixels as your object*. If you then subtract the sky frame from the object frame, you end up with a difference frame which has a much lower sky level as compared with your object. Consequently, the errors introduced by the flat-fielding errors go down. In fact, if the sky subtraction is perfect, you can reduce your object errors down to the percentage error in the flat field. However, remember that in doing this you introduce noise from photon statistics, since the subtracted frame has $\sqrt{2}$ the error of the original frame (if the sky dominates). So you don't necessarily win by doing this every time: only if the sky is brighter than the object. Depending on your requirements, you might be able to get away with observing the sky for somewhat less time than the object, if you want to keep the big systematic errors small and don't care about losing S/N at the faintest levels; but beware, you may pay something for your time saved. You can win this back if you average several sky frames for each object frame, but you have to beware for the sky frames changing with time.

Note if your object is small, you can do this without losing any observing time by dithering the object around the field. If your object is large, however, you increase your overhead substantially. But you don't have any choice.

Scattered light and flat fielding

Flats in the IR

Flats are additionally complicated in the infrared, because every frame you take, including flat field frames, have an *additive* extra component arising from the radiation coming from the telescope and dome structures. This is more of a problem at longer IR wavelengths, because at shorter ones, the sky dominates. Given this, how do you make a flat at all? One possibility is dome flats with and without lights: subtract the two for your flat. Of course, this has the potential problems associated with dome flats: incorrect illumination and color mismatches. But what can you do? Take both this kind and sky flats. If the variable spatial additive component from your telescope is small in the sky flats, you'll probably do better with these.

Since we are talking systematic errors in sky determination causing the biggest problems in your photometric accuracy, it is very difficult to estimate the errors analytically. Remeber, systematic errors are the most troublesome! Pretty much the only way you can check the accuracy of your flats is to take images of the same object at several different field locations and empirically determine how well the measurements repeat. This is highly advised in the IR, at least once with every instrument. I call it the "1000 points of light" test.

Drift-scanning

Flat fielding and distortion: pixel area effects

7.2 Things to do at the telescope

Biases

Flats

Darks

Checking detector behavior. Image histograms

Determining readout noise and gain.

Consider an ensemble of measurments taken at a light level L. The noise in this ensemble should be $\sigma^2 = LG + rn^2$, where G is the gain and rn is the readout noise in electrons, and σ is the noise in electrons.

If you do this at a lot of different light levels, then you can plot σ^2 vs L, and the slope should give you G and the intercept rn^2. However, remember that if you compute σ from the images, this gives σ in DN, so the slope will give you $1/G$. This test is also excellent for checking the basic performancer of a detector. Deviations from nonlinearity can also usually be seen on such a plot.

However, in its most straightforward application the test is very time consuming and hard to analyze: you have to take many exposures at each different light level, and then determine a gain and readout noise for each pixel and look at them all. It is much easier just to use the set of all pixels as your ensemble at each light level. However, you can't do this directly, because each pixel may have a different sensitivity and different fixed pattern noise, so you're not measuring a true ensemble. If there is significant variation of sensitivity than you can't use a whole area at all, because the noise properties will vary across the area. You can avoid these problems by working with the difference between pairs of observations: if the light level is the same in the two images, then you'll be left with an image that only has noise.

Specifically, take a pair of images and form the difference. The expected noise is

$$\sigma^2 = 2(LG + rn^2)$$

where L is the light level, G is the gain, and σ is the noise in electrons. Since $\sigma(electrons) = G\sigma_{DN}$, we have

$$\sigma_{DN}^2 = 2(\frac{L}{G} + \frac{rn^2}{G^2})$$

. You can directly measure $\sigma(DN)$ from your difference image. Make sure to do it over a region which doesn't vary significantly in light level. Now repeat the measurement at a variety of light levels and plot $\sigma^2(DN)$ vs L. If you fit a line through this, the slope is $2/G$ and the intercept is $2rn^2/G^2$. If a straight line doesn't fit the points, then there is some sort of problem, which you should probably track down.

You can abbreviate this test if you just want to get an estimate of the gain and readout noise. First, take a pair of bias frames. These have light level of zero, so the noise from the difference just gives you $\sqrt{2}rn$. (Note that you still need to take a pair in case there is superbias structure). Then take a pair at a high light level; at this level the readout noise is probably negligible, and you can determine the gain from

$$G = 2\frac{L}{\sigma_{DN}^2}$$

8 Basic astronomical image processing

8.1 What is involved in image processing?

Basic image processing tasks: I/O, display, image arithmetic, image statistics, image display (1D, 2D, 3D).

Astronomical image analysis: surface photometry, stellar photometry, image simulation, Fourier analysis, spectroscopic reduction and analysis.

Packages vs programming.

8.2 Astronomy packages

- IRAF (http://iraf.noao.edu)
- IDL (astronomy users library: http://idlastro.gsfc.nasa.gov/homepage.html)
- XVISTA (http://astro.nmsu.edu/ holtz/xvista)
- FIGARO
- MIDAS
- AIPS
- Add-on packages: STSDAS, PROS, DAOPHOT

Issues: availability, cost, philosophy, data handling (disk vs. memory), speed, ease of use (e.g., keywords vs. parm files), language and access to existing code, ability to add new code, scripts/ procedures (internal control language).

8.3 Image display

Much astronomical data is in the form of 2D images. It is critical to understand how to display such data and be able to see all of the information it contains. This is complicated by the fact that the *dynamic range* of most computer displays is smaller than that of most astronomical detectors. Most image displays provide only 8-bits of display range, giving only 256 possible colors. Any image with more dynamic range must somehow be compressed into 8-bits before it can be displayed. This can be done either by sampling the true image coarsely, which allows viewing of the whole dynamic range but can lead to the apparent loss of intensity detail, or by fully sampling only a part of the true image range, which leads to the loss of ability to view detail outside the chosen range. Beware of packages which make this choice automatically; note that this "convenience" can lead to loss of apparent information, and also that manual choice of scaling levels assures that the user knows values of numbers in his/her image: something you should know to make sure that values are somewhere around levels that you expect! Image scaling parameters are generally specified by a low and a high data value (or a low value and a range) which give the limits in the true data which will be scaled into 8-bits.

Note you can also use a nonlinear scaling to sample a larger (or smaller) range. Example: logarithmic or square root scaling.

Once an intensity subsection is chosen, it can be displayed with any choice of "color map", which specifies the colors to be assigned to each of the display levels.

These can be various shades of grey (greyscale) or some other color, or some arbitrary color scheme (pseudo-color). Note that most packages allow the user to manipulate the color table, allowing users to change the contrast and brightness of a displayed subsection; for this reason, it is usually reasonable to chose a range with a significantly larger range than 256 data values.

Most packages will allow the user to inspect individual data values based on a cursor location. Beware, however, of packages which give data readout based on scaling parameter and 8-bit display number only: these are unable to give correct values outside of the scaled region of the image.

Other useful display tools include zoom, blink, interactive image analysis (peak, valley, fwhm, etc), marking of objects, etc.

8.4 Image file formats

There are a number of ways one might consider storing a 2-D image in a disk file. Generally, since these images may be large, we are interested in formats which minimize the amount of disk space used and the amount of I/O time needed to read/write the images into memory. For this reason, images are rarely saved as formatted numbers; since the dynamic range of most detectors is at least 16 bits (-32768 to 32767), this would require at least 5 bytes per pixel, for data which is intrisically only 2 bytes of information. Consequently, images are generally stored as unformatted numbers. The binary representation of numbers can differ between machine types. Generally, however, most machines use 2's complement for integer and an IEEE definitiion for floating point numbers; machines differ, however, on the byte-order of these representations, with some architechtures putting the least significant byte first, while others put the most significant byte first.

Once a representation for the numbers is chosen, the values of individual pixels are simply dumped sequentially to a file. However, some additional information is needed to reconstruct these numbers into a 2d image. In particular, one needs to know the number of rows and columns in an image in order to properly wrap a sequence of numbers back into a 2D format. In addition, there may be additional information one wants to convey with an image, e.g., values identifying the scales and values along each axis, auxiliary information about how the image was obtained (exposure time, telescope used), etc. etc.

The standard file format for astronomical images is the FITS format. This format is described in a series of papers (refs: A&AS 44, 363; A&AS 44,371; A&AS 73, 359). Basically, a FITS file in its simplest form consists of two sections: a header section and a data section. The header section gives the required information to interpret and unwrap the data and also includes any optional information about the frame. This information is formatted in an arbitrary number of 80-byte ASCII records, each of which associates a value (integer, floating point, or character) with a predefined keyword; keywords can be up to 8 characters in length. Required keywords for FITS format are 1) SIMPLE = T or F depending on whether special extensions to the FITS format are used, 2) BITPIX, which gives the number of bits/pixel, telling the computer what format the data is stored in (16-bit integer, 32-bit integer, or 32-bit floating point). Note that FITS data by definition uses 2's complement or IEEE floating point with msb first. 3) NAXIS, which gives the total dimensionality of the image (since FITS can be used for 1-D images (spectra), 2-D images (single images), 3-D images (image stacks), etc. For image reconstruction from the file, the order of FITS pixels is that columns vary first, then rows. 4) a series of NAXIS[1-N] cards given the dimension of each axis., 5) an end card, which signifies the end of the header section.

Optional header cards include: CRVAL[1-n], CRPIX[1-N], CDELT[1-N], CTYPE[1-N] OBJECT, DATE-OBS (dd/mm/yy), INSTRUME, TELESCOP, OBSERVER, RA, DEC, EPOCH, COMMENT, HISTORY, etc. Note in particular, BZERO and BSCALE (true = disk* bscale + bzero.

Both the FITS header and the data section are padded with blanks (header) or zeros (data) to fill a multiple of 2880 bytes (which was done insure compatability with machine types of essentially all possible word lengths). Consequently, a FITS file contains a set of filled 80-card header records, a set of blank records after the END card to pad to the next larger multiple of 2880 bytes, then the raw data in a byte stream, padded with zero's to provide an even multiple of 2880 total bytes.

A variety of software packages read/write FITS images, and several subroutine libraries have been made publically available to make it easy for you to do I/O with FITS images in your own programs (FITSLIB, others?). IDL also contains FITS I/O routines.

Note that IRAF is almost unique in that it uses its own internally defined file format. It provides the capability for converting between FITS and IRAF format, but all IRAF calculations (which, remember, are disk based), are in the internal format. IRAF developers do not tell users what exactly the format of the IRAF files is so that they may reserve the right to change it in the future (with back-compatability). In order for users to be able to access IRAF images from their own software, they do provide access to subroutines which can be called from FORTRAN programs through the IMFORT library.

In recent versions of IRAF, it has become possible to use the FITS data format for all calculations instead of the IRAF internal format; I highly recommend using this (it is not the default unless you add something in your login.cl file).

8.5 Combining images

Sample mean as best estimator of parent population mean in least-squares, maximum likelihood sense. However, the mean is not especially *robust* in the case of outliers. Outliers occur in lots of astronomical examples, e.g., cosmic rays or filtering of stars. Combining images while doing the best job of rejecting outliers is a critical part of many data reduction/analysis tasks.

Some more robust estimators: median, but it produces larger error of the mean than does the mean. Often can make use of a priori knowledge about outliers: e.g., stars and CRs are always positive. This leads to routines like reject the maximum pixel. However, this leads to biases for all pixels without outliers. Hence, a better technique is min-max rejection; even this still leads to biases in pixels which have an outlier, and really throws away signal on others. Probably the best bet is to do n-sigma rejection, then recomputation of the mean. Problem here is that estimator of sigma can be very biased in the presence of outliers. So this really only is easily applied in the situation where error model is known. However, it can be applied more generally if you compute both mean and variance from the sample with the maximum value removed.

What if you don't have equal weight exposures, e.g. twilight flats? You need to normalize and potentially, consider the effect of different noise levels.

Note that the subtlety of combining frames comes into its full play when you are doing sky subtraction frames. Clearly these frames will have objects and these need to be filtered out. To do this, you must take many different sky frames with pointing shifts between them and combine the stack to make a sky frame. Be very wary of how you do this combination! It is very easy to do a straight median which to your eye will look superb; you will not see any residual sources. However, make sure to try the experiment of comparing several supersky frames by dividing them. Unless you are very careful in your combination, I'll place bets that when you divide your superskys that you will see remnants of wings of your objects. These will systematically corrupt your measurements of your program objects.

The method we've developed to optimize rejection of objects is the following. First, determine the shifts between all of the sky frames. Shift each of the frames to a common pointing and stack them up (we'll discuss how to shift images shortly). This provides a maximal S/N frame on which you can recognize faint sources. Identify all of the sources on this frame, probably automatically by some sort of positive peak detection. Make masks around each of the peaks, and make these masks larger than the peaks themselves so you also remove low level wings. Then shift each of these masks back to the pointings of each sky frame, and combine all of your sky frames at their original pointings, but ignore masked pixels in each frame. This method allows you to find and remove the very faintest objects you can see. If you were just to do peak rejection on each individual frame, you'll miss the faintest objects. If you just combine the frames without any masking at all, you'll get poor results. Note this technique is almost required if you find that the sky frames are changing with time. In this case, you'll just have to use the nearest neighbor sky frames (possibly just two frames!). However, you can find objects just fine using many more frames.

Measuring positions of objects and relative pointings between frames

It's clear that to do IR processing, you will be required to be shifting and coadding frames, because all of your object frames will not be at exactly the same pointing since you will be going off and observing sky and the telescope pointing won't repeat perfectly. In addition, you probably want to *plan* or making the pointings somewhat different for each observation so you can ignore bad pixels without losing pieces of your image. There are many other tasks for which you might need to shift or resample images, for example, if you want to combine images taken at two different telescopes, or two different wavelengths (maybe even with the same instrument if there are refractive elements!).

Before performing shifts, you need to determine how much to shift images. Generally, the quickest way to do this is to find some point sources on your image, measure their positions on each frame, and average the differences to get the frame to frame shifts. Beware if shifts are very large, because then you need to check to make sure distortion is not a problem; the presence of distortion leads to a different shift at different locations in the field. Aligning distorted images is much more complicated.

There are several ways to measure positions. A simple and accurate way to form the first moment of an image distribution:

$$< x >= \frac{\sum x I_{ij}}{\sum I_{ij}}$$

$$< y >= \frac{\sum y I_{ij}}{\sum I_{ij}}$$

Other possibilities are fits to the image shape. For example, fit a Gaussian to the image, solving for the brightness, width, and centroid. This requires a 2D non-linear least squares fit, which is somewhat more complicated, but still fairly simple. Other alternatives are Gaussian fits to the marginal distribution, a technique used the by the astrometry group at the US Naval Observatory, which has a large stake in determining accurate positions, since they are one of the main producers of quality astrometry in the world. Clearly, these methods only work if you have point sources with sufficient signal-to-noise.

An alternate way if you have extended objects or if you have lower S/N is to use cross correlation. Take your two images, multiply them together and sum the result. Do this multiple times with various shifts between the images. Fit a surface to the resulting sums as a function of row and column shifts, and find the peak. This is the desired shift. Of course, to do cross corrlations, you need to know how to shift images!

8.6 Shifting images

There are two issues you want to consider when shifting images: you want to know if they will conserve flux and you want to know if they will preserve spatial information. Generally, we probably want to *require* flux conservation, and we want to preserve as much spatial information as necessary. You may also be interested in carrying along error information for each pixel values.

If you know the underlying shape, you can determine the value at every location. In reality, however, we sample or bin the underlying function. There is a clear distinction between sampling and binning. Binning becomes equivalent to sampling when your function in well approximated by a linear function across the pixel, but not if there is any curvature.

First let's consider functions which are sampled (not binned). For many imaging applications, there is a physical process (e.g., seeing) which limits the amount of high spatial frequency information. When it is the case that there is an upper frequency cutoff in the power spectrum of the objects being observed, it is possible to recover the full function from samples of it if the samples are sufficiently fine. This is known as the sampling theorem. It says that if you have a band-limited function, you can recover the function if you have samples at spacing not exceeding $0.5/s_c$, where s_c is the frequency at which the power spectrum goes to zero. For example, if you say seeing wipes out scales less than 0.2 arcsec, you need to sample at 0.1 arcsec to recover the full function.

In practice, we don't measure a frequency cutoff and we don't rigorously define critical sampling, but we talk about it anyway. Generally, in astronomy, critical sampling is sampling at at least twice the full-width- half-maximum of the seeing disk (or the size of the Airy disk for diffraction limited applications). Note that this definition is pretty loose considering we have square, not circular pixels. Probably something like three samples is more reasonable.

If you can recover the function, then you can sample it at different locations. You can do this using the sampling theorem by something known as sinc interpolation. This works by recovering the original function; the process is done by filtering by a box function in transform space, which is equivalent to convolution with a sinc function in real space.

However, even at "critical" sampling, the binned function is not equivalent to the sampled function. So in fact, sinc interpolation may not be accurate unless you are better than critically sampled. If you are undersampled, it fails miserably. Near critical sampling, it can lead to non-flux conserving interpolation, which you generally really want to avoid. So it is not preferred unless you are very well sampled.

Perhaps the simplest method of shifting is nearest pixel shifts, and in fact, it turns out to be one of the very best ways to shift. The concept is simple: determine the shift between images, round it to the nearest integer, shift your image by that amount trivially, and add the results. This obviously conserves flux, and preserves full error information. It also does a surprisingly good job of preserving spatial information.

More sophticated routines are not necessarily better. Generally, these schemes use some sort of polynomial interpolation, where a nth order polynomial that runs through the n+1 points surrounding the desired location is determined. This is often known as Lagrangian interpolation. For $n = 1$, this is just linear interpolation, or, in two dimensions, bilinear interpolation. In 2D, you can do bilinear interpolation one dimension at a time. The formula is just:

$$F(x, y_1) = F(x_1, y_1) + (x - x_1) * (F(x_2, y_1) - F(x_1, y_1)$$
$$F(x, y_2) = F(x_1, y_2) + (x - x_1) * (F(x_2, y_2) - F(x_1, y_2)$$
$$F(x, y) = F(x, y_1) + (y - y_1) * (F(x, y_2) - F(x, y_1)$$

This can be extended to higher order. Often, bicubic is used. However, higher order doesn't necessarily buy you anything. One has to be very careful with higher order schemes, because in their simplest implementations, they do not conserve flux. It is straightforward to write higher order routines which fit *binned* functions and thus preserve flux, but many implementations you find will not do this.

In the case of very well sampled images, all of these routines do very well. As you get less well sampled, however, you need to be more careful. In fact, nearest pixel shifting does as well or better as more sophisticated routines, so I recommend it highly. Binned Lagrangian 4th order works very nicely, and I often use it.

Spline interpolation: guarantees continuity of *derivatives*.

9 Photometry

Consider techniques for stellar photometry, surface photometry, calibration of broad band and emission line photometry.

9.1 Aperture photometry

Find stars. Sum up total light. It is spread over several pixels in a form given by the point spread function $PSF(i, j)$. However, also need to estimate and subtract background.

Error estimation: we're going to sum over pixels. First just consider the object aperture:

$$S = \sum [Star(i,j) + B(i,j)] - NB$$

where the sum is over N pixels, and B is the background per pixel. To get the true signal S, N would need to approach infinity since the stellar profile continues out to great distances. In practice, however, we choose N such that a repeatable fraction of light falls within the N pixels independent of exposure. Next, we want to determine the error in our measurement of S. The noise is

$$\sigma_S^2 = \sum (Star(i,j) + B(i,j) + rn^2) + N^2\sigma_B^2$$

which is just

$$\sigma_S^2 = S + \sum (B(i,j) + rn^2) = S + N(B + rn^2) + N^2\sigma_B^2$$

So now we have to consider how we will estimate B. In the simplest case, we just go away from the star and take then mean level. Generally, we want to minimize effects of any nonuniform background, so we consider an annulus around the star. If this annulus contains N_a pixels, then

$$B = \sum_a B(i,j)/N_a$$

and

$$\sigma_B^2 = \frac{\sum [B(i,j) + rn^2]}{N_a^2}$$

From this we see that errors in determining the background are small contributors to the total error in the star if $N_a >> N$, and generally that is the case. If the errors in determining the background are negligible, then

$$S/N = \frac{S}{\sqrt{S + N(B + rn^2)}}$$

which is an *important result.*

Using this, one can consider the optimal choice of aperture. Clearly, as the aperture increses, the signal-to-noise decreases, so long as you have a reasonable fraction of the total starlight; this is especially true for fainter stars. The optimal choice of aperture will depend on the brightness of the star: a larger aperture will be better for brighter stars, a smaller aperture for fainter stars. But also recall that we don't want to use a too small aperture or else we won't be able to compare results from different frames because of changes in the PSF.

This leads to the commonly used technique of using small apertures for all of the stars on the frame, but large apertures for a few bright stars. The few bright stars are assumed to have representative PSFs for this frame, so all of the small aperture measurements can be *aperture corrected* to large aperture measurements without the increase in noise you would get if you actually used a large aperture. Note that you can't go arbitrarily small as you will eventually run into the problem of PSF variations across the frame if for no other reason than any small dependences on pixel centering.

In fact, you can do better for S/N if you use additional information. If you know the shape of the PSF, you can use this information to fit your stellar image, increasing the S/N in the process. Simple linear least squares argument, if you know PSF and position accurately, leads to

$$\sigma_S = S + N_{eff}(B + rn^2)$$

where

$$N_{eff} = \frac{1}{\sum PSF(i,j)^2}$$

Note that you'll improve S/N, but only if your assumption that your knowledge of the PSF is good is valid. This naturally leads into the next area of stellar photometry in crowding fields, when you are forced into fitting the PSF whether you like it or not.

9.2 Crowding

Clearly, the above technique breaks down as you have more than a few stars in your frame for two reasons: the stars may have overlapping light, and there may be stars in your sky annuli. This leads to techniques for *crowded field* photometry, a complicated subject which we'll just review quickly.

In crowded field photometry, the idea to that you have to solve for the brightnesses of overlapping stars simultaneously. The way you do this is to use information about the point spread function. Simply, the technique consists of: finding stars, finding the PSF, grouping the stars, and simultaneously fitting for stellar brightnesses and positions. You almost certainly have to do positions because your initial estimates will probably be biased by neighbors. You also might consider fitting for the background as well. We'll consider each one of the steps in order:

Finding stars

Need to consider automation because of completeness issues, not to mention tedium.

Look for peaks. More sublety: look for peaks which look like they have the right shape to be stars. Matched detector algorithm. Shape parameters for spurious object rejection.

Finding the PSF

Find bright isolated stars. Tabulate the PSF. However, rememeber that you're going to have to use this PSF to estimate brightnesses for other stars, so you'll need to be able to interpolate it accurately. Consider using a functional fit to PSF, possibly carry along residuals as well.

Grouping stars

Need to consider all stars that overlap. This can get troublesome, however. Possible solutions: consider central stars only, simultaneously advance all stars.

Fitting stars

For all pixels under consideration, compare observed values with first guess values. Use residuals to refine your guesses (nonlinear least squares). Continue iterating until parameters converge.

Subleties

Sky estimation. Mode approx: 3*median - 2 * mean.

Multiple iterations of entire procedure to improve PSF by subtracting neighbors, also to find new stars under wings of other stars.

Multiple colors/frames simultaneously.

Completeness. Function of crowding. Spurious detections as well as misses.

Averaging measurements. Averaging mags not the same as averaging counts. Beware of weighted means if you use estimated errors, because these can be biased.

9.3 Atmospheric extinction

We've already desribed the absorption/scattering effects of earth's atmosphere:

$$F = F_0 \exp(-\tau(z=0)X)$$

or

$$m = m_0 + 1.086\tau(z=0)X$$

$$m = m_0 + kX$$

$$m_0 = m - kX$$

where k is the *extinction coefficient.* As we've discussed, this can vary from night to night, so if you're interested in accurate photometry (better than a couple percent), you need to measure it on your night. Also remember that the extinction coefficient is wavelenghth dependent, so you need a separate number for each filter. It's also critical to measure this to check for photometric weather: even if you don't require a few percent, if there are clouds you might get a lot worse! So you need to check.

The extinction coefficient can be determined by making multiple observations of a star at different airmasses. Then you can solve for k and m_0 using least squares.

Note that you need to sample a good range of airmasses to get good leverage on the fit, and you must bracket the airmasses of all of your program objects.

When doing broad-band work, however, there's an additional subtlety because of the wavelength dependence. Two different stars might have different extinction coefficients in the *same* filter because the stars might have very different colors, which has an effect if the filter bandpass is broad. This problem can be solved by using *second order extinction coefficients*, where the extinction is a function not only of the airmass, but also of the stellar color:

$$m = m_0 + kX + k_2(color)X$$

where *color* can be either an observed or a known color for the star.

Using the above formalism, you have to solve for extinction coefficients plus a magnitude for every star you observe. This requires a fair bit of observing to constrain all of the parameters. Clearly, if you can observe stars of *known* brightnesses, you will have better constraints on the extinction coefficients. These leads us into the discussion of *standard stars*.

9.4 Standard stars

Standard stars are required so that different observers are able to compare results with each other. The reason this is true is because every observational setup is likely to have different response functions, so the same stars will *not* be observed to have the same brightnesses with each separate setup. Differences in response come from many factors: size and condition of the telescope optics, number and type of optics in the system, bandpass and quality of the filter, response function of the CCD, etc. In addition, it is very difficult to measure all of these responses in any absolute sense, so it is almost impossible to go straight from observations to inferences about physical fluxes in absolute units.

To get around these problems, systems of standard stars have been set up so observers can calibrate any new observations against the known brightnesses of the standard stars. A standard system is just a more-or-less arbitrary choice of one particular instrumental setup, but is generally one which is very stable and which someone has spent a lot of observing time establishing. Generally, the standard system is tied to absolute flux observations of some primary standard star so that observations on the standard systems can be converted to physical fluxes.

Let's consider some examples. First consider two identical systems. Clearly these will give the same observed flux as each other. Now let the systems differ only by a multiplicative factor in throughput. Now the observed fluxes will differ by a constant factor, or a constant additive factor in magnitudes. With observations of standard stars, it is straightforward to find the magnitude difference between the observed magntidue and the standard magntidue to calibrate out this constant factor. The additive constant in magnitudes is called the *zeropoint*.

A slight divergence is called for here to mention that different packages apply some zeropoints to the calculation of instrumental magnitude just to make the numbers (instrumental magnitudes) look "reasonable" (i.e., familiar). This is no problem since we determine a zeropoint on top of this to calibrate, but caution is required if talking about instrumental magnitudes measrued with different software systems!

OK, back to differing instrumental systems. Now let's consider the more usual case where the *shape* of the response curve as well as the absolute throughput differs from the standard system. Let's consider a system which has slightly different filters with a different wavelength cutoff. Now you observe a different number of counts from the standard system, and the difference depends on the spectrum of the object you are looking at. Now spectra can differ for lots of reasons, but the biggest effect is just from spectral slope differences. If your filter has a shorter wavelength cutoff (on the red side), then you'll observe less counts than the standard system, and preferentially less counts for a redder star. To first order, this can be calibrated out by solving for an additive constant which depends on the spectrum of the object being looked at. Since we don't in general know the spectrum, we have to parameterize it by something we can observe, and the best choice here is the stellar color, as inferred from observations taken through more than one filter. Clearly the best bet is also to observe a "color" taken near in wavelength to the filter we are trying to calibrate; generally one constructs a color from the desired filter and one nearby in wavelength (but sufficiently different to have real color information!).

Transformation coefficients.

$$m_0 = m - kX$$

$$M = m_0 + t(color) + z$$

where capital letters are the magnitude on the standard system, z is the zeropoint, and t is the *transformation coefficent*.

The color is generally parameterized by the ratio of the flux at two different wavelengths, or, in magnitudes, the difference between the magnitudes. The two wavelengths should be measured near in wavelength to the wavelength of the filter being corrected; generally, one uses the bandpass being corrected as one of the wavelenghts and an adjacent bandpass as the other. For example, when correcting V magnitudes, people usually use $B-V$, $V-R$, or $V-I$ for the color term.

There are two ways to define the color, either in terms of the observational system or in terms of the standard system. The latter is slightly preferred for using least-sqaures (small errors on the independent variable), and also because it allows observations from different nights to be combined. Note that this formulation does not require you to know the colors of your objects a priori, it's just algebra to figure them out as long as you have observations in both filters.

The use of these *first-order* transformation coefficients is accurate as long as your filter system does not differ much from the standard system, and additionally, that the spectrum of your program objects does not differ significantly from the spectrum of the standard objects. The more these conditions are not met, the less accurate the results. Some additional accuracy in the case of differing systems can be achieved by using higher order transformation coefficients. However, even in this case, it is always important to remember that if the spectrum of the program object differs significantly from the standards, derived fluxes can be significantly in error.

Certainly, you get to a point when the response of one system is so different than the response of another system that no transformation can be determined. In this case, you have two different photometric systems. In fact, there are several different photometric systems at use in astronomy today, and each has advantages and disadvantages.

It is common practice to combine the equations for extinction and transformation coefficients into a single equation:

$$M = m + kX + t(color) + z$$

where I have ignored second order coefficients. The advantage of combining the equations is that you can use the information about the known magnitudes of the standard stars for the extinction term, so you can combine observations of different standards at different airmasses to derive the extinction coefficient and do not need to observe the *same* star at multiple airmasses.

In practice, one observes a set of standards of different colors at different airmasses. Then one uses least squares techniques to solve for the values of k, t, and z which minimize the difference between the observed, transformed, magnitudes and the standard magnitudes. Then one applies these coefficients to measurements of your program objects to derive their magnitudes. Clearly, to do so, you must measure the colors of your program objects by making observations in more than one bandpass since, unlike the standard stars, the colors of your program objects are not known *a priori*.

9.5 Least squares fitting

We have encountered the concept of least-squares fitting several times: once to derive extinction and transformation coefficients and zero points, and again when we talked about crowded field photometry. Least squares fitting has a very large number of applications in science in general, as it is a commonly used technique to determine the parameters of some model which one wishes to fit to their data.

The concept of least-squares fitting has a statistical basis when one considers a set of observations which arise from a system which can be described by some mathematical model but where the observations have errors which are normally distributed around the predicted model values. In this case, we can write the probability of observing any particular data point given a model, $y(x)$:

$$P(y_i) \propto \exp \frac{-0.5(y_i - y(x_i))^2}{\sigma_i^2}$$

The total probability of observing *all* the data points in a data set is just the product of all of the probabilities:

$$P \propto \Pi_{i=1}^{N} \exp \frac{-0.5(y_i - y(x_i))^2}{\sigma_i^2}$$

We wish to ask the question: for what choice of model, $y(x)$, over all models which fulfill some functional form, is the total probability maximized. Maximizing the above function is the same as maximimizing its logarithm, or, alternatively, minimizing the negative of its logarithm, i.e., minimizing

$$\chi^2 \equiv \sum_{i=1}^{N} \frac{(y_i - y(x_i))^2}{\sigma_i^2}$$

This is the standard least-squares equation: one wishes to minimize the sum of the square of the differences between observation and model by adjusting the model parameters. We define this quantity as "chi-squared", or χ^2. For a good fit, one would expect that on average, points deviate from the model by roughly σ_i, so one would expect a χ^2 for a set of N observations to approach N if the model is a good model and the errors have been estimated properly. In fact, it is possible given an observed value of χ^2 to compute the probability that this value would be obtained if the model is correct; this allows one to judge the quality or likelihood that the model which minimizes χ^2 is actually the correct model. One often discussed the reduced χ^2 quantity, χ^2_ν, which is just χ^2 divided by $N + M$, where N is the total number of points, and M is the number of free parameters in the fits; the latter is there because any set of observations which does not have more data points than the number of free parameters can generally be fit perfectly even for an incorrect model. The total $N + M$ is called the degrees-of-freedom of the fit.

How does one do least-squares in practice? Basically, one wishes to minimize χ^2 with respect to one or more parameters of a model. Let us first consider models which can be written in the form:

$$y(x_i) = \sum_{i=1}^{M} a_k X_k(x_i)$$

In this case we can write

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - \sum_{k=1}^{M} a_k X_k(x_i)}{\sigma_i} \right]^2$$

We minimize by taking derivatives with respect to each parameter, a_k, and simulateously solving for the parameters with zero derivative with respect to each parameter. This gives us a system of M equations of the form:

$$\sum_{i=1}^{N} 2 \left[\frac{y_i - \sum_{k=1}^{M} a_k X_k(x_i)}{\sigma_i} \right] (-X_m(x_i)) = 0$$

for each parameter, m, (and you can get rid of the negative sign as well). From this we get

$$\sum_{i=1}^{N} \frac{X_m \sum_{k=1}^{M} a_k X_k(x_i)}{\sigma_i^2} = \sum_{i=1}^{N} y_i X_m(x_i)$$

If we define a matrix, α_{kl}, to be

$$\alpha_{km} \equiv \sum_{i=1}^{N} \frac{X_m(x_i) X_k(x_i)}{\sigma_i^2}$$

and a vector, β, to be

$$\beta_k \equiv \sum_{i=1}^{N} \frac{y_i X_k(x_i)}{\sigma_i^2}$$

then we can write:

$$\alpha_{km} a_m = \beta_k$$

This is a linear set of equations with M unknowns, the a_m, which can be solved by standard linear algebra techniques (see, for example, Numerical Recipes, chapter 2), e.g. matrix inversion:

$$a_m = \alpha^{-1} \beta$$

Associated with solving for the "best" parameters, one often wishes to compute the errors associated with the fit parameters. This is discussed in Numerical Recipes (chapter 15), with the result:

$$\sigma^2(a_m) = \alpha_{mm}^{-1}$$

One can also consider the situation where a model is *non-linear* in the parameters, a_m, i.e., the derivatives with respect to each parameter may depend on the parameter value. This leads to non-linear least squares techniques. These are more complex, as one can have situations where there are many minima in χ^2 and one needs to find the global minimum rather than a local one. Such problems require a starting guess of a reasonable solution and then iteration towards the best solution. The crowded field photometry problem falls into this category because the model is nonlinear in the position parameters.

9.6 Photometric systems

For broadband systems, the most common is the UBVRI system, originally established by Johnson in the 1960s. However, even the UBVRI system has undergone some revisions, and Johnson's original system is not quite the one used predominantly today.

UBVRI. Cousins RI. Subtleties of Johnson vs. Cousins vs. Landolt.

Gunn system, uvgriz.

HST system.

Sloan system.

Choice of photometric system for studies of stars: determination of temperatures, reddenings, surface gravities, metallicites.

Medium band: Stromgren system, uvbyβ. Others: DDO, Washington.

9.7 Emission line photometry

Observations need to subtract continuum if it is there. Neighboring filters. Beware of neighboring lines. Optimal choice of filter widths and exposure time ratios.

Calibration different. Use emission line sources of known brightness. Need to know filter profile if objects are at different redshifts. Alternatively use stars to get integral under filter, and use ratio of line throughput to total throughput to get flux.

Surface photometry vs. integrated photometry.

9.8 Surface photometry

What if we have extended (resolved) objects? Can talk about either integrated or surface brightnesses. If integrated brightnesses, use some sort of aperture, remember aperture corrections (though you probably can't estimate them very well, and they probably don't matter too much).

For surface brightnesses, just look at fluxes per area on sky. Remember that there will be some mixing from PSF/seeing; topic of extracting most spatial information will be discussed next time. Counts/pixel go directly to mag/square arcsec using stellar calibration, size of pixels.

If object is irregular, not much more you can do. Hard to observe faint irregular objects accurately because of S/N. However, if object has some degree of regularity (e.g., galaxies), one can average over regions to increase S/N substantially. Generally, many galaxies can be fairly well parameterized by elliptical isophotes, so if you can determine ellipse parameters, you can average along a given isophote to increase S/N by $\sim \sqrt{N_{pix}}$.

Note that if all isophotes are concentric and have same ellipticity, one could use elliptical aperture photometry. But many galaxies have twisting isophotes, so often one needs to solve for varying ellipticity and position angle as a function of semimajor axis.

Several methods for doing this have been presented by Kent (ApJ 266, 562) and Lauer (ApJ 311, 34). They are similar in that they solve for elliptical isophotes using nonlinear least squares. Both start with some a priori guess of isophote parameters. Kent method solves for Fourier coefficients which parametrize (observed-guess) ellipse; ellipse parameters modified to minimize A_1, B_1, A_2, B_2:

$$I = I_0 + A_1 cosE + B_1 sinE + A_2 cos2E + B_2 sin2E + \cdots$$

where E is the eccentric anomaly:

$$x = a_0 cosE$$

$$y = a_0 sinE(1 - \epsilon_0)$$

If image is perfectly elliptical, then all higher order terms vanish. For this method, one needs to choose a set of ellipses (semimajor axes) and sample them in the data. Generally, these are spaced by $\sim$ one pixel in semimajor axis: this is good for spatial resolution but can be poor for S/N.

Lauer describes galaxy as having power law surface brightness profile with constant ellipticity/position angle between some a priori specified semimajor axes.

$$G(x, y) = \mu_k (\frac{r}{d_k})^\gamma$$

$$\gamma = log(\frac{\mu_{k+1}}{\mu_k})/log(\frac{d_{k+1}}{d_k})$$

$$d_k = a_k(1 - \epsilon_k)[sin^2(\phi_k - \theta_k) + (1 - \epsilon_k)^2 cos^2(\phi_k - \theta_k)]^{-1/2}$$

All pixels are included in fit and parameters are adjusted by nonlinear least squares.

How to obtain. Calibration.

Surface photometry, integrated photometry and variable pixel area effects.

10 Spatial filtering

Various reasons why you might want to apply spatial filters: low pass filters to get rid of cosmic rays/bad pixels, high pass filters to look for features. Many techniques possible. Generally, in other fields, people often filter in Fourier domain, but doesn't seem to be so useful in astronomy.

Low pass filtering: smoothing, median filtering.

High pass filtering: unsharp masking.

Edge detection.

Object recognition.

Deconvolution

To remove known PSF. Problem: noise.

If you know expected profile of model, you're better off convolving and fitting, e.g. psf fitting. Sometimes, however, you don't. If you're going to fit a model to a deconvolved image, strongly consider fitting a convolved model to the data instead.

Fourier deconvolution: since observed function is convolution of true signal with response function, we have:

$$S(f) = R(f)U(f)$$

where U is true signal, R is response function, S is observed signal. Consequently, just FFT observed and response functions, divide, and IFFT.

However, in reality there is noise, so we can't directly get U.

$$O(f) = S(f) + N(f)$$

where O is the observed signal. Instead we want to get our best estimate of U:

$$U\prime(f) = \frac{O(f)\Phi(f)}{R(f)}$$

where Φ is some filter we apply to eliminate the noise. What should we choose for Φ? From least squares considerations, one can derive the *Wiener filter*:

$$\Phi(f) = \frac{|S(f)|^2}{|S(f)|^2 + |N(f|^2}$$

where the interpretation is easy to see graphically. However, to do this, we need to estimate $S(f)$. Either know noise properities a priori, or take power spectrum and estimate them.

Lucy deconvolution: guess image, convolve with PSF, take ratio of image to model, convolve ratio with PSF transpose, use this to update model. Iterate.

Maximum entropy.

Beware that some deconvolution schemes do not preserve flux.

11 Spectroscopic data

11.1 Spectroscopic data reduction

Review long-slit spectrograph. Describe what images look like: star, sky, galaxy. Use APO 3.5m/DIS 000110 data as example.

Data reduction:

1. normal CCD processing: overscan, bias, dark, preflash.

2. flat fielding. Note problem that dome flats have spectral energy distribution of light source. "Flatten" the flats. Discuss slit illumination (note slit contribution) and methods for determining it (dome and twilight flats). (call specflat 44 53).

3. wavelength calibration. Basic principles. Caveats: line curvature, flexure (note skylines). (call read 40 43. call wavecal 40 41 43.)

4. flux calibration: correction for throughput as a function of wavelength. Not always required, e.g. if measured strengths *relative* to nearby continuum. Spectrophotometric standards, e.g. Massey et al. ApJ 328, 315 (1988). (call ./flux)

5. Object reduction: extracting object spectrum and sky spectrum. Aperture extraction vs. optimal extraction. Caveats: spectral curvature.

11.2 Spectroscopic data analysis

Line positions: velocities. Note cross correlation techniques.

Line widths: rotation and velocity dispersion. Note various techniques (Fourier).

Line strengths: measure flux in line (usually emission lines).

Equivalent widths: gives strength of line (absorption or emission) *relative* to strength of continuum.

12 Astronomy databases

Astronomy databases include both homogeneous and inhomogeneous ones (i.e. collected from a variety of sources). Recall there is no such thing as perfect data, so one should always consider how the data were obtained and potential limitations. This may become more of an issue as we evolve towards larger and larger projects for which automated processing becomes essential.

Example: pipeline data quality: bogus SDSS quasars in the EDR

Example: selection effects: high-redshift rotation curves

12.1 General data bases

- ADS (http://adswww.harvard.edu): Publications

- ADC (http://adc.gsfc.nasa.gov): various catalogs
- CADC (http://cadcwww.dao.nrc.ca): nice front-end to lots of catalogs
- NED (http://nedwww.ipac.caltech.edu): extragalactic object front-end
- SIMBAD/CDS (http://cdsweb.u-strasbg.fr/Simbad.html: many published tables/catalogs

12.2 Specific project databases

- HST (http://archive.stsci.edu)
- 2MASS (http://www.ipac.caltech.edu/2mass/)
- SDSS (http://ganymede.nmsu.edu/sdss/)

12.3 SDSS

Survey overview: image roughly 1/4 celestial sphere using drift scanning in 5 colors (ugriz). Follow-up spectroscopy of roughly 1 million galaxies. Drift scanning produces stripes across sky, obtained in two strips; there is some overlap. Imaging camera has 6x5 CCDs, each 2Kx2K.

Data reduction by pipeline. Objects identified by run, rerun, camcol, field, object. A run is a single drift-scan across the sky.

Data products:

- SDSS database: includes position, flux, shape information on all objects, redshift, classification, and line strengths of those with spectra. Note data flags.
- Spectra
- Atlas images
- Corrected images

Data interfaces:

- Sky server
- MAST
- sdssQT
- FNAL image retrieval

Observing proposals

Proposal planning and writing:
scientific justification for non-specialists.
appropriate choice of observing time.
justification of time. Weather considerations.

Observing techniques

planning. Science planning. Instruction manuals. Contingency for weather.
Calibration data. More open shutter time often better, even during day!
Efficiency. Observing files and scripts.
Data quality checks: real-time reduction vs later reduction. Focus checks.