X-ray astronomy is different ..... 

- Problem 1: Photon counting with small number statistics
- Problem 2: Spectral line spread function is often broad and messy
  - forced to forward-folding approach
- Problem 3: Bands are very broad, so energy (wavelength) dependence more obvious (e.g. in PSF)
- Problem 4: Different optics - PSF degrades rapidly off axis
- Problem 5: The telescope is not pointing steadily like, say, HST - it's moving back and forth across the source.

- But:
- Advantage: We have more information on each photon (position, energy, arrival time)
Scope

- Caveat: will cover ACIS imaging data only
- Basics the same for HRC and gratings, but with extra wrinkles
• In optical astronomy, the primary data set is an image. In radio interferometry, it's a visibility array.

• In X-ray astronomy, the primary data set is an event list - a table of (putative) photons
  - Our software makes it easy to generate an image from the event list, so it's easy to forget that's what you have. But making the image loses information.
  - First cut way of thinking about the event list: it's a 4-dimensional array of x, y, time, energy. But most pixels are empty (we don't have many photons!) so it's more compact to just list the non-empty ones.
  - Complication: we actually have many more parameters for each photon, not just 4.
### Columns for Table Block EVENTS

<table>
<thead>
<tr>
<th>CoNo</th>
<th>Name</th>
<th>Unit</th>
<th>Type</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>time</td>
<td>s</td>
<td>Real18</td>
<td>154361559.6127299946:154436827.4158589973 S/C TT corresponding to mid-exposure</td>
</tr>
<tr>
<td>2</td>
<td>ccd_id</td>
<td>Int2</td>
<td></td>
<td>0:9 CCD reporting event</td>
</tr>
<tr>
<td>3</td>
<td>node_id</td>
<td>Int2</td>
<td></td>
<td>0:3 CCD serial readout amplifier node</td>
</tr>
<tr>
<td>4</td>
<td>expno</td>
<td>Int4</td>
<td></td>
<td>0:02147483647 Exposure number of CCD frame containing event</td>
</tr>
<tr>
<td>5</td>
<td>chip(chipx,chipy)</td>
<td>pixel</td>
<td>Int2</td>
<td>1:1024 Chip coords</td>
</tr>
<tr>
<td>6</td>
<td>tdet(tdetx,tdety)</td>
<td>pixel</td>
<td>Int2</td>
<td>1:8192 ACIS tiled detector coordinates</td>
</tr>
<tr>
<td>7</td>
<td>det(detx,de ty)</td>
<td>Real4</td>
<td></td>
<td>0.50:8192.50 ACIS detector coordinates</td>
</tr>
<tr>
<td>8</td>
<td>sky(x,y)</td>
<td>pixel</td>
<td></td>
<td>0.50:8192.50 sky coordinates</td>
</tr>
<tr>
<td>9</td>
<td>pha</td>
<td>adu</td>
<td>Int4</td>
<td>0:36855 total pulse height of event</td>
</tr>
<tr>
<td>10</td>
<td>pha_ro</td>
<td>adu</td>
<td>Int4</td>
<td>0:36855 total read-out pulse height of event</td>
</tr>
<tr>
<td>11</td>
<td>energy</td>
<td>eV</td>
<td>Real14</td>
<td>0:1000000.0 nominal energy of event (eV)</td>
</tr>
<tr>
<td>12</td>
<td>pi</td>
<td>chan</td>
<td>Int4</td>
<td>1:1024 pulse invariant energy of event</td>
</tr>
<tr>
<td>13</td>
<td>fltgrade</td>
<td>Int2</td>
<td></td>
<td>0:255 event grade, flight system</td>
</tr>
<tr>
<td>14</td>
<td>grade</td>
<td>Int2</td>
<td></td>
<td>0:7 binned event grade</td>
</tr>
<tr>
<td>15</td>
<td>status[4]</td>
<td>Bit(4)</td>
<td></td>
<td>event status bits</td>
</tr>
</tbody>
</table>
Energy slices through an event list, 0.1 - 10 keV
Level 1 Event List - Calibrated but Dirty

- Node boundaries
- Lots of background
- Bad columns
- Source!
- Bad pixels
Level 2 event list - cleaned and filtered

Energy filter 300-7000 eV removes background but not signal
Grade filter removes cosmic ray events etc
Good time filter removes times of high background, poor data quality

Sources fuzzy far off axis (PSF big)

Beware chip gaps!

More sources!
During an observation, Chandra's optical axis describes this 'dither pattern' on the sky, (Problem 5), smearing the image of a point source. The RA, Dec, roll angle of the telescope versus time is called the 'aspect solution'; the asol1.fits file provides this for each observation.

We record the motion of the guide stars in the star tracker so that we can calculate RA and Dec for EACH PHOTON and so reconstruct the image.
Chandra aspect-corrected data

This is what you get after calibration but before cleaning the data. Note the sharp point sources near the center.
In instrument space, the photons are spread out over 20 arcsec and have bad columns going through them - so be careful of the effective exposure time. If you didn't dither, you could lose the source entirely if it landed on a bad pixel.
Exposure map

Spatial Response: EXPOSURE MAP

The *Exposure Map*, \( E(\Delta h, \lambda, \hat{p}) \), retains spatial information at the expense of spectral. It has units of [cm\(^2\) counts photons\(^{-1}\)].

\[
\int d\lambda \ S(\lambda, \hat{p}) \approx \frac{C(\Delta h, \hat{p})}{E(\Delta h, \lambda, \hat{p})}
\]

\( C \) is the observed counts per spatial bin in a pulse-height bin. \( S \) is the source flux, with units of [phot cm\(^{-2}\) s\(^{-1}\) Å\(^{-1}\)].

**Instrument Map** — efficiency calibration information, band integrated. (create with `mkinstmap`)

**Exposure Map** — applies telescope aspect history and coordinate transformations (= area x time). (create with `mkexpmap`).

\[ = \text{mirror area x detector QE} \]

\[ = \text{Instmap Aspect} \]
Problem 3: Exposure map is energy dependent; must assume a spectrum if using a broad band.
Event analysis or binned analysis?

• Don't make an image too quickly. If you can get an answer directly from the event list, that's better - binning the data loses information, and collapsing the axes loses information.

• Spatial analysis: make an image (using dmcopy)
  - lose energy and time information

• Spectral analysis: make a 'PHA file' using dmextract (or a grating spectrum using tgextract )
  - lose spatial and time information

• Temporal analysis: make a light curve using dmextract
The fundamental equation of astronomy

\[ N(E) = A(E)F(E) \Delta T \]

Our instrument makes a spectrophotometric measurement; the sensitivity ("effective area") \( A(E) \) tells us how to convert from flux to instrumental counts for a given exposure time \( \Delta T \).

But, a real instrument doesn't measure the true energy, it measures instrumental energy \( E' \). The line spread function ("response matrix" in X-rays) \( R(E,E') \) describes how a monochromatic input spectrum is broadened by the instrument (Problem 2).

Let us further assume that the instrumental energy \( E' \) is measured in discrete channels (bins) \( E'_i \). Then

\[ N(E'_i) = \int A(E)R(E, E'_i)F(E)dE\Delta T \]

Of course, you may not be measuring all of the light from the source. Even if it's a point source, there may be an aperture correction. We need the PSF \( P(x-x',y-y') \) and the spatial dependence of the QE, \( q(x,y) \). Then at a given instrument position \( x',y' \)

\[ N(E'_i, x'_i, y'_i) = \int \int A(E)R(E, E'_i)F(E, x, y)P(x-x'_i, y-y'_i)q(E, x'_i, y'_i)dEdxdy\Delta T \]

The source may also be variable in time - we'll ignore this for the purposes of this talk. The detector sensitivity is time-variable on long timescales, but for a single observation you just have to worry about times when the data is filtered - the Good Time Intervals (GTIs)

\[ N(E'_i, x'_i, y'_i) = \int \int A(E)R(E, E'_i)F(E, x, y, t)P(x-x'_i, y-y'_i)q(E, x'_i, y'_i)dEdxdydt \]
Pulse height

When you plot an optical spectrum, the wavelength (or energy) axis is really an instrumental quantity. A spectral line is broadened by instrumental effects, so the energies plotted are not the true energies of the photon. However, the instrument is calibrated (i.e. the definition of instrumental energy is rescaled) such that the peak of a line is at the correct energy.

In X-ray astronomy, instead of using the instrumental energy $E'$, we work with the energy bin number. For historical reasons to do with long-forgotten instruments, this bin number is know as the PI channel (for 'pulse invariant' channel) - we'll denote it by $P$. So, for fixed energy bin widths $dE$,

$$E' = P \ dE = \text{[on average]} \ E$$

The instrument actually measures a raw energy bin number $p$, called the PHA channel, or 'pulse height analyser channel'. The scaling of the instrumental energy to real energy depends on position and time:

$$E'(\text{raw}) = p \ dE = g(x,y,t)P \ dE$$

This function $g$ (the gain) is usually assumed to obey

$$g(x,y,t) = g_{\text{spatial}}(x,y) \ g_{\text{t}}(t)$$

and we provide calibrations of both the spatial gain and the temporal gain.
We pick a parameterized $F(E)$ such as warm absorber models, lines, thermal plasma codes. Which $F(E)$? You must pick one based on expected physics, but match number of free parameters with quality of data.

With less than 100 counts, we usually just use count ratios (X-ray colors) for spectral analysis.

Does one model fit significantly better than another? Be careful that two physically different models may look quite similar in $F(E)$ space.

Incompletely calibrated instrumental features may show up in residuals, limiting factor in high S/N spectra – these features may include edges. Beware apparent science in regions where $A(E)$ is changing rapidly.
• The CALDB (Calibration Database) contains everything you need that's not part of your specific observation.

• It's designed as a multimission directory structure. The Chandra files are in $CALDB/data/chandra

• Within that, they are arranged by instrument and kind of calibration. But, with luck, the software will find the CALDB files you need automatically.

• Just make sure that you keep the CALDB up to date! But, be careful - if you start off processing with a given version of the CALDB and CIAO, then upgrade to a new CALDB and CIAO, things are sometimes incompatible. Check the release notes.