Massachusetts Institute of Technology

Physics 8.14 Spring 2009

Data analysis and curve fitting

Due at the start of session 5 (week of February 17, 2009)

Problem 1: Fitting a simple counting experiment

Cosmic ray detectors on the earth's surface detect showers with charged particles created by cosmic protons colliding with the atmosphere. A subset of 12 detectors in a much larger array detect N charged particles. The following table gives the measured data:

N	0	1	2	3	4	5	6	7	8	9	10	11	12
Events	5	10	24	40	42	36	22	14	3	4	1	2	4

a. Fit the data to a Poisson and a Gaussian distribution, respectively. In each case, state the mean, variance, standard deviation with errors and the reduced χ^2.

b. Determine confidence levels for each fit hypothesis (see, e.g. Bevington).

c. Create a histogram with fit curves.

d. What assumptions have you made about the physical processes that produce the charged particles?

Problem 2: Gaussian or Lorentzian?

In high-resolution spectroscopy experiments, such as Doppler-free spectroscopy, Mößbauer spectroscopy, Quantum Information Processing, and the Zeeman effect, you will have to fit lineshapes (or dips). From the Uncertainty Principle we know that $\Delta E \Delta t \geq \hbar/2$ for a wave-packet, which translates into $\Gamma\tau > \hbar$ for a resonant line with full width at half maximum (FWHM)of Γ from a de-excitation exponential lifetime τ. There is an interesting relation between Γ and τ, due to the fact that the energy Fourier transform of an exponential decay in time results into a Lorentzian non-relativistically (as is mostly the case in 8.14). An easy derivation follows:

The emission of a spectral line can be described as a damped oscillator with natural frequency ω_0, and damping constant γ. The time-dependent amplitude is represented by

$$f(t) = C\,\mathrm{e}^{-\gamma t}\,\mathrm{e}^{i\omega_0 t} \quad \text{with} \quad \int_{-\infty}^{\infty} |f(t)|^2\,dt = C^2 \left| \int_{-\infty}^{\infty} \mathrm{e}^{-2\gamma t}\,dt \right| = \frac{C^2}{2\gamma} = 1 \Rightarrow C = \sqrt{2\gamma} \qquad (1)$$

which leads to the complete amplitude

$$f(t) = \sqrt{2\gamma}\mathrm{e}^{-\gamma t}\mathrm{e}^{i\omega_0 t}, \qquad (2)$$

which is a decaying sinusoidal function, as expected. The Fourier transform of the amplitude is

$$F(\omega) = \sqrt{\frac{\gamma}{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-i\omega t}\mathrm{e}^{i\omega_0 t}\mathrm{e}^{-\gamma t}\,dt = \sqrt{\frac{\gamma}{\pi}} \frac{i}{\omega_0 - \omega + i\gamma}, \qquad (3)$$

which corresponds to an intensity in the frequency domain:

$$I(\omega) = |F(\omega)|^2 = \frac{\frac{\gamma}{\pi}}{(\omega_0 - \omega)^2 + \gamma^2} \qquad \text{with} \qquad \int |F(\omega)|^2 \, dt = 1. \tag{4}$$

Taking $\Gamma = 2\hbar\gamma$ and $E = \hbar\omega$, we get a **Lorentzian** intensity profile:

$$I(\omega) = I_0 \frac{\Gamma/(2\pi)}{(E_0 - E)^2 + (\Gamma/2)^2}. \tag{5}$$

Comparing this line shape to a **Gaussian** profile

$$I_G(E, E_0, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{E - E_0}{\sigma}\right)^2\right], \tag{6}$$

with FWHM $2.354\,\sigma = \Gamma$, shows a substantial difference.

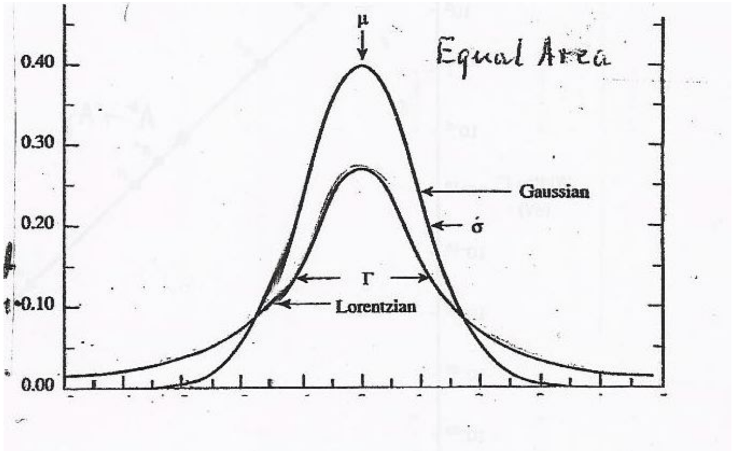

Figure 1: From Bevington, page 33. See also chapter 9.

What is the point? Well, the physically interesting quantity is the "natural" linewidth, Γ, whereas resolution effects like Doppler broadening and/or stochastic errors will make the line shape appear more Gaussian (recall the Central Limit Theorem). So, if we want to claim that we measured a natural line, we have to prove that the shape is right, and state the lifetime of the excited state.

Problem 2.1

Fit the data set 'lineshape1.txt' available from http://web.mit.edu/8.13/www/handouts.shtml to both a Lorentzian probability density function (PDF) and also a Gaussian PDF. Compare the fit results and χ^2 values to determine the correct fit hypothesis.
Hint: think about how to handle zero bins in the data set.

Problem 2.2

Fit the data set 'lineshapedata.txt' also available from http://web.mit.edu/8.13/www/handouts.shtml to both Lorentzian and Gaussian PDFs, this time with the possible addition of DC and linear background terms. What are the χ^2 values? Which fit hypothesis is justified?

For each problem above, carry out the following steps:

a. Produce a plot of the data with error bars, assuming Poisson statistics.
b. Make an educated guess for initial values for a fit to Gaussians plus background.
c. Perform the fit and retain all values. You may find the matlab routines at http://web.mit.edu/8.13/www/jlmatlab.shtml useful.
d. Plot the data along with the fit.
e. Make a subtraction plot (data – fit) to show the residuals. Comments on the result.
f. As a check, repeat with different starting values. Compare the results.
g. Optional: Give confidence limits for each hypothesis.

Problem 3: Fitting multiple peaks

In this exercise, you wish to fit data with more than one resonant feature, where the determining the resonant frequencies is important. You will find this useful for most spectroscopic data.

a. Download the data file 'dataex2.txt' from http://web.mit.edu/8.13/www/handouts.shtml.
b. Plot the data and fit it to your best guess for its functional form.
c. Fit the data. Again, you might find the matlab routines fittemplate09.m and gradsearch.m at http://web.mit.edu/8.13/www/jlmatlab.shtml useful. Make sure you understand the matlab scripts, and replace filenames as needed.
d. Plot the data and fit on the same graph.
e. Plot the residuals.
f. What is the separation of the frequency of the two peaks (in arbitrary frequency units)? Give your answer with uncertainties to 1-σ.