chooch
Determination of Anomalous Scattering Factors
from X-ray fluorescence data

Gwyndaf Evans

April 14, 2004

Introduction

The effects of anomalous scattering are described mathematically by two correction terms which are applied to the normal atomic form factor or Thompson scattering factor f_o. The modified scattering factor is given by $f = f_o + f' + if''$ where f' is the real part and f'' the imaginary part of the anomalous scattering correction term.

These anomalous scattering factors vary most rapidly near characteristic absorption edges of atoms where the energy of the incident X-rays is similar to the binding energy of the absorbing electrons. Thought of classically anomalous scattering is analogous to any resonance effect such as an electrical LC circuit.

The optical theorem [4] says that the imaginary term f'' is directly related to the atomic absorption coefficient for an atom by

$$f'' = mc\epsilon_o E \mu_a / e^2 \hbar \tag{1}$$

where μ_a is the atomic absorption coefficient, E the X-ray energy and all other symbols take there usual meaning. As in other resonance phenomena such as dielectric susceptibility, the real part of the dispersive term is related to the imaginary part by a Kramers-Kronig (K-K) transformation. In the case of X-ray scattering the K-K transform takes the following form

$$f'(Eo) = \frac{2}{\pi} \oint_o^\infty \frac{(Ef''(E))}{(E_o^2 - E^2)} dE \tag{2}$$

Purpose

Why do we need to know f'' and f'? When performing Multiple or Single wavelength Anomalous Diffraction (MAD/SAD) experiments a crucial prerequisite is knowing at which wavelength(s) to measure diffraction data. This can only be determined at the time of the experiment for two main reasons

1. For a particular heavy atom element the X-ray energies to be measured are largely dependent on the environment of that element within the protein sample and its orientation with respect to the polarization vector of the incident X-ray beam.

2. The calibration of the incident X-ray energy at different X-ray beam lines will rarely be the same and as yet no adequate calibration standards have been established which are common to all crystallographic facilities.

In addition the calibration of each beam line may vary over time. As previously stated the f'' value is directly related to the atomic absorption coefficient for an atom. For a discussion of the difficulties and solutions associated with X-ray energy calibration for MAD see [2].

The X-ray fluorescence from the atom is a result of, and is directly proportional to, the absorption of the incident X-rays. This provides the experimenter with a means of determining the dependence of f''

on the X-ray energy since f'' is in turn related to the absorption cross-section. f' may then be determined computationally using the Kramers-Kronig relationship. This provides the necessary information with which to make a rational choice of which wavelengths to measure for the experiment. Clearly we also establish the magnitudes of the anomalous scattering factors as a function of X-ray energy. These values are potentially useful as starting points for heavy atom refinement during the latter stages of data analysis and phasing.

Determination of f'' and f'

Obtaining f'' from fluorescence data

Fluorescence spectra are generally measured directly from the same frozen protein crystal sample from which the diffraction data is to be measured. The spectra are typically recorded using a photo-multiplier (e.g. Bicron tube) or an energy resolving photo-diode type detector (e.g. Amptek). In both cases the fluorescence signal is recorded on an arbitrary scale. Determination of the corresponding f'' spectra is done via two stages.

Firstly the raw fluorescence spectrum must be background subtracted and corrected to subtract out any additional scattering effects which may be energy dependent. This procedure is typically very straight forward for data measured using a good energy resolving detector such as the Amptek since the measured signal is essentially dominated by fluorescent X-ray counts. However photo-multiplier tubes which have poorer energy discrimination will typically measure some elastic scattering component of the X-rays as well as the fluorescence signal and will therefore usually require a more careful background subtraction.

The procedure involves applying a low order polynomial or linear fit separately to the below edge region of the spectrum and the above edge region of the spectrum. The fits should be generated away from the absorption edge where the XANES effects are smallest. These fits are then applied to the raw spectrum such that it is normalized to be zero far below the absorption edge and unity above the edge. The normalized signal $N(E)$ is obtained by

$$\mathcal{N}(E) = \mathcal{R}(E) \left\{ f''_{above}(E) - f''_{below}(E) \right\} + f''_{below}(E) \tag{3}$$

where $\mathcal{R}(E)$ is the raw data, f''_{below} is the polynomial fit in the below edge region and f''_{above} the fit for the above edge region. All are functions of the X-ray energy E. Theoretical values of f'' have been calculated by Cromer & Libermann [1]. The calculations however take no account of the effects of coordination of anomalous scattering atoms to other atoms. The effects of coordination are most visible in the near edge region which also happens to be a region of interest for MAD and SAD experiments. Therefore the Cromer & Libermann tables are not applicable in the near edge region. However, away from the absorption edge, above and below in energy, the tables provide a good estimate of the true values of f''. This provides a way to convert the normalized fluorescence data into a f'' spectrum. The theoretical spectrum is essentially multiplied into the experimentally determined spectrum to produce an experimentally determined f'' spectrum.

Obtaining f' from f''

Given a f'' spectrum the K-K transformation may be used to directly obtain a f' spectrum. An algorithm has been described [3] which allows this to be carried out computationally. Complications arise in the calculation because of the singularity in the integrand of Equation 2 arising when E is equal to Eo and also because of the impractical limits of integration. The singularity is dealt with conveniently by the above algorithm and the integration limits are chosen such that the calculation remains possible but does not become inaccurate. Integration limits which extend only a few keV above and below the absorption edge will usually provide an accurate estimate of the X-ray energy corresponding to the minimum value of f' but the magnitude of the f' curve will in general be incorrect. To obtain highly accurate magnitudes integration limits are chosen which extend up to $50\times$ absorption edge energy and to very low energies of say 1 keV.

Organization of the program

chooch is run from the command line with various control options being specified through the use of command line switches.

The various steps performed by **chooch** are outlined below.

1. data input and checking

   - the fluorescence data is read from a file (see below for a description of the format) and basic sanity checks are performed on the data. The program attempts to guess which edge has been measured for a specified element by assuming that the middle of the scanned energy range is near the absorption edge of interest.

2. normalization of input spectrum

   - Normalization is performed as described above. A linear model is used to perform the fitting.

3. determination of f''

   - Theoretical values of f'' are obtained using the `mucal.c` (http://ixs.csrri.iit.edu/database/programs/mcmaster.html) routine written by Pathikrit Bandyopadhyay which uses the absorption cross-section values as published by McMasters [5]. .

4. smoothing and calculation of derivatives.

   - Smoothing is done with a Savitsky-Golay filter with a window width which is determined from the monochromator energy resolution. The resolution may be supplied by the user with the '`-r <resol>`' option

5. Kramers-Kronig transformation to obtain f'

   - The program uses numerical integration routines supplied with the Gnu Scientific Library (http://www.gnu.org/software/gsl/) to perform the K-K transformation.

6. analsysis and output of results

   - The program automatically selects the peak f'' energy and the minimum f' energy and outputs them. A PostScript plot of the f' and f'' spectrum is generated if requested by the user with the '`-p <psfile>`' option.

Input files

`<name>.raw` The raw input fluorescence data file. The first line should contain a title (upto 80 characters) which will be used for the output data and PostScript plot. The second line contains the number of data points (integer). Each subsequent line should contain two values referring to one data point - the X-ray energy **in electron-Volts (eV) (NOT keV)** and the measured fluorescence signal on an arbitrary scale.

e.g.

```
Fluor. spectrum for element Qu  ; Title (a80)
101                             ; No. data points (free format)
12300.0  2002                   ; Energy (eV), Flu. Signal (free format)
12300.5  2030
etc
.
12700.0  6700
```

Output files

`<name>.efs` [default = output.efs] Output ascii file containing calculated anomalous scattering factors.

`<name>.ps` (optional) [no default] PostScript output with a plot of anomalous scattering factors.

Installation

Requirements

Ensure that you have all the required libraries installed:

1. Gnu Scientific Library version 1.1 or later available from `http://www.gnu.org/software/gsl/gsl.html`

2. Cgraph version 2.04 PostScript plotting library (`http://neurovision.berkeley.edu/software/A_Cgraph.html`)

3. (optionally) PGPLOT graphics library (`http://www.astro.caltech.edu/ tjp/pgplot`)

Building the binary

1. Download source distribution chooch.tar.gz

2. Unpack the file in a directory of your choice e.g. usrlocalsrc using tar xzvf chooch.tar.gz for example.

3. Edit the top of the Makefile to select the correct machine architecture, point to the correct libraries and choose installation directories for binaries.

4. Type 'gmake' for a version without PGPLOT capability or 'gmake chooch-pg' for a version with PGPLOT linked. N.B. Both versions allow the generation of PostScipt output.

5. Type 'gmake install'

Running the program

Chooch has a number of options which can be controlled through command line switches. All the options may be display by typing '`chooch -h`'. An example of simple use of chooch to process a data file spec.raw measured around the K edge of selenium would be

```
chooch -e Se -a K -o spec.efs -p spec.ps spec.raw
```

This would generate a PostScript file spec.ps with the resulting anomalous scattering factors and an ascii file spec.efs of the same.

Citing the program

Use of **chooch** should be cited as G. Evans and R. F. Pettifer *J. Appl. Cryst.* **34**, 82—86, (2001).

References

[1] D. T. Cromer and D. Liberman. Relativistic calculation of anomalous scattering factors for X-rays. *J. Chem. Phys.*, 53:1891–1898, 1970.

[2] G. Evans and R. F. Pettifer. Stabilisation and calibration of x-ray wavelengths for anomalous diffraction experiments using synchrotron radiation. *Rev. Sci. Instr.*, 67(10):3428–3433, October 1996.

[3] J. J. Hoyt, D. de Fontaine, and W. K. Warburton. Determination of the anomalous scattering factors for Cu, Ni and Ti using the dispersion relation. *J. Appl. Cryst.*, 17:344–351, 1984.

[4] R. W. James. *The Optical Principles of the Diffraction of X-rays.* G. Bell and sons Ltd, London, 1969.

[5] W. H. McMasters, D. N. K. Grande, J. H. Mallet, and J. H. Hubbell. Compilation of X-ray cross sections. Technical Report UCRL-50174, Lawrence Radiation Laboratory (Livermore), 1969.