

Time-of-flight data reduction within the Bastille project

Björn Fåk
Institut Laue-Langevin

24-Nov-2015

The basic aim of data reduction for time-of-flight spectrometers such as IN4, IN5, and IN6 is to transform measured intensities as function of angles and time of flight into the dynamic structure factor as function of wave vector and energy, and related physical quantities. The sign convention used in this document for energy and wave vector transfer is

$$E = E_i - E_f \quad (1)$$

$$\mathbf{Q} = \mathbf{k}_i - \mathbf{k}_f, \quad (2)$$

respectively, where E_i (E_f) is the incoming (outgoing) neutron energy and \mathbf{k}_i (\mathbf{k}_f) the incoming (outgoing) neutron wave vector. Let $I(\Omega_{ij}, t)$ be the number of counts recorded in a time channel t and a detector pixel Ω_{ij} . A detector pixel Ω_{ij} is defined by two angles ϕ_i and θ_j , where ϕ_i and θ_j could be scattering angle and azimuthal angle, scattering angle projected on the horizontal scattering plane and out-of-the-plane angle, or some analogous definition. The extension of a detector pixel is $\Delta\phi_i$ and $\Delta\theta_j$ in angular space, and the distance between the sample and a pixel is L_{ij} . In this document, it is assumed that data are collected in histogram mode, but the extension to event mode is straightforward. The basic aim of the data reduction is thus to transform $I(\Omega, t)$ to $S(\phi, \theta, E)$, and to save this in a suitable data format, in order to allow more specialized programs to perform the last step to $S(\mathbf{Q}, E)$ and similar physical quantities. However, for diagnostics purposes and also to get an on-line handle on the actual results of the measurements, transformation within the standard data reduction of $S(\phi, \theta, E)$ to $S(Q, E)$ is required for powder samples and to $S(\mathbf{Q}, E)$ for single crystals. The latter requires the UB sample orientation matrix, and data from many sample orientations need in general to be combined. Derived quantities need also to be calculated for diagnostics and preliminary analysis, such as constant-energy and constant- \mathbf{Q} cuts, the dynamic susceptibility $\chi''(\mathbf{Q}, E)$, phonon density of states, etc.

Three basic principles guide the data reduction described in this note:

- (1) Never remove any pixels. Pixels that should not be retained for treatment should instead be listed in a mask file. Such a mask file could already be provided by the instrument, and updated during the treatment.
- (2) Introduce as little correlations as possible between pixels as late as possible in the data treatment. This doctrine may pragmatically be modified for reasons of poor statistics.
- (3) The different routines that constitute the data reduction should be independent, well documented, and do *just* what they are supposed to do.

Various widely used data reduction program employs the “flow diagram” given below to some extent and with some (not always well-defined) order of points. Examples of commonly used programs are `crossx`, `inx`, `lamp`, `[fr]ida`, `homer`, `mantid`, and `dave`, but there exists also a large variety of personal development or programs for restricted use developed by e.g. the Delft group, Karlsruhe, Jens Suck, Niels van Dijk, Körfer, Stunault, Fåk, etc. The algorithms used by these programs, or included in the “flow diagram” below, are in general scarcely, if at all, documented.

Efforts are also required in order to assure complete scripting of the data reduction, in order to reduce the influence of human errors on the result. Scripting is best done using simple ascii files that can be edited using any text editor, but graphical interfaces (that produces such ascii scripts) are probably also needed. Of equal importance is that the data reduction produces a log file that tells what was done and which parameters were used, to assure a complete record of the data reduction.

A basic outline of a typical data reduction is:

1. Gather run numbers for different Conditions: Sample (F), Empty (E), Vanadium (V), Empty-Vanadium (G), and possibly Cadmium (Cd). The latter is not treated systematically in this treatise.
2. Read and add data files for each Condition. Information in headers should be added or averaged, depending on the type of parameter. Counting times should clearly be added, while e.g. temperatures should be averaged. It is useful to calculate max and min values as well as standard deviations for quantities that should be averaged, such as the temperature. Ideally, the temperature (and similar sample environment variables) should be recorded periodically during a run, or at least at the beginning and at the end.
3. Check consistency for added files for each Condition. The aim is two-fold:
 - a. To verify that different runs are consistent. It is straightforward to see if part of the sample disappears or transforms or moves in the beam or if the temperature has changed due to out-of-equilibrium conditions, etc.
 - b. To identify inconsistent detector pixels or other instrumental problems, such as a radial collimator that stops oscillating during the experiment.

The consistency method is extremely powerful and is described in detail in Refs. [1, 2]. The information obtained can be used to remove inconsistent runs, prevent that runs taken under different conditions are inadvertently added, and to update the detector mask.

4. Extract integrated intensity and position of the peak in the monitor spectrum. The former is used to normalize the intensities for each Condition. The actual normalization of the data can be done here or, perhaps advantageously, postponed until later (cf. point 11.) The peak position is used on some instruments to calibrate the incoming energy. Various levels of sophistication exist to extract these quantities (moment methods, fits, or σ/I).
5. Extract integrated intensity and position of the elastic peak (if present) from the sample for each pixel i, j . For liquid samples or samples with small elastic scattering, it is better to use a vanadium run done under the same conditions as the sample. In unambiguous cases, the sample position x, y can be extracted by a fit to the peak position of all pixels. In such cases, the sample-detector distance for each pixel can be determined. By comparing (normalized) intensities with some specified criteria, the detector mask could be updated.
6. Determine the “flat” background away from the elastic peak for each pixel. This background is typically a purely instrumental (sample-independent) background, originating from electronic noise, cosmic radiation, and other ambient sources such as neighboring instruments or neutron guides. One method to determine this “flat” background is to take the minimal value of running averages (that can be parameterized) over the time frame. Detectors with high “flat” background could be “added” to the detector mask. The background can also be subtracted from each Condition, before standard background subtraction is done. Since the background determined in this way may be affected by thermal population, it is best determined at low temperatures, and stored for future use. A typical usage is to subtract the low- T Sample background from the Sample runs at all temperatures. The meaning of “Flat” could be extended to a sloping line, a simple polynomial, or some other simple shape characteristic of the instrument used.
7. Correct for frame overlap (if necessary). This is a difficult task in the general case, as frame overlap depends on the physics of the sample.
8. Determine the energy transfer for each pixel and time channel. [This is not $t2e$, it is just to label each time channel with an energy value. The Jacobian transformation of $t2e$ is done under point 15]. It is important to keep the time channel information along with the data reduction. There are various algorithms and approximations that can be used here. If the true/measured sample-pixel distance has been determined as outlined above, a more sophisticated algorithm can be used (see e.g. Ref. [3]).

9. Calculate statistical counting errors σ (one standard deviation) based on Poisson statistics. Computing time can be saved if variances σ^2 are used throughout the treatment rather than σ .
10. Absorption or self-shielding correction of data (and vanadium, and possibly empty can) as a function of pixel and time (or energy transfer). This can be done before or after the basic data reduction described in point 12, but it might be best before (i.e., here). Correction for standard sample geometries such as slabs (both transmission and reflection) and cylinders (including annular geometry) are required. Since some of these calculations are non-analytical and hence time consuming, interpolation procedures may be necessary. Computing time may be substantially reduced if the corrections can be saved for subsequent data sets. The transmission factors e (and g) used in point 12 could also be calculated here. Extensions of the absorption correction to irregular sample geometries as often encountered in single crystal experiments would be an interesting enhancement. Maybe some routines from the diffraction group could be adapted for inelastic scattering.
11. Normalize data to the beam monitor (this can also be done under point 4). In some cases, typically for diagnostics or instrument performance evaluation, normalization to measurement time (or something else) might be better.
12. Perform the basic data reduction, i.e., $(F - eE)/(V - gG)$. Here, e (and g) are pre-factors that grossly account for the sample transmission (and which could be calculated under point 10). These constants are here assumed to be independent of the angle (pixel). This equation needs to be modified when a Cd run is used.
13. Correct for the detector efficiency, which is a function of the final energy. It is customary to use a correction factor that is normalized to unity at elastic energy transfer, $\eta(E)/\eta(E = 0)$. It is here assumed that the width of the elastic peak is negligible for the vanadium run (which might not be true on D7?).
14. Absolute normalization of data. This requires a clean vanadium run under essentially identical conditions to the sample runs, in particular as to the size and shape in order to compensate for beam inhomogeneities, which can be particularly severe for instruments with strongly focusing beam optics such as IN4 or IN6. Note the importance of using a clean vanadium which has recently been annealed at high temperature in high vacuum in order to remove interstitial hydrogen that easily goes into the bcc structure of vanadium. In addition to knowing the mass of the vanadium (or number density and thickness) that is illuminated by the beam, the number density and thickness of the sample are also required (they are already needed for point 10).
15. Convert the intensities from time of flight $I(\Omega_{ij}, t)$ to energy transfer $I(\Omega_{ij}, E)$. The data are essentially multiplied by a factor t^3 to obtain cross-sections $d^2\sigma/d\Omega dE$ (standard in mantid?) and by t^4 to obtain $S(\Omega_{ij}, E)$ (convention at the ILL). Note that in the general case, different detector pixels will not have a common energy array \mathbf{E} , but each pixel will have its own array. Interpolations are required to obtain a common energy array, and should be deferred to as late as possible, if at all.
16. Optionally, the data $S(\Omega_{ij}, E)$ could be rebinned to equidistant (or smart [user-provided] non-equidistant energy bins E) so that all spectra have a common energy array. Since interpolation correlates neighboring energy bins, it should be deferred to as late as possible in the treatment. Care need to be taken as binning conventions are different for ISIS histogram data and ILL data.
17. On some instruments, detectors with the same ϕ but different θ can be grouped together for powder samples. This is traditionally done by some kind of `sumbank` command, either here or even earlier in the treatment. It is advantageous to do this as late as possible in the treatment, if the statistics allows so, and perhaps even better done in point 18. This operation will in general not correlate neighboring pixels. However, on instruments where the detectors are overlapping and unevenly spaced, such as IN4, an interpolation scheme could be considered.

18. For powder samples, optionally transform $S(\Omega_{ij}, E)$ to $S(Q, E)$. This will correlate neighboring Q values, and depending on the algorithm used, also neighboring energy bins. For coarse-grained data, i.e., large pixel size (IN6) or irregular pixel distribution (IN4), the quality of the results depend quite strongly on the algorithm used (its sophistication and amount of cross-correlation), while for fine-grained data (such as from IN5) the algorithm is less critical. Data with poor statistics is also very sensitive to the algorithm. An alternative method is successfully employed in `mslice` for visualization purposes, where $S(\Omega_{ij}, E)$ is not transformed to $S(Q, E)$ but instead is plotted on a regular Q – E grid along curved detector trajectories.
19. For single crystals, transform $S(\Omega_{ij}, E)$ to $S(\mathbf{Q}, E)$. This requires knowledge of the sample orientation in the beam (described by the 3×3 UB matrix). The calculations require in general a huge amount of computer memory, and may be too specialized for a general time-of-flight data reduction program.
20. Optionally transform $S(Q, E)$ to the imaginary part of the dynamic susceptibility $\chi''(Q, E)$ or to the symmetrized (classical) $\tilde{S}(Q, E)$, or simply check the detailed balance condition. Phonon-density-of-states calculations and a few other special transformations [such as $S(Q, E) \cdot E/Q$] are also of interest.
21. Cuts are of great physical and diagnostics importance. For diagnostics, such tools need to be available for checking various steps of the data reduction. For the physics, most useful are cuts at constant Q (or \mathbf{Q}), sometimes called energy scans, i.e., $S_Q(E)$ for some Q range, and cuts at constant energy E , sometimes called Q scans, $S_E(Q)$. The calculation of such quantities requires some care in the normalization of the cuts, and both `lamp` and `mslice` have been (or are still) incorrect in some cases.
22. Multiple scattering corrections are outside the scope of this treatise. The `mecat` and `discuss` fortran codes released in 1974 are becoming obsolete, `McStas` appears to have severe teething problems, and other codes do not seem to be in much general use.
23. Resolution calculations are also outside the scope of this treatise. They belong to more specific data analyses.
24. Smoothing should be avoided as much as possible, as it severely correlates neighboring data in an often uncontrolled way. Binning is in most cases a better alternative.
25. Visualization of both 2D and 1D data using standard graphical interfaces (e.g. `matlab`) is an absolute requisite. Graphics should also be easily printed and saved in files that are as small as possible. It is worth the effort to push the fine-tuning of the visualization in the direction to produce figures that are essentially ready for publication.

References

- [1] A. Stunault, K.H. Andersen, Y. Blanc, B. Fåk, H. Godfrin, K. Guckelsberger, and R. Scherm, *Time-of-flight spectroscopy: Energy calibration and consistency check*, Physica B 180–181 (1992) 926–928.
- [2] H. R. Glyde, B. Fåk, N. H. van Dijk, H. Godfrin, K. Guckelsberger, and R. Scherm, *Effective mass, spin fluctuations, and zero sound in liquid ^3He* , Phys. Rev. B 61 (2000) 1421–1432.
- [3] A. Stunault, private communication.