Why direct and post-refinement determinations of absolute structure may give different results

David John Watkin and Richard Ian Cooper

Overview of the equations used in CRYSTALS, based on material extracted from the original text of Watkin & Cooper, Acta Cryst. (2016). B72, 661–683.

Introduction

The introduction by Rogers (1981) of a new parameter, η, as a refineable multiplier onto f" in the least squares optimisation of a crystal structure (equation 1) was the first attempt to directly determine absolute structures as part of the refinement process (hereafter called *direct* determination).

$$F_h = \sum_j (f_j^0 + f_j' + i\eta f_j'') \times (\cos(2\pi hx_j) + i\sin(2\pi hx_j)) \qquad (1)$$

Flack (1983) recognised that the η parameter had no physical significance except for values of ± 1, and introduced a new formulation of the problem. He proposed that a given sample be regarded as a twin by inversion, and that refining the twin faction would reveal the absolute structure. Representing $|F|^2_h$ by I^+ and $|F|^2_{\bar{h}}$ by I^-

$$I_o^+ \approx I_c^+ = (1-x)I_s^+ + xI_s^- \qquad (2)$$

where the subscripts s indicates a quantity computed from the atomic model with the Flack parameter x set to zero (*i.e.* a non-twinned single crystal), c a quantity computed from a twinned model (*i.e.* Flack parameter not necessarily zero) and o an observed quantity. Like the Rogers method, this proposal refined the absolute structure parameter using all the reflection data as part of the normal structure optimisation, but had the advantage that the parameter had a real physical significance throughout the whole range from zero to one. This innovation increased awareness of the existence of twinning by inversion and fears that samples may not have been enantio-pure. For convenience we will use the term *Flack parameter* to imply x determined by an unspecified method, and Flack(x) to imply its determination as part of the main structure refinement.

The 1993 release of SHELXL included a post-refinement method for determining the Flack parameter, by a method which came to be known as "Hole in One". Equation (2) can be re-arranged to give the Flack parameter directly from observed structure factors and structure factors computed from the atomic model and its inverse (Sheldrick, 2014).

$$I_o^+ - I_s^+ \approx x(I_s^- - I_s^+) \qquad (3)$$

In spite of fears that post-refinement determinations of absolute structure might be compromised because of neglect of potential covariance with the other refineable parameters, Hooft *et al.* (2008) devised a method based on a Bayesian analysis of Friedel differences (see Müller, 1988 for an interpretation of Friedel pair). These authors recast equation (3) to treat Friedel pairs of reflections simultaneously.

$$D_o \approx D_c = (1-2x)D_s \qquad (4)$$

where $D_s = (I_s^+ - I_s^-)$ and similar for D_o and D_c [footnote: this equation first appears in this form in Thompson & Watkin (2011)]. For convenience later, we have called x computed from equation 4 the Bijvoet(d) parameter.

Their process, which used weights derived from the variances of the observed intensities modified by information obtained from the normal probability plot (npp) of the Friedel residuals (Abrahams & Keve, 1971), yielded values of the absolute structure parameter, Hooft(y), not unlike those from the Flack(x) method. The method further enabled one to estimate the probabilities of the correctness of absolute structure assignments for enantio-pure or 50:50 racemically twinned samples.

Parsons *et al.* (2013) examined the use of equation (4) and its quotient form, equation (5), which we will call Parsons(q), both for post-refinement determination of the Flack parameter, and as restraints during the direct refinement of Flack(x).

$$Q_o \approx Q_c = (1 - 2x)Q_s \qquad\qquad (5)$$

where $Q_s = (I_s^+ - I_s^-)/2A_s$ etc. and $A_s = (I_s^+ + I_s^-)/2$

These authors showed that the Hooft(y), Parsons(q) and Bijvoet(d) estimates of the absolute structure parameter were usually similar to direct refinement of the Flack(x), but with significantly lower standard uncertainties. They also observed that using equations (4) or (5) as restraints on the least squares refinement gave values of the Flack(x) in close agreement with post-refinement estimates of absolute structure.

Direct refinement of the Flack(x) parameter

Direct refinement of the Flack(x) parameter simultaneously with the other structural parameters is now commonplace. Flack *et al.* (2006) recommend that a full set of Friedel pairs be measured on an area detector instrument, preferably with high redundancy in order to optimise empirical intensity scaling, and that refinement be started with the Flack parameter set to 0.5 to minimise the risk of refinement to a false minimum. This is particularly important in the case of space groups with floating origins, in which the structure may distort to accommodate an incorrectly assigned absolute structure - the polar dispersion error (Cruickshank and McDonald, 1967).

Post-refinement determination of the Flack parameter

The relation between the absolute structure of a crystalline material and the measured Friedel pairs is given in equation (2). The worryingly high standard uncertainty of the Flack(x) parameter determined for many materials of known enantiopurity and absolute structure has led to a search for methods for determining Flack parameter more robustly than simply including it in the main least squares refinement, especially in cases where the resonant signal is likely to be weak. Not infrequently, these methods involve the use of selected sub-sets of the original or new data.

Given a reasonably well refined model, the absolute structure can be estimated by solving equations (2), (3), (4) or (5) for x by conventional least squares. The disagreement sometimes seen between the Hole-in-One method (equation (3)) and Bijvoet difference method (equation (4)) might, in part, be due to the additional information introduced by pairing up reflections for the differences, with the possibility that certain kinds of errors in the model or in the data might be correlated and tend to cancel out.

The denominators in the Parsons (q) expression (5) were based (Parsons *et al.*, 2013) on an extension of the earlier recognition that on a serial four circle diffractometer setting angles could be chosen so that the absorption effect for reflections $\boldsymbol{h}$ and $\bar{\bar{\boldsymbol{h}}}$ would be similar (Le Page *et al.*, 1990). On an area detector diffractometer these conditions are rarely satisfied, and in any case the final intensity of each reflection is usually the average of several measurements made with quite different setting angles.

Equation (5) can be rewritten as:

$$D_o/2A_o \approx D_c/2A_c = (1 - 2x)D_s/2A_s \qquad\qquad (6)$$

Here the A_o and A_s seem to be scale factors down weighting the contribution of strong reflections to the absolute structure parameter. However, when each reflection pair is weighted by the inverse of the variance of the observed quantities, this down weighting disappears.

If equation (6) is rewritten as:

$$D_o \approx (1 - 2x)D_s(A_o/A_s) \qquad\qquad (7)$$

we can see that if A_o can be regarded as $A_s \pm error$, the ratio $A_o : A_s$ could take large values when the calculated structure amplitude is very small – such reflections must be excluded from any quotient calculation. In fact, if A_o is not very similar to A_s then there is a reasonable probability that there is something wrong with the model, the data or both. We can also see that the A_o/A_s terms act as per-reflection scale factors and should be counted as independent variables.

Just as plots of F_o versus F_c can be of diagnostic value in a normal structure refinement, so plots of D_o versus D_s and $2A_o$ versus $2A_s$ can give insight into absolute structure determination (Parsons *et al.* 2012a). The $2A_o$-$2A_s$ plot should have unit gradient and might identify outliers in which the quotient in equation (7) lies far from unity. For enantio-pure materials correctly assigned, the D_o versus D_s scatter plot should also have unit gradient, and for materials with a large Friedif (Flack & Bernardinelli, 2011) this is usually clearly evident. For materials with a Friedif less than 100 the linear relationship is always less clear (Cooper, Watkin & Flack, 2016).

The importance of a given datum on its own fitted value is measured by its leverage (Prince, 2004). Since the mean values of D_o and D_s (and the corresponding quotients) are close to zero, fitting a straight line can be regarded as a one-parameter model, so that the leverage of each data point is given by:

$$P_{ii} = \frac{w_i d_i^2}{\sum w_i d_i^2} \qquad (8)$$

where d_i are the values of either D_s or Q_s. The data with greatest leverage are those with large absolute values of D_s or Q_s. Remember that although D_s does not depend directly on A_s, large D_s can only be possible for large A_s. If each observation in the post-refinement determination of the Flack parameter is weighted by the inverse of its variance, P_{ii} is proportional to the square of the signal:noise . To a first approximation, $\sigma^2(I) \propto I$ (Evans, 2006), (but see also section 5.2.6) so that the resonant difference originating from strong reflections will have large standard uncertainties, and be down weighted. The most useful reflections are likely to be those of intermediate intensity and with large resonant difference. This is in agreement with the leverage analysis for the Flack(x) parameter in the least squares refinement of all structural parameters (Parsons *et al.*, 2012b).

Equation (4) can be made to yield values of the Flack parameter on a per Friedel pair basis:

$$x = (D_s - D_o)/2D_s \qquad (9)$$

Plotting x from equation (9) against D_s (Fig 2) should give a horizontal line at the value of the Flack parameter. If $|D_s|$ is very small compared to $|D_o|$, the value of x can take extreme values. For a structure with low resonant scattering, individual x can be ill-determined, and even for good data many extreme values can be seen. The massive vertical distribution near the centre of the plot (which includes both positive and negative estimates of x) corresponds to small values of the denominator in (4) and (9), and it is only the data lying distant from $|D_s| = 0$ which contain useful information.

Iterative reweighting

The Le Page algorithm (Le Page *et al.*, 1990) in effect assigns a value of ±1 to the Rogers' η value of the selected reflections on a one-by-one-basis as opposed to direct refinement of η from all the reflections in the main least squares calculations. It tacitly assumes that the material is enantiopure. Equation 9 enables us to also evaluate the Flack parameter on a reflection-by-reflection basis - the data used in creating Fig 2. We could in principal evaluate the Flack parameter from each pair of carefully selected and re-measured reflections – or even from just one very carefully selected and **very *very*** carefully measured pair. Because x is a continuously meaningful parameter in the range 0 to 1, it is not necessary to assign it an integer value. Now, rather than re-measuring selected reflection pairs to estimate x, we can use all the pairs measured in the original data collection to give individual estimates of x. With the exception of unknown correlations introduced during the measurement process, these estimates of x will be experimentally independent (or at least as independent as the measurements of the original data were). The values of x can take values wildly outside of the 0 to 1 range - these are physically impossible and correspond to outliers originating either from large experimental errors, or are artefacts of a small denominator in equation 9. Following the arguments of Blessing & Langs (1987) for the merging of equivalent reflections, we can merge these individual x-values, and since each x-value has an associated experimental variance, we can compute both the external variance

$$x' = \frac{\sum w_i x_i}{\sum w_i}$$

$$(10)$$

$$\sigma_{ext}^2 = \frac{\sum w_i \sigma^2(x_i)}{\sum w_i}$$

(11)

and the internal variance (Appendix A)

$$\sigma_{int}^2 = \frac{n \sum w_i (x_i - x')^2}{(n-1) \sum w_i}$$

(12)

The probability of an individual x_i can be estimated from

$$p_i = \exp\left[\frac{-(x_i - x')^2}{2\sigma_{ext}^2}\right]$$

(13)

Friedel pairs yielding a value of x differing from the average value of x by several variances have a low probability. This probability can be used as a modifier for the weight ($w'_i = w_i p_i$) used to compute a new weighted average value of x, and the process repeated (Blessing & Langs, 1987). Since the distribution of the computed Flack parameters may be dispersed, skewed or long-tailed, the process is started using the median value of the x_i as an initial estimate of x. Thus, rather slack values can be set for the various initial filter thresholds used in selecting reflections, and a smoothly varying function be used to down-weight suspect data. Friedel pairs with a probability p_i greater than a user-adjustable threshold (typically 0.001) are counted to provide an indication of the number of 'useful' reflections in the data. The process is terminated when the number of 'useful' reflections is the same for two successive iterations, or until ten iterations are completed. In this latter case, the process is regarded as being unconverged and unsuccessful. This situation seems to arise when the resonant scattering is small compared to the errors in the intensity measurements. The standard uncertainty on the final value of x' is estimated from the weighted external variance:

$$\sigma_{mean}^2 = \frac{\sum w'_i \sigma^2(x_i)}{n \sum w'_i}$$

(14)

Ratios of Averages and Averages of Ratios
Letting ($1-2x$) in equation (4) be represented by c, then for each Friedel pair we have

$$c_i = Do_i / Ds_i$$

(15)

An average value of c_i can be computed as a least squares estimate (see appendix A)

$$\langle c \rangle = \frac{\sum Do_i Ds_i}{\sum Ds_i Ds_i}$$

(16)

or as a simple mean

$$c' = \frac{\sum c_i}{n}$$

(17)

leading to $<x>$ and x'. Equation 16 is a ratio of averages (there is a *1/n* term in both the numerator and the denominator), equation 17 is the average of the individual ratios, c_i. In general, if all the summations are made over the same number of data points and there are no wildly eccentric outliers, the values of $<x>$ and x' are similar. An indication of the presence of outliers can be obtained by computing these coefficients using all the measured Friedel pairs. If they are substantially different, the distribution of the errors in *Do* may be skewed, there may be outliers, the errors may swamp any signal or there may be contributors to (15) where the $D_s i$ are tiny. Weighted versions of equations 16 and 17 can be recomputed during the Blessing and Langs (1987) process, where outliers are progressively down-weighted. If convergence is achieved before the maximum number of cycles is reached, $<x>$ and x' are usually very similar. Both values are output by CRYSTALS.

Influence of Weighting Schemes
In the discussion so far it has been assumed that the weights for the post-refinement analyses have been derived from the observed variances of the original diffraction data via equations 40, 41 and 42. However, it has long been established practice to use more complex weighting schemes in the main structure refinement. These weights are computed from

empirical formulae with coefficients selected to give a flat distribution of weighted residuals. This process is intended to allow for unidentified errors in the data and shortcomings in the model (Cruickshank, 1961). Weights computed in this way have an influence on the Flack(x) parameter and its su as determined during the main refinement (Bernardinelli & Flack, 1987). In order to see the influence of these weights on the post-refinement determination of absolute structure, they can be converted to observational pseudo-variances by

$$\sigma^2(I_{lsq}) = 1/\text{weight}_{lsq} \qquad (18)$$

where weight $_{lsq}$ is the weight assigned to the reflection during refinement.

Ratios of Averages and Averages of Ratios.

For an individual Freidel pair we can write equation 4 as

$$x = 0.5 - Do/(2Ds) \qquad (19)$$

Defining

$$c_i = Do_i / Ds_i \qquad (20)$$

we get

$$x_i = 0.5 - 0.5c_i \qquad (21)$$

Ratio of Averages:

Writing (20) as

$$[Ds]\langle c\rangle = [Do] \qquad (22)$$

where the terms in square brackets are column vectors of the model and observed Friedel differences, the least squares estimate of c from a set of Friedel pairs is

$$[Ds]^T[Ds]\langle c\rangle = [Ds]^T[Do] \qquad (23)$$

from which a weighted value for <c> can be obtained as

$$\langle c\rangle = \frac{\sum w_i Do_i Ds_i}{\sum w_i Ds_i Ds_i} \qquad (24)$$

with

$$w_i = 1./\sigma^2 Do_i \qquad (25)$$

Equation (24) can be rewritten as a ratio of averages

$$\langle c\rangle = \frac{\dfrac{1}{n}\sum w_i Do_i Ds_i}{\dfrac{1}{n}\sum w_i Ds_i Ds_i} \qquad (26)$$

Letting

$$\Delta_i = Do_i - \langle c\rangle Ds_i \qquad (27)$$

gives

$$\sigma^2(c) = \frac{\sum w_i \Delta_i^2}{n\sum w_i Ds^2} \qquad (28)$$

from which

$$\langle x\rangle = (1 - \langle c\rangle)/2 \qquad (29)$$

and

$$\sigma(\langle x\rangle) = \sigma(\langle c\rangle)/2 \qquad (30)$$

Average of Ratios.

Alternatively, we can evaluate individual c_i from equation (20) and x_i from (21), and form the (weighted) average of these ratios.

$$x_i = (Ds_i - Do_i)/2Ds_i \qquad (31)$$

$$x' = \sum w_i x_i / \sum w_i \qquad (32)$$

$$\sqrt{w_i} = 1/\sigma(x_i) = 2Ds_i/\sigma(Do_i) \qquad (33)$$

Following Blessing and Langs (1987) we can form the internal and external estimates of the variance of the sample, and hence the variance of the average.

$$\sigma^2_{int}(x) = \frac{n}{n-1} \frac{\sum w_i (x_i - x')^2}{\sum w_i} \qquad (34)$$

$$\sigma^2_{ext}(x) = \frac{\sum w_i \sigma^2(x_i)}{\sum w_i} \qquad (35)$$

$$\sigma^2(x') = \frac{\sigma^2_{int}}{n} \ or \ \frac{\sigma^2_{ext}}{n} \qquad (36)$$

For a list of paired observations, the ratio of averages and the average of ratios will be the same if there is a linear relationship between the observations and the error distributions are similar. A difference between these two statistics indicates a problem that should be investigated.

Standard Uncertainties of Differences and Quotients.
The relation between the absolute structure of a crystalline material and the measured Friedel pairs was given in equation (2)

$$Io \approx Ic = (1-x)Is^+ + xIs^- \qquad (37)$$

Writing

$$Do = Io^+ - Io^-$$

$$\mathbf{Ds = Is^+ - Is^-}$$

$$Ao = 0.5(Io^+ + Io^-)$$

$$As = 0.5(Is^+ + Is^-)$$

$$Qo = Do/2Ao$$

$$Qs = Ds/2As$$

we obtain the difference expression

$$Do \approx Dc = (1-2x)Ds \qquad (38)$$

and the quotient expression

$$Do/2Ao \approx (1-2x)Ds/2As$$

$$Qo \approx Qc = (1-2x)Qs \qquad (39)$$

from which

$$\sigma^2(Do) = \sigma^2(Io^+) + \sigma^2(Io^-) \qquad (40)$$

$$\sigma^2(Ao) = 0.25\sigma^2(Do) \qquad (41)$$

$$\sigma^2(Qo) = \left[\frac{2}{(Io^+ + Io^-)^2} \right]^2 \left[Io^{-2}\sigma^2(Io^+) + Io^{+2}\sigma^2(Io^-) \right] \qquad (42)$$

If $Io^+ \approx Io^- = Io = Ao$ and $\sigma^2(Io^+) \approx \sigma^2(Io^-) = \sigma^2(Io)$, we get

$$\sigma^2(Qo) = \sigma^2(Io) \Big/ \left(2Io^2\right)$$

and hence

$$Qo / \sigma(Qo) = Do / \sigma(Do)$$

and

$$Qo / \sigma^2(Qo) = 2IoDo / \sigma^2(Do)$$