

# *The Zürich School of Crystallography*

*University of Zürich*

*June 9 - 22, 2013*

## **Optimizing the model of a crystal structure**

Hans-Beat Bürgi  
Institute of Organic Chemistry, University of Zürich, Switzerland  
hans-beat.buergi@krist.unibe.ch

### **Contents**

- 1 Model of the diffraction experiment
- 2 Optimization conditions
- 3 Linear regression
- 4 Non-linear least-squares optimization
- 5 Restraints and constraints
- 6 Quality of the model
- 7 Weighting schemes

# 8 Optimizing the model of a crystal structure

## 1 Model of the diffraction experiment

Structure solution by direct methods, interpretation of the Patterson function or other procedures provides assignments of atom types and estimates of coordinates for some or all atoms in the unit cell. The reliability of these results is limited and depends on the quality of the initial estimates of the phases of the structure factors. In order to improve the reliability of the structure determination, a mathematical model of the electron density and – more generally – of the entire diffraction experiment is required. Its parameters must be adjusted to reproduce optimally the experimental observations.

The electron density in a unit cell is approximated as a sum of atomic densities, as described earlier, but – to be physically realistic – it has to also take into account atomic motion. The motionless atomic densities are usually assumed to be spherically symmetric, but may also be chosen to account for deviation from spherical symmetry due to chemical bonding. In standard structure analysis spherically symmetric density functions calculated by Hartree-Fock methods for isolated atoms are used. The details of atomic and molecular vibrations are usually very complex. It is therefore assumed that each atom moves independently of all others in a quadratic potential which may be isotropic or anisotropic. The corresponding probability density function takes the form of a 3D Gaussian distribution. Mathematically the smearing of the static atomic density  $\rho_{0,i}(\mathbf{r} - \mathbf{r}_i)$  by thermal motion is expressed in terms of a convolution with the Gaussian probability density:

$$\rho(\mathbf{r}) = \sum_{i=1}^q \left[ \rho_{0,i}(\mathbf{r} - \mathbf{r}_i) \otimes \frac{1}{(2\pi)^{3/2} \|\mathbf{U}_i\|^{1/2}} \exp\left[-\mathbf{r}^T (\mathbf{U}_i)^{-1} \mathbf{r}\right] / 2 \right]$$

(see footnote 3 in ‘Introduction to crystals, electromagnetic radiation and interference’ by Hans-Beat Bürgi and ‘Math refresher’ by Gervais Chapuis). The elements of the symmetric matrix  $\mathbf{U}$  are the mean-square expectation values of the atomic displacements  $\Delta x$ ,  $\Delta y$ ,  $\Delta z$  from the mean atomic position:  $U^{11} = \langle \Delta x^2 \rangle a^2$ ,  $U^{22} = \langle \Delta y^2 \rangle b^2$ ,  $U^{33} = \langle \Delta z^2 \rangle c^2$ ,  $U^{12} = \langle \Delta x \Delta y \rangle ab$ ,  $U^{13} = \langle \Delta x \Delta z \rangle ac$ ,  $U^{23} = \langle \Delta y \Delta z \rangle bc$ ;  $\|\mathbf{U}\|$  is the determinant of  $\mathbf{U}$ . Fourier transformation leads to a structure-factor expression, which is now somewhat more complicated

$$\begin{aligned} F(hkl) &= \sum_{i=1}^q f_{0,i}(\mathbf{S}) \text{pop}_i \exp\left[-2\pi^2 (\mathbf{r}^{*T} \mathbf{U}_i \mathbf{r}^*)\right] \cdot \exp 2\pi i \mathbf{r}_i \cdot \mathbf{r}^* \\ &= \sum_{i=1}^q f_{0,i}(\mathbf{S}) \text{pop}_i \exp\left[-2\pi^2 (h^2 a^2 U_i^{11} + k^2 b^2 U_i^{22} + l^2 c^2 U_i^{33} \right. \\ &\quad \left. + 2hka^* b^* U_i^{12} + 2hla^* c^* U_i^{13} + 2klb^* c^* U_i^{23})\right] \\ &\quad \cdot \exp 2\pi i (hx_i + ky_i + lz_i). \end{aligned}$$

The atomic scattering factor  $f_i$  of my previous lectures (see ‘Scattering from crystals: diffraction’ by Hans-Beat Bürgi) has been divided into three parts,  $f_{0,i}$  being the Fourier transform of the static density  $\rho_{0,i}(\mathbf{r} - \mathbf{r}_i)$ ,  $\text{pop}_i$  being an occupation factor and  $\exp\left[-2\pi^2 (\mathbf{r}^{*T} \mathbf{U}_i \mathbf{r}^*)\right]$  being the Fourier transform of the 3D Gaussian distribution function describing atomic motion. In the early stages of model optimization an isotropic approximation to atomic displacement parameters is usually sufficient

$$\begin{aligned}
F(hkl) &= \sum_{i=1}^q f_{0,i}(\mathbf{S}) \text{pop}_i \exp - 2\pi^2 (U_i^{iso} \mathbf{r}^* \cdot \mathbf{r}^*) \cdot \exp 2\pi i \mathbf{r}_i \cdot \mathbf{r}^* \\
&= \sum_{i=1}^q f_{0,i}(\mathbf{S}) \text{pop}_i \exp - 2\pi^2 U_i^{iso} (2 \sin \theta / \lambda)^2 \cdot \exp 2\pi i (hx_i + ky_i + lz_i).
\end{aligned}$$

For every atom there are thus three positional coordinates,  $x_i, y_i, z_i$ , to be optimized, at least one, but usually six displacement parameters  $U_i^{kl}$  and a population parameter  $\text{pop}_i$  to be determined, up to ten parameters in total.

An appropriate model for a diffraction experiment has to take into account factors other than diffraction by spherically symmetric electron density distributions. These include a factor  $k$  scaling the structure factors calculated for one unit cell to the volume of the crystal specimen, twin parameters (*BASF*), a Flack parameter (if the absolute structure of the crystal specimen needs to be specified), extinction parameter(s) (*EXTI*, accounting approximately for the failure of the kinematic scattering theory used throughout) and several others. The resulting model for the scattered intensity may be quite complex:

$$I_{model} = \frac{1}{k} |F_{model}(hkl)|^2 g(\text{BASF}, \text{Flack}, \text{EXTI}, \text{etc.}).$$

## 2 Optimization conditions

The condition used in crystal structure least-squares optimization is to minimize the sum  $S$  of the squares of the differences between observed and suitable model intensities. The uncertainty of the measurements is accounted for by a weight factors  $w(hkl)$

$$\begin{aligned}
S &= \sum_{hkl} w(hkl) [I_{obs}(hkl) - I_{model}(hkl)]^2 \\
&= \sum_{hkl} w(hkl) \left[ I_{obs}(hkl) - \left[ \frac{1}{k} |F_{model}(hkl)|^2 g(\text{BASF}, \text{Flack}, \text{EXTI}, \text{etc.}) \right] \right]^2.
\end{aligned}$$

The minimization conditions are

$$\frac{\partial S}{\partial x_i} = \frac{\partial S}{\partial y_i} = \frac{\partial S}{\partial z_i} = \frac{\partial S}{\partial U_i^{11}} = \frac{\partial S}{\partial U_i^{22}} = \frac{\partial S}{\partial U_i^{33}} = \frac{\partial S}{\partial U_i^{12}} = \frac{\partial S}{\partial U_i^{13}} = \frac{\partial S}{\partial U_i^{23}} = \dots = \frac{\partial S}{\partial k} = 0, \quad (i = 1, \dots, q).$$

These calculations imply some straightforward, but complicated algebra. The optimization process is therefore illustrated first with a much simpler example, linear regression.

## 3 Linear regression

Suppose that consecutive measurements of a quantity  $A$  may be expected to increase an initial value  $q$  by a constant amount  $p$

$$A_{model}(n) = pn + q \quad (n = 0, \dots, N).$$

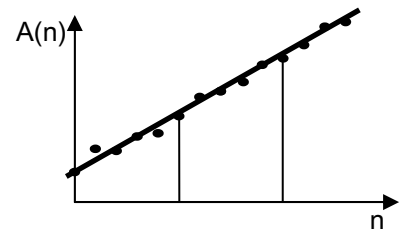


Fig. 1. Results of consecutive measurements of  $A_{obs}(n)$  expected to increase linearly with  $n$ .

Given a series of measurements  $A_{obs}(n)$  with reliability  $w(n)$ , what are the best estimates of  $p$  and  $q$ ? The quantity to be minimized and the optimization conditions are

$$S = \sum_{n=0}^N w(n)[A_{obs}(n) - A_{model}(n)]^2$$

$$= \sum_{n=0}^N w(n)[A_{obs}(n) - pn - q]^2$$

$$\frac{\partial S}{\partial p} = \frac{\partial S}{\partial q} = 0.$$

The derivative calculations lead to two equations in the unknowns  $p$ ,  $q$ . The equations are called linear because  $p$  and  $q$  occur to the first power

$$p \sum_{n=0}^N w(n)n^2 + q \sum_{n=0}^N w(n)n = \sum_{n=0}^N w(n)nA_{obs}(n)$$

$$p \sum_{n=0}^N w(n)n + q \sum_{n=0}^N w(n) = \sum_{n=0}^N w(n)A_{obs}(n).$$

These equations are more conveniently expressed in matrix notation:

$$\begin{bmatrix} \sum_{n=0}^N w(n)n^2 & \sum_{n=0}^N w(n)n \\ \sum_{n=0}^N w(n)n & \sum_{n=0}^N w(n) \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} \sum_{n=0}^N w(n)nA_{obs}(n) \\ \sum_{n=0}^N w(n)A_{obs}(n) \end{bmatrix}$$

abbreviated as  $\mathbf{Np} = \Delta$ .

The optimized parameters are

$$\mathbf{p} = \mathbf{N}^{-1}\Delta.$$

A graphical representation of this procedure is given in Fig. 2. The function  $S$  being quadratic in  $p$  and  $q$ , is represented by a 3D parabolic basin. The coordinates of its minimum represent the optimal values of the parameters. Note that the calculation of  $p_{opt}$  and  $q_{opt}$  require the known quantities  $A_{obs}(n)$ ,  $n$  and  $w(n)$  only.

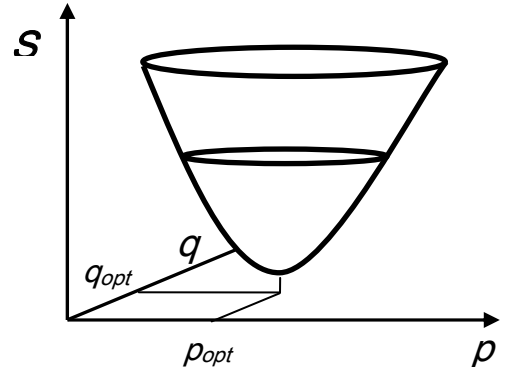


Fig. 2. Parabolic dependence of  $S$  on the parameters  $p$  and  $q$  for a linear least-squares problem. The coordinates of the minimum of the parabola are  $p_{opt}$  and  $q_{opt}$ .

#### 4 Non-linear least-squares optimization

The structure factor  $F(hkl)$  and thus  $I_{model}(hkl)$  is clearly a non-linear function in the parameters  $x_i, y_i, z_i, U_i^{kl}$ . The quantity  $S$  is therefore no longer parabolic and – most importantly – may show several minima implying that the minimization conditions

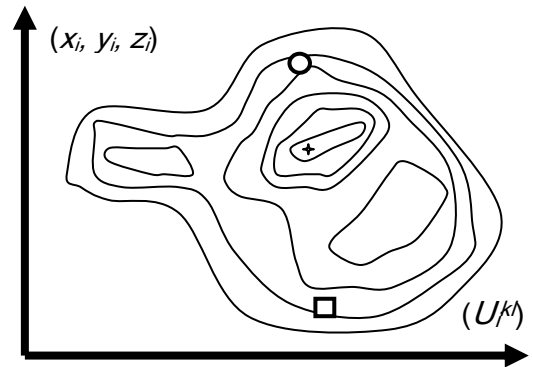


Fig. 3. Schematic contour-line representation of  $S$  as a function of the parameters  $x_i, y_i, z_i, U_i^{kl}$  for a non-linear least-squares problem. The absolute minimum is indicated by a cross. The circle indicates a set of approximate structural parameters likely to refine to the absolute minimum of  $S$ ; the square indicates a set of approximate structural parameters likely to refine to the nearby subsidiary minimum of  $S$ .

may lead to more than one solution. This is shown schematically in Fig. 3, where  $S$  is represented by contour lines. The problem is to find the absolute minimum of  $S$  (cross in Fig. 3). From the initial solution of the phase problem approximate values of the parameters  $x_i, y_i, z_i, U_i^{kl}$  are available. If these parameters are close to the absolute minimum (circle in Fig. 3) a modified form of linear regression is highly likely to find this minimum. If the initial parameters are close to one of the subsidiary minima (e.g. the square in Fig. 3), the optimization process may well end up in the subsidiary minimum and produce a deficient or even erroneous description of the crystal structure. From the point of view of actual calculations, the non-linearity of the structure-factor equation implies that the simple matrix approach given for linear regression cannot be used as it stands.

The solution to the problem is to linearize it. The model intensities  $I_{model}(hkl)$  are expanded into a Taylor series up to first order terms in the changes of the structural parameters  $\Delta p_i$  relative to the point characterized by the set of  $P$  approximate parameter values  $x_{i,0}, y_{i,0}, z_{i,0}, U_{i,0}^{kl}, k, etc.$  abbreviated below as the vector  $\mathbf{p}_0$  (circle in Fig. 3)

$$I_{model}(hkl) = I_{model}(hkl; \mathbf{p}_0) + \sum_{i=1}^P \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \Delta p_i.$$

Instead of determining the parameters, only the shifts of the parameters  $\Delta p_i$  need to be found. From the shifts, improved parameters  $\mathbf{p}_1 = \mathbf{p}_0 + \Delta \mathbf{p}$  are obtained. The shape of  $S$  around  $\mathbf{p}_0, \mathbf{p}_1$  and its absolute minimum is usually not parabolic, not even approximately. Linearization and determination of the parameter shifts must therefore be repeated until convergence is reached, *i.e.* until the parameter shifts are smaller than a specified limit. The function to be minimized is

$$\begin{aligned} S(\mathbf{p}) &= \sum_{hkl} w(hkl) \left[ I_{obs}(hkl) - I_{model}(hkl; \mathbf{p}_0) - \sum_{i=1}^P \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \Delta p_i \right]^2 \\ &= \sum_{hkl} w(hkl) \left[ \Delta I(hkl; \mathbf{p}_0) - \sum_{i=1}^P \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \Delta p_i \right]^2 \end{aligned}$$

and the minimization conditions are

$$\frac{\partial S}{\partial \Delta p_i} = 0 \quad (i = 1, \dots, P).$$

The resulting  $P$  linear equations in  $\Delta p_i$  are

$$\begin{aligned} \sum_{j=1}^P \sum_{hkl} w(hkl) \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_j} \right)_{\mathbf{p}_0} \Delta p_j \\ = \sum_{hkl} w(hkl) \Delta I(hkl; \mathbf{p}_0) \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \quad (i = 1, \dots, P). \end{aligned}$$

In matrix notation, these equations are,

$$\mathbf{N} \Delta \mathbf{p} = \Delta$$

as previously. The element of  $\mathbf{N}$  in row  $i$  and column  $j$  is

$$\sum_{hkl} w(hkl) \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_j} \right)_{\mathbf{p}_0}.$$

The element of  $\Delta$  in row  $i$  is

$$\sum_{hkl} w(hkl) \Delta I(hkl; \mathbf{p}_0) \left( \frac{\partial I_{model}(hkl; \mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0}.$$

The parameter shifts and the new parameters become

$$\begin{aligned} \Delta \mathbf{p} &= \mathbf{N}^{-1} \Delta \\ \mathbf{p}_1 &= \mathbf{p}_0 + \Delta \mathbf{p}. \end{aligned}$$

The simplicity of the final result is deceiving as may be seen from the following consideration. For a relatively small structure with 11 atoms a model with anisotropic displacement parameters contains at least  $11 \cdot 9 + 1 = 100$  parameters. The number of elements of  $\mathbf{N}$  and  $\Delta$  to be calculated is  $100 \cdot 101/2 + 100 = 5150$ . Assuming a typical ratio of parameters to observation of 1:10 there will be 1000 observations implying that 1000 values of  $I_{model}(hkl; \mathbf{p})$  and  $1000 \cdot 100 = 100'000$  values of  $\partial I_{model}(hkl; \mathbf{p})/\partial p_i$  need to be computed and properly summed up 5150 times. Keep these numbers in mind when you do your own least-squares refinement, probably with many more parameters and observations and in as little time as a few seconds to a few minutes. Experience this blessing of modern-day computing power knowingly!

## 5 Restraints and constraints

It sometimes happens that the structural model chosen is physically unreasonable or that the information contained in the diffraction intensities is insufficient to determine all parameters of a structural model. In such cases the matrix  $\mathbf{N}$  of normal equations may become singular or near-singular, *i.e.* the determinant  $\|\mathbf{N}\|$  is zero or very small (on the order of the numerical accuracy of the processor in your computer). As a consequence the inverse matrix  $\mathbf{N}^{-1}$  is undetermined and the parameter shifts  $\Delta \mathbf{p}$  cannot be obtained. Oscillating values of one or more elements  $\Delta p_i$  and unreasonably large standard uncertainties of the parameters (see below) are signs of such problems. The problem must be remedied either by choosing a more reasonable structural model or by coding independent external information on the crystal structure into restraints or constraints.

Restraints are conditions  $q_j^{res}$  imposed on features  $q_j^{model}(\mathbf{p})$  of the structural model that must be fulfilled within a certain tolerance. Such conditions are called pseudo-observational equations and are added to the minimization function  $S$

$$S' = S + \sum_j w_j [q_j^{model}(\mathbf{p}) - q_j^{res}]^2.$$

They affect  $\mathbf{N}$  and - if properly chosen - may remove the singularity.

### 5.1 Restraints

In asymmetric units containing several chemically identical units (sometimes related by pseudo-translation or other pseudo-symmetry operations), one may wish to restrain the geometry of these units to be closely similar. For distances  $d_i$  in units 1 and 2 the restraints take the form

$$S' = S + \sum_i w_i [d_{i,1}^{model}(\mathbf{p}) - d_{i,2}^{model}(\mathbf{p})]^2.$$

If an interatomic distance or other geometrical parameter  $d_j$  is poorly determined by the diffraction data, but known from other experimental or theoretical sources, its value  $d_j^{res}$  may also be introduced into the model of refinement in terms of a restraint. Examples include C-H, Cl-O or B-F distances, the latter in disordered  $\text{ClO}_4^-$  or  $\text{BF}_4^-$  ions, deviations from planarity in fragments with more than three atoms, *etc.*

$$S' = S + \sum_j w_j [d_j^{model}(\mathbf{p}) - d_j^{res}]^2.$$

For a nearly singular matrix of normal equations  $\mathbf{N}$  the parameter shifts between cycles  $k$  and  $k+1$  from  $\mathbf{p}_k$  to  $\mathbf{p}_{k+1}$  may become unreasonably large, thus catapulting the model from an area near the absolute minimum to one near a subsidiary minimum (Fig. 3). This effect can sometimes be avoided by shift-limiting restraints

$$S' = S + \sum_m w_m [p_{m,k+1} - p_{m,k}]^2.$$

This restraint translates into adding  $w_m$  to the diagonal element of  $\mathbf{N}$  corresponding to  $p_m$ .

In some space groups the origin in one or more directions can not be fixed on symmetry elements. Examples include  $P1$ ,  $P2$ ,  $P2_1$ ,  $P3$ ,  $Pna2_1$ , *etc.* Most software in common use fix the origin with a restraint described by H.D. Flack & D. Schwarzenbach (1988) (*Acta Cryst. A* **44**, 499-506).

## 5.2 Constraints

A constraint is an exact relationship between features of a structural model,  $q_j^{res} = q_j^{model}(\mathbf{p})$ . A constraint may be simulated in terms of a restraint by increasing its weight  $w_j$  to a very large value, *i.e.* decreasing the tolerance for deviations of  $q_j^{model}(\mathbf{p})$  from  $q_j^{res}$ . Charge neutrality in ionic compounds is an example that is conveniently treated in this way.

Exact constraints may be expressed in two ways. On one hand a constraint parameter may be assigned a fixed value which is not refined, *i.e.* it is not considered in the derivative calculations of the minimization conditions. Examples include special values of atomic coordinates and atomic displacement parameters required for certain special positions in some space groups. Alternatively and more generally constraints are expressed with the help of LaGrange multipliers (see any advanced text book of mathematics).

## 6 Quality of the model

There are several ways to gauge the quality of a structural model. Some quality indicators, such as  $R$ -factors and Goodness-of-fit ( $GOF$ ) are global quantities. They give only limited information on the quality of a structure determination. Others, such as the standard uncertainties of the refined parameters, which will be introduced below, refer to specific details of the structural model.

### 6.1 Global quality indicators: $R$ -factors and Goodness-of-fit

Many different  $R$ -factors have been invented. Here we mention only those in general use:  $R1$ ,  $wR2$  and  $GOF$

$$R1 = \frac{\sum_{hkl} \left| |F_{obs}(hkl)| - |F_{model}(hkl; \mathbf{p})| \right|}{\sum_{hkl} |F_{obs}(hkl)|},$$

$$wR2 = \left\{ \frac{\sum_{hkl} w(hkl) [I_{obs}(hkl) - I_{model}(hkl; \mathbf{p})]^2}{\sum_{hkl} w(hkl) [I_{obs}(hkl)]^2} \right\}^{1/2} = \left\{ \frac{S(\mathbf{p})}{\sum_{hkl} w(hkl) [I_{obs}(hkl)]^2} \right\}^{1/2},$$

$$GOF = \left\{ \frac{\sum_{hkl} w(hkl) [I_{obs}(hkl) - I_{model}(hkl; \mathbf{p})]^2}{n_{obs} - n_{par}} \right\}^{1/2} = \left\{ \frac{S(\mathbf{p})}{n_{obs} - n_{par}} \right\}^{1/2}.$$

The simplest of them,  $R1$ , measures the fraction of the sum of the structure-factor amplitudes that is *not* explained by the model with parameters  $\mathbf{p}$ . It is often multiplied by 100 and given as a percentage value. The quantity  $wR2$  is based on a similar idea; it compares the sum of the squared differences between observed and model intensities with the sum of squared intensities, both weighted with the reliability factors  $w(hkl)$ . The two quantities  $wR2$  and  $GOF$  are closely related to the function  $S$  minimized during refinement and represent reliable measures of model quality.

The values of  $R1$ ,  $wR2$  and  $GOF$  to be expected in a structure analysis depend on several factors: the quality of the crystal and thus of the diffraction data, on one hand, and the adequateness of the model, on the other hand. For molecular crystals (excluding proteins)  $R1 \sim 0.05$  is not unreasonable; as a rule of thumb  $wR2 \sim 2-3 R1$ . The expectation value of  $GOF$  is 1, but depends strongly on the reliability factors  $w(hkl)$  as discussed in chapter 7 on weighting schemes.

## 6.2 Reliability of structural model, standard uncertainties and correlation

The results of a least-squares refinement are incomplete without a quantitative statement of its uncertainty. Such a statement reflects a lack of knowledge due to random and systematic defects in the observed data and to deficiencies in the model. Least-squares refinement provides not only optimized atomic coordinates and displacement parameters but also their standard uncertainties (s.u.) and correlations between s.u.'s in terms of an uncertainty matrix  $\mathbf{U}$

$$\mathbf{U} = GOF \mathbf{N}^{-1}.$$

The diagonal elements of  $\mathbf{U}$  are  $u_i^2$ , the off-diagonal elements are  $u_i u_j cor_{ij}$ . The  $u_i$ 's are the s.u.'s of the parameters  $p_i$  (previously called estimated standard deviations). The convergence of the least-squares optimization may be expressed in terms of the ratios between parameter changes and their s.u.s,  $\Delta p_i / u_i$ . Optimization is generally continued until all  $\Delta p_i / u_i < 0.001$ . The quantities  $cor_{ij}$  are the mutual uncertainty coefficients (previously called correlation coefficients). Their meaning will be explained below.

Since atomic coordinates are dimensionless fractions of the cell constants, so are their s.u.'s. Thus the accuracy of atomic positions is the product of the coordinate s.u.'s with the cell constants. For an average structure analysis it is a few thousands of an Å. Due to their large number of electrons, heavy atoms contribute more to the total scattering than the lighter atoms with fewer electrons. Therefore the accuracy in position tends to be higher for the heavy atoms. The positions of hydrogen atoms with their single electron are only poorly defined in X-ray structure analysis. Their positional uncertainties are usually a few hundredths



to tenths of an Å. Much more precise locations of hydrogen atoms are obtained from neutron diffraction experiments. On decreasing the temperature, the s.u.'s of the parameters  $U^{kl}_i$  generally decrease because the diffraction intensities at high values of  $\sin\theta/\lambda$  become stronger and thus easier to measure reliably.

Note that the uncertainty matrix  $\mathbf{U}$  depends on the model, including constraints and restraints. If a parameter  $p_i$  is constrained, its uncertainty is necessarily zero. If a parameter or a derived quantity is restrained, its uncertainty  $u(p_i)$  or  $u(d_j)$  is approximately given by the weight of the restraint, namely as  $\sim w^{-1/2}$ . It is therefore important to report all restraints and constraints used in the structure model explicitly.

The meaning of the mutual uncertainty coefficients  $cor_{ij}$  is best illustrated graphically. Fig. 4 shows projections of  $S(\mathbf{p})$  onto the plane  $p_i, p_j$  for two different situations. On the left an example of positive mutual uncertainty is shown. The partial uncertainty matrix

$$\mathbf{U}(p_i, p_j) = \begin{bmatrix} u_i^2 & u_i u_j cor_{ij} \\ u_i u_j cor_{ij} & u_j^2 \end{bmatrix}$$

is represented as an ellipsoid. In the left half of Fig. 4 the uncertainty in  $p_i$  and  $p_j$  is assumed to be about equal, whereas it is larger in the direction  $p_i + p_j$  and smaller in the direction  $p_i - p_j$ . This is due to a positive mutual uncertainty coefficient  $cor_{ij}$ . On the right of Fig. 4 the case of negative mutual uncertainty is illustrated; the uncertainty is larger in the direction  $p_i - p_j$  and smaller in the direction  $p_i + p_j$ .

Particularly strong mutual uncertainty exists between parameters related by a symmetry operation. Consider  $p_i$  and  $p_j = -p_i$ , which are assumed to be related by a centre of symmetry or a twofold axis. Any change in  $p_i$  implies a change of the same magnitude, but of opposite sign in  $p_j$ . The mutual uncertainty coefficient between the two quantities is  $-1$ ! Note that large s.u. and mutual uncertainty coefficients are often associated with a near-singular, ill-determined matrix of normal equations  $\mathbf{N}$ . The general expression for multidimensional ellipsoids analogous to those in Fig. 4 is

$$S = \Delta\mathbf{p}^T \mathbf{U}^{-1} \Delta\mathbf{p}$$

where  $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_{opt}$  are deviations from optimized parameters.

Quite generally, mutual uncertainty has consequences on quantities derived from  $p_i$  and  $p_j$ . Suppose a bond distance depends on the difference between  $p_i$  and  $p_j$ . For positive correlation such a distance will be more accurately determined than indicated by  $u(p_i)$  and  $u(p_j)$  alone. The converse holds for negative mutual uncertainty. In general the influence of mutual uncertainty on the s.u. of a quantity  $d$  depending on the set of parameters  $\mathbf{p}$  is given by

$$u^2[d(\mathbf{p})] = (\partial d / \partial \mathbf{p})^T \mathbf{V}(\mathbf{p}) (\partial d / \partial \mathbf{p}).$$

The elements of the vector  $\partial d / \partial \mathbf{p}$  are  $\partial d / \partial p_i$ , i.e. derivatives of the derived quantity  $d$  with respect to the parameters  $p_i$  that  $d$  depends on.

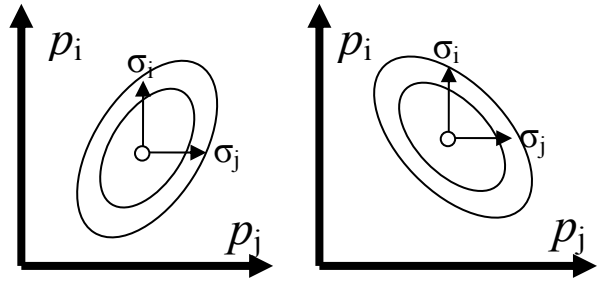


Fig. 4. Parameter correlation. Left: positive correlation coefficient, the difference between  $p_i$  and  $p_j$  is better defined than their sum and than the individual parameters. Right: negative correlation coefficient, the sum of  $p_i$  and  $p_j$  is better defined than their difference and than the individual parameters.

## 7 Weighting schemes

The quality of experimental observations varies widely depending on the quality of the crystal, the quality of the measuring equipment, experimental design and on the time and effort invested in an experiment. This is the reason for introducing a reliability or weighting factor  $w(hkl)$  in structure-factor least-squares calculations. The quantities finally used in such calculations are usually obtained in several steps.

The first step considers individual reflections. The quantity  $I_{obs}(hkl)$  is not directly measured. It is the difference between the total measured intensity  $I_{tot}$  and a suitably scaled background intensity  $sI_B$  corrected for Lorentz- and polarization effects  $Lp$  (see ‘Data processing’ by Bernhard Spingler and Michael Wörle):

$$I_o(hkl) = \frac{I_{tot} - sI_B}{Lp}.$$

The outcomes of many measurements of the same quantity under the same conditions are never exactly the same, but always affected by some random error. The distribution of measurements and thus their reliability may either be determined experimentally by multiple measurements or approximated by a theoretical model. The distribution of intensity measurements or number of photons registered ( $k$ ) is usually approximated by a Poisson distribution

$$P(k) = \frac{I^k}{k!} \exp(-I).$$

The expectation or mean value of  $k$  is  $\langle k \rangle = I$ , and the standard deviation of the mean is

$$\sigma(I) = (I)^{1/2}.$$

With this assumption and taking into account the propagation of uncertainties as described in the preceding section the s.u. of  $I_{obs}(hkl)$  becomes

$$u[I_{obs}(hkl)] = \left( \frac{I_{tot} + s^2 I_B}{Lp} \right)^{1/2}.$$

When multiple measurements  $I_{obs}(hkl)_i$  are available the mean intensity and its standard uncertainty become

$$\langle I_{obs}(hkl) \rangle = \frac{\sum_{i=1}^N I_{obs}(hkl)_i / u^2 [I_{obs}(hkl)_i]}{\sum_{i=1}^N 1 / u^2 [I_{obs}(hkl)_i]},$$

$$u[\langle I_{obs}(hkl) \rangle] = \left( \frac{N \sum_{i=1}^N [I_{obs}(hkl)_i - \langle I_{obs}(hkl) \rangle]^2 / u^2 [I_{obs}(hkl)_i]}{N-1 \sum_{i=1}^N 1 / u^2 [I_{obs}(hkl)_i]} \right)^{1/2}.$$

If there is a choice between the two, the larger of the two quantities is used to define the weight

$$w(hkl) = \text{Min of } 1/u^2[\langle I_{obs}(hkl) \rangle] \text{ or } 1/u^2[I_{obs}(hkl)].$$

With modern area detectors some reflections and their symmetry equivalents may have been measured several times in different runs, whereas others may have been observed only once or twice. In such cases the multiple measurements are used to generate a model  $u[I_{obs}(hkl)]$  which should be applicable to all observations irrespective of their frequency of observation. The frequently-observed reflections are divided into groups according to their intensity  $I_{obs}(hkl)$ . Standard uncertainties  $u[\langle I_{obs}(hkl) \rangle]$  and  $u[I_{obs}(hkl)]$  are averaged within each group. The group uncertainty  $u^2[I_{obs}(hkl)]$  is expressed in terms of these averages and of a correction term proportional to the group average  $\langle I_{obs}(hkl) \rangle^2$  according to

$$u^2[\langle I_{obs}(hkl) \rangle] = k\{u^2[I_{obs}(hkl)] + [g\langle I_{obs}(hkl) \rangle]^2\}.$$

The constants  $k$  and  $g$  are determined by least-squares optimization. If the  $u^2[I_{obs}(hkl)]$  are *bona fide* estimates, the constant  $g$  should equal 0 and  $k$  should equal 1. Acceptable values of  $g$  are in the range of 0.0-0.1, of  $k$  in the range 0.7-1.3. Deviations from these approximate values indicate problems with the intensity measurements, scaling and absorption correction of the data or with some other systematic deficiency of the data.

The weights in the structure-factor least squares are usually chosen as

$$w(hkl) = 1 / \left[ u_M^2 + \left( a \frac{\langle I_{obs}(hkl) \rangle + 2I_{model}(hkl)}{3} \right)^2 + bI_{obs}(hkl) \right].$$

To determine the constants  $a$  and  $b$  the data are divided into *ca.* 10 equally populated ranges of resolution and of the ratio  $|F_c|/|F_c(\max)|$ . The constants are chosen such that in every range the *GOF* is as close to 1 as possible. Ideally  $a$  and  $b$  should be zero. In practice values of  $a \sim 0.05$  and  $b \leq 1$  indicate inadequacies of the model of the electron density. For very good data sets the deviation from ideal behaviour may be due to the use of spherical atomic scattering factors which neglect effects of chemical bonding. For disordered structures the model of disorder may be incomplete or insufficiently flexible. Finally one should also consider the possibility that  $a$  and  $b$  compensate for inadequacies of  $u^2$  from the error model and thus for unrecognized systematic errors in the experiment and the data processing.

If results of the highest accuracy and precision attainable from a given data set are aspired to, the questions of weighting should be given careful consideration. It is nearly always possible to choose values of  $a$  and  $b$  which produce a *GOF* of 1. However, if this condition is fulfilled at the price of unreasonable values of the constants  $a$ ,  $b$ ,  $k$  and  $g$  the standard uncertainties of the atomic coordinates, anisotropic displacement parameters and of derived quantities such as bond lengths and angles should be considered with some suspicion. Comparison of distances and angles obtained from many independent studies of the same chemical compounds in the same crystal modification indicate indeed that their reported standard uncertainties are underestimated by a factor of 2-3. To give the reader an appreciation of the reliability of standard uncertainties from a least-squares refinement, it is recommended to report the values of the constants  $a$ ,  $b$ ,  $k$  and  $g$  together with *R*-factors and *GOF*-values. Unfortunately this is not a general practice.

*Acknowledgment:* I thank H.D. Flack for improving the above text and correcting many misprints (July 2009).