

## Computational aspects of PDF refinement: development of the program MolPDF

J. Rodriguez-Carvajal and Alexei Bytchkov

(Inspired from the manual of PDFFIT 1.4 by Thomas Proffen and Simon Billinge)

These notes have been written as an extension of the appendix of the manual of PDFFIT 1.4 by Thomas Proffen and Simon Billinge. We have added some additional formulae to clarify the relations of thermal parameters with the width of Gaussian peaks used to model the  $G(r)$  function defined below, as well as the derivatives when crystallographic constraints are used. The expression of  $\sigma_{ij}$  is also different of the corresponding expression used now in PDFgui because we use a special sharpening coefficient applied to specific distances. Contrary to PDFgui, which transforms a crystallographic description to  $P1$  before applying eventual symmetry constraints, we use the full crystallographic symmetry from the beginning and, in case we need to reduce the symmetry, we use subgroups of the parent space group for the analysis of real data. This avoids the application of wrong constraints inadvertently.

The ideal model for PDF analysis is based in the function  $G(r)$  that for a static distribution of atoms is provided by the formula:

$$G(r) = \frac{1}{r} \sum_{i,j} \frac{b_i b_j}{\langle b \rangle^2} \delta(r - r_{ij}) - 4\pi r \rho_0 \quad [1]$$

The variable  $r$  is the distance from the origin in real space and  $r_{ij}$  is the distance between atoms  $(i, j)$ . The delta function should be replaced by Gaussians in the case the atoms  $(i, j)$  are moving and characterized by displacement parameters. To compare the PDF with a experimentally measured quantity, the formula [1] should be convoluted with the function  $w(r)$ , which is the Fourier transform of the step function equal to 1 up to  $Q=Q_{\max}$  and zero above  $Q_{\max}$ . The function is

$$w(r) = FT[step\{0, Q_{\max}\}] = \frac{\sin(Q_{\max} r)}{r} \quad [2]$$

For a data set  $s$  and  $P$  contributing phases the total PDF is calculated as

$$G(r, s) = f_s B(r, s) \sum_{p=1}^P f_p G_p(r, s) \quad [3]$$
$$B(r, s) = \exp\left[-\frac{(rQ_{damp}(s))^2}{2}\right];$$

Where  $f_s$  is the global scale factor for pattern  $s$  and  $f_p$  is phase scale factor (their sum is equal to 1). The  $B(r, s)$  function takes into account the experimental resolution factor (it depends on the data set and not on the phase). The convolution with [2] is dependent of the particular data set. The quantity that has to be compared with the experimental PDF for set  $s$  is:

$$G_{obs}(r, s) \Rightarrow G_{calc}(r, s) = G(r, s) * w(r, s) = G(r, s) * \frac{\sin(Q_{s,max} r)}{r} \quad [4]$$

To simplify the expressions we will use a single phase and a single PDF pattern (single data set) in most of the formulae. In the following we will drop the  $s$  and  $p$  indices when convenient (we replace  $G_p(r, s)$  by  $G(r)$ ). The extension to multiple phases and multiple data sets is straightforward.

The effective way of calculating the PDF of a single phase is to use the formula:

$$G(r) = \frac{1}{Nr} \sum_{i,j} A_{ij} T(r - r_{ij}, \sigma_{ij}) - 4\pi r \rho_0 \quad [5]$$

$$A_{ij} = \frac{c_i c_j b_i b_j}{\langle b \rangle^2}; \quad T(r - r_{ij}, \sigma_{ij}) = \frac{1}{\sqrt{2\pi} \sigma_{ij}} \exp \left[ -\frac{(r - r_{ij})^2}{2\sigma_{ij}^2} \right]; \quad N = \sum_i c_i; \quad \rho_0 = \frac{N}{V}$$

The normalized Gaussian contains the interatomic distances and the width is dominated by the thermal displacement parameters as used in Crystallography. However the correlated motion can be taken into account by using the expression:

$$\sigma_{ij} = s_{mol} \sigma'_{ij} \sqrt{1 - \frac{\delta_1}{r_{ij}} - \frac{\delta_2}{r_{ij}^2} + Q_B^2 r_{ij}^2} \quad [6]$$

The sharpening factor  $s_{mol}$  is normally equal to 1.0 but it may be refined for intramolecular distances, or for distances for which a restraint or a rigid body constraint has been provided. An eventual dependency of  $s_{mol}$  on the distance is being investigated. Where  $\sigma'_{ij}$  is calculated taking into account the thermal displacement parameters of atom  $i$  and  $j$ , as follows:

$$\sigma'_{ij} = \frac{1}{r_{ij}} \sqrt{\sum_{l,m=1}^3 r_{ij}^l r_{ij}^m \{U_{lm}(i) + U_{lm}(j)\}} \quad [7]$$

### Relations between different definitions of thermal parameters

$$T(\mathbf{h}) = \exp\{-2\pi^2 \langle (\mathbf{h}\mathbf{u})^2 \rangle\} = \exp\{-2\pi^2 \mathbf{h}^t \mathbf{U}^* \mathbf{h}\} = \exp\{-\mathbf{h}^t \boldsymbol{\beta} \mathbf{h}\} = \exp\{-2\pi^2 \mathbf{H}^t \mathbf{U}_{cart} \mathbf{H}\}$$

$$\mathbf{N} = \begin{pmatrix} a^* & 0 & 0 \\ 0 & b^* & 0 \\ 0 & 0 & c^* \end{pmatrix} \quad T(\mathbf{h}) = \exp\{-2\pi^2 \mathbf{h}^t \mathbf{N} \mathbf{U}_{CIF} \mathbf{N}^t \mathbf{h}\} = \exp\left\{-2\pi^2 \begin{pmatrix} ha^* & kb^* & lc^* \end{pmatrix} \mathbf{U}_{CIF} \begin{pmatrix} ha^* \\ kb^* \\ lc^* \end{pmatrix}\right\}$$

$$\mathbf{x}_{cart} = \mathbf{A}\mathbf{x}_{frac} \Rightarrow \mathbf{U}_{cart} = \mathbf{A}\mathbf{U}^*\mathbf{A}^t; \quad \mathbf{U}^* = \mathbf{A}^{-1}\mathbf{U}_{cart}(\mathbf{A}^t)^{-1}; \quad \mathbf{U}_{CIF} = \mathbf{N}^{-1}\mathbf{U}^*(\mathbf{N}^t)^{-1}; \quad \mathbf{U}^* = \mathbf{N}\mathbf{U}_{CIF}\mathbf{N}^t$$

$$\beta_{\mu\nu} = 2\pi^2 \langle u_\mu u_\nu \rangle = 2\pi^2 a_\mu^* a_\nu^* U_{\mu\nu} \quad \mathbf{u} = \sum_\mu u_\mu \mathbf{a}_\mu = \sum_\mu U_\mu a_\mu^* \mathbf{a}_\mu = \sum_\mu U_\mu \mathbf{e}_\mu \Rightarrow \mathbf{e} = (a^* \mathbf{a} \quad b^* \mathbf{b} \quad c^* \mathbf{c})$$

$$\boldsymbol{\beta} = 2\pi^2 \mathbf{N}\mathbf{U}_{CIF}\mathbf{N}^t \Rightarrow \mathbf{U}_{CIF} = \frac{1}{2\pi^2} \mathbf{N}^{-1} \boldsymbol{\beta} \mathbf{N}^{-1}$$

The Cartesian components of the thermal displacement parameters are needed for determining the thermal ellipsoid principal axes and orientation. The mean-square displacement along an arbitrary direction given by the unitary vector  $\mathbf{s}$  is given by:

$$\langle u_s^2 \rangle = \mathbf{s}^t \mathbf{U}_{cart} \mathbf{s} \quad \text{with} \quad T(\mathbf{H}) = \exp\{-2\pi^2 \mathbf{H}^t \mathbf{U}_{cart} \mathbf{H}\}$$

The relations are independent of the reference system we use, so  $T(\mathbf{H})=T(\mathbf{h})$  and then:

$$\mathbf{H}^t \mathbf{U}_{cart} \mathbf{H} = \mathbf{h}^t \mathbf{U}^* \mathbf{h} \Rightarrow \mathbf{H}^t \mathbf{A} \mathbf{U}^* \mathbf{A}^t \mathbf{H} \Rightarrow \mathbf{h} = \mathbf{A}^t \mathbf{H} \Leftrightarrow \mathbf{H} = (\mathbf{A}^t)^{-1} \mathbf{h}$$

$$\mathbf{H}^t \mathbf{U}_{cart} \mathbf{H} = \mathbf{h}^t \mathbf{N} \mathbf{U}_{CIF} \mathbf{N} \mathbf{h} = \mathbf{H}^t \mathbf{A} \mathbf{N} \mathbf{U}_{CIF} \mathbf{N} \mathbf{A}^t \mathbf{H} \Rightarrow \mathbf{U}_{cart} = \mathbf{A} \mathbf{N} \mathbf{U}_{CIF} \mathbf{N} \mathbf{A}^t$$

$$\langle u_s^2 \rangle = \mathbf{s}^t \mathbf{U}_{cart} \mathbf{s} = \mathbf{s}^t \mathbf{A} \mathbf{N} \mathbf{U}_{CIF} \mathbf{N} \mathbf{A}^t \mathbf{s} = \mathbf{u}^t \mathbf{A}^t \mathbf{A} \mathbf{N} \mathbf{U}_{CIF} \mathbf{N} \mathbf{A}^t \mathbf{A} \mathbf{u} = \mathbf{u}^t \mathbf{G} \mathbf{N} \mathbf{U}_{CIF} \mathbf{N} \mathbf{G} \mathbf{u}$$

$$\langle u_s^2 \rangle = \mathbf{s}^t \mathbf{U}_{cart} \mathbf{s} = \mathbf{h}^t \mathbf{U}^* \mathbf{h} = \mathbf{h}^t \mathbf{N} \mathbf{U}_{CIF} \mathbf{N} \mathbf{h} = \frac{1}{r_{ij}^2} \mathbf{r}_{ij}^t \mathbf{G} \mathbf{N} \mathbf{U}_{CIF} \mathbf{N} \mathbf{G} \mathbf{r}_{ij}$$

$$\mathbf{v} = \frac{\mathbf{r}_{ij}}{r_{ij}} \text{ (unitary vector in direct basis)} \Rightarrow \mathbf{v}^* = \mathbf{G} \frac{\mathbf{r}_{ij}}{r_{ij}} \text{ (unitary vector in reciprocal basis)}$$

The components of  $\mathbf{r}_{ij}$  are fractional coordinates of the interatomic vector.  $\mathbf{G}$  is the direct metric tensor (symmetric matrix). In practice the calculation of [7] is done properly by using:

$$\sigma_{ij}'^2 = \frac{1}{r_{ij}^2} \mathbf{r}_{ij}^t \mathbf{G} \mathbf{N} \{ \mathbf{U}_{CIF}(i) + \mathbf{U}_{CIF}(j) \} \mathbf{N} \mathbf{G} \mathbf{r}_{ij} = \frac{1}{r_{ij}^2} \mathbf{r}_{ij}^t \mathbf{U}_n(i, j) \mathbf{r}_{ij} \quad [8]$$

$$\mathbf{U}_n(i, j) = \mathbf{G} \mathbf{N} \{ \mathbf{U}_{CIF}(i) + \mathbf{U}_{CIF}(j) \} \mathbf{N} \mathbf{G}$$

The quantity  $\mathbf{U}_n(i, j)$  for each type of pair is previously stored and the calculation of [8] is done at the same time of calculating the interatomic distances.

Let us establish the relation between  $\sigma'_{ij}$  and  $B_{iso}(i)$  and  $B_{iso}(j)$  as used in crystallography. The relation between  $\boldsymbol{\beta}$  and  $B_{iso}$  (or a diagonal matrix,  $\mathbf{B}_{iso}$ , having  $B_{iso}$  as terms) is given by the expressions:

$$\boldsymbol{\beta} = \frac{1}{4} \mathbf{O}_C \mathbf{O}_C^t \mathbf{B}_{iso} = \frac{1}{4} \mathbf{G}^* \mathbf{B}_{iso} \Rightarrow U_{ij} = \frac{\beta_{ij}}{2\pi^2 a_i^* a_j^*} \Rightarrow \mathbf{U}_{CIF} = \frac{1}{2\pi^2} \mathbf{N}^{-1} \boldsymbol{\beta} \mathbf{N}^{-1}$$

$$\mathbf{U}_{CIF} = \frac{1}{2\pi^2} \mathbf{N}^{-1} \boldsymbol{\beta} \mathbf{N}^{-1} = \frac{1}{8\pi^2} \mathbf{N}^{-1} \mathbf{G}^* \mathbf{N}^{-1} B_{iso}$$

In which the matrix  $\mathbf{O}_C$  is the transformation matrix from Cartesian coordinates to fractional coordinates (just the inverse of the  $\mathbf{A}$ -matrix used above:  $\mathbf{O}_C = \mathbf{A}^{-1}$ ). Substituting  $\mathbf{U}_{CIF}$  into the expression [8] we obtain:

$$\begin{aligned}\sigma'_{ij} &= \frac{1}{8\pi^2 r_{ij}^2} \mathbf{r}'_{ij} \mathbf{G} \mathbf{N} \left\{ \mathbf{N}^{-1} \mathbf{G}^* \mathbf{N}^{-1} B_{iso}(i) + \mathbf{N}^{-1} \mathbf{G}^* \mathbf{N}^{-1} B_{iso}(j) \right\} \mathbf{N} \mathbf{G} \mathbf{r}_{ij} \\ \sigma'_{ij} &= \frac{B_{iso}(i) + B_{iso}(j)}{8\pi^2 r_{ij}^2} \mathbf{r}'_{ij} \mathbf{G} \left\{ \mathbf{G}^* \right\} \mathbf{G} \mathbf{r}_{ij} = \frac{B_{iso}(i) + B_{iso}(j)}{8\pi^2 r_{ij}^2} \mathbf{r}'_{ij} \mathbf{G} \mathbf{r}_{ij} = \frac{B_{iso}(i) + B_{iso}(j)}{8\pi^2}\end{aligned}\quad [8']$$

Notice that  $\sigma'_{ij}$  does not depend on cell parameters when using  $B_{iso}$  parametrization.

### Refinement of the observed PDF pattern using Least-Squares

When we have a single observed PDF pattern and a single phase contributing to it, the function to be minimized by least squares (having an initial crystallographic model) is the following:

$$\chi^2 = \sum_r \left[ \frac{G_{obs}(r) - G_{calc}(r, \boldsymbol{\beta})}{\sigma(r)} \right]^2 \quad [9]$$

Where  $\sigma(r)$  is the standard deviation of  $G_{obs}(r)$  and the vector  $\boldsymbol{\beta}$  contains the free parameters of the model. In our case they are: scale factor(s)  $f_p$ , resolution damping factor(s)  $Q_{damp}(s)$ , sharpening/broadening factors:  $\delta_1, \delta_2, Q_B$ , cell parameters  $(a, b, c, \alpha, \beta, \gamma)$ , positional parameters  $(x, y, z)_{i=1, \text{Natoms}}$  (with constraints of the space group or not), occupation factors  $c_i$  and the thermal parameters  $\{B_{iso}\}_{i=1, \text{Niso}}$  for isotropic atoms or  $\{U_{CIF}(1,1), U_{CIF}(2,2), U_{CIF}(3,3), U_{CIF}(1,2), U_{CIF}(1,3), U_{CIF}(2,3)\}_{i=1, \text{Naniso}}$  for anisotropic atoms. For doing properly the refinement the gradient of the function 9 with respect to the free parameters is needed. The derivatives of the function  $G_{calc}(r, \boldsymbol{\beta})$  will be discussed in the forthcoming paragraphs. One has to keep in mind that:

$$\frac{\partial G_{calc}(r, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{\partial \{G(r, \boldsymbol{\beta}) * w(r)\}}{\partial \boldsymbol{\beta}} = \frac{\partial G(r, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} * w(r) \quad [10]$$

So, what will be discussed for each possible parameter to be refined are the derivatives:  $\frac{\partial G(r, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$  and we have to take into account that they have to be convoluted with  $w(r)$  for getting the full derivatives.

### Derivative of $G(r)$ with respect to the scale factors and resolution damping factors

From equations [3] we have:

$$\frac{\partial G(r, s)}{\partial f_s} = B(r, s) \sum_{p=1}^P f_p G_p(r, s) = \frac{G(r, s)}{f_s}; \quad \frac{\partial G(r, s)}{\partial f_p} = f_s B(r, s) G_p(r, s) \quad [11]$$

$$\frac{\partial G(r, s)}{\partial Q_{damp}(s)} = -r^2 Q_{damp}(s) f_s B(r, s) \sum_{p=1}^P f_p G_p(r, s) = -r^2 Q_{damp}(s) G(r, s) \quad [12]$$

### Derivative of $G(r)$ with respect to the molecular sharpening factor $s_{mol}$ .

From equations [5] we see that  $\sigma_{ij}$  parameters appear only in the Gaussian function  $T$ . We have first to make the derivative with respect to sigma and then  $\sigma_{ij}$  with respect to  $s_{mol}$ . This factor is different for each phase so for a particular phase we have:

$$\begin{aligned} \frac{\partial G(r, s)}{\partial s_{mol}(p)} &= f_s B(r, s) f_p \frac{\partial G_p(r, s)}{\partial s_{mol}(p)} = f_s f_p B(r, s) \frac{1}{N_p r} \sum_{i,j} A_{ij} \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} \frac{\partial \sigma_{ij}}{\partial s_{mol}} \\ T(r - r_{ij}, \sigma_{ij}) &= \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left[-\frac{(r - r_{ij})^2}{2\sigma_{ij}^2}\right] \\ \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} &= -\frac{1}{\sqrt{2\pi}\sigma_{ij}^2} \exp\left[-\frac{(r - r_{ij})^2}{2\sigma_{ij}^2}\right] + \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left[-\frac{(r - r_{ij})^2}{2\sigma_{ij}^2}\right] \frac{(r - r_{ij})^2}{\sigma_{ij}^3} \\ \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} &= \frac{T(r - r_{ij}, \sigma_{ij})}{\sigma_{ij}} \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \\ \frac{\partial \sigma_{ij}}{\partial s_{mol}} &= \sigma'_{ij} \sqrt{1 - \frac{\delta_1}{r_{ij}} - \frac{\delta_2}{r_{ij}^2} + Q_B^2 r_{ij}^2} = \frac{\sigma_{ij}}{s_{mol}} \end{aligned}$$

So, finally:

$$\frac{\partial G(r, s)}{\partial s_{mol}(p)} = \frac{f_s f_p B(r, s)}{N_p r} \sum_{i,j} A_{ij} \frac{T(r - r_{ij}, \sigma_{ij})}{s_{mol}} \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \quad [13]$$

If we consider  $s_{mol}$  as a function depending on a parameter  $s_h$ , and the distance  $r$ , like  $s_{mol} = 1 - \exp(-s_h r)$ , the derivative with respect to the  $s_h$  parameter is given by:

$$\begin{aligned} \frac{\partial G(r, s)}{\partial s_h} &= \frac{f_s f_p B(r, s)}{N_p r} \sum_{i,j} A_{ij} \frac{T(r - r_{ij}, \sigma_{ij})}{s_{mol}} \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \frac{\partial s_{mol}}{\partial s_h} = \\ &= \frac{f_s f_p B(r, s)}{N_p r} \sum_{i,j} A_{ij} \frac{T(r - r_{ij}, \sigma_{ij})}{s_{mol}} \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] r_{ij} \exp(-s_h r) \end{aligned} \quad [13']$$

In the code, the intermediate variables used for representing these formulae are:

$$\text{Gval}(r): T(r - r_{ij}, \sigma_{ij}), \text{yc}(r) = \text{Intens} * \text{Gval}(r): A_{ij} T(r - r_{ij}, \sigma_{ij}), \text{rmij}(r): \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right]$$

### Derivative of $G(r)$ with respect to the linear sharpening (dynamic correlation) factor $\delta_1$ .

The derivative with respect to  $\delta_1$  is done similarly to the previous one except that the final derivative of  $\sigma_{ij}$  with respect to  $\delta_1$  is different:

$$\frac{\partial G(r, s)}{\partial \delta_1(p)} = f_s B(r, s) f_p \frac{\partial G_p(r, s)}{\partial \delta_1(p)} = f_s f_p B(r, s) \frac{1}{N_p r} \sum_{i,j} A_{ij} \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} \frac{\partial \sigma_{ij}}{\partial \delta_1}$$

$$\frac{\partial \sigma_{ij}}{\partial \delta_1} = s_{mol} \sigma'_{ij} \frac{1}{2 \sqrt{1 - \frac{\delta_1}{r_{ij}} - \frac{\delta_2}{r_{ij}^2} + Q_B^2 r_{ij}^2}} \left( -\frac{1}{r_{ij}} \right) = -\frac{s_{mol} \sigma_{ij}'^2}{2 r_{ij} \sigma_{ij}}$$

So, finally:

$$\frac{\partial G(r, s)}{\partial \delta_1(p)} = -\frac{f_s f_p B(r, s)}{2 N_p r} \sum_{i,j} A_{ij} \frac{T(r - r_{ij}, \sigma_{ij})}{\sigma_{ij}^2} \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \frac{s_{mol} \sigma_{ij}'^2}{r_{ij}} \quad [14]$$

### Derivative of $G(r)$ with respect to the quadratic sharpening (dynamic correlation) factor $\delta_2$ .

The derivative with respect to  $\delta_2$  is equal to the previous one except that each term in the sum must be divided again by  $r_{ij}$ , because:

$$\frac{\partial \sigma_{ij}}{\partial \delta_2} = s_{mol} \sigma'_{ij} \frac{1}{2 \sqrt{1 - \frac{\delta_1}{r_{ij}} - \frac{\delta_2}{r_{ij}^2} + Q_B^2 r_{ij}^2}} \left( -\frac{1}{r_{ij}^2} \right) = -\frac{s_{mol} \sigma_{ij}'^2}{2 r_{ij}^2 \sigma_{ij}}$$

So the final derivative has a similar expression to that of  $\delta_1$ :

$$\frac{\partial G(r, s)}{\partial \delta_2(p)} = -\frac{f_s f_p B(r, s)}{2 N_p r} \sum_{i,j} A_{ij} \frac{T(r - r_{ij}, \sigma_{ij})}{\sigma_{ij}^2} \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \frac{s_{mol} \sigma_{ij}'^2}{r_{ij}^2} \quad [15]$$

### Derivative of $G(r)$ with respect to the broadening factor $Q_B$ .

Similarly to the previous derivatives with respect to  $\delta_1$  and  $\delta_2$ , the derivative with respect to  $Q_B$  is given by:

$$\frac{\partial \sigma_{ij}}{\partial Q_B} = s_{mol} \sigma'_{ij} \frac{1}{2 \sqrt{1 - \frac{\delta_1}{r_{ij}} - \frac{\delta_2}{r_{ij}^2} + Q_B^2 r_{ij}^2}} 2 Q_B r_{ij}^2 = \frac{s_{mol} Q_B r_{ij}^2 \sigma_{ij}'^2}{\sigma_{ij}}$$

$$\frac{\partial G(r, s)}{\partial Q_B(p)} = \frac{f_s f_p B(r, s)}{N_p r} \sum_{i,j} A_{ij} \frac{T(r - r_{ij}, \sigma_{ij})}{\sigma_{ij}^2} \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] s_{mol} Q_B r_{ij}^2 \sigma_{ij}'^2 \quad [16]$$

### Derivative of $G(r)$ with respect to the cell parameters.

The cell parameters an (for  $a, b, c, \alpha, \beta, \gamma$ ) appear in the expressions of  $r_{ij}$ ,  $\sigma_{ij}$  and cell volume  $V_p$ , so that only the terms  $T(r-r_{ij}, \sigma_{ij})$  and  $\rho_0$  have to be explicitly considered.

The general expression is:

$$\frac{\partial G_p(r, s)}{\partial a_n(p)} = f_s f_p B(r, s) \left\{ \frac{1}{N_p r} \sum_{i,j} A_{ij} \frac{\partial T(r-r_{ij}, \sigma_{ij})}{\partial a_n(p)} - 4\pi r \frac{\partial \rho_0}{\partial a_n(p)} \right\} \quad [17]$$

The derivatives inside the braces, dropping the index  $p$ , are calculated as follows:

$$\begin{aligned} \frac{\partial T(r-r_{ij}, \sigma_{ij})}{\partial a_n} &= \frac{\partial T(r-r_{ij}, \sigma_{ij})}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial a_n} + \frac{\partial T(r-r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} \frac{\partial \sigma_{ij}}{\partial a_n} \\ \frac{\partial T(r-r_{ij}, \sigma_{ij})}{\partial a_n} &= \frac{T(r-r_{ij}, \sigma_{ij})}{\sigma_{ij}} \left\{ \left[ \frac{r-r_{ij}}{\sigma_{ij}} - \frac{r\sigma_{ij}}{r_{ij}^2} \right] \frac{\partial r_{ij}}{\partial a_n} + \left[ \frac{(r-r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \frac{\partial \sigma_{ij}}{\partial a_n} \right\} \end{aligned} \quad [18]$$

In the code, the variable  $rsij(r)$  corresponds to:  $\left[ \frac{(r-r_{ij})}{\sigma_{ij}} - \frac{r\sigma_{ij}}{r_{ij}^2} \right]$ .

If we write:

$$\sigma_{ij} = s_{mol} \sigma'_{ij} \sqrt{1 - \frac{\delta_1}{r_{ij}} - \frac{\delta_2}{r_{ij}^2} + Q_B^2 r_{ij}^2} = s_{mol}(r_{ij}) \sigma'_{ij}(r_{ij}) C(r_{ij})$$

Now we need the derivatives of  $r_{ij}$  and  $\sigma_{ij}$  with respect to  $a_n$ .

We notice that for anisotropic case we have:

$$\begin{aligned} \frac{\partial \sigma_{ij}}{\partial a_n} &= \frac{\partial \sigma_{ij}}{\partial s_{mol}} \frac{\partial s_{mol}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial a_n} + \frac{\partial \sigma_{ij}}{\partial \sigma'_{ij}} \frac{\partial \sigma'_{ij}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial a_n} + \frac{\partial \sigma_{ij}}{\partial C} \frac{\partial C}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial a_n} \\ \frac{\partial \sigma_{ij}}{\partial a_n} &= \sigma'_{ij}(r_{ij}) C(r_{ij}) \frac{\partial s_{mol}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial a_n} + s_{mol} C(r_{ij}) \frac{\partial \sigma'_{ij}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial a_n} + s_{mol} \sigma'_{ij}(r_{ij}) \frac{\partial C}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial a_n} \\ \frac{\partial \sigma_{ij}}{\partial a_n} &= \left( \frac{\sigma_{ij}}{s_{mol}} \frac{\partial s_{mol}}{\partial r_{ij}} + \frac{\sigma_{ij}}{\sigma'_{ij}} \frac{\partial \sigma'_{ij}}{\partial r_{ij}} + \frac{\sigma_{ij}}{C} \frac{\partial C}{\partial r_{ij}} \right) \frac{\partial r_{ij}}{\partial a_n} \end{aligned} \quad [19]$$

The middle term in parenthesis and the derivative of w.r.t the cell parameters will be calculated later.

For inorganic compounds for which  $s_{mol}$  is constant and equal to 1, the first term is zero. If, moreover, both atoms are isotropic the middle term is also zero because  $\sigma'_{ij}$  is independent of the distance. In such a case

$$\frac{\partial \sigma_{ij}}{\partial a_n} = \frac{\sigma_{ij}}{C} \frac{\partial C}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial a_n} = \frac{\sigma_{ij}}{C^2} \left( \frac{\delta_1}{2r_{ij}^2} + \frac{\delta_2}{r_{ij}^3} + r_{ij} Q_B^2 \right) \frac{\partial r_{ij}}{\partial a_n} = \frac{\sigma'_{ij}}{\sigma_{ij}} \left( \frac{\delta_1}{2r_{ij}^2} + \frac{\delta_2}{r_{ij}^3} + r_{ij} Q_B^2 \right) \frac{\partial r_{ij}}{\partial a_n} \quad [19']$$

The last expression is the most general for isotropic case. For molecular compounds considering isotropic case and a simple expression of  $s_{mol}$ , like  $s_{mol}=1-\exp(-s_h r_{ij})$ , we have:

$$\begin{aligned} \frac{\partial \sigma_{ij}}{\partial a_n} &= \left( \frac{\sigma_{ij}}{s_{mol}} \frac{\partial s_{mol}}{\partial r_{ij}} + \frac{\sigma_{ij}}{C} \frac{\partial C}{\partial r_{ij}} \right) \frac{\partial r_{ij}}{\partial a_n} = \left( s_h \exp(-s_h r_{ij}) \frac{\sigma_{ij}}{s_{mol}} + \frac{\sigma_{ij}}{C^2} \left( \frac{\delta_1}{2r_{ij}^2} + \frac{\delta_2}{r_{ij}^3} + r_{ij} Q_B^2 \right) \right) \frac{\partial r_{ij}}{\partial a_n} \\ \frac{\partial \sigma_{ij}}{\partial a_n} &= \left( s_h \exp(-s_h r_{ij}) \frac{\sigma_{ij}}{s_{mol}} + \frac{\sigma_{ij}}{C^2} \left( \frac{\delta_1}{2r_{ij}^2} + \frac{\delta_2}{r_{ij}^3} + r_{ij} Q_B^2 \right) \right) \frac{\partial r_{ij}}{\partial a_n} \end{aligned} \quad [19'']$$

The term  $\frac{\partial r_{ij}}{\partial a_n}$  is represented in the code by the variable `drrd`. Taking into account the expression [7], we

obtain, by deriving with respect to  $a_n$ , the following expressions:

$$\begin{aligned} \frac{\partial \sigma'_{ij}}{\partial a_n} &= r_{ij} \sigma'_{ij} \frac{\partial}{\partial a_n} \left( \frac{1}{r_{ij}} \right) + \frac{1}{r_{ij}} \frac{1}{2r_{ij} \sigma'_{ij}} \frac{\partial}{\partial a_n} \left( \sum_{l,m} r_{ij}^l r_{ij}^m \{U_{lm}(i) + U_{lm}(j)\} \right) \\ \frac{\partial \sigma'_{ij}}{\partial a_n} &= -\frac{\sigma'_{ij}}{r_{ij}} \frac{\partial r_{ij}}{\partial a_n} + \frac{1}{2r_{ij}^2 \sigma'_{ij}} \sum_{l,m} \frac{\partial}{\partial a_n} (r_{ij}^l r_{ij}^m) \{U_{lm}(i) + U_{lm}(j)\} \end{aligned}$$

$$\begin{aligned} \sum_{l,m} \frac{\partial}{\partial a_n} (r_{ij}^l r_{ij}^m) \{U_{lm}(i) + U_{lm}(j)\} &= \sum_{l,m} \{\delta_{nl} d_{ij}^l r_{ij}^m + r_{ij}^l \delta_{nm} d_{ij}^m\} \{U_{lm}(i) + U_{lm}(j)\} = \\ &= \sum_m d_{ij}^n r_{ij}^m \{U_{nm}(i) + U_{nm}(j)\} + \sum_l d_{ij}^n r_{ij}^l \{U_{nl}(i) + U_{nl}(j)\} = 2d_{ij}^n \sum_l a_l d_{ij}^l \{U_{nl}(i) + U_{nl}(j)\} \end{aligned}$$

So that

$$\frac{\partial \sigma'_{ij}}{\partial a_n} = -\frac{\sigma'_{ij}}{r_{ij}} \frac{\partial r_{ij}}{\partial a_n} + \frac{d_{ij}^n}{r_{ij}^2 \sigma'_{ij}} \sum_l a_l d_{ij}^l \{U_{nl}(i) + U_{nl}(j)\} \quad [20]$$

This derivative is represented by the variable `dsp`. In the case of both isotropic temperature factors the contribution given by [19] is zero.

We have taken into account that the components  $r_{ij}^l = d_{ij}^l a_l$ , where  $d_{ij}^l = x_j^l - x_i^l$  are the fractional components of the interatomic vector. The derivatives with respect to the angles make that the second term is zero because  $r_{ij}^l$  do not depend on them. We need now to calculate the derivatives of  $r_{ij}$  with respect to  $a_n$ , which are explicitly given by the following relations:



$$\begin{aligned}
r_{ij} &= \sqrt{\mathbf{d}_{ij}^T \mathbf{G} \mathbf{d}_{ij}} \Rightarrow \frac{\partial r_{ij}}{\partial a_n} = \frac{1}{2r_{ij}} \mathbf{d}_{ij}^T \frac{\partial \mathbf{G}}{\partial a_n} \mathbf{d}_{ij} \\
\frac{\partial r_{ij}}{\partial a} &= \frac{1}{r_{ij}} \left( ad_{ijx}^2 + b \cos \gamma d_{ijx} d_{ijy} + c \cos \beta d_{ijx} d_{ijz} \right) & \frac{\partial r_{ij}}{\partial \alpha} &= -\frac{bc}{r_{ij}} \sin \alpha d_{ijy} d_{ijz} \\
\frac{\partial r_{ij}}{\partial b} &= \frac{1}{r_{ij}} \left( bd_{ijy}^2 + a \cos \gamma d_{ijx} d_{ijy} + c \cos \alpha d_{ijy} d_{ijz} \right) & \frac{\partial r_{ij}}{\partial \beta} &= -\frac{ac}{r_{ij}} \sin \beta d_{ijx} d_{ijz} \\
\frac{\partial r_{ij}}{\partial c} &= \frac{1}{r_{ij}} \left( cd_{ijz}^2 + a \cos \beta d_{ijx} d_{ijz} + b \cos \alpha d_{ijy} d_{ijz} \right) & \frac{\partial r_{ij}}{\partial \gamma} &= -\frac{ab}{r_{ij}} \sin \gamma d_{ijx} d_{ijy}
\end{aligned} \tag{21}$$

The derivatives with respect to angles are given in radians and they have to be multiplied by  $180/\pi$  to work with degrees. The derivatives of the density  $\rho_0 = N_p/V_p$  with respect to the cell parameters are given by:

$$\frac{\partial \rho_0}{\partial a_n} = -\frac{N_p}{V_p^2} \frac{\partial V_p}{\partial a_n} = \begin{cases} -\frac{\rho_0}{a_n} & \text{for } a_n = a, b, c \\ -\frac{\rho_0}{V_p^2} a^2 b^2 c^2 (\sin \alpha \cos \alpha - \sin \alpha \cos \beta \cos \gamma) & \text{for } a_n = \alpha \\ -\frac{\rho_0}{V_p^2} a^2 b^2 c^2 (\sin \beta \cos \beta - \sin \beta \cos \alpha \cos \gamma) & \text{for } a_n = \beta \\ -\frac{\rho_0}{V_p^2} a^2 b^2 c^2 (\sin \gamma \cos \gamma - \sin \gamma \cos \alpha \cos \beta) & \text{for } a_n = \gamma \end{cases} \tag{22}$$

We have to use the expressions [20] to explicitly calculate [17], once [18] has been calculated by inserting [19] and then into [17]. Finally [16] has to be calculated using [17] and [21].

### Derivative of $G(r)$ with respect to the thermal parameters.

Thermal parameters appear only in the function  $T$  and they determine the width of the Gaussians. The derivatives are calculated as:

$$\frac{\partial G_p(r, s)}{\partial U_{lm}(i)} = \frac{f_s f_p B(r, s)}{N_p r} \sum_{i,j} A_{ij} \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} \frac{\partial \sigma_{ij}}{\partial \sigma'_{ij}} \frac{\partial \sigma'_{ij}}{\partial U_{lm}(i)} \tag{23}$$

Where

$$\frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} = \frac{T(r - r_{ij}, \sigma_{ij})}{\sigma_{ij}} \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \text{ and } \frac{\partial \sigma_{ij}}{\partial \sigma'_{ij}} = \frac{\sigma_{ij}}{\sigma'_{ij}}$$

$$\frac{\partial \sigma'_{ij}}{\partial U_{lm}(i)} = \frac{1}{2r_{ij}^2 \sigma'_{ij}} 2r_{ij}^l r_{ij}^m$$

$$\begin{aligned}
\frac{\partial G_p(r,s)}{\partial U_{lm}(i)} &= \frac{f_s f_p B(r,s)}{N_p r} \sum_{i,j} A_{ij} \frac{T(r-r_{ij}, \sigma_{ij})}{\sigma_{ij}} \left[ \frac{(r-r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \frac{\sigma_{ij}}{\sigma'_{ij}} \frac{1}{2r_{ij}^2 \sigma'_{ij}} 2r_{ij}^l r_{ij}^m \\
\frac{\partial G_p(r,s)}{\partial U_{lm}(i)} &= \frac{f_s f_p B(r,s)}{N_p r} \sum_{i,j} A_{ij} \frac{T(r-r_{ij}, \sigma_{ij})}{r_{ij}^2 \sigma'^2_{ij}} \left[ \frac{(r-r_{ij})^2}{\sigma_{ij}^2} - 1 \right] r_{ij}^l r_{ij}^m
\end{aligned} \tag{24}$$

*Derivative with respect to  $B_{iso}$*

The derivative with respect to the  $B_{iso}(i)$  can be performed easily as:

$$\begin{aligned}
\frac{\partial G_p(r,s)}{\partial B_{iso}(i)} &= \frac{f_s f_p B(r,s)}{N_p r} \sum_{i,j} A_{ij} \frac{\partial T(r-r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} \frac{\partial \sigma_{ij}}{\partial \sigma'_{ij}} \frac{\partial \sigma'_{ij}}{\partial B_{iso}(i)} \\
\frac{\partial \sigma'_{ij}}{\partial B_{iso}(i)} &= \frac{\partial}{\partial B_{iso}(i)} \sqrt{\frac{B_{iso}(i) + B_{iso}(j)}{8\pi^2}} = \frac{1}{2\sigma'_{ij}} \frac{1}{8\pi^2} \\
\frac{\partial G_p(r,s)}{\partial B_{iso}(i)} &= \frac{f_s f_p B(r,s)}{16\pi^2 N_p r} \sum_{i,j} A_{ij} \frac{T(r-r_{ij}, \sigma_{ij})}{\sigma'^2_{ij}} \left[ \frac{(r-r_{ij})^2}{\sigma_{ij}^2} - 1 \right]
\end{aligned}$$

*Additional note taking into account the symmetry*

In this paragraph we explicitly use the symmetry constraints relating the atom  $j$  to a corresponding atom  $k$  in the asymmetric unit. In the following expression (in which we have removed the CIF label) we can write the dependency of  $\mathbf{U}(j)$  on  $\mathbf{U}(k)$ . The atom  $i$  is always in the asymmetric unit and the atom  $j$  is related to the atom  $k$  in the asymmetric unit by:

$$\mathbf{r}_j = S^{(j)} \mathbf{r}_k + \mathbf{t}^{(j)} \Rightarrow \mathbf{U}(j) = S^{(j)} \mathbf{U}(k) S^{t(j)}$$

$$\begin{aligned}
\sigma'^2_{ij} &= \frac{1}{r_{ij}^2} \mathbf{r}_{ij}^t \mathbf{G} \mathbf{N} \left\{ \mathbf{U}(i) + S^{(j)} \mathbf{U}(k) S^{t(j)} \right\} \mathbf{N} \mathbf{G} \mathbf{r}_{ij} = \frac{1}{r_{ij}^2} \mathbf{r}_{ij}^t \mathbf{U}_n(i, j) \mathbf{r}_{ij} \\
\sigma'_{ij} &= \frac{1}{r_{ij}} \sqrt{\mathbf{r}_{ij}^t \mathbf{U}_n(i, j) \mathbf{r}_{ij}} \Rightarrow \frac{\partial \sigma'_{ij}}{\partial p} = \frac{1}{2r_{ij} \sqrt{\mathbf{r}_{ij}^t \mathbf{U}_n(i, j) \mathbf{r}_{ij}}} \mathbf{r}_{ij}^t \frac{\partial \mathbf{U}_n(i, j)}{\partial p} \mathbf{r}_{ij} \\
\frac{\partial \sigma'_{ij}}{\partial p} &= \frac{1}{2r_{ij}^2 \sigma'_{ij}} \mathbf{r}_{ij}^t \frac{\partial \mathbf{U}_n(i, j)}{\partial p} \mathbf{r}_{ij} = \frac{1}{2r_{ij}^2 \sigma'_{ij}} \mathbf{r}_{ij}^t \mathbf{G} \mathbf{N} \left\{ \frac{\partial \mathbf{U}(i)}{\partial p} + S^{(j)} \frac{\partial \mathbf{U}(k)}{\partial p} S^{t(j)} \right\} \mathbf{N} \mathbf{G} \mathbf{r}_{ij} \\
\frac{\partial \sigma'_{ij}}{\partial p} &= \frac{1}{2r_{ij}^2 \sigma'_{ij}} \mathbf{r}_{ij}^t \mathbf{G} \mathbf{N} \left\{ \mathbf{M}(i) + S^{(j)} \mathbf{M}(k) S^{t(j)} \right\} \mathbf{N} \mathbf{G} \mathbf{r}_{ij}
\end{aligned}$$

Where the matrix  $\mathbf{M}$  is formed by the multipliers of the U-type, and they are of the form:

$$M_{\alpha\beta}(i) = \frac{\partial U_{\alpha\beta}(i)}{\partial p} = m_{\alpha\beta}^{(i)} \delta_{p[\alpha\beta]}$$

In which the delta function means the following:  $\delta_{p[\alpha\beta]}^{(i)}=0$  if the component  $\alpha\beta$  does not depend on  $p$  and  $\delta_{p[\alpha\beta]}^{(i)}=1$  if the component  $\alpha\beta$  depends on  $p$ . The final derivative can be written in component forms as:

$$\begin{aligned}\frac{\partial \sigma'_{ij}}{\partial p} &= \frac{1}{2r_{ij}^2 \sigma'_{ij}} \mathbf{r}_{ij}' \mathbf{G} \mathbf{N} \left\{ \mathbf{M}(i) + S^{(j)} \mathbf{M}(k) S^{t(j)} \right\} \mathbf{N} \mathbf{G} \mathbf{r}_{ij} \\ \frac{\partial \sigma'_{ij}}{\partial p} &= \frac{1}{2r_{ij}^2 \sigma'_{ij}} \sum_{\alpha, \beta} (\mathbf{r}_{ij}' \mathbf{G} \mathbf{N})_{\alpha} \left\{ M_{\alpha\beta}(i) + \sum_{\gamma, \delta} S_{\alpha\gamma}^{(j)} M_{\gamma\delta}(k) S_{\beta\delta}^{(j)} \right\} (\mathbf{N} \mathbf{G} \mathbf{r}_{ij})_{\beta}\end{aligned}$$

### Derivative of $G(r)$ with respect to the atomic positions.

The atomic coordinates appear only in the function  $T$  through  $r_{ij}$  and  $\sigma_{ij}$ . So the derivatives are calculated similarly to those of the cell parameters. If we use the symbol  $x_n(i)$  ( $n=1, 2, 3$ ) for one of the fractional coordinates of the atom  $i$ , we have:

$$\frac{\partial G_p(r, s)}{\partial x_n(i)} = \frac{f_s f_p B(r, s)}{N_p r} \sum_{i, j} A_{ij} \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial x_n(i)} \quad [25]$$

$$\begin{aligned}\frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial x_n(i)} &= \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial x_n(i)} + \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial \sigma_{ij}} \frac{\partial \sigma_{ij}}{\partial x_n(i)} \\ \frac{\partial T(r - r_{ij}, \sigma_{ij})}{\partial x_n(i)} &= \frac{T(r - r_{ij}, \sigma_{ij})}{\sigma_{ij}} \left\{ \left[ \frac{r - r_{ij}}{\sigma_{ij}} - \frac{r \sigma_{ij}}{r_{ij}^2} \right] \frac{\partial r_{ij}}{\partial x_n(i)} + \left[ \frac{(r - r_{ij})^2}{\sigma_{ij}^2} - 1 \right] \frac{\partial \sigma_{ij}}{\partial x_n(i)} \right\}\end{aligned} \quad [26]$$

The derivatives to be inserted in [25] are:

$$\frac{\partial \sigma_{ij}}{\partial x_n(i)} = \frac{\partial \sigma_{ij}}{\partial \sigma'_{ij}} \frac{\partial \sigma'_{ij}}{\partial x_n(i)} + \frac{\partial \sigma_{ij}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial x_n(i)} = \frac{\sigma_{ij}}{\sigma'_{ij}} \frac{\partial \sigma'_{ij}}{\partial x_n(i)} + \frac{\partial r_{ij}}{\partial x_n(i)} \left( \frac{\delta_1}{r_{ij}^2} + \frac{2\delta_2}{r_{ij}^3} + 2r_{ij} Q_B^2 \right) \quad [27]$$

$$\begin{aligned}\frac{\partial \sigma'_{ij}}{\partial x_n(i)} &= r_{ij} \sigma'_{ij} \frac{\partial}{\partial x_n(i)} \left( \frac{1}{r_{ij}} \right) + \frac{1}{r_{ij}} \frac{1}{2r_{ij} \sigma'_{ij}} \frac{\partial}{\partial x_n(i)} \left( \sum_{l, m} r_{ij}^l r_{ij}^m \{U_{lm}(i) + U_{lm}(j)\} \right) \\ \frac{\partial \sigma'_{ij}}{\partial x_n(i)} &= -\frac{\sigma'_{ij}}{r_{ij}} \frac{\partial r_{ij}}{\partial x_n(i)} + \frac{a_n}{r_{ij}^2 \sigma'_{ij}} \sum_m a_m d_{ij}^m \{U_{nm}(i) + U_{nm}(j)\}\end{aligned} \quad [28]$$

Finally the only unknown we have to calculate is the derivative of  $r_{ij}$ :

$$r_{ij} = \sqrt{\mathbf{d}_{ij}^T \mathbf{G} \mathbf{d}_{ij}} \Rightarrow \frac{\partial r_{ij}}{\partial x_n(i)} = \frac{1}{2r_{ij}} \left\{ \frac{\partial \mathbf{d}_{ij}^T}{\partial x_n(i)} \mathbf{G} \mathbf{d}_{ij} + \mathbf{d}_{ij}^T \mathbf{G} \frac{\partial \mathbf{d}_{ij}}{\partial x_n(i)} \right\}$$

$$\frac{\partial \mathbf{d}_{ij}}{\partial x_n(i)} = \frac{\partial}{\partial x_n(i)} [x_1(j) - x_1(i), x_2(j) - x_2(i), x_3(j) - x_3(i)] = \begin{cases} [-1, 0, 0] & \text{for } x_1(i) \\ [0, -1, 0] & \text{for } x_2(i) \\ [0, 0, -1] & \text{for } x_3(i) \end{cases}$$

Explicitly we have:

$$\begin{aligned} \frac{\partial r_{ij}}{\partial x_1(i)} &= -\frac{1}{r_{ij}} (a^2 d_{ij1} + ab \cos \gamma d_{ij2} + ac \cos \beta d_{ij3}) \\ \frac{\partial r_{ij}}{\partial x_2(i)} &= -\frac{1}{r_{ij}} (b^2 d_{ij2} + ab \cos \gamma d_{ij1} + bc \cos \alpha d_{ij3}) \\ \frac{\partial r_{ij}}{\partial x_3(i)} &= -\frac{1}{r_{ij}} (c^2 d_{ij3} + ac \cos \beta d_{ij1} + bc \cos \alpha d_{ij2}) \end{aligned} \quad [29]$$

Inserting [27] into [26] and then [26] and [27] into [25], etc., then we finally can calculate the final expression given by [23].

*Additional note taking into account the symmetry*

In this paragraph we explicitly use the symmetry constraints relating the atom  $j$  to a corresponding atom  $k$  in the asymmetric unit. The vector  $\mathbf{d}_{ij}$  can be written explicitly as:

$$\begin{aligned} r_{ij} &= \sqrt{\mathbf{d}_{ij}^T \mathbf{G} \mathbf{d}_{ij}} \Rightarrow \mathbf{d}_{ij} = \mathbf{r}_j - \mathbf{r}_i = S^{(j)} \mathbf{r}_k + \mathbf{t}^{(j)} - \mathbf{r}_i \\ \frac{\partial r_{ij}}{\partial p} &= \frac{1}{2r_{ij}} \left( \frac{\partial \mathbf{d}_{ij}^T}{\partial p} \mathbf{G} \mathbf{d}_{ij} + \mathbf{d}_{ij}^T \mathbf{G} \frac{\partial \mathbf{d}_{ij}}{\partial p} \right) \\ \frac{\partial \mathbf{d}_{ij}}{\partial p} &= S^{(j)} \frac{\partial \mathbf{r}_k}{\partial p} - \frac{\partial \mathbf{r}_i}{\partial p} \Rightarrow \frac{\partial d_{ij\alpha}}{\partial p} = \sum_{\beta} S_{\alpha\beta}^{(j)} \frac{\partial x_{k\beta}}{\partial p} - \frac{\partial x_{i\alpha}}{\partial p} \end{aligned}$$

It is more convenient to write the last equation in components, as shown. Whatever coordinate  $(x, y, z)$ , may depend on the free parameter  $p$ . The crystallographic constraints are linear in the parameter  $p$ , so that  $x_{\alpha} = x_{0\alpha} + m_{\alpha} p \delta_{\alpha p}$ , where  $x_{0\alpha}$  is a constant,  $\delta_{\alpha p}$  is zero if there is no dependency on  $p$  or  $\delta_{\alpha p} = 1$  if  $x_{\alpha}$  effectively depends on  $p$ , and  $m_{\alpha}$  is called the multiplier, which is generally equal to 1, 2 or to a simple rational number. The last derivative can be written as:

$$\frac{\partial d_{ij\alpha}}{\partial p} = \sum_{\beta} S_{\alpha\beta}^{(j)} \frac{\partial x_{k\beta}}{\partial p} - \frac{\partial x_{i\alpha}}{\partial p} = \sum_{\beta} S_{\alpha\beta}^{(j)} m_{\beta}^{(k)} \delta_{p\beta}^{(k)} - m_{\alpha}^{(i)} \delta_{p\alpha}^{(i)}$$

For instance if the position of one atom is of the form  $\mathbf{r} = (x, y, z) = (p, 2p, z)$ , the derivative with respect to  $p$  will be (1, 2, 0). For a position of the form  $\mathbf{r} = (x, y, z) = (p, -p+1/4, z)$ , the derivative with respect to  $p$  will be (1, -1, 0).

### Derivative of $G(r)$ with respect to the site occupancies.

The site occupancy for a given phase appears in the terms  $A_{ij}$ ,  $N_p$  and  $\rho_0$ , so we have:

$$\frac{\partial G_p(r, s)}{f_p \partial c_n} = \frac{\partial}{\partial c_n} \left( \frac{1}{N_p r} \sum_{i,j} A_{ij} T(r - r_{ij}, \sigma_{ij}) - 4\pi r \frac{N_p}{V_p} \right)$$

$$\frac{\partial G_p(r, s)}{f_p \partial c_n} = -\frac{1}{N_p^2 r} \frac{\partial N_p}{\partial c_n} \sum_{i,j} A_{ij} T_{ij} + \frac{1}{N_p r} \sum_{i,j} T_{ij} \frac{\partial A_{ij}}{\partial c_n} - \frac{4\pi r}{V_p} \frac{\partial N_p}{\partial c_n}$$

Taking into account the expressions [5] we obtain:

$$\frac{\partial G_p(r, s)}{f_p \partial c_n} = \frac{1}{N_p r} \left[ \frac{1}{c_n} \sum_j A_{nj} T_{nj} - \frac{1}{N_p} \sum_{i,j} A_{ij} T_{ij} \right] - \frac{4\pi r}{V_p} \quad [30]$$

### Spherical and platelet-like nanoparticles

In nanoparticles the correlation in  $G(r)$  at distances that are larger than the size of the particles disappears. The effect in  $G(r)$  is that the intensity of peaks at high  $r$  is lower than the calculated  $G_\infty(r)$  of the bulk. The observed  $G(r)$  is obtained from  $G_\infty(r)$  by multiplying by a form factor  $f(r)$ :  $G(r) = f(r) G_\infty(r)$ .

The form factor for different kind of nanoparticles has been calculated by K.Kodama, S. Iikubo, T. Taguchi and S. Shamoto (*Acta Cryst.* **A62**, 444–453 (2006)). Here we provide the form factors for two kinds of nano particles, sheets of thickness  $t$  and spheres of diameter  $d$ :

$$f_{sph}(r) = \begin{cases} \frac{1}{2} \left( \frac{r}{d} \right)^3 - \frac{3}{2} \frac{r}{d} + 1 & r \leq d \\ 0 & r > d \end{cases} \quad f_{sheet}(r) = \begin{cases} 1 - \frac{r}{2t} & r \leq t \\ \frac{t}{2r} & r > t \end{cases} \quad [31]$$

For a distribution of diameters the following general expressions have been derived by R.C. Howell, T. Proffen and S.D. Conradson (*Phys.Rev.* **B73**, 094107 (2006)):

$$f_{DE}(r) = \int_0^\infty f_{sph}(r, d') P(d') dd' \quad P(d', D, n) = \frac{1}{n! D} \left( \frac{d'}{D} \right)^n e^{-d'/D} \quad [32]$$

The average diameter  $d$  and width of the distribution are related to the parameters  $(n, D)$  as:

$$d = \langle d' \rangle = \int_0^{\infty} d' P(d') dd' = (n+1)D \quad \sigma = \sqrt{\langle d'^2 \rangle - d^2} = \frac{d}{\sqrt{n+1}} \quad [33]$$

Using the above equations one can derive a closed form for  $n \geq 3$ :

$$f_{DE}(r, D, n) = e^{-r/D} \sum_{k=0}^{n-2} \frac{1}{k!} \left( 1 - \frac{3k}{2n} + \frac{k(k-1)(k-2)}{2n(n-1)(n-2)} \right) \left( \frac{r}{D} \right)^k \quad [34]$$

The function  $f_{DE}(r, D, n)$  should be multiplied by  $G_{\infty}(r)$  for obtaining the observed  $G(r)$  before convoluting with the sinc function (see equations [2, 4]). The user should select the value of  $n$ , and  $D$  can be refined for obtaining the average diameter and widths from equations [33]. The derivative of [34] with respect to  $D$  is:

$$\frac{\partial f_{DE}(r, D, n)}{\partial D} = \frac{r}{D^2} f_{DE}(r, D, n) - \frac{e^{-r/D}}{D} \sum_{k=1}^{n-2} \frac{1}{(k-1)!} \left( 1 - \frac{3k}{2n} + \frac{k(k-1)(k-2)}{2n(n-1)(n-2)} \right) \left( \frac{r}{D} \right)^k \quad [35]$$

The general expression of  $G(r)$  is

$$G(r, s) = f_s B(r, s) \sum_{p=1}^P f_p f_{pDE}(r, D_p, n_p) G_p(r, s) \quad [36]$$

The derivative of  $G(r)$  with respect to  $D_p$  is given by the expression:

$$\frac{\partial G(r, s)}{\partial D_p} = f_s B(r, s) \sum_{p=1}^P f_p \frac{\partial f_{pDE}(r, D_p, n_p)}{\partial D_p} G_p(r, s) \quad [37]$$

### The use of restraints together with the refinement of $G(r)$

For molecular systems the (internal) interatomic distances and bond angles are practically fixed even in the case of amorphous or nanocrystalline materials. Only the torsion angles giving rise to different conformations of a same molecule change with the state of the sample. Crystalline molecular systems are mostly or relatively low symmetry and the number of free parameters may be much higher than those corresponding to ionic or metallic systems. The refinement of a large number of parameters is unstable and if not constrained (or restraints) are applied, the molecule is unphysically distorted. The use of restraints does not diminish the number of free parameters it only increases the number of observations in a very convenient way. The cost function to be minimized is now the following:

$$\chi^2 = \sum_r \left[ \frac{G_{obs}(r) - G_{calc}(r, \beta)}{\sigma(r)} \right]^2 + \sum_{i=1}^{N_{rest}} \left[ \frac{V_{obs}(i) - V_{calc}(i, \alpha)}{\sigma(i)} \right]^2 \quad [38]$$

Where  $V_{obs}(i)$  is the value of the prescribed variable (interatomic distance or bond angle) for observation  $i$  and  $V_{calc}(i, \alpha)$  is the corresponding calculated value that depends on a subset  $\alpha$  of the parameter vector  $\beta$ . The standard deviation  $\sigma(i)$  is provided by the user and it is a measure of the strength of the prescribed restraint. The smaller the value of  $\sigma(i)$ , the stronger the weight of the restraint. The most important restraints in molecular systems are the internal (inside the molecule) distances and bond angles. Other restraints may be: anti-bump prescriptions that avoid the interpenetration of atoms of the same or different molecules, planarity of group of atoms, special torsion angles, etc.

Presently, the only restraints fully implemented concern bond-distances and bond-angles.

(To be completed)