# 8.3. Constraints and restraints in refinement

By E. Prince, L. W. Finger, and J. H. Konnert

In Chapter 8.1, the method of least squares is discussed as a technique for fitting a theoretical model that contains adjustable parameters to a set of observations. The discussion is very general and contains very little mention of what sorts of quantities the observations are or what the model represents. In crystallography, the model is a crystal, which is constructed from identical unit cells that contain atoms, and which diffracts X-rays, neutrons or electrons in a manner that is characteristic of the arrangement of those atoms. The sample may be either a single crystal or a polycrystalline powder, and the observations are diffracted intensities, which may be fitted directly, as in the Rietveld method for powders (see Chapter 8.6; also Rietveld, 1969), or converted to derived quantities such as integrated intensities, squared moduli of structure amplitudes, or the structure amplitudes themselves. The model generally contains a scale factor and may contain parameters describing other experimental effects, such as extinction. Each atom in the unit cell requires three parameters to describe its mean position and various parameters to describe random deviations from that position owing to thermal motion or disorder. Models that treat each atom independently, however, do not allow for the fact that a great deal more is known about a crystal initially than simply its chemical composition. Atoms have fairly definite sizes and tend to occupy sites whose surroundings conform to a rather limited set of common configurations. In this chapter, we discuss ways of using this additional information. First, we shall discuss the use of constraints to reduce the number of parameters that must be varied and account for relationships among parameters that are dictated by the laws of chemistry and physics. Then we shall discuss the use of restraints, which effectively add to the number of observations that must be fitted by the model.

## 8.3.1. Constrained models

The techniques of least squares are applicable for refining almost any model, but the question of the suitability of the model remains. The addition of parameters may reduce the residual disagreement, but lead to solutions that have no physical or chemical validity. Addition of constraints is one method of constricting the solutions.

### 8.3.1.1. *Lagrange undetermined multipliers*

The classical technique for application of constraints is the use of Lagrange undetermined multipliers, in which the set of $p$ parameters, $x_j$, is augmented by $p-q$ $(q<p)$ additional unknowns, $\lambda_k$, one for each constraint relationship desired. The problem may be stated in the form: find the minimum of

$$S=\sum_{i=1}^{n} w_i[y_i-M_i(\mathbf{x})]^2, \tag{8.3.1.1a}$$

subject to the condition

$$f_k(\mathbf{x})=0 \quad (k=1,2,\ldots,p-q). \tag{8.3.1.1b}$$

This may be shown (Gill, Murray & Wright, 1981) to be equivalent to the problem: find a point at which the gradient of

$$S'=\sum_{i=1}^{n} w_i[y_i-M_i(\mathbf{x})]^2+\sum_{k=1}^{p-q}\lambda_k f_k(\mathbf{x}) \tag{8.3.1.2}$$

vanishes. Solving for the stationary point leads to a set of simultaneous equations of the form

$$\partial S'/\partial x_j=\partial S/\partial x_j+\sum_{k=1}^{p-q}\lambda_k\partial f_k(\mathbf{x})/\partial x_j=0 \tag{8.3.1.3a}$$

and

$$\partial S'/\partial\lambda_k=f_k(\mathbf{x})=0. \tag{8.3.1.3b}$$

Thus, the number of equations, and the number of unknowns, is increased from $p$ to $2p-q$. In cases where the number of constraint relations is small, and where it may be difficult to solve the relations for some of the parameters in terms of the rest, this method yields the desired results without too much additional computation (Ralph & Finger, 1982). With the large numbers of parameters, and large numbers of constraints, that arise in many crystallographic problems, however, the use of Lagrange multipliers is computationally inefficient and cumbersome.

### 8.3.1.2. *Direct application of constraints*

In most cases encountered in crystallography, constraints may be applied directly, thus reducing rather than increasing the size of the normal-equations matrix. For each constraint introduced, one of the parameters becomes dependent on the remaining set, and the rank of the remaining system is reduced by one. For $p$ parameters and $p-q$ constraints, the problem reduces to $q$ parameters. If the Gauss–Newton algorithm is used (Section 8.1.4), the normal-equations matrix is $\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A}$, where

$$A_{ij}=\partial M_i/\partial x_j, \tag{8.3.1.4}$$

and $\boldsymbol{W}$ is a weight matrix. A constrained model, $M_i(\mathbf{z})$, maybe constructed using relations of the form

$$x_j=g_j(z_1,z_2,\ldots,z_q). \tag{8.3.1.5}$$

Applying the chain rule for differentiation, the normal-equations matrix for the constrained model is $\boldsymbol{B}^T\boldsymbol{W}\boldsymbol{B}$, where

$$B_{ik}=\partial M_i(\mathbf{x})/\partial z_k=\sum_{j=1}^{p}[\partial M_i(\mathbf{x})/\partial x_j](\partial x_j/\partial z_k). \tag{8.3.1.6}$$

This may be written in matrix form $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{C}$, where $C_{jk}=\partial x_j/\partial z_k$ defines a $p\times q$ constraint matrix. The application of constraints involves (*a*) determination of the model to be used, (*b*) calculation of the elements of $\boldsymbol{C}$, and (*c*) computation of the modified normal-equations matrix.

The construction of matrix $\boldsymbol{C}$ by a procedure known as the *variable reduction method* may be presented formally as follows: Designate by $\boldsymbol{Z}$ the matrix whose elements are

$$Z_{jk}=\partial g_j(\mathbf{x})/\partial x_k, \tag{8.3.1.7}$$

and partition $\boldsymbol{Z}$ in the form $\boldsymbol{Z}=(\boldsymbol{U},\boldsymbol{V})$, where $\boldsymbol{V}$ is composed of $(p-q)$ columns of $\boldsymbol{Z}$ chosen to be linearly independent, so that $\boldsymbol{V}$ is nonsingular. [$\boldsymbol{V}$ is shown as the last $(p-q)$ columns only for convenience. Any linearly independent set may be chosen.] The rows of $\boldsymbol{Z}$ form a basis for a $(p-q)$-dimensional subspace of the $p$-dimensional parameter space, and we wish to construct a basis for $\mathbf{z}$, a $q$-dimensional subspace that is orthogonal to it, so that all shifts within *that* subspace starting at a point where the constraints are satisfied, a *feasible point*, leave the values of the constraint relations unchanged. This basis is used for the columns of $\boldsymbol{C}$, which is given by

$$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{I}_q\\ -\boldsymbol{V}^{-1}\boldsymbol{U}\end{pmatrix}. \tag{8.3.1.8}$$





references