$$E = w_X \sum_{i=1}^{n} (\Delta_X^i)^2 + w_Y \sum_{i=1}^{n} (\Delta_Y^i)^2 + w_Z \sum_{i=1}^{n} (\Delta_Z^i)^2.$$

The residuals between the calculated (X_i, Y_i, Z_i) and observed spot centroids are

$$\Delta_X^i = X_i - X_i' = X_0 + F\mathbf{S}_i \cdot \mathbf{d}_1 / \mathbf{S}_i \cdot \mathbf{d}_3 - X_i'$$
$$\Delta_Y^i = Y_i - Y_i' = Y_0 + F\mathbf{S}_i \cdot \mathbf{d}_2 / \mathbf{S}_i \cdot \mathbf{d}_3 - Y_i'$$
$$\Delta_Z^i = Z_i - Z_i' = \varphi_0 + \Delta_\varphi \sum_{j=-\infty}^{\infty} (j - 1/2) R_j^i - Z_i'.$$

Let s_μ $(\mu = 1, \ldots, k)$ denote the k independent parameters for which initial estimates are available. Expanding the residuals to first order in the parameter changes δs_μ gives

$$\Delta(s_\mu + \delta s_\mu) \approx \Delta(s_\mu) + \sum_{\mu=1}^{k} \frac{\partial \Delta}{\partial s_\mu} \delta s_\mu.$$

The parameters should be changed in such a way as to minimize $E(\delta s_\mu)$, which implies $\partial E / \partial \delta s_\mu = 0$ for $\mu = 1, \ldots, k$. The δs_μ are found as the solution of the k normal equations

$$\sum_{\mu'=1}^{k} \left(w_X \sum_{i=1}^{n} \frac{\partial \Delta_X^i}{\partial s_\mu} \frac{\partial \Delta_X^i}{\partial s_{\mu'}} + w_Y \sum_{i=1}^{n} \frac{\partial \Delta_Y^i}{\partial s_\mu} \frac{\partial \Delta_Y^i}{\partial s_{\mu'}} + w_Z \sum_{i=1}^{n} \frac{\partial \Delta_Z^i}{\partial s_\mu} \frac{\partial \Delta_Z^i}{\partial s_{\mu'}} \right) \delta s_{\mu'}$$
$$= - \left(w_X \sum_{i=1}^{n} \Delta_X^i \frac{\partial \Delta_X^i}{\partial s_\mu} + w_Y \sum_{i=1}^{n} \Delta_Y^i \frac{\partial \Delta_Y^i}{\partial s_\mu} + w_Z \sum_{i=1}^{n} \Delta_Z^i \frac{\partial \Delta_Z^i}{\partial s_\mu} \right).$$

The parameters are corrected by δs_μ and a new cycle of refinement is started until a minimum of E is reached. The weights

$$w_X = 1 / \sum_{i=1}^{n} (\Delta_X^i)^2, \quad w_Y = 1 / \sum_{i=1}^{n} (\Delta_Y^i)^2, \quad w_Z = 1 / \sum_{i=1}^{n} (\Delta_Z^i)^2$$

are calculated with the current guess for s_μ at the beginning of each cycle.

The derivatives appearing in the normal equations can be worked out from the definitions given in Sections 11.3.2.2 and 11.3.2.4, and only the form of the gradient of the Z residuals is shown. Assuming $\sigma_i = \sigma_M / |\zeta_i|$ $(i = 1, \ldots, n)$ is constant for each reflection, the gradients of the Z residuals are obtained from the chain rule and the relation $\mathrm{d}\,\mathrm{erf}(z)/\mathrm{d}z = [2/(\pi)^{1/2}] \exp(-z^2)$.

$$\frac{\partial \Delta_Z^i}{\partial s_\mu} = \frac{\partial \Delta_Z^i}{\partial \varphi_i} \frac{\partial \varphi_i}{\partial s_\mu}$$
$$\frac{\partial \Delta_Z^i}{\partial \varphi_i} = \frac{\Delta_\varphi}{(2\pi)^{1/2} \sigma_i} \sum_{j=-\infty}^{\infty} \exp[-(\varphi_0 + j\Delta_\varphi - \varphi_i)^2 / 2\sigma_i^2]$$
$$\frac{\partial \varphi_i}{\partial s_\mu} = \cos\varphi_i \frac{\partial \sin\varphi_i}{\partial s_\mu} - \sin\varphi_i \frac{\partial \cos\varphi_i}{\partial s_\mu}.$$

Obviously, $\partial \Delta_Z^i / \partial s_\mu$ is small for a fully recorded reflection because of the small values of all exponentials appearing in $\partial \Delta_Z^i / \partial \varphi_i$. In contrast, the gradient for a partial reflection, equally recorded on two adjacent images, is most sensitive to parameter variations because one of the exponentials assumes its maximum value. In the limiting case of infinitely fine-sliced data, it can be shown that $\lim_{\Delta_\varphi \to 0} \partial \Delta_Z^i / \partial \varphi_i = 1$. Thus, the refinement scheme based on observed Z centroids, as described here and implemented in *XDS*, is applicable to fine-sliced data – and to data recorded with a large oscillation range as well.

11.3.3. Integration

A fundamental requirement for a general integration method is that it should distinguish carefully between signal and background points within its integration domain. For weak reflections, this distinction cannot be made reliably because of the errors superimposed on the signal. The problem can be solved, however, provided that both weak and strong reflections share the same profile shape – an assumption that has been adopted by most data-processing packages.

The intensity distribution of a reflection can be modelled analytically or derived from the observed profiles of neighbouring strong spots. For the rotation method, the profile shape depends strongly on the specific path of the reflection through the Ewald sphere and on variations in the angle of incidence of the diffracted beam on a flat detector. These geometrical distortions can be eliminated by mapping the reflections onto the coordinate system defined in Section 11.3.2.3, which simplifies the task of modelling the expected intensity distribution as all reflection profiles become similar.

11.3.3.1. *Spot extraction*

The region around a spot is defined by the two parameters δ_D and δ_M, which represent spot diameter and reflecting range, respectively. It is assumed that the coordinates of all image pixels contributing to the intensity of a spot satisfy $|\varepsilon_1| \leq \delta_D/2, |\varepsilon_2| \leq \delta_D/2$ and $|\varepsilon_3| \leq \delta_M/2$ when mapped to the profile coordinate system $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ defined in Section 11.3.2.3. Regions of neighbouring reflections may overlap. As implemented in *XDS*, potential overlap is dealt with by a simple strategy: pixels within the overlap region are assigned to the nearest spot. This is carried out in two steps. First, reflections predicted to occur on a given rotation image are found by generating and testing all possible indices h, k, l up to the highest resolution recorded by the detector. Reflection indices, coordinates of the diffracted beam wave vector and the expected fraction of spot intensity recorded on the image are saved in a table. In the second step, each reflection boundary is traced in the image and corrected to exclude pixels belonging to overlapping reflections, which are rapidly located in the table by the hash technique. The image scaling factor obtained from the mean image background and the neighbourhood pixel values belonging to the reflections recorded in the image are saved on a scratch file dedicated to the currently processed data image.

At regular intervals, these files are merged such that all pixel values belonging to a spot found in the contributing images follow each other. Reflections for which contributing pixels are expected further ahead in data processing are just copied to a scratch output file. The other reflections are mapped to the Ewald sphere, as described below, and their three-dimensional profiles and accompanying information are routed to the main output file of the spot-extraction step. After the file-merging procedure, spot extraction continues.

11.3.3.2. *Background*

The region around a spot is assumed to have been chosen to be large enough to include a sufficient number of pixels which can be used for determination of the background. Background determination, as implemented in *XDS*, begins by sorting all pixels belonging to a reflection by increasing intensity. For weak or absent reflections, these values should represent a random sample drawn from a normal distribution. If this is not the case, the pixel with the largest intensity is removed until the sampling distribution of the remaining smaller items satisfies the expected distribution. This method will also exclude pixels with unexpectedly high values, such as ice reflections. The background, determined as the mean value of the accepted pixels, is systematically overestimated for strong spots because of some residual intensity extending into the accepted background pixels. This residual intensity is estimated

references