Table 3.4.1
Expressions for $Q(hkl)$ for different types of symmetry

Symmetry	$Q(hkl)$
Cubic	$(h^2 + k^2 + l^2)A_{11}$
Tetragonal	$(h^2 + k^2)A_{11} + l^2A_{33}$
Hexagonal	$(h^2 + hk + k^2)A_{11} + l^2A_{33}$
Orthorhombic	$h^2A_{11} + k^2A_{22} + l^2A_{33}$
Monoclinic	$h^2A_{11} + k^2A_{22} + l^2A_{33} + hlA_{13}$
Triclinic	$h^2A_{11} + k^2A_{22} + l^2A_{33} + hkA_{12} + hlA_{13} + klA_{23}$

diffraction angles by

$$d_{hkl} = \lambda/(2\sin\theta_{hkl}).$$

d_{hkl}, the spacing of the planes (hkl) in the direct lattice, is obtained by the dot products of the reciprocal-lattice vectors with themselves:

$$(\mathbf{r}^*_{hkl})^2 = \frac{1}{d^2_{hkl}} = h^2a^{*2} + k^2b^{*2} + l^2c^{*2} + 2hka^*b^*\cos\gamma^* + 2hla^*c^*\cos\beta^* + 2klb^*c^*\cos\alpha^*, \quad (3.4.1)$$

where α^* is the angle between $\mathbf{b}^*$ and $\mathbf{c}^*$, β^* is the angle between $\mathbf{c}^*$ and $\mathbf{a}^*$, and γ^* is the angle between $\mathbf{a}^*$ and $\mathbf{b}^*$. If we introduce

$$Q(hkl) = \frac{10^4}{d^2_{hkl}}$$

[where $Q(hkl)$ differs from $\sin^2\theta_{hkl}$ by a scale factor of $(200/\lambda)^2$], (3.4.1) becomes

$$Q(hkl) = h^2A_{11} + k^2A_{22} + l^2A_{33} + hkA_{12} + hlA_{13} + klA_{23}, \quad (3.4.2)$$

where

$$A_{11} = 10^4a^{*2},\ A_{22} = 10^4b^{*2},\ A_{33} = 10^4c^{*2},$$
$$A_{12} = 2\times10^4a^*b^*\cos\gamma^*,\ A_{13} = 2\times10^4a^*c^*\cos\beta^*,$$
$$A_{23} = 2\times10^4b^*c^*\cos\alpha^*.$$

The number of parameters A_{ij} in (3.4.2) depends on the type of symmetry: from 1 in the case of cubic symmetry to 6 for triclinic symmetry (see Table 3.4.1).

The quadratic form (3.4.2) relates the observed $Q(hkl)$ values to the reciprocal cell parameters and, consequently, to the direct cell. It is the basic equation used in powder-indexing procedures. Therefore the indexing problem (Werner, 2002) is to find A_{ij} and, for each observed $Q(hkl)$ value, three crystallographic indices (hkl) satisfying (3.4.2) within a suitable tolerance parameter Δ:

$$Q(hkl) - \Delta < h^2A_{11} + k^2A_{22} + l^2A_{33} + hkA_{12} + hlA_{13} + klA_{23} < Q(hkl) + \Delta. \quad (3.4.3)$$

The importance of using accurate $Q(hkl)$ values in (3.4.3) is obvious. Moreover, it is worth noticing that (3.4.3) must lead to physically reasonable indexing – low-angle peaks should correspond to small integer values for h, k and l and the values of the cell parameters and cell volume should be reasonable.

3.4.2.1. *Figures of merit*

An important task is the introduction of a figure of merit (FOM) that is able to (*a*) describe the physical plausibility of a trial cell and its agreement with the observed pattern, and (*b*) select the best cell among different possible ones. de Wolff (1968) made an important contribution in this direction. He developed the M_{20} figure of merit defined by

$$M_{20} = \frac{Q_{20}}{2\langle\varepsilon\rangle N_{20}}, \quad (3.4.4)$$

where Q_{20} is the Q value corresponding to the 20th observed and indexed peak, N_{20} is the number of different calculated Q values up to Q_{20}, and $\langle\varepsilon\rangle$ is the average absolute discrepancy between the observed and the calculated Q values for the 20 indexed peaks; the factor 2 is a result of statistics, explained by the larger chance for an observed line to sit in a large interval as compared with sitting in a small interval. The rationale behind M_{20} is as follows: the better the agreement between the calculated and the observed peak positions (the smaller the $\langle\varepsilon\rangle$ value) and the smaller the volume of the unit cell (the smaller the N_{20} value), the larger the M_{20} value and consequently the confidence in the proposed unit cell. A rule of thumb for M_{20} is that if the number of unindexed peaks whose Q values are less than Q_{20} is not larger than 2 and if $M_{20} > 10$, then the indexing process is physically reasonable (de Wolff, 1968; Werner, 2002). This rule is often valid, but exceptions occur. The use of the first 20 peaks is a compromise (coming from experience) between introducing a quite large number of observed peaks (depending on the number of parameters of the unit cell) and avoiding the use of high-angle peak positions, which are more affected by errors. M_{20} is statistically expected to be 1 in case of completely arbitrary indexing. It has no upper limit (it can be very large when $\langle\varepsilon\rangle$ is very small).

Smith & Snyder (1979) proposed the F_N criterion in order to overcome the limits of M_{20} with respect to its dependence on the 20 lines and on crystal class and space group. The F_N figure of merit is given by

$$F_N = \frac{1}{\langle|\Delta2\theta|\rangle}\frac{N}{N_{\rm poss}},$$

where $\langle|\Delta2\theta|\rangle$ is the average absolute discrepancy between the observed and calculated 2θ peak position values and $N_{\rm poss}$ is the number of possible diffraction lines up to the Nth observed line. The values of $\langle|\Delta2\theta|\rangle$ and $N_{\rm poss}$, $(\langle|\Delta2\theta|\rangle, N_{\rm poss})$, are usually given with F_N. With respect to M_{20}, F_N is more suitable for ranking the trial solutions and less for indicating their physical plausibility (Werner, 2002).

Both M_{20} and F_N, being based on the discrepancies between observed and calculated lines, are less reliable if there are impurity peaks; if the information about the unindexed lines is not taken into account, the risk of obtaining false solutions increases. Alternative FOMs based on joint probability have also been proposed (Ishida & Watanabe, 1967, 1971). Among the recently developed FOMs, we mention:

(1) $Q_{\rm partial}$ (Bergmann, 2007):

$$Q_{\rm partial} = \sum_i \min\left[w_i, \left(\frac{x_i - \hat{x}_i}{\delta_i}\right)^2\right],$$

where the summation is over the number of observed lines, w_i is the observed weight of line i, x and $\hat{x}_i$ are the observed and simulated line positions, respectively, and δ_i is the observed random error of line i. $Q_{\rm partial}$ is multiplied by a factor that depends on the symmetry of the simulated lattice (triclinic, ..., cubic), on the unit cell volume and on the number of ignored peaks.

(2) McM_{20} (Le Bail, 2008):

$$McM_{20} = [100/(R_pN_{20})]B_rS_y,$$

where N_{20} is the number of possible lines that might exist up to the 20th observed line (for a primitive P lattice). R_p is