Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 3 November 2008
Accepted 30 July 2009

Dedicated to our friend and colleague Professor Dr Armin Kirfel (Bonn) on the occasion of attaining the status Emeritus Professor

Structure determination without Fourier inversion. V. A concept based on parameter space

Helmuth Zimmermann[a] **and Karl F. Fischer**[b]*

[a]Lehrstuhl für Kristallographie und Strukturphysik, Universität Erlangen-Nürnberg, Staudtstrasse 3, 91058 Erlangen, Germany, and [b]Technische Physik der Universität des Saarlandes, 66041 Saarbrücken, Germany. Correspondence e-mail: karl.fischer@mx.uni-saarland.de

The parameter-space concept for solving crystal structures from reflection amplitudes (without employing or searching for their phases) is described on a theoretically oriented basis. Emphasis is placed on the principles of the method, on selecting one of three types of parameter spaces discussed in this paper, and in particular on the structure model employed (equal-atom point model, however usually reduced to one-dimensional projections) and on the system of 'isosurfaces' representing experimental 'geometrical structure amplitudes' in an orthonormal parameter space of as many dimensions as unknown atomic coordinates. The symmetry of the parameter space as well as of the imprinted isosurfaces and its effect on solution methods is discussed. For point atoms scattering with different phases or signs (as is possible in the case of X-ray resonant or of neutron scattering) it is demonstrated that the 'landscape' of these isosurfaces remains invariant save certain shifts of origin known beforehand (under the condition that all atomic scattering amplitudes have been reduced to 1 thus meeting the requirement of the structure model above). Partly referring to earlier publications on the subject, measures are briefly described which permit circumventing an analytical solution of the system of structure-amplitude equations and lead to either a unique (unequivocal) approximate structure solution (offering rather high spatial resolution) or to all possible solutions permitted by the experimental data used (thus including also all potential 'false minima'). A simple connection to Patterson vectors is given, also a first hint on data errors. References are given for practical details of various solution techniques already tested and for reconstruction of three-dimensional structures from their projections by 'point tomography'. We would feel foolish if we tried to aim at any kind of 'competition' to existing methods. Having mentioned 'pros and cons' of our concept, some ideas about potential applications are nevertheless offered which are mainly based on its inherent resolution power though demanding rather few reflection data (use of optimal intensity contrast included) and possibly providing a result proven to be unique.

1. Introduction

A procedure for the determination of a crystal structure shall be described which avoids the intermediate step of a Fourier synthesis of an approximate electron density and instead attempts to obtain the m_u atomic coordinates $X_j = (x_j, y_j, z_j)^t$, $j = 1, \ldots, m_u$ directly from the structure amplitudes. The solution of the phase problem thus becomes obsolete. Such a procedure may be helpful in cases where only a restricted number of Fourier coefficients is available by means of experimental limitations; their Fourier sum gives no sufficient approximation of the scattering density or if the local resolution of the density map is too poor or if multiple solutions have to be explored and discussed ('homometry', 'false minima'). Procedures of this kind already exist since long ago. Extensive citations may be found in the references of Fischer *et al.* (2005) (hereafter FKZ I). Closest to the idea presented here is the proposal by Navaza & Silva (1979) discussed in short in §3.1.

There is apparently no urgent need of a 'new' concept for crystal structure determination thanks to the overwhelming success of the 'direct methods' and the perfectly working program systems based on them which permit routinely assessing crystal structures. However, any solution principle based on Fourier inversion sometimes leaves unsatisfactory answers to one or more questions. In particular, owing to the strategy of direct methods, there is generally no guarantee that a solution found is unique (unequivocal). Also the spatial resolution [notoriously limited by $|\mathbf{h}|_{\max} = (2\sin\theta/\lambda)_{\max}$, *i.e.* series-termination errors plus errors in the data] may be insufficient, *e.g.* in some cases of pseudo-symmetry. Searching for a method providing unique solution and higher resolution (perhaps with even fewer data) led one of us to refer to the concept of Ott (1927) and to develop it a little further on the basis of an equal-atom point structure (Knof, 1989; Pilz, 1996; Fischer & Pilz, 1997; Pilz & Fischer, 1998, 2000; see also §3.2). As expected, both higher resolution and unique results could be attained (in principle as well as in practice) by solving for the roots of a polynomial; however, this is at the expense of a sincere drawback caused by three 'weak points':

(i) The coefficients of the polynomial contain powers of point structure amplitudes increasing up to the number of unknown parameters, thus becoming sensitive to all kinds of data errors for structures with even moderate numbers of atoms.

(ii) For establishing the polynomial, the first reflection order amplitude is essential.

(iii) The determination of signs (or phases) could become uncertain for problems with many atoms (although in these cases all possible sign combinations are found at the same time).

Looking around for a concept using all data at the first power only, the parameter-space idea was born.

From amplitudes alone, however, merely interatomic vectors can be determined, at least in principle, *i.e.* one is restricted to a 'kind of Patterson' solution type. How this problem is overcome will also be discussed.

In this paper the task of structure solution is simplified to equal-point-atom structures (§2 and §3) and two ideas are combined: the orthonormal 'parameter space' of the vectors of the independent variables with the conventional scalar product in $\mathbf{R}^m$ (§4.1), and the 'isosurfaces' (§4.2). Properties and applications are discussed in §4–§8. In §4.3 we shall relax a few of the above restrictions.

Some ideas of this concept including examples of practical solution techniques have either been published before, *e.g.* FKZ I, Kirfel *et al.* (2006) (herafter FKZ II), Fischer *et al.* (2008) (hereafter FKZ III) and Kirfel & Fischer (2009) (hereafter FKZ IV), or have appeared as meeting abstracts. In this paper we offer a condensed theoretical background with a scope wider than just aiming at direct practical use. Therefore, we shall not discuss different solution techniques in detail. Emphasis is rather placed on the basis of our concept and deducing from it different options for reaching or approaching a structure solution in the parameter space which hitherto was considered too cumbersome from a numerical point of view [see, for example, the highly esteemed book of Stout & Jensen (1970), pp. 300–301]. A few of these options are based on "our forefathers' old knowledge" (which nowadays can be neglected for daily application because much of it is incorporated in present routine programs, perhaps unknown to some users). A reader of this paper may possibly also find some new insights into structure-determination principles seen from the parameter-space point of view. We also want to note explicitly that no kind of 'contest' with present-day's most efficient program systems is intended. Some of the advantages of our concept may, however, offer chances in certain critical cases of structure research which will be mentioned in the summary.

2. The structure factor

By means of the structure factor the reduction to the least number of dimensions of the 'parameter space' introduced above will be shown. This number is always denoted by m, independent of whether the crystal structure is one-, two- or three-dimensional or a projection is considered.

2.1. Symmetry

The usual reduction of the m_u atoms in the unit cell to the representative atoms in the asymmetric unit reduces the dimension m of the parameter space to $3m_0$,

$$F(H) = \sum_{j=1}^{m_0} \frac{f_j(|H|)}{|\mathrm{PtS}(X_j)|} \sum_{n=1}^{N} \exp\big[2\pi i H\big(R_n X_j + T_n\big)\big], \qquad (1)$$

where N denotes the multiplicity of the general position of the space group and $|\mathrm{PtS}(X_j)|$ gives the order of the site symmetry group of the atom in position X_j.

Reducing all atoms to identical points of scattering power $f = 1$ leads to the so-called 'geometrical structure factor'

$$G(H) = \sum_{j=1}^{m_0} \frac{1}{|\mathrm{PtS}(X_j)|} \sum_{n=1}^{N} \exp\big[2\pi i H\big(R_n X_j + T_n\big)\big]. \qquad (2)$$

For space groups containing $P\bar{1}$ as a subgroup with a centre of symmetry in the origin, the problem can be simplified further. In the case of $P\bar{1}$ the structure-factor equation takes the form

$$F(H) = 2\sum_{j=1}^{m_0} f_j(|H|)\cos\big(2\pi H X_j\big) + \sum_{x,y,z\in\{0,\,1/2\}} f_{xyz}(|H|)\,\mathrm{ocf}(x,y,z)(-1)^{2(x+y+z)}, \qquad (3)$$

where ocf(x, y, z) denotes the occupation factor (usually 0 or 1) of the special position x, y, z with $x, y, z \in \{0, 1/2\}$ and m_0 counts the points in the asymmetric unit. Correspondingly,

$$G(hkl) = 2\sum_{j=1}^{m_0} \cos\big[2\pi\big(hx_j + ky_j + lz_j\big)\big] + \sum_{x,y,z\in\{0,\,1/2\}} \mathrm{ocf}(x,y,z)(-1)^{2(x+y+z)}. \qquad (4)$$

The centrosymmetric case will be preferred in this paper and usually the occupation of special positions will be neglected. (Special positions were discussed for the one-dimensional centrosymmetric case in FKZ I, p. 654.)

2.2. Projections

A further reduction of the dimension can be achieved using projections. [The reconstruction of the three-dimensional information from the projections is a tomographic problem and will be discussed in a separate paper. For intermediate results see Fischer & Pilz (1997), Zimmermann *et al.* (2005, 2006) and FKZ IV.] For example, in the structure factors of the

$h00$ reflections all x coordinates of the atoms in the unit cell are encrypted, according to

$$F(h00) = \sum_{j=1}^{m_u} f_j(|h00|)\exp\big(2\pi i h x_j\big), \tag{5}$$

and in the same way for the geometric structure factor,

$$G(h00) = \sum_{j=1}^{m_u} \exp\big(2\pi i h x_j\big). \tag{6}$$

If the list of the x coordinates $(x_j)_{j\,=\,1,\ldots,m}$ and the list of y coordinates $(y_j)_{j\,=\,1,\ldots,m}$ have been found by means of any procedure, the net for possible (x, y) positions is fixed provided a common origin is given or found. The number of (x, y) combinations for a two-dimensional structure is roughly estimated by $2^m(m!)$. Thus it is useful to think about further restrictions even if one-dimensional solutions are known. In principle, however, such a procedure is possible. The step into the third dimension can be carried out in the same way. This approach is the main reason why especially one-dimensional structure factors and structure amplitudes are discussed below.

Further, this approach results in a special preference for the one-dimensional centrosymmetric situation, as only three of the 17 plane groups ($p1$, pm, $p3$) have acentric one-dimensional projections on the axes of the conventional basis. Also, only 54 of the 230 space groups have an acentric projection on at least one conventional axis.

2.3. The one-dimensional structure-factor and structure-amplitude equation

From the one-dimensional structure-factor equation,

$$F(h) = \sum_{j=1}^{m} f_j(|h|)\exp\big(2\pi i h x_j\big), \tag{7}$$

follows the reduced form of the geometric structure-factor equation,

$$G(h) = \sum_{j=1}^{m} \exp\big(2\pi i h x_j\big) = \sum_{j=1}^{m} e_j(h) = G_r + iG_i. \tag{8}$$

The geometrical structure amplitude is defined by

$$g(h) = \Big|\sum_{j} \exp\big(2\pi i h x_j\big)\Big| = (G_r^2 + G_i^2)^{1/2}. \tag{9}$$

In the one-dimensional acentric case, no special positions exist.

In the centrosymmetric case without special positions the structure-factor equation can be simplified further,

$$G(h) = 2s(h)g(h) = 2\sum_{j=1}^{m} \cos\big(2\pi h x_j\big), \tag{10}$$

where $s(h)$ denotes the sign and

$$g(h) = \Big|\sum_{j=1}^{m} \cos\big(2\pi h x_j\big)\Big|$$

denotes the modulus of the geometric structure factor. (m counts the atomic positions in the asymmetric unit only as in §4.1.) If special positions (centres of inversion) are occupied the term for the special positions has to be added,

$$2s(h)g(h) = 2\Big[\sum_{j=1}^{m} \cos\big(2\pi h x_j\big) + (1/2)\sum_{x\in\{0,\,1/2\}} \mathrm{ocf}(x)(-1)^{2x}\Big] \tag{11}$$

(*cf.* FKZ I, p. 654).

3. Parameter spaces

In spite of the fact that the practical treatment of the problem will be handled using one-dimensional structure factors, the following discussion treats the general case and will be reduced to the one-dimensional case afterwards. The question asked in this section is: how can the structure-amplitude equation be resolved towards the atomic coordinates? Different possibilities shall be discussed.

3.1. The space of parallel hyperplanes F^m

In a first step the structure-factor equation can be considered as a linear equation of the parameters $e_j(hkl) = \exp[2\pi i(hx_j + ky_j + lz_j)]$,

$$F(hkl) = \sum_{j=1}^{m_u} f_j(|hkl|)\, e_j(hkl). \tag{12}$$

The structure factor thus can be considered as the scalar product of an m_u-dimensional vector of atomic scattering factors with an m_u-dimensional vector consisting of points of the complex unit circle. If, for any j and three linear independent reflections $h_1k_1l_1$, $h_2k_2l_2$, $h_3k_3l_3$, the $e_j(h_p,k_p,l_p) = \exp[2\pi i \varphi_j(h_p,k_p,l_p)]$ ($p = 1, 2, 3$) are known, the 'structure is solved', as from this information the matrix relation follows,

$$\begin{pmatrix} h_1 & k_1 & l_1 \\ h_2 & k_2 & l_2 \\ h_3 & k_3 & l_3 \end{pmatrix}\begin{pmatrix} x_j \\ y_j \\ z_j \end{pmatrix} = \frac{1}{2\pi i}\begin{pmatrix} \ln(e_1) \\ \ln(e_2) \\ \ln(e_3) \end{pmatrix} = \begin{pmatrix} \varphi_j(h_1k_1l_1) \\ \varphi_j(h_2k_2l_2) \\ \varphi_j(h_3k_3l_3) \end{pmatrix}. \tag{13}$$

It can be resolved by matrix inversion. This idea uses an m_u-dimensional complex parameter space $\mathbf{E}^{mu}$ which has the topology of the direct product of m_u unit circles $\mathbf{E}$, *i.e.* the topology of a complex torus in the m_u-dimensional space.

The difficulty with this approach is that different structure factors determine different (parallel) hyperplanes and the 'solution points' on the different hyperplanes can only be compared after the reduction to the unknown coordinates. To find them, the calculation of the logarithms of the summands of the structure-factor equation is needed. Thus, this linear approach is not apt for a direct solution of the structure.

However, just this approach was used by Navaza & Silva (1979) for the analysis of Karle–Hauptman determinants. They realized that a scalar product of two solution vectors for **h** and **k** in $\mathbf{E}^m$ can be defined such that a normalized structure factor $E(\mathbf{h} - \mathbf{k})$ results. The metric tensor for this scalar product can be shown to be a Karle–Hauptman matrix. The inversion of this matrix can be used to solve the structure.

3.2. The space of polynomials $\mathbf{X}^m$

In the second stage the structure factor can be considered as a polynomial in the variables $e_{jq} = \exp(2\pi i x_{jq})$, $j = 1, \ldots, m_u$, $q = 1, 2, 3$. The general form is

$$F(hkl) = \sum_{j=1}^{m_u} f_j \, e_{j1}^h \, e_{j2}^k \, e_{j3}^l. \tag{14}$$

In the corresponding m-dimensional solution space $\mathbf{X}^m$ any structure factor defines a hyperface ('polynomial isostructure-factor surface', for short: 'polyface'). This space also has the topology of the surface of an m-dimensional torus. The solution of a structure is the intersection point of at least m polyfaces.

In the centrosymmetric case the structure factor is a real number but the equation becomes more complex because the powers of the exponential terms turn into trigonometric polynomials with a lot of terms.

In the one-dimensional case with

$$F(h) = \sum_{j=1}^{m} f_j \, e_j^h, \quad G(h) = g(h) \exp[i\varphi(h)] = \sum_{j=1}^{m} e_j^h, \tag{15}$$

we find the equations which are the initial step of the method of Ott (1927) and Avrami (1938, 1939) on the basis of Carathéodory's solution (Caratheodory, 1911). Further attempts to solve systems of structure-factor equations were made by Kutschabsky *et al.* (1971). For a new mathematical development see DiPippo & Howe (2000), and for details see also Pilz & Fischer (1998, 2000).

3.3. The space of trigonometric polynomials $\mathbf{Q}^m$

In a third step the m-dimensional space of coordinates x_{jq} can be chosen for a parameter space ($m = dm_u$, where $d = 1, 2$ or 3 is the dimension of the crystal space and m_u is the number of atoms). The structure factor becomes a trigonometric polynomial in this case,

$$\begin{aligned} F(H) &= \sum_{j=1}^{m_u} f_j(|H|) \exp\big(2\pi i H X_j\big) \\ &= \sum_{j=1}^{m_u} f_j(|H|) \cos\big(2\pi H X_j\big) + i \sum_{j=1}^{m_u} f_j(|H|) \sin\big(2\pi H X_j\big). \end{aligned} \tag{16}$$

The parameters x_{jq}, in principle, span the real space $\mathbf{R}^m$. As the structure factor is periodic it is sufficient for each parameter to consider the m-dimensional unit cell, $0 \le x_{jq} < 1$. The remaining parameter space (again using the orthonormal basis of $\mathbf{R}^m$) is an m-dimensional 'cube' $[0,1[^m$ which is denoted by $\mathbf{Q}^m$. Again, the hyperfaces defined by a special value of the structure factor, isosurfaces for short (in FKZ I denoted by $M[h;c]$), intersect in the solution point. At least m isosurfaces are necessary to fix a solution point. Some remarks with respect to the algebraic structure of the parameter space are as follows.

(i) The description of a structure as a vector X in the parameter space means that the parameter space is the direct sum of m_u copies of the crystal space.

(ii) The addition of a fixed vector T to the coordinates in crystal space corresponds to an addition of a vector $T^{mu} := (t_1,t_2,t_3,\ t_1,t_2,t_3,\ \ldots,\ t_1,t_2,t_3)$ to X. In this way a shift of origin works in the parameter space. Especially in the case of one-dimensional structures all copies of a structure which only differ in the choice of origin form a straight line in the parameter space parallel to the main symmetry axis $[1, 1, \ldots, 1]$. (For acentric structures, the 'inverse solutions' are arranged on a second line.) A further application will be shown below in the context of negative and complex scattering factors.

(iii) Any space-group operation $(R|T)$ modulo $\mathbf{Z}^3$ corresponds to a permutation of the indices of the components of X. Thus the space group operates on the parameter space as a subgroup of the symmetric group $\mathbf{S}_{mu}$, the group of all permutations of m_u elements. Thus it will be useful to choose the sequence of the list of atoms in groups of symmetrically equivalent atoms as the parameter space will be decomposed into symmetry subspaces.

By specializing the case of the section above, the situation becomes easier in the centrosymmetric case,

$$F(H) = 2 \sum_{j=1}^{m_u/2} f_j(|H|) \cos\big(2\pi H X_j\big), \tag{17}$$

and the parameter space can be reduced from $\mathbf{Q}^m := [0, 1[^m$ to $\mathbf{P}^m := [0, 1/2[^m$. This form of the structure factor is thus the starting point for our investigations, especially in the one-dimensional crystal space. By using it in the form of the geometric structure amplitude the maximum permutation symmetry in the parameter space can be achieved.

4. Symmetry and some properties of isosurfaces in the parameter space

For easy understanding of the following we found it useful to refer to some well known trivial features of the structure-factor equation.

4.1. Symmetry and asymmetric unit

The parameter spaces have the conventional orthonormal metric of $\mathbf{R}^m$. The symmetry discussion can easily be performed for the general case of three-dimensional structures. As we will use the results in the one-dimensional case only, it shall be developed for this case below.

The symmetry of the structure factor in the parameter space has two sources:

(i) The symmetry properties of the exponential and trigonometric functions,

(ii) the commutativity of the addition.

The basis for case (i) is the frieze symmetry $pm(g2)$ [the aperiodic parts of the symbol are written in parentheses according to Bohm & Dornberger-Schiff (1967)] of the graphs of trigonometric functions, *i.e.* the fact that the values of all trigonometric functions can be reconstructed from the values of $\cos(2\pi x)$, with $x \in [0, 1/4[$: From, for example,

$$\begin{aligned}\cos(2\pi x) = \sin\left[2\pi\left(x+\tfrac{1}{4}\right)\right] &= -\cos\left[2\pi\left(\tfrac{1}{2}-x\right)\right]\\ &= \cos[2\pi(1-x)],\end{aligned} \qquad (18a)$$

it follows for the general exponential function

$$\exp[2\pi i(x+p/4)] = i^p \exp(2\pi i x) \quad \text{with integer } p. \qquad (18b)$$

In both cases the main symmetry is the periodicity of these trigonometric functions,

$$t(x) : t(x+n) = t(x), \qquad (18c)$$

for any integer n, which remains true for sums of such terms.

The results for the one-dimensional geometric structure factor are:

(i) For acentric structures we use

$$G_r(h|X) + i\,G_i(h|X) = G(h|X) = \sum_{j=1}^{m} \exp\left(2\pi i h x_j\right)$$

in the parameter space $\mathbf{Q}^m$, and (for centrosymmetric cases) in the parameter space $\mathbf{P}^m$ with a volume reduced by a factor 2^m {because from equation (18*a*) the asymmetric unit of

$$G(h|X) = 2\sum_{j=1}^{m} \cos\left(2\pi h x_j\right)$$

is [0, 1/2] for any component x_j of X}.

(ii) In both cases (centric and acentric) the geometrical structure factor is invariant under all permutations of the indices j of the x_j. Thus the symmetric group $\mathbf{S}_m$ of order $m!$ operates on the parameter space. The asymmetric unit $\mathbf{A}^m$ in $\mathbf{Q}^m$ and $\mathbf{A}^m \cap \mathbf{P}^m$ in $\mathbf{P}^m$ can be fixed by means of the condition $x_1 \geq x_2 \geq \ldots \geq x_{m-1} \geq x_m$. According to this rule $\mathbf{A}^m$ is a convex region in $\mathbf{Q}^m$. It forms a simplex with a set of $(m+1)$ corners: $\{(0, \ldots, 0); (1, 0, \ldots, 0); \ldots (1, 1, \ldots, 1, 0); (1, \ldots, 1)\}$. The volume $|\mathbf{A}^m|$ is $|\mathbf{Q}^m|/m!$ or $|\mathbf{P}^m|/m!$, respectively. As one-dimensional projections of three-dimensional structures are considered, the overlap of atoms causing the case $x_j = x_{j+1}$ cannot be excluded. Solutions of this kind will be found on (lower-dimensioned) surfaces of the asymmetric unit. (The above-mentioned reduction factor of $m!$ is essential for the actual CPU-time demand in structure determination.) *Cf.* also DiPippo & Howe (2000).

(iii) Further, $g(h|X)$ has translation symmetry because of equation (18*c*). The translation lattice is $\Gamma_h := \{\mathbf{t}|\ \mathbf{t} = \mathbf{z}/h,\ \mathbf{z} \in \mathbf{Z}^m\}$. The volume of the asymmetric unit for this single h is thus reduced by a factor $1/h^m$.

(iv) From equation (18*b*) it results that $I'_2 := (1/2h)(1, \ldots, 1)$ is an antisymmetry translation in $\mathbf{Q}^m$ for any $G(h|X)$, *i.e.* $G(h|X) = -G(h|X + I'_2)$. For the modulus $g(h|X)$ of $G(h|X)$ it is thus a true centring translation. The volume of the asymmetric part is reduced by a further factor 1/2. In the centrosymmetric case, it follows from the same argument that $I'_4 := (1/4h)(1, \ldots, 1)$ is an antisymmetry translation; it is a true translation, however, for the structure amplitude.

4.2. Geometric properties of *g*(*h*)-isosurfaces

4.2.1. Scaling symmetry (negligently called self-symmetry in FKZ I). Comparing

$$g(h|X) = \left|\sum_{j=1}^{m} \cos\left(2\pi h x_j\right)\right| \quad \text{and} \quad g(1|X) = \left|\sum_{j=1}^{m} \cos\left(2\pi x_j\right)\right|$$

shows that the graphs of both functions only differ by a scaling factor h, or $1/h$ respectively, applied to the coordinates x_j [*cf.* Figs. 3(*a*)–3(*d*) of FKZ I]. Thus all scale-invariant geometrical properties of $g(h|X)$ can be investigated considering $g(1|X)$. The same holds for the noncentrosymmetric case.

4.2.2. Tangent hyperplanes (for centrosymmetric structures only). For the tangent planes it is easier to consider $G(1|X)$ and to discuss the behaviour at the points with $G(1|X) = 0$ separately. The tangent hyperplane of $G(1|X)$ in X_0 is given by the equation

$$\sum_{j=1}^{m} \sin\left(2\pi x_{0j}\right)\left(x_j - x_{0j}\right) = 0. \qquad (19)$$

Points of the half-integral lattice $\Gamma_{1/2} = \{X | X = \frac{1}{2}Z, Z \in \mathbf{Z}^m\}$, *e.g.* the corners of the asymmetric unit, are singular points and have no well defined tangent hyperplanes. For the edges of the asymmetric unit given by $\{(x_0, 0, \ldots, 0), (x_0, x_0, 0, \ldots, 0), (x_0, x_0, x_0, 0, \ldots, 0), \ldots\}$ for the coordinates $(x_1, \ldots, x_m)$ of the points in the tangent hyperplane we find the equations $x_1 - x_0 = 0$; $(x_1 - x_0) + (x_2 - x_0) = 0$; $(x_1 - x_0) + (x_2 - x_0) + (x_3 - x_0) = 0$; .... As we have chosen an orthonormal metric in the parameter space, these equations show that the tangent hyperplanes are always orthogonal with respect to the edge vectors of the asymmetric unit. This result may be helpful if linear approximations of the isosurfaces are used. The scaling symmetry shows that this is true for any $G(h|X)$. Except for the points where $G(h|X)$ is zero, the same arguments are true for all $g(h|X)$.

4.2.3. Zero values of $G(1\,|\,X)$. *Centrosymmetric case.* For $X = (1/4, 1/4, \ldots, 1/4)$: $G(1|X) = 0$ for any m. Also the straight lines given by $(x, 1/2 - x, 1/4, \ldots, 1/4)$ and its permutations consist of zeros of $G(1|X)$. For $m > 2$ the zero isosurface of $G(1|X) = 0$ is curved but it contains these straight lines. The importance of structure factors with value zero is that they need no phase. Each $G(h|0)$ or $g(h|0)$ thus principally reduces the dimension of the problem from m to $m - 1$ (Ott, 1927) so that m such observations uniquely define the solution. For small deviations ($g \simeq 0$) see the end of §6.1.

Noncentrosymmetric case. Isosurfaces for $G(1|X) = 0$ are $(m-2)$-dimensional because both G_r and G_i must be zero.

4.2.4. Extremal points. Maxima of $G(1|X)$ are the points $X \in \mathbf{Z}^m$, *i.e.* the points with integral coordinates. The maximal value is m. For centrosymmetric structures, the minimum $-m$ lies in the point (1/2, 1/2, ..., 1/2). Considering only amplitudes $g(1|X)$ in an 'elementary cell' $\mathbf{Q}^m$ for centrosymmetric as well as for noncentrosymmetric structures, we find these maxima in points of a primitive lattice (noncentrosymmetric) or in a body-centred one (centrosymmetric), respectively. Each $G(h|m)$ or $g(h|m)$ thus reduces the structure solution to a known number of singular points in $\mathbf{A}^m$ (or half of them after origin definition).

4.2.5. Approximate behaviour. For centrosymmetric structures, from the approximation $\cos(2\pi x_j) \simeq 1 - 2\pi^2 x_j^2$ we obtain

$$G(1|X) \simeq \Big(m - 2\pi^2 \sum_{j=1}^{m} x_j^2\Big),$$
$$g(1|X) \simeq \Big|m - 2\pi^2 \sum_{j=1}^{m} x_j^2\Big| = m\Big|1 - 2\pi^2(1/m) \sum_{j=1}^{m} x_j^2\Big|. \qquad (20)$$

This means that close to lattice points of the parameter space the isosurfaces are hyperspheres of radius $r = (1/\pi)t^{1/2}$ for $g = m - t$ $(0 < t \ll 1)$, and can thus be described by a much 'simpler' function.

For the acentric case with the approximation $\exp(2\pi i x_j) = 1 + 2\pi i x_j - 2\pi^2 x_j^2$ we find

$$G(1|X) \simeq \Big(m + 2\pi i \sum_{j=1}^{m} x_j - 2\pi^2 \sum_{j=1}^{m} x_j^2\Big), \qquad (21)$$

thus on the hyperplane defined by $\sum_{j=1}^{m} x_j = 0$ we find the same condition as in the centrosymmetric case.

4.3. Different atomic scattering factors

The permutation symmetry of the parameter space of order $m!$ (imprinted on the symmetry of the system of isosurfaces) is a consequence of the equal-point-atom assumption for our structure model (simplified to $f_j = 1$). Whenever this prerequisite is violated, the permutation symmetry is broken. This causes less transparency of the problem and demands more CPU time for solving a structure, both disadvantages increasing with increasing m (see last paragraph of 'Early and recent concepts' in FKZ I, p. 644). For this to happen it suffices to introduce different f_j being constant in reciprocal space (different point scatterers without 'thermal parameters'). Thus, in §4.3.2 and §4.3.3, the scattering amplitudes of different atoms remain standardized to 1 and the change is restricted to their scattering phase.

4.3.1. Different scattering amplitudes, all positive. Assume a centrosymmetric one-dimensional structure of two different atomic species having, *e.g.*

$$f_j = f_1 = 1 (j = 1, \ldots, m') \qquad \text{and}$$
$$f_j = f_2 \neq 1 (j = m' + 1, \ldots, m). \qquad (22)$$

The geometrical structure amplitude then reads as

$$g(h) = \Big|f_1 \sum_{j=1}^{m'} \cos(2\pi h x_j) + f_2 \sum_{j=m'+1}^{m} \cos(2\pi h x_j)\Big| \qquad (23)$$

or

$$g(m, h) = \big|f_1 s_1 g(m', h) + f_2 s_2 g(m - m', h)\big|. \qquad (24)$$

All isosurfaces are changed. Merely three special positions within $\mathbf{P}^m$, namely at (0, 0, …, 0) and at (1/2, 1/2, …, 1/2) representing $g_{\max} = \sum_{j=1}^{m} f_j$ plus the antisymmetry centre at (1/4, 1/4, …, 1/4) representing $g(1) = 0$, retain their $g(1)$ in these positions. A simple case ($m = 2$, $f_1 \neq f_2$, centrosymmetric) is depicted in FKZ I, p. 654, Fig. 10.

Applying these changed isosurfaces would eliminate some of the errors in the data-reduction process (namely those connected with, for example, different neutron scattering amplitudes having the same sign). Whether this outweighs the increased computing demand owing to the reduced order of permutation symmetry must be decided in each case. If, for example, m' atoms have $f_j = 1$ and $m - m'$ have $f_j \neq 1$, the order of permutation symmetry is reduced from $m!$ to $(m')! \times (m - m')!$. Consequently, given a problem with m independent atoms, the binomial coefficient $\nu_{\rm A} = m!/[m'!(m - m')!]$ as factor for increasing the volume of the asymmetric unit $|\mathbf{A}^m|$ has a minimum of $\nu_{\rm A} = m$ for $m' = 1$ (or $m - m' = 1$ alternatively) and attains a maximum for $m' = m/2$ (m even), $\nu_{\rm A} = (2m')!/(m'!)^2$. Even for moderate m, perhaps a single heavy atom [or at most two, entailing a factor of $m(m - 1)/2$] can perhaps be tolerated. Therefore, and also because of its analogy, we refrain from treating the acentric case in more detail.

4.3.2. Different scattering signs (*e.g.* for neutron scattering). From FKZ I (p. 654 and Fig. 11, again for the simple $m = 2$ case with inversion centre) we learn that the overall picture of the system of isosurfaces for $g(h)$ appears unchanged save a shift of origin by π/h in the direction of one principal axis. Generalized to $m = m' + m^-$ point atoms having $|f| = 1$ and counting the m' positive scatterers prior to the negative ones, we obtain

$$\begin{aligned} &g(h; m'|x_1, \ldots, x_m) \\ &\quad := \Big|\sum_{j=1}^{m'} \cos\big(2\pi h x_j\big) - \sum_{j=m'+1}^{m} \cos\big(2\pi h x_j\big)\Big| \\ &\quad = \Big|\sum_{j=1}^{m'} \cos\big(2\pi h x_j\big) + \sum_{j=m'+1}^{m} \cos\big\{2\pi h[x_j + (1/2h)]\big\}\Big| \\ &\quad = g[h|x_1, \ldots, x_{m'}, x_{m'+1} + (1/2h), \ldots, x_m + (1/2h)]. \end{aligned} \qquad (25)$$

The isosurfaces for $g(1; m')$ are thus shifted by (0, …, 0, −1/2, …, −1/2) (m' times 0, m^- times −1/2) compared with the $g(1)$ with all $f = +1$. This means that no new types of isosurfaces occur.

Though this shift of origin disturbs the permutation symmetry 'in principle', the latter may, however, be retained if the shifted origin (a separate one for each h) is used thus including a corresponding shift of the respective atomic coordinates.

4.3.3. Different scattering phases [*e.g.* for resonant ('anomalous') X-ray scattering]. The same arguments can be used, this time in the acentric case, if some simple conditions are observed. Assume $f_j = 1$ ($j = 1, \ldots, m'$) for the normal scatterers and $f_j = f' + if'' = \exp(2\pi i\varphi_j)$ ($j = m' + 1, \ldots, m$) with

$$(f')^2 + (f'')^2 = 1 \quad \text{and} \quad \arctan(f''/f') = 2\pi\varphi \qquad (26)$$

for the resonant atoms ('anomalous scatterers'). Then

$$\begin{aligned} g(h;m'|x_1,\ldots,x_m) &= \Big|\sum_{j=1}^{m'}\exp\big(2\pi ihx_j\big) \\ &\quad+\sum_{j=m'+1}^{m}\exp\big\{2\pi ih\big[x_j+\big(\varphi_j/h\big)\big]\big\}\Big| \\ &= g\Big[h|x_1,\ldots,x_{m'},x_{m'+1} \\ &\quad+(\varphi_{m'+1}/h),\ldots,x_m+(\varphi_m/h)\Big]. \qquad (27) \end{aligned}$$

The isosurfaces for $g(h)$ are thus shifted by $(0, \ldots, 0, -\varphi_{m'+1}/h, \ldots, -\varphi_m/h)$ (m' times 0, $m - m'$ times $-\varphi/h$) compared with $g(h)$ with all $f = +1$. The complete geometry of the isosurfaces and the asymmetric unit thus remains unchanged save a shift of origin which depends on h. Compare Figs. 1(*a*) and 1(*b*) for $g(1)$ (without and with resonant scattering) and Figs. 1(*b*) and 1(*c*) for $g(1)$ and $g(2)$ (resonant scattering for different h). It is worth noting that from the derivation above it follows that even different kinds of resonant scatterers do not change the overall 'landscape' provided correspondingly different φ are considered. (How the 'anomalous-dispersion contrast' of Bragg scattering amplitudes can be used within the parameter-space concept will be treated in a separate paper.)

The net result of this section is a consequence of §3.3, remark (ii), and may be summarized as follows. The 'landscape' of the isosurface system in the m-dimensional parameter space is independent of the phase φ_j (including sign) of the scattering factor f_j of all independent point atoms under the restriction that this scattering factor (may it be real, negative or complex) has the same amplitude (usually $|f_j| = 1$). Compared with the isosurface system of a one-dimensional structure with all $f_j = +1$, the origin is shifted by a corresponding $-\varphi_j/h$ in each of the m directions. This result is parallel to that of Cervellino & Ciccariello (2005), who proved that different signs of neutron scattering amplitudes (which in their derivation may also be different) do principally not prevent a structure solution (thus the condition of 'non-negative scattering density' for direct-methods-based phase determination is overcome in general). By the above findings, the scope of Cervellino & Ciccariello is even widened to atoms exhibiting complex scattering factors, however restricted to $|f_j| = 1$. (The above-mentioned shift of origin by $-\varphi_j/h$ is no fundamental obstacle for solving a structure, though it adds, of course, to the computing demand.)

5. Principles of structure solution

5.1. By intersection of isosurfaces

A point in the parameter space, representing a structure, is fixed by the intersection of m isosurfaces. Analytically, this means solving a system of at least m equations with trigonometric or exponential summands. The attempt to solve them directly leads to difficulties even in the simple case of $m = 4$, as the procedure involves ambiguous logarithms and inverse trigonometric functions such that the variables cannot be

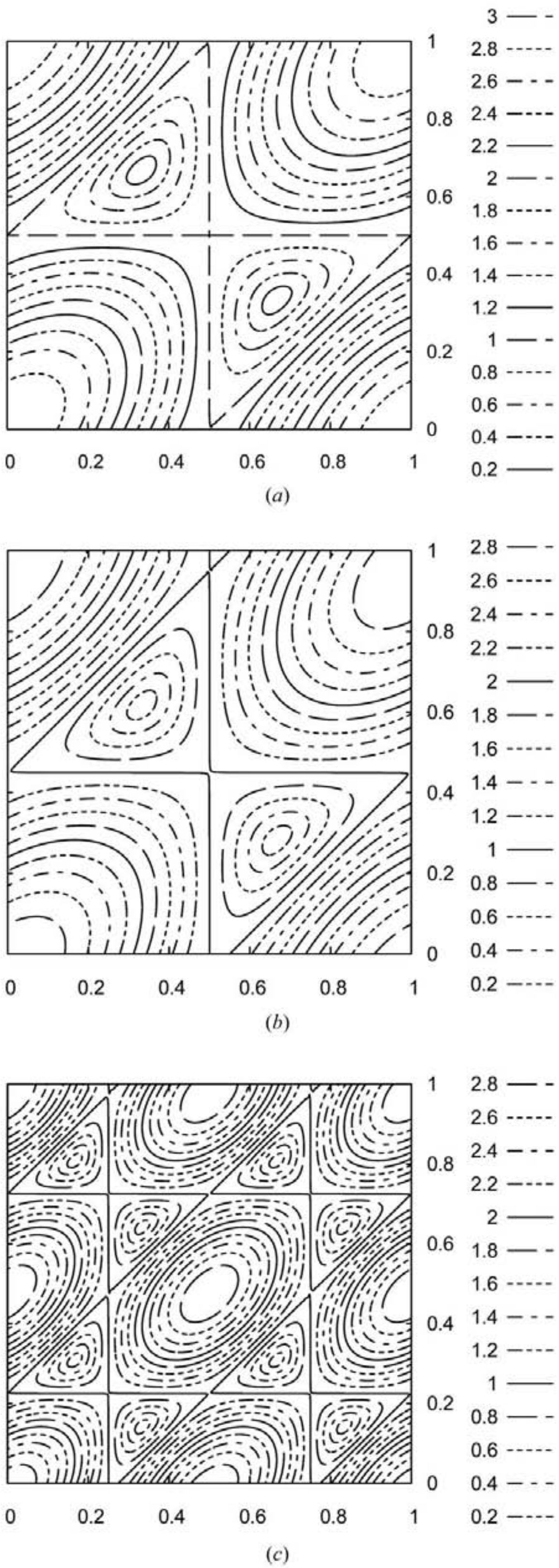

Figure 1
Isosurfaces of $g(1|X)$ for a one-dimensional acentric $m = 3$ structure, third atom in $x_3 = 0$, thus $\mathbf{P}^3 \rightarrow \mathbf{P}^2$. Coordinate axis x_1 horizontal, x_2 vertical. Contours at 0.2, isosurface for $g(1|X) = 1.0$ depicted as straight lines. (*a*) All atoms $f = 1$. (*b*) $f_1 = f_3 = 1$, second atom: $f_a' = 0.95$, $f_a'' = 0.31 \rightarrow \varphi_a = 0.050_5$: observe vertical shift of all isosurfaces by -0.05. (*c*) Same as (*b*) but for isosurfaces of $g(2|X)$: observe the vertical shift is halved.

separated easily. It is thus useful to consider geometric and numeric approaches.

(*a*) *Linear approximation*. Pairs of parallel $(m - 1)$-dimensional hyperplanes are calculated such that they completely enclose the isosurface of a given 'experimental' $g(h)$, the region between them thus holding the solution point X. Intersection with corresponding regions leads to a 'solution region' with linear boundaries. If this solution region is small enough, its centre of gravity is a good approximation of X. Some examples are given in FKZ I.

(*b*) *Grid calculation*. The direct procedure is to use a set of known structure amplitudes $g(h_1), \ldots, g(h_n)$ and to choose a system of grid points $\{X_j | j = 1, \ldots, q\}$ and a threshold $t, 0 < t$, in the asymmetric unit of the parameter space and to calculate the structure amplitudes $g(h_1|X_j), \ldots, g(h_n|X_j)$ for all grid points. If for a $g(h_k|X_j)$ the difference $|g(h_k|X_j) - g(h_k)| > t$, the point j is eliminated and the next grid point is calculated. If after this procedure there are no grid points left, the procedure has to be repeated with a refined grid. If there are points left, a refined grid is introduced in the neighbourhood of these points and the procedure is repeated. If there are too many grid points which fulfil the condition, the threshold t has to be diminished. After using some stopping rule a set of 'permitted' solution points should be left.

Such a grid procedure is (in principle) an extended 'trial-and-error' technique and may need a lot of calculation time. It is thus worthwhile to reduce the possible solution space. In addition, up to now absolute values of $g(h)$ were a prerequisite which can hardly be met for a small batch of experimental reflection data unable to provide a statistically sound Wilson plot (see §5.2.1 and §5.2.2).

5.2. By reduction of solution regions

Two additional types of isosurfaces are introduced in this chapter. Both are based on data combinations from a single 'central reciprocal-lattice line' and are on the 'absolute' scale. Because of their geometrical complexity (compared with the g-isosurfaces), they are of limited use for finding X directly by intersection or approximation techniques; they are, however, rather useful for confining possible solution regions in $\mathbf{A}^m$, thus approximating a solution X.

5.2.1. Quotients of structure amplitudes and their isosurfaces. (i) *Definition and general features.* Products of structure factors are thoroughly investigated objects in crystallography, especially those whose reflection indices form closed paths on the reciprocal lattice, as their phase sums are invariant under the shift of origin in direct space. No such algebraic properties are known about quotients of structure factors; they play a certain role if the scale factor cannot be fixed with sufficient precision.

While the product $F(H)F(-H) = |F(H)|^2$ gives the square modulus of the structure factor (which is also invariant under origin shifts), the quotient $F(H)/F(-H) = \exp\{2\pi i[2\varphi(H)]\}$ marks an arbitrary point on the complex unit circle. The quotients of structure factors thus do not seem to be a useful tool for structure determination. However, indeed the quotients of structure amplitudes are used already in crystallography, namely for the results of powder diffraction. For use in the JCPDS database the intensities of the reflections are divided by the intensity of the strongest reflection and the result is given as a percentage of the maximal intensity. This sequence was recognized to be characteristic for the structure and used for the identification in the database.

Again we will consider the one-dimensional centrosymmetric case. For any structure vector $X = (x_1, \ldots, x_m) \in \mathbf{P}^m$ the quotient $q(h, k|X)$ of the structure amplitudes (first introduced by Kirfel, see Kirfel & Fischer, 2005) is given by

$$q(h, k|X) = \frac{g(h|X)}{g(k|X)} = \Big|\sum_{j=1}^{m} \cos(2\pi h x_j)\Big| \Big/ \Big|\sum_{j=1}^{m} \cos(2\pi k x_j)\Big|. \qquad (28)$$

The quotient is well defined at all points of the parameter space where $g(k|X) \neq 0$. One advantage of using quotients is that the problem of an unknown scale factor is eliminated.

The most interesting property of the quotients is that the inequalities $q(h, k|X) \geq c + \varepsilon$ and $q(h, k|X) \leq c - \varepsilon$ (for some $\varepsilon > 0$) define regions in the parameter space within which the solution of the structure cannot be found ('forbidden regions' in contrast to 'permitted' ones). As, except for the case discussed above, the quotients are smooth functions, these regions are rather regular and can be approximated using a relatively coarse grid [although the corresponding isosurfaces for given $q(h, k|X)$ appear geometrically more 'complicated' than do those for the $g(h|X)$]. It is then a question of a clever organization of the procedure to reduce the numerical effort.

(ii) *Sequences of $q(h, k)$*. Assume n data $g(h)$ from a 'central reciprocal-lattice row', *i.e.* a series of harmonic reflections, sorted according to decreasing g values. Call the strongest one $g(k_1)$ and the smallest $g(k_n)$ and keep the 'true indices' h. Then, two special sequences of independent $q(h, k) \leq 1$ $(n - 1$ each) can be formed: either a series with $h = k_2, \ldots, k_n$ and $k = k_1$ or with $h = k_1, \ldots, k_{n-1}$ and $k = k_2, \ldots, k_n$. Applying the above-mentioned inequalities $q(h, k) \leq 1$ (separately for each sequence) and intersecting the corresponding 'permitted solution regions' provides a 'confined solution region' whose centre (however selected) may serve as an approximate solution. (For more details, see FKZ II and FKZ III.)

(iii) *Dividing the parameter space by zero isosurfaces into 'fields'*. If (for centric structures only) any $g(h) = 0$ is introduced as denominator, *e.g.* for calculating q, it causes an $(m - 1)$-dimensional jump discontinuity along its isosurface [where, if signs are taken into consideration, the function calculated (or a part of it) jumps from $+\infty$ to $-\infty$ normal to this isosurface]. For instance, the parameter space of a given $q(h, k)$ using an experimental $g(k) = 0$ is thus divided into m-dimensional regions (to be called 'fields') surrounded by these $(m - 1)$-dimensional jump discontinuities. This is one of the reasons why q-isosurfaces do not 'invite' simple approximations over a large range of parameter space. Within each of these fields, no such discontinuity can occur and simple approximations to isosurfaces are basically possible. [For superimposing isosurfaces for different h, see Kirfel *et al.* (2006) (FKZ II) and Kirfel & Fischer (2009) (FKZ IV).]

5.2.2. Quasi-normalized structure amplitudes and their isosurfaces. A second (and more preferable) option for obtaining 'experimental' data called $e(h, n)$ from unscaled $g(h)$ is offered through dividing $g(h)$ by any average of a given batch of n of those $g(h)$, taken, for example, from a 'central reciprocal-lattice line' and defined by

$$e(h,n) := \frac{g(h)}{\left[(1/n)\sum_{k=1}^{n} g^2(k)\right]^{1/2}} = \frac{n^{1/2} g(h)}{\left[\sum_{k=1}^{n} g^2(k)\right]^{1/2}} \qquad (29)$$

['quasi-normalized amplitudes', the above being the best option of various types of averaging; see Kirfel & Fischer (2007) and FKZ IV for details]. Their corresponding isosurfaces look even more 'complicated in detail' as are the corresponding ones of $q(h, k)$ and consequently ask for a finer grid than the latter in 'trial-and-error' approximations. Sequences of $e(h, n)$ may be used as those of $q(h, k)$.

5.2.3. Some general results of solution-space reduction by inequalities. At the beginning it should be acknowledged that Kirfel was the first to introduce both $q(h, k)$ (Kirfel & Fischer, 2005) and $e(h, n)$ (Kirfel & Fischer, 2007) as well as the new inequalities [Kirfel *et al.*, 2006 (FKZ II)] related to them.

From §5.2.1 and §5.2.2 it follows that each inequality between different $g(h)$ or between $q(h, k)$ or between $e(h, n)$ defines $(m - 1)$-dimensional boundaries for a 'permitted solution region' within $\mathbf{A}^m$. Consecutive overlay of these regions confines more and more a 'combined solution region' (or a number of them). Whenever the region(s) obtained appear(s) small enough, an approximate solution has been found (or possible ones to be discriminated later). To attain this, neither $n \geq m$ data nor $g(1)$ are necessary (though low-order reflections may ease the confinement by reducing the number of solution regions). Evidently, the higher the contrast between the $g(h)$ involved the fewer data may lead to a 'refinable' approximate X. These findings again confirm the spatial resolution power of the parameter-space concept whose use is not restricted by series-termination problems.

5.2.4. Use of a few known signs. In centrosymmetric structures, origin definition by assuming a given sign for a selected $g(h)$ with h odd automatically excludes half of $\mathbf{A}^m$ for possible solutions independent of the numerical value of $g(h)$ [see Figs. 2(*a*)–2(*d*) for an $m = 2$ case]. The boundaries between permitted and forbidden solution regions are taken from 'field boundaries' of §5.2.1, (iii). Corresponding (however different) fractions of $\mathbf{A}^m$ are obtained for additional known signs $s(h)$ of any $g(h)$. By intersecting a few of these 'halves', a drastic reduction of permitted solution regions is obtained. For the sign combination $s(2) = s(4) = +1$ and $s(3) = -1$, for example, only 1/72 of $\mathbf{A}^2$ is permitted, resulting in an error of $\pm\Delta x_{1,2} \leq 0.03$ (Fig. 2*e*).

While the general result (which neglects knowledge of the amplitudes) can just as well be assumed true for higher m, two warnings are mandatory:

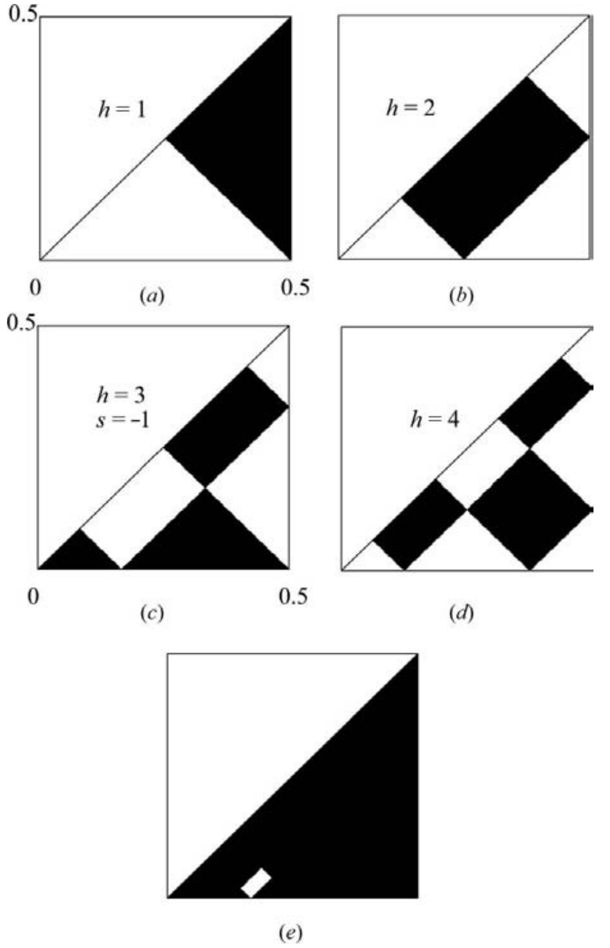

Figure 2
Permitted (white) and forbidden (black) solution regions in $\mathbf{A}^2$ (lower left triangle) of a one-dimensional centric $m = 2$ structure for four reflections (coordinate axes as in Fig. 1): (*a*), (*b*), (*d*) $s = +1$; (*c*) $s = -1$; (*e*) overlay of $h = 1, \ldots, 4$. Observe the small solution area.

(i) All field boundaries owing to $g(h) = 0$ (in Fig. 2) are 'weak' ones owing to uncertainties of the zero values (see §6 below).

(ii) For $m > 2$, these boundaries are neither linear nor equidistant and not yet explored in detail.

If in addition to signs the amplitude is also considered, the dimension m of the problem is reduced by the number of reflections with known signs and amplitudes as described above for zero-intensity observations (§4.2.3). [In the example of Fig. 2(*e*), the solution would thus be overdetermined.]

5.3. General result of a one-dimensional solution

Each asymmetric unit $\mathbf{A}$ of $\mathbf{Q}$ or $\mathbf{P}$ houses at least two inverse (thus crystallographically equivalent) solutions as a consequence of centring translations [see §4.1, (iv)]:

(i) In the acentric case, two solutions are related by inversion either at the origin or at $(1, 1, \ldots, 1)$ causing $x_{j'} = -x_j$ or $1 - x_j$, respectively. For centrosymmetric structures, the point $(1/4, 1/4, \ldots, 1/4)$ acts as inflection point, causing $-x_{j'} = 1/2 - x_j$. Consequently, only one half of the asymmetric part must be searched through. The boundary separating the two halves cuts normal to the main symmetry axis [1, 1, ..., 1] at (1/2, 1/2, ..., 1/2) for acentric and at (1/4, 1/4, ..., 1/4) for centric **A**, respectively. (In the latter case, the isosurface for $g = 0$ may also be used instead, see also §5.2.)

(ii) Additional equivalent acentric solutions exist, if one of the point atoms (*e.g.* the mth one) defines the origin. The dimension of the parameter space is thus reduced by one to $\mathbf{Q}^{m-1}$, and $\mathbf{A}^{m-1}$ as its asymmetric part contains $m - 1$ additional solution points owing to selecting each of the remaining atoms for origin [a consequence of §3.3, statement (iii)].

The combination of (i) and (ii) for the acentric unit $\mathbf{A}^{m-1}$ results in $2m$ solution points (see Fig. 3). Having safely located any of them as the first one, the search can be terminated thanks to the equivalence of the others.

Table 1
Vector set matrix for an acentric structure with $x_1 \geq x_2 \geq \ldots \geq x_{m-1} \geq x_m = 0$ and $u_{i,j} := x_j - x_i$.

0	$u_{1,2}$	$u_{1,3}$	...	$u_{1,j}$	...	$u_{1,m-1}$	$u_{1,m} = -x_1$
$u_{2,1}$	0	$u_{2,3}$	...	$u_{2,j}$	...	$u_{2,m-1}$	$u_{2,m} = -x_2$
$u_{3,1}$	$u_{3,2}$	0	...	$u_{3,j}$	...	$u_{3,m-1}$	$u_{3,m} = -x_3$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
$u_{i,1}$	$u_{i,2}$	$u_{i,3}$	...	$u_{i,j}$	...	$u_{i,m-1}$	$u_{i,m} = -x_i$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
$u_{m-1,1}$	$u_{m-1,2}$	$u_{m-1,3}$	...	$u_{m-1,j}$	...	0	$u_{m-1,m} = -x_{m-1}$
$u_{m,1} = x_1$	$u_{m,2} = x_2$	$u_{m,3} = x_3$	...	$u_{m,j} = x_j$	...	$u_{m,m-1} = x_{j-1}$	$0 = x_m$

Table 2
Vector set matrix for a centrosymmetric structure (full unit cell, Harker vectors indicated in bold).

$u_{1,j}$	0	$x_2 - x_1$	...	$x_m - x_1$	$-x_m - x_1$	...	$-x_2 - x_1$	$\mathbf{-2x_1}$
$u_{2,j}$	$x_1 - x_2$	0	...	$x_m - x_2$	$-x_m - x_2$	...	$\mathbf{-2x_2}$	$-x_2 - x_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$u_{m,j}$	$x_1 - x_m$	$x_2 - x_m$	...	0	$\mathbf{-2x_m}$	...	$-x_2 - x_m$	$-x_1 - x_m$
$u_{-m,j}$	$x_1 + x_m$	$x_2 + x_m$	...	$\mathbf{2x_m}$	0	...	$-x_2 + x_m$	$-x_1 + x_m$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$u_{-2,j}$	$x_1 + x_2$	$\mathbf{2x_2}$	...	$x_m + x_2$	$-x_m + x_2$	...	0	$-x_1 + x_2$
$u_{-1,j}$	$\mathbf{2x_1}$	$x_2 + x_1$	...	$x_m + x_1$	$-x_m + x_1$	...	$-x_2 + x_1$	0

(iii) As to the question 'how can atomic coordinates be obtained directly from reflection amplitudes (instead of interatomic vectors) without employing their phases?', let us assume a one-dimensional acentric m-atom problem with $x_1 \geq x_2 \geq \ldots \geq x_{m-1} \geq x_m$. Its Patterson vectors are defined as $u_{ij} \equiv x_j - x_i$ $(i, j = 1, \ldots, m)$ forming the conventional $m \times m$ vector set matrix (see Table 1) whose rows and columns represent pictures of the structure as seen from different origins. Defining the origin by the mth atom ($x_m = 0$) thus simply turns the mth row of the matrix into a solution (and the mth column provides the inverse solution). Any other x_j taken as origin-definer trivially 'produces' an equivalent pair of solutions as row and column numbered $i = j$ thus resulting in the $2m$ equivalent solutions [see (i) and (ii) above]. For a simple $m = 3$ case (the parameter space reduced to two dimensions, see above), each of all six equivalent solutions depicted in Fig. 3 represents the coordinate components of one of the six non-zero elements of a (3 × 3)-vector set (which, if generalized to $m \times m$, answers the question above). For a one-dimensional centrosymmetric structure, Table 2 demonstrates an analogue solution by subtracting from any row (or column) the corresponding x_j. As in any attempt to solve Patterson functions, the key step is defining or finding an origin for the picture of the structure. (Harker vectors, being very helpful in deconvoluting a Patterson, do not show up explicitly in our concept, however.)

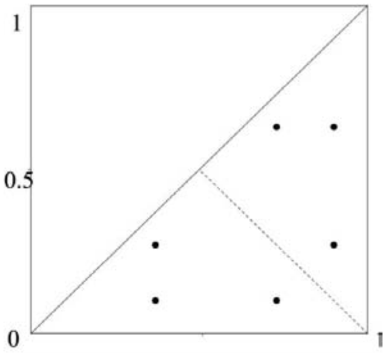

Figure 3
All six equivalent solution points for a one-dimensional acentric $m = 3$ case, depicted as dots in $\mathbf{A}^2$. Observe their coordinate and symmetry relations.

6. Data errors

For easy formulae, we assume inversion symmetry (summation over m independent atoms j) and a few data from a 'central reciprocal-lattice line' [abbreviated by $|F_{\rm obs}(h)|$], thus Wilson statistics cannot be applied. (Generalization to any set of two- or three-dimensional data or to acentric projections is trivial.)

The discussion shall be restricted to errors of $g(h)$ caused by the 'data-reduction' process $|F(h)| \rightarrow g(h)$ necessary for applying the parameter-space concept. From the underlying equal-point-atom model it follows that neutron scattering amplitudes from any chemical compound need two approximations for obtaining $g(h)$ (different nuclear scattering amplitudes and different thermal displacement parameters), while, for X-ray structure amplitudes $|F(h)| \rightarrow g(h)$, a correction for the 'theta fall-off' of the atomic scattering factor has to be added to the above.

The reduction of measured intensities $I(h)$ into $|F(h)|$ as well as the derivation of standard experimental errors shall not be addressed. Also, the problem of 'absolute scaling' is omitted as it is solved according to §5.2.1 and §5.2.2, and standard error-propagation routines can be applied there. [For analogies of and differences between, for example, 'quasi-normalized' $e(h, n)$-values and the $|E(h)|$ of direct methods, see FKZ IV.]

6.1. Neutron scattering amplitudes

As a first step, the observed amplitudes $|F_{\rm obs}(h)|$ have to be converted into 'zero-Kelvin' data $|F_{{\rm obs},0}(h)|$ by using an 'overall (Debye–Waller) temperature factor' $B_{\rm ov}$ which itself is obtained from 'individual estimated temperature factors' B_j for each atom,

$$\begin{aligned}|F_{{\rm obs},0}(h)| &\simeq |F_{\rm obs}(h)|\exp\big[+B_{\rm ov}\big(\sin^2\theta/\lambda^2\big)\big]\\ &= |F_{\rm obs}(h)|\exp[+B_{\rm ov}({\rm sq1})h^2],\end{aligned} \tag{30}$$

using

$$B_{\rm ov}(h) = -\ln\big[|F_{t,\max}(h)|/|F_{0,\max}(h)|\big]/h^2({\rm sq1}) \tag{31}$$

with

$$|F_{t,\max}(h)| = \textstyle\sum f_j \exp\big[-B_j h^2({\rm sq1})\big] \tag{32}$$

and

$$({\rm sq1}) = \sin^2\theta/\lambda^2 \qquad \text{for reflection } h = 1, \tag{33}$$

and averaging all $B_{\rm ov}(h)$ *de gusto*, *e.g.* by using h^2 for weighting. B_j can be estimated fairly well from chemical experience. Assumed errors $\Delta B_j < B_j$ are transferred into $\Delta B_{\rm ov}$ by conventional error propagation.

Second, converting $|F_{{\rm obs},0}(h)| = 2|\sum_{j=1}^{m} f_j\cos(2\pi h x_j)|$ into $g(h)$ comprises averaging different nuclear scattering amplitudes (here called f_j in analogy to the following section on X-ray data) providing, owing to the 'unitary structure amplitude' $U_{\rm obs}(h) = |F_{{\rm obs},0}(h)|\,/\,2\sum f_j$,

$$\begin{aligned}\Big|\textstyle\sum f_j\cos\big(2\pi h x_j\big)\Big| &= U_{\rm obs}(h)\textstyle\sum f_j\\ &\simeq \big(\textstyle\sum f_j/m\big)\Big|\textstyle\sum\cos\big(2\pi h x_j\big)\Big|\\ &= \big(\textstyle\sum f_j/m\big)g(h).\end{aligned} \tag{34}$$

Thus, a 'socializing' of both f_j and their individual geometrical terms $\cos(2\pi h x_j)$ takes place, indicated by '$\simeq$' in the above expression. Errors introduced by this measure depend, of course, both on the difference(s) of f_j (constant for a given structure in neutron diffraction data) and on the relative share of the individual $\cos(2\pi h x_j)$ within $g(h)$ including their different signs, these shares being different for each reflection h [and their sign differences becoming less important the more $g(h)$ approaches m because for $g(h) = m$ then $\cos(2\pi h x_j)$ must be either all +1 or all −1]. Thus, the errors introduced into $g(h)$ by this 'data reduction' can be more severe the smaller $g(h)$ so that for structures with different f_j any $g(h) \simeq 0$ obtained this way should be used with caution.

6.2. X-ray scattering amplitudes

The 'data reduction' follows the same scheme as for neutron diffraction. The different dependence of $f_j(\sin\theta/\lambda)$ and thus $f_j(h)$ must be observed for each reflection separately in equation (34). The above-mentioned 'socializing' of both f_j and their individual geometrical terms $\cos(2\pi h x_j)$ is different for each h for chemically different atoms. Thus, data-reduction errors not only depend on the relative share of the individual $\cos(2\pi h x_j)$ in $g(h)$ and on the difference(s) of f_j, the latter may also change for each h (owing to different 'size' of atoms) giving rise to an additional source of errors in the resulting $g(h)$.

7. Homometry and quasi-homometry ('false minima')

Two structures are called *homometric* if their Patterson functions, *i.e.* the sets of interatomic distances, coincide. Homometric structures were a fundamental problem for early X-ray crystallography as it raised the question as to whether the results of a structure determination were unique. It proved, however, that in practice homometry was a rather exotic phenomenon. However, dealing with, for example, pseudosymmetric crystal structures, *quasi-homometric* structures had to be considered, *i.e.* structures which cannot be distinguished on the basis of a limited set of reflection amplitudes with limited accuracy.

In these cases, a structure model found by conventional methods is normally refined reaching a single minimum in the χ^2 landscape of the actual parameter space. Without a complete search of this multidimensional space (its asymmetric part, and perhaps reduced to the geometrical parameters), a risk remains that this minimum might be a 'false' one. Any uncertainty is avoided by the parameter-space concept discussed here, as it will provide all possible (geometrical) solutions compatible with the structure amplitudes. (In applying it to one-dimensional projections it should be kept in mind that homometry, being extremely scarce in three-dimensional structures, occurs statistically more frequently upon each reduction of dimension by projection of structures.)

8. Discussion and summary

As mentioned above, this parameter-space concept is based on using Bragg reflection amplitudes [after data reduction to $g(h)$] for assessing one-dimensional structures (or structure projections) without searching for nor applying signs or phases. As a consequence, the primary information which can be obtained consists of interatomic distances, thus similar to that from a Patterson function [see §5.3, (iii), for details]. Compared with a usual Patterson map, our results offer two advantages, however. First, thanks to the point-atom model, the spatial resolution is much higher (being infinite for error-free data in principle) and according to 'old knowledge' any perfectly resolved point Patterson map can always be 'deconvoluted' into the atomic arrangement. Second, a noncentrosymmetric one-dimensional structure with one of its m atoms in the origin automatically provides all possible interatomic distances together with the coordinates, however with $2m$ equivalent solutions in $\mathbf{A}^{m-1}$ [see §5.3, (i), (ii): all solutions come into play during the reconstruction of two- and/or three-dimensional structures].

In centrosymmetric one-dimensional projections, the origin is fixed thanks to the Harker vectors (which are, however, only indirectly contained in the solution), with a choice of two origins at $x = 0$ and/or 1/2 resulting in two 'inverse' solutions contained in $\mathbf{A}^m$ [see §4.1, (iv)].

It may be worth mentioning in this context that the reconstruction of a three-dimensional structure from one-dimensional projections also requires first finding or defining an origin and then preserving it throughout the reconstruction process (as is well known in tomography and other two- or three-dimensional imaging). Details will be dealt with in a separate paper.

Compared with Fourier methods, a few advantages and also some disadvantages exist.

Advantages:

(i) The concept 'as such' appears rather simple, if not naïve.

(ii) The phase problem is avoided.

(iii) The solution is complete: the coordinates of all atoms appear 'at once'.

(iv) Fewer reflections are needed than for a Fourier: neither series-termination effects nor convergence problems will occur.

(v) Even fewer than m reflections suffice for obtaining an approximate solution (see below under applications).

(vi) The region(s) for possible solutions can be strongly reduced by using few reflections, especially those with maximum experimental contrast.

(vii) Zero observations can play an important role in the solution process, in particular by permitting use of optimal experimental contrast.

(viii) The first-order reflection has not necessarily to be used.

(ix) The concept offers high spatial resolution: with at least $m + 1$ (or $3m + 1$) error-free relative amplitudes, the resolution is 'infinite' (with errors, see below and 'Disadvantages').

(x) The solution is either unique or, if not, all possible solutions which satisfy the experimental data are presented at the same time ('homometry', extremely rare in three-dimensional structures; however, statistically increasing for two-dimensional and particularly for one-dimensional structure projections). For data with experimental or other errors ('quasi-homometry'), all 'comparable false solutions' are found at the same time (including, of course, the correct one).

(xi) The procedure is strongly deductive, no probability functions are employed.

(xii) It can also be applied to noncentrosymmetric structures.

(xiii) The phases of reflections defining one-dimensional projections may, after solution by this method, increase an insufficient collection of phases for direct methods (Kirfel's idea).

Disadvantages:

(i) The exponential growth of computing time with the number of atoms reduces the application of the method to rather small structures (see FKZ IV for present limits).

(ii) One of the consequences of the inherent high spatial resolution is a correspondingly high vulnerability to data errors from which the effective resolution may suffer.

(iii) The chemical composition (thus including m) and the occupation of special positions must be known. (The latter may be found by 'trial and error' as described in FKZ I.)

(iv) The ambition to use only few experimental structure amplitudes $|F(h)|$ for structure solution gives a weak statistical basis for the normalization of structure factors to structure amplitudes of point scatterers.

(v) The same argument holds for determination or estimation of the Debye–Waller factor.

(vi) Different atoms with equal or almost equal coordinates will produce overlap in one-dimensional projections resulting often in complications in the structure-search procedure. [If, in a one-dimensional projection, n point atoms (of m) overlap exactly, the solution point X is located at a boundary of $\mathbf{A}^m$ having an *eigensymmetry* of order n while 'almost exact' overlap results in X being close to that boundary.]

(vii) The development of this concept has not yet reached a state that a program system could be conceived permitting complete three-dimensional structure solutions without user interference.

The parameter-space concept describes the well known problem of structure solution as seen from another point of view. Hence, quite a few minor points representing 'trivial old knowledge' will be found in this context (*e.g.* the symmetry in and other features of the structure-factor equation mentioned in §2). This unconventional aspect gives rise to the 'pros and cons' mentioned above. Possible applications thus emerge from its main advantages *in praxi* (which may, of course, also be combined) and may be subdivided into four categories related to the following.

(i) Results can be obtained from rather few data:

(*a*) Incomplete data set (for whatever reason).

(*b*) Even an insufficient number of data (less than m) may lead to an approximate solution region in the parameter space small enough as basis for further work, *e.g.* by least-squares refinement [see FKZ III ('series of g, q or e . . . ') and FKZ IV ('solving one-dimensional test structures . . . ')].

(ii) Higher spatial resolution compared with Fourier methods:

(*a*) Split-atom positions can be found employing only those data providing optimal resolution at the small solution region of the parameter space in question, thus omitting unwanted 'noise' caused by unimportant data, perhaps additionally including a reduction of m (*cf.* test example by Kirfel & Fischer, 2004).

(*b*) A relatively small number of 'heavy atoms', whose partial structure amplitudes have been separated from those of the complete many-atom structure thanks to 'anomalous-dispersion contrast', may be located with a better resolution than by, for example, deconvoluting a Patterson function (to be used after refinement as the basis for completing the structure). Partial structure amplitudes conform well to the equal-atom requirement (save perhaps varied Debye–Waller factors and/or different occupation densities). In case of 'resonant-scattering contrast' one may additionally consider point-like 'atoms'.

(*c*) Some types of so-called 'pseudo-symmetric structures', whose 'solution point' X in $\mathbf{A}^m$ is located, for example, close to an extreme point (§4.2.4) where the gradient of the field of isosurfaces is rather small (for h not too large), may benefit from the resolution power of the parameter-space concept (and also from the next feature).

(iii) All possible solutions are provided:

(*a*) Uniqueness of result may be proven (hitherto not common).

(*b*) All (quasi-)homometric solutions meeting the same data set can be checked using other criteria in order to reduce their number and eventually find a unique one.

(iv) Use of maximal experimental contrast (in analogy to ‘optical’ representation of objects): combining data with $g \simeq m$ and $g \simeq 0$ may already provide (an) approximate solution(s) which appears impossible to obtain by Fourier analysis.

Of course, other parameter spaces may be conceived or are already in use [*e.g.* for locating the centre of a given molecule and finding its orientation parameters; see, for example, Stout & Jensen (1970), p. 340]. In our concept, the high order of permutation symmetry in particular permits the geometry of a (equal-atom) point structure to be determined, within a limit given by present computer technology and perhaps numerical developments coming to our attention in the near future.

We authors want to express our sincere and warm appreciation to Armin Kirfel’s untiring cooperation both as professional colleague and personal friend. He contributed most of the countless test calculations and also numerous key ideas, in particular for the solution praxis. Some of them are mentioned above. We also wish to thank Professor H. Burzlaff, Erlangen-Uttenreuth, for reminding us that neutron scattering data suffer less from data-reduction errors than do X-ray reflection amplitudes. Last but not least we are grateful to an unknown referee whose constructive criticism let us omit some defects and find hints to theoretical literature unknown to us.

References

Avrami, M. (1938). *Phys. Rev.* **54**, 300–303.
Avrami, M. (1939). *Z. Kristallogr.* **100**, 381–393.
Bohm, J. & Dornberger-Schiff, K. (1967). *Acta Cryst.* **23**, 913–933.
Caratheodory, C. (1911). *Circ. Matem. Palermo*, **32**, 93.
Cervellino, A. & Ciccariello, S. (2005). *Acta Cryst.* A**61**, 494–500.
DiPippo, St. & Howe, E. (2000). arXiv:Math/9803097v3.
Fischer, K. F., Kirfel, A. & Zimmermann, H. (2005). *Z. Kristallogr.* **220**, 643–656.
Fischer, K. F., Kirfel, A. & Zimmermann, H. (2008). *Croat. Chem. Acta*, **81**, 381–389.
Fischer, K. F. & Pilz, K. (1997). *Acta Cryst.* A**53**, 475–483.
Kirfel, A. & Fischer, K. F. (2004). *Z. Kristallogr. Suppl.* **21**, 101.
Kirfel, A. & Fischer, K. F. (2005). *Z. Kristallogr. Suppl.* **22**, 101.
Kirfel, A. & Fischer, K. F. (2007). *Abstr. DGK Meeting*, 199-07-id61.
Kirfel, A. & Fischer, K. F. (2009). *Z. Kristallogr.* **224**, 325–339.
Kirfel, A., Fischer, K. F. & Zimmermann, H. (2006). *Z. Kristallogr.* **221**, 673–680.
Knof, W. (1989). PhD thesis, University of Saarbruecken, Germany.
Kutschabsky, L., Kulpe, S. & Reck, G. (1971). *Acta Cryst.* A**27**, 29–35.
Navaza, J. & Silva, A. M. (1979). *Acta Cryst.* A**35**, 266–275.
Ott, H. (1927). *Z. Kristallogr.* **66**, 136–153.
Pilz, K. (1996). PhD thesis, , University of Saarbruecken, Germany.
Pilz, K. & Fischer, K. F. (1998). *Acta Cryst.* A**54**, 273–282.
Pilz, K. & Fischer, K. F. (2000). *Z. Kristallogr.* **215**, 640–649.
Stout, G. H. & Jensen, L. H. (1970). *X-ray Structure Determination.* New York: Wiley.
Zimmermann, H., Kirfel, A. & Fischer, K. F. (2005). *Z. Kristallogr. Suppl.* **22**, 122.
Zimmermann, H., Kirfel, A. & Fischer, K. F. (2006). *Z. Kristallogr. Suppl.* **23**, 107.