Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 30 April 2005
Accepted 23 June 2005

The algebraic approach to the phase problem

A. Cervellino[a]‡ and S. Ciccariello[b]*

[a]Laboratory for Neutron Scattering, PSI Villigen and ETH Zurich, CH-5232 Villigen PSI, Switzerland, and [b]Università di Padova, Dipartimento di Fisica ''G. Galilei'' and Unità INFM, Via Marzolo 8, I-35131 Padova, Italy. Correspondence e-mail: salvino.ciccariello@pd.infn.it

A rather detailed report is presented on the present status of the algebraic approach to the phase problem in the case of an ideal crystal in order to make clear that some points must still be proven for it to apply to neutron scattering. To make this extension, the most important results that were previously obtained in the case of X-ray scattering are derived again by a different procedure. By so doing, the three-dimensional case is treated explicitly, the polynomial equations in a single variable whose roots determine the positions of the scattering centres are explicitly reported and the procedure is shown to generalize to neutron scattering, overcoming the difficulty related to the non-positivity of the scattering density. In this way, it is fully proven that the atomicity assumption removes the phase ambiguity in the sense that the full diffraction pattern of an ideal crystal can uniquely be reconstructed from a suitable finite portion of it in both X-ray and neutron scattering. The procedures able to isolate these portions that contain the pattern's full information are also given.

1. Introduction

The main problem of crystallography consists in determining the electron or the scattering-length density of a crystal knowing its X-ray or neutron diffraction pattern as well as its chemical composition. The problem is brought to its essence by assuming that the atoms are point-like objects, which is known as the atomicity assumption. Under this assumption, the unknown quantities to be extracted from scattering data are the positions of the atoms present in the unit cell. The number of atoms being finite, the number of unknown quantities that determine the scattering pattern is also finite. Consequently, on the basis of implicit function theory, the knowledge of a (sufficiently large) portion of the diffraction pattern should yield a number of equations sufficient to determine the positions of the atoms within the unit cell (see *e.g.* Hauptman & Karle, 1950). One concludes that the phase problem is solvable under the atomicity assumption and that the full diffraction pattern can be reconstructed from a finite portion of it (Sayre, 2002). For this claim to be considered a mathematical statement, it is necessary: to know a procedure able to single out from the observed pattern a set of equations having a Jacobian different from zero; to find all the different solutions from the isolated set of equations; and to show that each of these solutions reproduces the full diffraction pattern. As these three points have not been fully clarified, the aforesaid claim can only be considered a conjecture. In fact, this conjecture is so widely shared by the crystallographic community as to give the impression that it is a well established mathematical result. In reality, the full proof of the conjecture was recently obtained by Cervellino & Ciccariello (1996, 2001) for the case of X-ray scattering only (these two papers will be referred to as I and II in the following).

The aim of this paper is to generalize the proof to the case of neutron scattering. The proof is based on the so-called algebraic approach put forth by Ott (1927) a long time ago. The plan of the paper is as follows. The next section (§2) reviews Ott's algebraic approach and its generalization, obtained in the late seventies, on the basis of the vectorial space formalism of Goedkoop (1950). By so doing, we point out what must still be proven for the conjecture to become a mathematical property. Since the analysis of these points look rather mathematical, at least to practical crystallographers, it is reported in the companion deposited part of this paper.[1] Hence in the last section (§3), we simply detail our results and draw our final conclusions calling attention to one point that, in our opinion, deserves further investigation because it could be practically relevant. Concerning the deposited part, in its first section (§1), we report a proof of the conjecture different from that given in papers I and II. This proof explicitly deals with the three-dimensional case and gives the polynomial equations in a single variable to be solved in order to determine the coordinates of the scattering centres. In §2, we generalize the proof to the case of neutron scattering. This

---

‡ On leave from Istituto di Cristallografia (CNR-IC), Via Amendola 122/O, I-70126 Bari, Italy.

[1] Supplementary data for this paper, including Appendices *A* and *B*, are available from the IUCr electronic archives (Reference: AU0349). Services for accessing these data are described at the back of the journal.

generalization requires that some of the results reported in I and II are obtained under weaker assumptions. These technical points are left to deposited Appendices *A* and *B*.

2. The algebraic approach

In the case of ideal crystals (*i.e.* infinitely large crystals made up of atoms forming, in real space, a mathematical lattice denoted as $\mathcal{Z}^3$), the observed scattering intensities $I_{\mathbf{h}}$, leaving aside some normalization factors, are simply equal to the square moduli of the so-called structure factors $F_{\mathbf{h}}$. These are the Fourier transform values (in appropriate units) of the unit-cell scattering density (*i.e.* the electron or the scattering-length density depending on whether one respectively considers X-ray or neutron scattering), associated with a set of scattering vectors $\mathbf{h}$ that form a lattice in reciprocal space still denoted by $\mathcal{Z}^3$. The *atomicity* assumption states that the atoms can be treated as point-like objects. Hence, $F_{\mathbf{h}}$ takes the simple expression

$$F_{\mathbf{h}} = \sum_{j=1}^{N} \hat{\mathcal{Z}}_j \exp(i2\pi\mathbf{h}\cdot\mathbf{r}_j), \tag{1}$$

where N is the number of atoms in the unit cell and $\hat{\mathcal{Z}}_j$ denotes the electron number or the scattering length of the jth atom located at the point having position vector $\mathbf{r}_j$. The knowledge of the chemical composition of the sample ensures that N and the $\hat{\mathcal{Z}}_j$s are known. On the contrary, the knowledge of the $F_{\mathbf{h}}$ values requires the knowledge of $\mathbf{r}_1, \ldots, \mathbf{r}_N$. Equation (1) specifies however the dependence of $F_{\mathbf{h}}$ on the atomic position vectors so that its left hand side (l.h.s.) should more completely be written as $F_{\mathbf{h}}(x_1, y_1, z_1, \ldots, x_N, y_N, z_N)$. Assume for a moment one knows the value of the structure factor (s.f.) relevant to a particular reflection $\mathbf{h}$. Then, one has the complex equation

$$F_{\mathbf{h}}(x_1, y_1, z_1, \ldots, x_N, y_N, z_N) = F_{\mathbf{h}}. \tag{2}$$

This corresponds to two real equations that, on the basis of implicit function theory, determine two of the variables $x_1, \ldots, z_N$ in terms of the remaining ones. Hence, if one knows the s.f. values relevant to m different reflections and $m \geq 3N/2$, in principle all the coordinates of the atomic positions could be determined [Karle & Hauptman (1950) and, more recently, Fischer *et al.* (2005)]. For this to happen, it is necessary that the equations considered in (2) be functionally independent. Therefore, the m considered reflections must be properly chosen. Moreover, one must also find an algorithm able to find the solution because the implicit function theory only gives the conditions for the existence of a local solution.

2.1. Ott's formulation

Ott's (1927) algebraic approach gives a procedure able to overcome the second and, to some extent, the first of the aforesaid difficulties. The main feature of Ott's analysis can more simply be presented by restricting ourselves to the one-dimensional case. Thus, we shall consider the reflections $\mathbf{h} = h00$ that lie on axis $\mathbf{a}^*$ of reciprocal space. The associated s.f.s read $F_{\mathbf{h}} = F_h = \sum_{j=1}^{N} \hat{\mathcal{Z}}_j \exp(i2\pi h x_j)$. For greater simplicity, we also assume that $x_1 \neq x_2 \neq \ldots \neq x_N$. After putting $\xi_j \equiv \exp(i2\pi x_j)$, the ξ_js form a set of N different unimodular complex numbers that can be seen as the roots of the polynomial equation

$$P(z) = \prod_{j=1}^{N}(z - \xi_j) = z^N + a_1 z^{N-1} + a_2 z^{N-2} + \ldots + a_N = 0, \tag{3}$$

whose coefficients a_j are related to the roots ξ_j by the relations

$$a_j = (-1)^j \sum_{1\leq l_1 < l_2 < \ldots < l_j \leq N} \xi_{l_1}\xi_{l_2}\ldots\xi_{l_j}, \quad j = 1, \ldots, N, \tag{4}$$

and are therefore independent of the $\hat{\mathcal{Z}}_j$s. If we know the coefficients a_j, the ξ_js can simply be obtained by solving equation (3), which will be called the *resolvent* equation. Ott showed how to determine the a_js from an appropriate set of $F_{\mathbf{h}}$s using a suggestion of Carathéodory (1911). In fact, he observed that the condition $P(\xi_j) = 0$ can be written as

$$\xi_j^N = -a_N - a_{N-1}\xi_j - a_{N-2}\xi_j^2 - \ldots - a_1\xi_j^{N-1} \tag{5}$$

and, after multiplying it by ξ_j^p with p integer, as

$$\begin{aligned} \xi_j^{N+p} &= -a_N\xi_j^p - a_{N-1}\xi_j^{p+1} - a_{N-2}\xi_j^{p+2} - \ldots - a_1\xi_j^{p+N-1}, \\ &\qquad j = 1, 2, \ldots, N, \quad p = 0, \pm1, \pm2, \ldots \end{aligned} \tag{6}$$

The substitution of these equalities in the F_{N+p} expression (1) gives

$$\begin{aligned} F_{N+p} &= -a_N F_p - a_{N-1}F_{p+1} - a_{N-2}F_{p+2} - \ldots - a_1 F_{p+N-1}, \\ &\qquad p = 0, \pm1, \pm2, \ldots \end{aligned} \tag{7}$$

By letting p range over the set of values $[(-N+1), (-N+2), \ldots, 0]$, one obtains the system of linear equations

$$\begin{array}{llllll} a_1F_0 & +a_2F_{-1} & +a_3F_{-2} & +\ldots+a_NF_{-N+1} & = -F_1 \\ a_1F_1 & +a_2F_0 & +a_3F_{-1} & +\ldots+a_NF_{-N+2} & = -F_2 \\ & & \vdots & & \vdots \\ a_1F_{N-1} & +a_2F_{N-2} & +a_3F_{N-3} & +\ldots+a_NF_0 & = -F_N, \end{array} \tag{8}$$

having the coefficients of the resolvent equation (3) as its solution. One concludes that the knowledge of the 'first' $(N+1)$ s.f.s $F_0, F_1, \ldots, F_N$ (the reason why we do not consider the s.f.s relevant to the negative reflections will become clear in a moment) determines: the coefficients $a_1, \ldots, a_N$ of the resolvent equation by solving equations (8); the ξ_js by solving the resulting resolvent equation; and, finally, the $\hat{\mathcal{Z}}_j$s by solving the linear equation system

$$\begin{aligned} \sum_{j=1}^{N} \hat{\mathcal{Z}}_j &= F_0 \\ \sum_{j=1}^{N} \hat{\mathcal{Z}}_j \xi_j &= F_1 \\ \vdots &= \vdots \\ \sum_{j=1}^{N} \hat{\mathcal{Z}}_j \xi_j^{N-1} &= F_{N-1}. \end{aligned} \tag{9}$$

Four important consequences of this result are worth stressing. Equation (7) implies that: (i) each F_{N+m}, with $m \geq 0$, is a linear combination of its N preceding s.f.s; (ii) the coefficients of the linear combination are independent of m and are simply related to those of the resolvent equation; (iii) each F_{N+m} with $m \geq 0$ can also be expressed as a linear combination of the 'first' N s.f.s (*i.e.* $F_0, F_1, \ldots, F_{N-1}$) and the coefficients involved in this new linear combination depend on $a_1, \ldots, a_N$ in a way determined by the value of m.[2] This amounts to saying that all the s.f.s F_{N+m} with $m \geq 0$ can recursively be determined from $F_0, F_1, \ldots, F_{N-1}$. Moreover, the s.f.s associated with negative reflections are known in terms of those associated with positive reflections thanks to the Friedel property $F_{-m} = \bar{F}_m$, where the overbar denotes the complex conjugate. One concludes that (iv) all the s.f.s relevant to the reflections lying on the chosen crystallographic axis are known linear combinations of $F_0, F_1, \ldots, F_{N-1}$.[3] Thus, these are the *basic* s.f.s, in the sense that the diffraction pattern along the considered crystallographic axis is determined and can be reconstructed from its portion related to reflections $0, 1, \ldots, (N-1)$ and to reflection N. The reason why one has to include the last reflection follows from equation (8). In fact, this system is determined if all the s.f.s $F_0, F_1, \ldots, F_N$ are known because these determine the remaining F_{-m}s by the Friedel relation. In conclusion, the knowledge of the first $(N+1)$ s.f.s fully solves the problem of reconstructing the one-dimensional structure of the crystal under analysis because it determines the basic coefficients $a_1, \ldots, a_N$ defining the resolvent equation, all the atomic positions x_j by the resolvent equation and, finally, all the $\hat{\mathcal{Z}}_j$s by system (9).

These results show the power and the elegance of Ott's algebraic method although the involved mathematical analysis is far from being complete for two reasons. First, the s.f.s are not the experimentally known quantities. Second, the one-dimensionality assumption must be removed. The restrictiveness of this assumption is made evident by the remark that the number of different x projections of the spatial configuration of the N atoms, present in the unit cell, generally is not equal to N, although it cannot exceed this value. Hence, it is an unknown integer that must be determined by the solution procedure. This important point will be analysed in the following sections. We conclude this subsection by completing the presentation of the algebraic approach along Ott's line.

In fact, we shall now discuss how Avrami (1938) generalized Ott's approach so that this involves observable quantities instead of s.f.s and becomes a mathematical algorithm. Based partly on a paper by Patterson (1935), Avrami worked out from the experimentally known $I_{\mathbf{h}}$s other quantities that have the algebraic form of a s.f. To this aim, starting from the $I_{\mathbf{h}}$'s definition, one writes

$$I_{\mathbf{h}} = |F_{\mathbf{h}}|^2 = \sum_{j=1}^{N} \hat{\mathcal{Z}}_j^2 + \sum_{1 \leq j \neq k \leq N} \hat{\mathcal{Z}}_j \hat{\mathcal{Z}}_k \exp[i2\pi \mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_k)] \tag{10}$$

and, after setting

$$\mathcal{I}_{\mathbf{h}} \equiv I_{\mathbf{h}} - \sum_{j=1}^{N} \hat{\mathcal{Z}}_j^{\,2}, \tag{11}$$

one can say that the $\mathcal{I}_{\mathbf{h}}$s are known quantities since they are obtained by subtracting the known term $\sum_{j=1}^{N} \hat{\mathcal{Z}}_j^2$ from the observable $I_{\mathbf{h}}$s. Denoting by l the pair of index values (j, k) with $j \neq k$, one puts $\mathbf{R}_l \equiv \mathbf{r}_j - \mathbf{r}_k$ and $f_l \equiv \hat{\mathcal{Z}}_j \hat{\mathcal{Z}}_k$, and from (10) and (11) one finds that

$$\mathcal{I}_{\mathbf{h}} = \sum_{l=1}^{N(N-1)} f_l \exp(i2\pi \mathbf{h} \cdot \mathbf{R}_l). \tag{12}$$

In this way, the algebraic structure of 'observable' $\mathcal{I}_{\mathbf{h}}$s is similar to that of s.f.s and – as Avrami first found out – Ott's procedure can be applied to them in order to determine the $\mathbf{R}_l$s and the f_ls. Strictly speaking, the defined $\mathcal{I}_{\mathbf{h}}$s cannot however be considered s.f.s because some of the involved $\mathbf{R}_j$s do not lie within the unit cell. Hence, in order to write $\mathcal{I}_{\mathbf{h}}$ as a s.f. one must proceed as follows. Each vector $(\mathbf{r}_j - \mathbf{r}_k)$ is brought within the unit cell by adding to it a vector $\mathbf{m}_{j,k}$ with components equal to 0 or -1, so that each component of the vector $\boldsymbol{\delta}$, defined as

$$\boldsymbol{\delta} = \mathbf{r}_j - \mathbf{r}_k + \mathbf{m}_{j,k}, \tag{13}$$

is non-negative and smaller than one. As (j, k) runs over its $N(N-1)$ values, we label by $\hat{\jmath}$ the different $\boldsymbol{\delta}$s defined by (13) and we denote by $\overline{\mathcal{N}}'$ the number of different $\boldsymbol{\delta}_{\hat{\jmath}}$s. Moreover, we denote by $\mathcal{L}_{\hat{\jmath}}$ the set of pairs (j, k) such that $(\mathbf{r}_j - \mathbf{r}_k)$ defines the same $\boldsymbol{\delta}_{\hat{\jmath}}$ after applying (13). Then, the second sum on the right-hand side (r.h.s.) of (10) becomes

$$\sum_{\hat{\jmath}=1}^{\overline{\mathcal{N}}'} \exp(i2\pi \mathbf{h} \cdot \boldsymbol{\delta}_{\hat{\jmath}}) \sum_{(j,k) \in \mathcal{L}_{\hat{\jmath}}} \hat{\mathcal{Z}}_j \hat{\mathcal{Z}}_k. \tag{14}$$

After setting

---

[2] This property is immediately proven by recursion. The property is true for F_N. Consider the next s.f. F_{N+1}. Owing to (7), F_{N+1} is a linear combination of $F_1, F_2, \ldots, F_{N-1}, F_N$ and it becomes a linear combination of $F_0, F_1, \ldots, F_{N-1}$ using the linear dependence of F_N in terms of $F_0, F_1, \ldots, F_{N-1}$. The coefficients of the final combination will depend on those present in (7) and on those that relate F_N to $F_0, F_1, \ldots, F_{N-1}$ and will therefore depend on $a_1, \ldots, a_N$ in a way determined by m (equal to 1 in this case). In this way, step by step, one determines the linear expression of any F_{N+m}, with $m > 0$, in terms of $F_0, F_1, \ldots, F_{N-1}$.

[3] In this respect, it is not useless to observe that, after setting $p = -m$ in equation (7), this can be written as $F_{-m} = [-a_{N-1}F_{-m+1} - a_{N-2}F_{-m+2} - \ldots - a_1 F_{-m+N-1} - F_{-m+N}]/a_N$. Owing to the fact that $a_N \neq 0$, one concludes that the s.f. relevant to a negative reflection is a linear combination of its next N s.f.s. By recurrence, similarly to what we did before, one shows that F_{-m} is a linear combination of $F_0, F_1, \ldots, F_{N-1}$. Combining this relation with the Friedel relation, one finds that coefficients a_j must obey a set of constraints, not unexpected because resolvent equation (3) must be such that its roots are all distinct and unimodular.

$$\nu_{\hat{j}} \equiv \sum_{(j,k)\in\mathcal{L}_{\hat{j}}} \hat{\mathcal{Z}}_j \hat{\mathcal{Z}}_k, \quad \hat{j} = 1, 2, \ldots, \overline{\mathcal{N}}', \tag{15}$$

equation (12) becomes

$$\mathcal{I}_{\mathbf{h}} = \sum_{\hat{j}=1}^{\overline{\mathcal{N}}} \nu_{\hat{j}} \exp(i2\pi\mathbf{h}\cdot\boldsymbol{\delta}_{\hat{j}}), \tag{16}$$

where $\overline{\mathcal{N}}$ is the number of $\nu_{\hat{j}}$s different from zero. It is stressed that these formulae apply in both X-ray and neutron scattering. In the second case, some of the $\nu_{\hat{j}}$s can be negative and $\overline{\mathcal{N}}$ can be smaller than $\overline{\mathcal{N}}'$ because the negativeness of some $\hat{\mathcal{Z}}_j$s can make some of the $\nu_{\hat{j}}$s, defined by (15), equal to zero. [In the last case, it is assumed that index $\hat{j}$ is assigned in such a way that $\nu_{\hat{j}} = 0$ if $\hat{j} = (\overline{\mathcal{N}}+1), \ldots, \overline{\mathcal{N}}'$.] Moreover, equation (16) shows that the 'subtracted' peak intensities $\mathcal{I}_{\mathbf{h}}$, defined by equation (11), are the Fourier transforms of the scattering density relevant to the infinitely resolved Patterson (1939) map

$$\rho_{\rm Pat}(\mathbf{r}) = \sum_{\hat{j}=1}^{\overline{\mathcal{N}}} \nu_{\hat{j}} \delta(\mathbf{r} - \boldsymbol{\delta}_{\hat{j}}), \tag{17}$$

corresponding to a set of $\overline{\mathcal{N}}$ scattering centres located at $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_{\overline{\mathcal{N}}}$ with 'charges' $\nu_1, \ldots, \nu_{\overline{\mathcal{N}}}$. This scattering density generally is not a positive density in the case of neutron scattering owing to the possible negativeness of some $\nu_{\hat{j}}$s. Even though each subtracted intensity is mathematically well represented by both (12) and (16), only the second has the algebraic expression of a s.f. Hence, the algebraic approach by Ott must be applied using algebraic expression (16) for subtracted intensities with the consequence that we do not know the value of $\overline{\mathcal{N}}$, *i.e.* the number of terms involved in the sum defining a s.f., as well as the values of the charges $\nu_{\hat{j}}$. [Note that $\overline{\mathcal{N}}$ and the $\nu_{\hat{j}}$s would have been known whenever we were allowed to use (12)]. Besides this complication, the remaining difficulty in the application of the Ott–Avrami approach consists in removing the one-dimensionality assumption. This will be done in the following sections. In closing this subsection, we stress that the application of the algebraic approach to the $\mathcal{I}_{\mathbf{h}}$s only determines $\overline{\mathcal{N}}$, the $\boldsymbol{\delta}_{\hat{j}}$s and the $\nu_{\hat{j}}$s or, equivalently, the quantities defining the density of the infinitely resolved Patterson map (17). The determination of the atomic positions requires the subsequent deconvolution of this map, *i.e.* to solve equations (13) and (15). In principle this can be done by a finite number of mathematical operations (see *e.g.* Appendix *E* of I). In this way, all the atomic configurations that reproduce the observed diffraction pattern are determined.

2.2. Vectorial space formulation

In the early fifties, an important series of papers pointed out the existence of inequality relations among the s.f.s relevant to different reflections (Harker & Kaspar, 1948; Karle & Hauptman, 1950; Sayre, 1952; Hauptman & Karle, 1952). In most cases, these relations were obtained by exploiting the positiveness of the scattering density since the authors explicitly analysed the case of X-ray scattering. Goedkoop (1950), investigating in more detail the inequalities obtained by Karle & Hauptman (1950), introduced a vector-space formalism very useful to reformulate the phase problem for the X-ray scattering from a crystal. Now we show how Ott's formulation can be obtained by the vectorial space formalism. Following Goedkoop, if the crystal contains N atoms in its unit cell, one considers an N-dimensional Hilbert space, denoted by $\mathcal{H}(N)$, and one associates with each reflection $\mathbf{h}$ of $\mathcal{Z}^3$ a vector of $\mathcal{H}(N)$, denoted by $|\mathbf{h}\rangle$ and defined as follows[4]

$$|\mathbf{h}\rangle \equiv \sum_{j=1}^{N} \hat{\mathcal{Z}}_j^{1/2} \exp(i2\pi\mathbf{h}\cdot\mathbf{r}_j)|e_j\rangle. \tag{18}$$

The set of vectors $|\mathbf{h}\rangle$, with $\mathbf{h} \in \mathcal{Z}^3$, forms a lattice of vectors, denoted as $\mathcal{G}(N)$ and referred to as Goedkoop's lattice. Because $\langle e_i|e_j\rangle = \delta_{i,j}$,[4] one immediately verifies that

$$\begin{aligned}\langle\mathbf{k}|\mathbf{h}\rangle &= \sum_{i,j=1}^{N} \overline{\hat{\mathcal{Z}}_i^{1/2} \exp(i2\pi\mathbf{k}\cdot\mathbf{r}_i)} \hat{\mathcal{Z}}_j^{1/2} \exp(i2\pi\mathbf{h}\cdot\mathbf{r}_j)\langle e_i|e_j\rangle \\ &= \sum_{i=1}^{N} \hat{\mathcal{Z}}_i \exp[i2\pi(\mathbf{h}-\mathbf{k})\cdot\mathbf{r}_i] \\ &= F_{\mathbf{h}-\mathbf{k}}.\end{aligned} \tag{19}$$

Equation (19) shows that the set of scalar products of any pair of vectors of $\mathcal{G}(N)$ coincides with the set of s.f. values. It is now remarked that the algebraic manipulations reported in equation (19) are correct only if all the $\hat{\mathcal{Z}}_j$s are positive so that definition (18) only works in the case of X-ray scattering. The geometrical structures of $\mathcal{H}(N)$ and $\mathcal{G}(N)$ reflect into a series of relations that must be obeyed by the s.f.s. These relations were mainly used to reduce the phase ambiguities of the s.f.s, the latter moduli being experimentally known (Castellano *et al.*, 1973; de Rango *et al.*, 1974; Podjarny *et al.*, 1976; Knossow *et al.*, 1977; Navaza & Silva, 1979; Rothbauer, 1970, 1974, 1980; Silva & Navaza, 1981; Podjarny, 1981).[5]

However, as was first found by Navaza & Silva, they also allow us to recover the results of Ott described in the previous subsection. To show this point, it is first observed that $\mathcal{G}(\overline{\mathcal{N}})$ is a subset of $\mathcal{H}(N)$ and it exactly contains N linearly independent vectors (see Appendix *A* of I). Consider now the vectors of $\mathcal{G}(N)$ associated with a set of M reflections $\mathcal{B}_M \equiv \{\mathbf{h}_1, \ldots, \mathbf{h}_M\}$ and denote by $(\mathcal{D}_M)$ the $M\times M$ matrix whose (i,j) element is $\mathcal{D}_{M;i,j} = \langle\mathbf{h}_i|\mathbf{h}_j\rangle = F_{\mathbf{h}_j-\mathbf{h}_i}$. The matrix $(\mathcal{D}_M)$ is known as the Karle–Hauptman (KH) matrix generated by the vectors $|\mathbf{h}_1\rangle, \ldots, |\mathbf{h}_M\rangle$ or by the reflection set $\mathcal{B}_M$.[6] The positivity of the metric of the Hilbert space implies that $\det(\mathcal{D}_M) \geq 0$ and, more precisely, that the vectors

---

[4] We find it convenient to adopt here the quantum-mechanical notation $|\cdot\rangle$ and $\langle\cdot|\cdot\rangle$ for vectors and scalar products. Moreover, the $|e_j\rangle$s with $j = 1, \ldots, N$ define an orthonormal complete basis of $\mathcal{H}(N)$ in the sense that $\langle e_i|e_j\rangle = \delta_{i,j}$ ($\delta_{i,j}$ being Kronecker's symbol) and that each vector $|v\rangle$ of $\mathcal{H}(N)$ can be written as $|v\rangle = \sum_{j=1}^{N} v_j|e_j\rangle$ with $v_j = \langle e_j|v\rangle$.

[5] Navaza & Silva and Rothbauer also exploited, along the line first considered by Goedkoop on the basis of Bertaut's (1958) results, the constraints arising from possible crystallographic symmetries that generally lead to a resolvent equation of lower degree.

[6] A KH matrix has the property of remaining invariant if it is evaluated using the vectors associated with the set of reflections $\mathbf{h}_1+\mathbf{p}, \mathbf{h}_2+\mathbf{p}, \ldots, \mathbf{h}_M+\mathbf{p}$ obtained by translating the set $\mathcal{B}_M$ by an arbitrary reflection $\mathbf{p}$. This property will be referred to as the translation invariance of $(\mathcal{D})$ and will be very useful later.

$|\mathbf{h}_1\rangle, \ldots, |\mathbf{h}_M\rangle$ are linearly independent if $\det(\mathcal{D}_M) \neq 0$ and linearly dependent if $\det(\mathcal{D}_M) = 0$. By extension, in the two cases, reflections $\mathbf{h}_1, \ldots, \mathbf{h}_M$ will be said to be linearly independent and dependent, respectively. Assume that $\det(\mathcal{D}_M) = 0$ and that the rank of $\mathcal{D}_M$ be $M - 1$. Then, $|\mathbf{h}_M\rangle$ can be written as

$$|\mathbf{h}_M\rangle = \sum_{j=1}^{M-1} \alpha_{\mathbf{h}_M,j} |\mathbf{h}_j\rangle \tag{20}$$

and the $\alpha_{\mathbf{h}_M,j}$s are the solutions of the system of linear equations

$$F_{\mathbf{h}_M - \mathbf{h}_i} = \sum_{j=1}^{M-1} \alpha_{\mathbf{h}_M,j} F_{\mathbf{h}_j - \mathbf{h}_i} = \sum_{j=1}^{M-1} \mathcal{D}_{M-1;i,j} \alpha_{\mathbf{h}_M,j}, \quad i = 1, \ldots, M-1, \tag{21}$$

resulting from the scalar products of (20) with $|\mathbf{h}_1\rangle, \ldots, |\mathbf{h}_{M-1}\rangle$, while, on the r.h.s. of (21), $\mathcal{D}_{M-1;i,j}$ denotes the (i, j)th element of $(\mathcal{D}_{M-1})$, the principal minor of $(\mathcal{D}_M)$ obtained by discarding the last row and the rightmost column of $(\mathcal{D}_M)$. The solution of (21) is

$$\alpha_i = \sum_{j=1}^{M-1} \mathcal{D}^{-1}_{M-1;i,j} F_{\mathbf{h}_M - \mathbf{h}_j}, \quad i = 1, \ldots, M-1. \tag{22}$$

If we translate the previous set $\mathcal{B}_M$ by an arbitrary reflection $\mathbf{p}$, we obtain the new set of linearly dependent reflections $\mathcal{B}_{M,\mathbf{p}} \equiv \{\mathbf{h}_1 + \mathbf{p}, \ldots, \mathbf{h}_M + \mathbf{p}\}$ owing to the property reported in footnote **5**. Then, equation (20) becomes

$$|\mathbf{h}_M + \mathbf{p}\rangle = \sum_{j=1}^{M-1} \alpha_{\mathbf{h}_M,j} |\mathbf{h}_j + \mathbf{p}\rangle. \tag{23}$$

The scalar product of equation (20) by $|\mathbf{h}_M + \mathbf{k}\rangle$, $\mathbf{k}$ being an arbitrary reflection, yields

$$F_{\mathbf{k}} = \sum_{j=1}^{M-1} \alpha_{\mathbf{h}_M,j} F_{\mathbf{k}-(\mathbf{h}_M - \mathbf{h}_j)}. \tag{24}$$

Finally, the scalar product of (20) by $|e_s\rangle$, the orthonormality of the $|e_j\rangle$s and (18) give

$$\exp(i2\pi \mathbf{h}_M \cdot \mathbf{r}_s) - \sum_{j=1}^{M-1} \alpha_{\mathbf{h}_M,j} \exp(i2\pi \mathbf{h}_j \cdot \mathbf{r}_s) = 0, \qquad s = 1, 2, \ldots, N. \tag{25}$$

Putting

$$\xi_s \equiv \exp(i2\pi x_s), \quad \eta_s \equiv \exp(i2\pi y_s), \quad \zeta_s \equiv \exp(i2\pi z_s), \qquad s = 1, \ldots, N, \tag{26}$$

and denoting the ath component of $\mathbf{h}_m$ by $h_{m;a}$ (with $a = 1, 2, 3$), equation (25) becomes

$${\xi_s}^{h_{M;1}} {\eta_s}^{h_{M;2}} {\zeta_s}^{h_{M;3}} - \sum_{j=1}^{M-1} \alpha_{\mathbf{h}_M,j} \xi_s^{h_{j;1}} \eta_s^{h_{j;2}} \zeta_s^{h_{j;3}} = 0, \quad s = 1, \ldots, N. \tag{27}$$

Equation (27) shows that the (ξ_s, η_s, ζ_s)s must be chosen among the roots of the following polynomial equation

$$P(x, y, z) \equiv x^{h_{M;1}} y^{h_{M;2}} z^{h_{M;3}} - \sum_{j=1}^{M-1} \alpha_{\mathbf{h}_M,j} x^{h_{j;1}} y^{h_{j;2}} z^{h_{j;3}} = 0, \tag{28}$$

involving now three variables and determined by the same α_js that are present in (20) and (22) and are determined by the linearly dependent reflection set $\mathcal{B}_M$. We already stressed that $\mathcal{G}(N)$ contains N linearly independent vectors. Assuming we know a set of such vectors denoted by $|\mathbf{h}_1\rangle, \ldots, |\mathbf{h}_N\rangle$, the set $\mathcal{B}_N \equiv \{\mathbf{h}_1, \ldots, \mathbf{h}_N\}$ of the associated reflections is a set of N linearly independent reflections. Then any set $\mathcal{B}_{N+1}$ obtained by 'enlarging' $\mathcal{B}_N$ to contain a further reflection $\mathbf{h}_{N+1}$ different from $\mathbf{h}_1, \ldots, \mathbf{h}_N$ is a set of linearly dependent reflections. Hence, equations(20)–(28) hold true with $M = (N + 1)$. These equations represent a partial generalization of the results reported in items (i)–(iii) of §2.1. In fact, equation (24) shows that, once a KH matrix with determinant equal to zero has been found, each s.f. can be expressed as a linear combination, with fixed coefficients, of $(M - 1)$ s.f.s related, as specified on the r.h.s. of equation (24), to the reflections that generate the singular KH matrix. This property generalizes items (i) and (ii) to the three-dimensional case. At the same time, from equation (20) with $M = N + 1$, it follows that

$$|\mathbf{h}\rangle = \sum_{j=1}^{N} \alpha_{\mathbf{h},j} |\mathbf{h}_j\rangle \tag{29}$$

and

$$F_{\mathbf{h}} = \sum_{j=1}^{N} \alpha_{\mathbf{h},j} F_{\mathbf{h}_j}. \tag{30}$$

The above equation implies that each s.f. $F_{\mathbf{h}}$ is a linear combination, with coefficients dependent on $\mathbf{h}$, of N fixed s.f.s. This property is the three-dimensional generalization of the first part of item (iii). In conclusion, equations (20)–(30), first obtained by Navaza & Silva (1979) within the vectorial formalism approach, generalize most of Ott's results to the three-dimensional case, since they coincide with equations (3), (7) and (8) when one restricts oneself to the one-dimensional case. It is however underlined that, while properties (i), (ii) and the first part of (iii) have been generalized to the three-dimensional case, nothing has been said about the second part of (iii) and property (iv), *i.e.* the reconstruction of the full diffraction pattern from a finite portion of it. At the same time, no statement was made in the three-dimensional case about the roots of the system of polynomial equations of the kind (28) that emerges by considering different and non-superposable sets of linearly dependent reflections. The two questions are intimately related. In fact, if we are able to reconstruct the full diffraction pattern from a portion of it, then we can determine all the s.f.s and, by these, the scattering density of the unit cell through the relation

$$\rho_{\rm cell}(\mathbf{r}) = \sum_{\mathbf{h} \in cZ^3} F_{\mathbf{h}} \exp(-i2\pi \mathbf{h} \cdot \mathbf{r}) = \sum_{j=1}^{N} \mathcal{Z}_j \delta(\mathbf{r} - \mathbf{r}_j).$$

In this way, all the scattering centre positions will be uniquely determined and it will be possible to get a system of polynomial equations of the kind (28) having the atomic positions as unique roots. Hence, the two questions: how to choose a finite portion of the diffraction pattern that allows us to reconstruct the full pattern and how to single out a set of polynomial equations having as unique root the set of atomic positions? These two questions were not answered by Navaza & Silva (1979), Navaza & Navaza (1992) or Rothbauer (1998) in his more recent polynomial approach to solve the phase problem for a crystal with point-like atoms.[7]

So far, we have considered the case of X-ray scattering. The extension of these results to the case of neutron scattering is possible proceeding along the line considered by Navaza & Navaza (1992) and Rothbauer (1998). Very briefly, one introduces now two Goedkoop lattices, denoted by $\mathcal{G}_u(N)$ and $\mathcal{G}_w(N)$, and respectively formed by the vectors

$$|\mathbf{h}_u\rangle \equiv \sum_{j=1}^{N} \exp(i2\pi\mathbf{h}\cdot\mathbf{r}_j)|e_j\rangle, \quad |\mathbf{h}_w\rangle \equiv \sum_{j=1}^{N} \mathcal{Z}_j \exp(i2\pi\mathbf{h}\cdot\mathbf{r}_j)|e_j\rangle \quad \text{with} \quad \mathbf{h}\in\mathcal{Z}^3. \tag{31}$$

Again, each of these vector lattices contains N linearly independent vectors and each s.f. is now equal to the scalar product of a vector of $\mathcal{G}_u(N)$ and a vector of $\mathcal{G}_w(N)$ since

$$\langle\mathbf{k}_w|\mathbf{h}_u\rangle = \sum_{j=1}^{N} \mathcal{Z}_j \exp[i2\pi(\mathbf{h}-\mathbf{k})\cdot\mathbf{r}_j] = F_{\mathbf{h}-\mathbf{k}}. \tag{32}$$

Also in this case a set $\mathcal{B}_M \equiv \{\mathbf{h}_1, \ldots, \mathbf{h}_M\}$ of reflections determines an $M\times M$ KH matrix $(\mathcal{D}_M)$ with its (i, j)th element equal to $\langle\mathbf{h}_{w,i}|\mathbf{h}_{u,j}\rangle = F_{\mathbf{h}_j-\mathbf{k}_i}$. But, in this case, the possible negativeness of some $\mathcal{Z}_j$s implies that the property $\det(\mathcal{D}_M)\geq 0$ no longer applies. Consequently, in contrast to the case of X-ray scattering, if one finds that $\det(\mathcal{D}_M)=0$ one can no longer conclude that the vectors associated with the reflections of $\mathcal{B}_M$ are linearly dependent. These considerations show that, in the case of neutron scattering, the vanishing of the determinant of a KH matrix is not as predictive as in the case of X-ray scattering.

This rather detailed review of the algebraic approach shows that four difficulties must still be overcome before one can state that the atomicity assumption removes the phase problem in the case of ideal crystals. The difficulties are the following. (*a*) To work with the subtracted intensities defined by (16) and this implies that, beside the $\boldsymbol{\delta}_j$s, $\overline{\mathcal{N}}$ and the ν_js are further unknowns to be determined. (*b*) To devise a procedure that isolates a finite set of reflections such that the full diffraction pattern can be reconstructed from it or, equivalently, that the knowledge of this finite set allows us to get a system of polynomial equations that has no spurious roots. For this reason, the reflections of this portion of the full pattern will be said to form a *complete* set of reflections. (*c*) To find a complete set of reflections lying as close as possible to the origin (in this way it will be possible to say whether the considered limiting sphere is large enough to solve the crystal's structure). (*d*) The previous difficulties must be removed in cases of both X-ray and neutron scattering.

3. Results and conclusion

The procedure to be followed in order to remove the noted difficulties (*a*)–(*d*) is detailed in the first two sections of the deposited data. Actually, the procedure allows us to say whether the observed limiting sphere is large enough to contain a complete set of reflections and, in the affirmative case, to obtain from the relevant subtracted intensities an infinite resolved Patterson map, *i.e.* the values of $\overline{\mathcal{N}}$, the $\boldsymbol{\delta}_j$s and the ν_js that represents the maximum information content of a diffraction pattern. Clearly, this result is based on the assumption that the atoms are treated as point-like objects. Therefore, the statement that the atomicity assumption removes any ambiguity in the phase problem for an ideal crystal for both X-ray and neutron scattering is now fully proven. To avoid any confusion, we stress that the aforesaid conclusion has been reached on a purely theoretical basis. In practice, of course, it is not new because many structures have been solved, using X-ray and neutron scattering data [for the latter see Hauptman (1976) and Hauptman & Langs (2003)] by direct methods that, still based on analytic expression (1) for the s.f.s, use probabilistic considerations and 'trial-and-error' techniques. One naturally wonders whether the algebraic approach, theoretically so powerful, might also be of some practical use. So far, the reported applications concern crystals containing up to ten atoms within the unit cell (see *e.g.* Fischer & Pilz, 1997; Rothbauer, 1998; Pilz & Fischer, 2000). The small number of atoms is hard to overcome because the size of the involved $(\mathcal{D})$ matrices increases so quickly with N to saturate computer capabilities and, at the same time, because experimental uncertainties on observed intensities tend to make the algebraic approach mathematically unstable as the size of $(\mathcal{D})$ increases. Nonetheless, in our opinion, this is a point still worth investigating. As shown by Cervellino & Ciccariello (1999) in a one-dimensional application of the algebraic approach, subtracted noisy intensities roughly determine $\overline{\mathcal{N}}$. The same analysis is feasible in three dimensions with experimental intensities forcing $\overline{\mathcal{N}}$ to be smaller than 100. In fact, this value is well within modern computer capabilities while the condition $\overline{\mathcal{N}}\leq 100$ tends to increase the degeneracy of the interatomic vectors that, as N increases, tend to form clusters. The resulting infinitely resolved Patterson map must be deconvoluted in terms of 20–50 point-like clusters of atoms that are known to occur. (Since each cluster can contain a large number of atoms in this way, the number of atoms of the unit cell is no longer confined to be small.) This point is the crucial one that needs further investigation.

References

Avrami, M. (1938). *Phys. Rev.* **50**, 300–309.
Bertaut, E. F. (1958). *Fortschr. Mineral.* **36**, 119.
Carathéodory, C. (1911). *Rend. Circ. Matem. Palermo*, **32**, 193–217.

---

[7] To be more specific, the problem of the uniqueness was explicitly mentioned by Navaza & Silva on p. 272 and by Navaza & Navaza on p. 697. The problem is also present in Rothbauer's formulation because on p. 197 the author mentions the necessity of solving his equation (2.3.3) for a 'sufficient number' of additional cases without specifying this number.

Castellano, E. E., Podjarny, A. D. & Navaza, J. (1973). *Acta Cryst.* A**29**, 609–615.

Cervellino, A. & Ciccariello, S. (1996). *Riv. Nuovo Cim.* **19/8**, 1–59.

Cervellino, A. & Ciccariello, S. (1999). *Z. Kristallogr.* **214**, 739–750.

Cervellino, A. & Ciccariello, S. (2001). *J. Phys. A: Math. Gen.* **34**, 731–755.

Fischer, K. F. & Pilz, K. (1997). *Acta Cryst.* A**53**, 475–483.

Fischer, K. F., Kirfel, A. & Zimmermann, H. (2005). *Z. Kristallogr.* **220**, 643–656.

Goedkoop, J. B. (1950). *Acta Cryst.* **3**, 374–78.

Harker, D. & Kasper, J. S. (1948). *Acta Cryst.* **1**, 70–75.

Hauptman, H. (1976). *Acta Cryst.* A**32**, 877–882.

Hauptman, H. & Karle, J. (1950). *Phys. Rev.* **80**, 244–248.

Hauptman, H. & Karle, J. (1952). *Acta Cryst.* **5**, 48–59.

Hauptman, H. & Langs, D. A. (2003). *Acta Cryst.* A**59**, 250–254.

Karle, J. & Hauptman, H. (1950). *Acta Cryst.* **3**, 181–187.

Knossow, M., de Rango, C., Mauguen, Y., Sarrazin, M. & Tsoucaris, G. (1977). *Acta Cryst.* A**33**, 119–125.

Navaza, A. & Navaza, J. (1992). *Acta Cryst.* A**48**, 695–701.

Navaza, J. & Silva, A. M. (1979). *Acta Cryst.* A**35**, 266–275.

Ott, H. (1927). *Z. Kristallogr.* **66**, 136–143.

Patterson, A. L. (1935). *Z. Kristallogr.* **90**, 517–523.

Patterson, A. L. (1939). *Phys. Rev.* **55**, 682–691.

Rango, C. de, Tsoucaris, G. & Zelwer, Ch. (1974). *Acta Cryst.* A**30**, 342–353.

Pilz, K. & Fischer, K. F. (2000). *Z. Kristallogr.* **215**, 640–649.

Podjarny, A. D. (1981). *Acta Cryst.* A**37**, 662–668.

Podjarny, A. D., Yonath, A. & Traub, W. (1976). *Acta Cryst.* A**32**, 281–292.

Rothbauer, R. (1970). *Neues Jahrb. Mineral. Monatsh.* H**10**, 473–481.

Rothbauer, R. (1974). *Neues Jahrb. Mineral. Monatsh.* H**1**, 14–21.

Rothbauer, R. (1980). *Acta Cryst.* A**36**, 27–32.

Rothbauer, R. (1998). *Z. Kristallogr.* **213**, 195–201.

Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.

Sayre, D. (2002). *Struct. Chem.* **13**, 81–96.

Silva, A. M. & Navaza, J. (1981). *Acta Cryst.* A**37**, 658–661.