TOPICAL REVIEW

Direct Methods and Powder Data: State of the Art and Perspectives

CARMELO GIACOVAZZO

Istituto di Ricerca per lo Sviluppo di Metodologie Cristallografiche, CNR, c/o Dipartimento Geomineralogico, Campus Universitario, Via Orabona 4, 70125 Bari, Italy, and Dipartimento Geomineralogico, Campus Universitario, Via Orabona 4, 70125 Bari, Italy. E-mail: giacovazzo@area.ba.cnr.it

Abstract

Solving crystal structures by applying direct methods to single-crystal data is a relatively easy task for structures with up to 100 atoms in the asymmetric unit. Their successful application to powder data is still a challenge unless the size of the structure is moderate. The rate of success depends on several factors like the efficiency of the full-pattern-decomposition programs, the peak over-lapping, the presence of preferred orientation in the powder specimen, the nature of the background, the type of radiation used for collecting experimental data *etc.* The main factors are analysed in order to provide a clear description of the specific problems that have to be forced when direct phasing from powder data is attempted.

1. Introduction

It is usual to say that direct methods solved in practice the phase problem for small molecules. This is not true when only powder data are available: in this case, even a crystal structure with few atoms in the asymmetric unit may constitute a considerable challenge. The difficulties begin with the indexing of the reflections and the definition of the space group. These problems are far from being trivial: their solution requires modern computational programs and keenness in their use. They are not treated here: we will assume that indexing and space-group problems have already been solved. Our interest will be focused on the consequences, on direct-methods effectiveness, of a characteristic feature of the powder diffraction, the collapse of the three-dimensional

reciprocal space of the individual crystallites on the one-dimensional 2θ axis. The possible effects may be schematized as:

(*a*) overlapping, accidental or systematic, of the diffraction peaks;

(*b*) considerable background, not easily definable with accuracy;

(*c*) preferred-orientation effects when the powder is not ideal (*i.e.* because of the non-random distribution of crystallites).

In such conditions, the quality of the information provided by a diffraction experiment deteriorates. While, for single-crystal data, the diffraction experiment loses the phases but preserves the knowledge of the diffraction moduli, in a powder diffraction experiment this last information is partially lost. Estimates of the structure-factor moduli may be obtained by decomposing the powder diffraction pattern; however, the unavoidable uncertainty of the estimates can strongly influence direct-methods efficiency. Recent advances in instrumentation (high-resolution neutron and synchrotron) techniques provide more informative diffraction patterns and require similar advances in the methods for treating the data.

In the following sections, we will schematically describe the various special difficulties that must be faced when powder data are used, including the characteristic problems arising when neutron data are employed. In order to show the practical implications of the various problems, some applications will also be described. We will use as test data the experimental diffraction patterns of the test structures quoted in Table 1.

2. Extracting structure-factor moduli

The recovery of the structure-factor moduli from a powder diffraction pattern is usually performed by a full-pattern-decomposition program: a modulus is associated with each component reflection so making possible the use of direct methods. The efficiency of the

---

Carmelo Giacovazzo became full Professor of Mineralogy at Bari University in 1975. He has contributed to direct-methods phasing techniques in crystallography by developing probabilistic theories for the estimation of structure invariants and seminvariants. Phasing proteins by exploiting derivative data and/or anomalous-dispersion effects is one of his fields of interest. Recently, he contributed to the phasing problem via powder data.

Table 1. *Code name and crystallochemical data for the test structures*

X: data collected by home diffractometer; N: neutron data; S: synchrotron data.

Code	Space group	Cell content	Z
AGPZ (X)	Pbca	$AgN_2C_3H_3$	8
BACO (N)	C2/m	$BaC_2O_4 \cdot D_2O$	4
BAER (X)	$R\bar{3}m$	$Si_{54}O_{108}3[C_{10}H_{15}N_2]$	1
BENZ (S)	$P2_1/a$	$C_6H_6 \cdot C_6F_6$	4
CABA (X)	P4/mmm	$Nd_{1.33}Ca_{0.25}Ba_{1.42}Cu_{2.92}O_{6.8}$	1
CFBR3 (N)	Pnma	$CFBr_3$	4
CF3BR (N)	$P2_1/a$	$CBrF_3$	4
CFCL (N)	Fdd2	CF_2Cl_2	8
CFI (N)	Cmca	CF_3I	8
CIM (X)	$P2_1/n$	$C_{10}N_6S$	4
CROP (X)	Imma	$CrPO_4$	12
CROX (X)	$P\bar{1}$	Cr_8O_{21}	1
CUPZ (X)	Pbca	$CuN_2C_3H_3$	8
ERVO (N)	$I4_1/amd$	$ErVO_4$	4
FECOAS (N)	P4/nmm	$Fe_{1.6}Co_{0.4}As$	2
GLYCIN (X)	$P2_1/c$	$N_2C_4O_3$	4
HOBA (S)	Pmmm	$HoBa_2Cu_{2.90}O_{6.85}$	1
HOFE (N)	$I4_1/mmm$	$HoFe_5Al_7$	2
LAMO (X)	$P2_1/a$	$LaMo_5O_8$	4
LASI (N)	$P2_1/c$	$La_2Si_2O_7$	4
LEV (S)	$R\bar{3}m$	$[Si_{54}O_{108}] \cdot 3C_8NH_{16}$	1
MES (X)	$P2_1/c$	$C_6NO_4SH_{13}$	4
METYL (S)	I222	$NaCD_3$	16
NBPO (S)	C2/c	$Nb_3(NbO)_2(PO_4)_7$	2
NDBA (X)	P4/mmm	$NdSrBaCu_{2.88}O_{6.8}$	1
NIZR (X)	$P2_1/n$	$Ni_{0.5}Zr_2(PO_4)_3$	4
PBS (S)	Pbca	PbS_2O_3	8
PICO (X)	$P\bar{1}$	$Pb_2C_4O_8$	1
RUCO (X)	Ibam	RuC_4O_4	4
RUHG (X)	$P2_1/c$	$HgRuC_4O_4$	8
SAPO (X)	Pmmm	$[(Si,Al,P)_{32}O_{64}] \cdot 2TPAOH$	1
SBPO (S)	$P2_1/n$	$Sb_2(PO_4)_3$	4
SIGMA2 (X)	$I4_1/amd$	$Si_{64}O_{128}4[C_{10}H_{17}N]$	1
STRIA (N)	C2/c	$C_3H_3N_3$	4
SULPH (N)	Pbcm	D_2S	8
T83 (X)	Imma	$Cs(Al,Ti)O_4$	4
TEAS (N)	$I4_1/amd$	$TbAsO_4$	4
TENI (N)	Amm2	$TbNiC_2$	2
TIPORF (X)	I4/m	$TiCl_2N_4C_{44}H_{28}$	2
TRAN (X)	$R\bar{3}c$	$Li_{1.15}Ti_{1.85}In_{0.15}P_3O_{12}$	6
YONO (S)	$P2_1$	$Y_4O(OH)_9NO_3$	2

decomposition program first depends on the quality of the pattern: if high-resolution neutron or X-ray (synchrotron) techniques are used, and the crystal structure is not too complex, a reliable intensity could be associated with each component reflection. Such an ideal situation does not occur for more complicated structures or when systematic (because of the symmetry) overlapping occurs. Special care in the extraction of the structure-factor moduli should be taken when non-conventional experimental devices are used (see *e.g.* Louër, Louër & Touboul, 1992).

Among the programs dedicated to pattern decomposition, we quote *ALLHKL* (Pawley, 1981), *XRS*-82 (Baerlocher, 1982), *LSQPROF* (Jansen, Peschar & Schenk, 1992), *PROFIT* (Toraya, 1986), *FULLPROF* (Rodrigues-Carvajal, 1990), *BER-HKL* (Berar, 1990),

M-PROFIL (Fitch, Murray & Jouanneaux, 1995), *FULFIT* (Jansen, Schäfer & Will, 1988).

The efficiency of any decomposition program also depends on the mathematical techniques on which it is based. We will make explicit reference to two prototype approaches [but alternative methods do exist, *e.g.* Bricogne (1991), Gilmore, Henderson & Bricogne (1991)]: the Pawley (1981) technique and the Le Bail algorithm (Le Bail, Duray & Fourquet, 1988). *ALLHKL* (Pawley, 1981), a widely used program based on the Pawley technique, performs a least-squares fitting of the diffraction pattern in order to evaluate integrated intensities. Such intensities, together with background, peak shape and cell dimensions, are parameters to be refined in the least-squares approach. When substantial overlap occurs, the intensity values are highly correlated and the technique can provide negative intensity values. Pawley was aware of the problem and introduced slack constraint terms to avoid such undesirable effects. The problem was not obviated and has been revisited by Toraya (1986), Jansen, Peschar & Schenk (1992) and Sivia & David (1994), who tackled the problem of highly correlated positive and negative intensities.

In the programs based on the Le Bail algorithm, the problem of wild intensity variations is surmounted since the algorithm intrinsically tends to equipartition intensities when peaks overlap. In our applications, we will use the program *EXTRA* (Altomare, Cascarano, Giacovazzo, Guagliardi, Moliterni, Burla & Polidori, 1995), a Le Bail algorithm based decomposition program. Programs based on the Pawley technique will probably provide sets of diffraction magnitudes substantially different from those based on the Le Bail algorithm, as will be shown in the next section.

It is worthwhile stressing that pattern-decomposition programs play a crucial role in the success of direct methods. Any error, casual or systematic, in the estimate of the structure-factor moduli will weaken the efficiency of the methods. Since they extract the phase values from the moduli, wrong phases will probably be derived from biased moduli. In order to reduce the ambiguity caused by the overlap, some procedures based on the Patterson squaring method (David, 1987; Cascarano, Giacovazzo, Guagliardi & Steadman, 1991; Estermann, McCusker & Baerlocher, 1992; Estermann & Gramlich, 1993) can be used in combination with profile-fitting procedures.

3. The structure-factor statistics

In §2, we emphasized the fact that, when two or more reflections heavily overlap, the Pawley technique often provides negative intensities for some of the reflections overlapping: to compensate, single individual intensities larger than the overall intensity clump could be provided. On the contrary, the Le Bail algorithm intrinsically tends to equipartition the overall intensity among the severely overlapping reflections. The general statistical conse-

quence is that the dispersion of the structure-factor amplitudes is expected to be higher for methods based on the Pawley technique than for methods based on the Le Bail algorithm. In order to give some numerical detail, we have applied *ALLHKL* and *EXTRA* to the experimental diffraction patterns of CROX and YONO (see Table 1). The structure-factor amplitudes derived were transformed by *SIRPOW*.92 (Altomare, Cascarano,

Fig. 1. Observed $|E|$ distributions obtained when *ALLHKL* and *EXTRA* are applied to the experimental diffraction pattern of CROX. The theoretical Wilson distribution for a centrosymmetric space group is also shown.
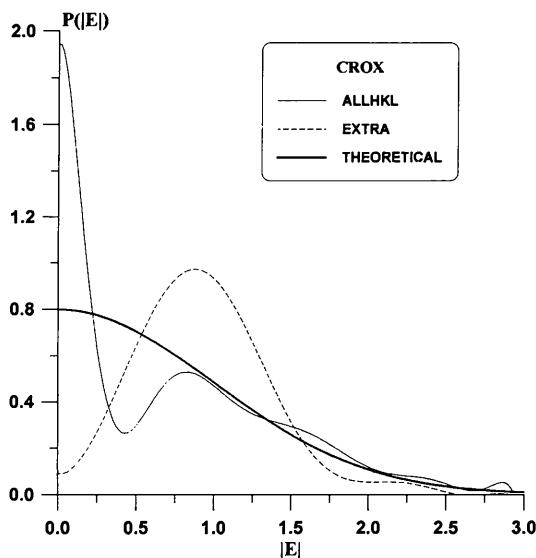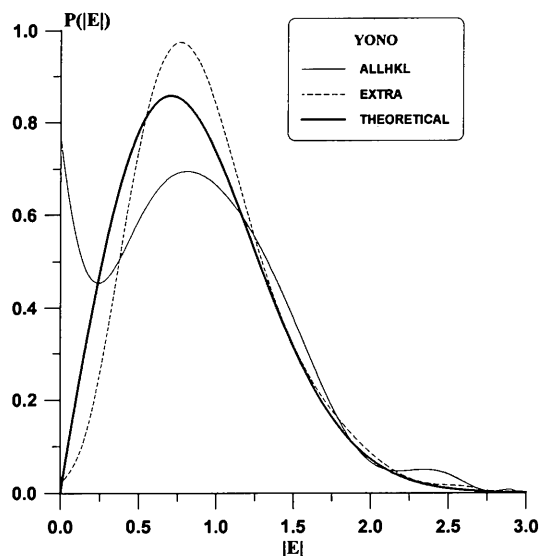

Fig. 2. Observed $|E|$ distributions obtained when *ALLHKL* and *EXTRA* are applied to the experimental diffraction pattern of YONO. The theoretical Wilson distribution for an asymmetric space group is also shown.

Table 2. *Moments of distributions for CROX and YONO as examples of centric and acentric structures*

The $\langle|E|^2\rangle$ and $\langle||E|^2 - 1|\rangle$ moments are given for the data sets extracted by *ALLHKL* and *EXTRA*, respectively. Some values of the cumulative $N(|E|^2)$ distributions are also shown.

		CROX	YONO
$\langle|E|\rangle$			
	ALLHKL	0.698	0.834
	EXTRA	0.903	0.898
	Theoretical	0.798	0.886
$\langle||E|^2 - 1|\rangle$			
	ALLHKL	1.076	0.824
	EXTRA	0.656	0.703
	Theoretical	0.968	0.736
$N(0.2)$			
	ALLHKL	0.434	0.262
	EXTRA	0.136	0.145
	Theoretical	0.345	0.181

Giacovazzo, Guagliardi, Burla, Polidori & Camalli, 1994) into normalized amplitudes $|E|$; their distributions are shown in Figs. 1 and 2, and some moments of the distributions are given in Table 2. We observe: (a) *ALLHKL* pattern decomposition always ends with a too high percentage of negative intensities, which are assumed equal to zero when input into *SIRPOW*.92. This explains the too high values of $P(|E|)$ for $|E|$ close to zero; (b) *ALLHKL* tends to produce centrosymmetric $|E|$ distributions while *EXTRA* (as all the Le Bail algorithm based programs) tends towards non-centric distributions.

The final suggestion coming from the above observations is that the Wilson statistics may not be so informative about the presence of the inversion centre in the space group as occurs for single-crystal data. Indeed, most of the features of the experimental distribution function are pattern-decomposition-method dependent rather than crystal-structure dependent (see also Cascarano, Favia & Giacovazzo, 1992; Estermann, McCusker & Baerlocher, 1992).

4. The role of the background in the normalization process

The background may be divided into a smooth and a structured part; the first is due to inelastic and resonant scattering, the second to TDS and, when present, to the amorphous content. The structureless part may be modelled by physically based equations (Suortti & Jennings, 1977): TDS, to a good approximation, is equal to the part of the Bragg scattering due to thermal motion. Very recently, a significant result has been obtained by Riello, Fagherazzi, Clemente & Canton (1995), who obtained a relationship which, on a physical basis, models the three most important background scattering components: air scattering, incoherent scatter-

ing and average diffuse scattering due to thermal disorder plus first-kind lattice disorder.

In the usual applications for crystal structure determination, the background is modelled by a polynomial of a suitably high order (usually up to order ten). Errors in the model of the peak shape will cause errors in the background and consequently errors in the estimate of the structure-factor amplitudes. We pay attention here to the high-angle region of the pattern, where, for X-ray data, the decay of the scattering-factor magnitudes with $\sin\theta/\lambda$ makes the errors in the background definition critical. Since the structure-factor amplitudes are small, even small errors in the background modelling can determine non-negligible relative changes in the amplitude estimation.

Unfortunately, in this high-angle section of the pattern, peaks are often broad; then peak tails are long and the separation of the peak contribution from the background is difficult.

The normalization of the structure factors is critically dependent on the background definition at high angles. Indeed, the value of the overall temperature factor, as determined by a Wilson plot, is strongly influenced by the average $|F|^2$ value in that region. Quite frequently, different full-pattern-decomposition programs provide sets of structure-factor amplitudes that, submitted to a Wilson-plot process, produce different thermal-factor values. For example, CROX structure-factor amplitudes derived by ALLHKL show an overall thermal factor $B = 1.77 \text{\AA}^2$, while $B = 0.85 \text{\AA}^2$ if EXTRA is used. In an analogous way, PBS data show $B = 1.150 \text{\AA}^2$ when processed by ALLHKL, $B = 1.25 \text{\AA}^2$ when processed by EXTRA.

Often, the Wilson scaling process may end with a negative overall thermal factor (Cascarano, Favia & Giacovazzo, 1992). Such a result has no physical meaning and is often caused by (a) omission of weak reflections from the set of processed data (Cascarano, Favia & Giacovazzo, 1992); (b) truncation of experimental data to $\sin\theta/\lambda$ values where Debye effects are important (Hall & Subramanian, 1982; Cascarano, Giacovazzo & Guagliardi, 1992); (c) imperfect modelling of the background and of the peak shape (Lutterotti & Scardi, 1990). Even the imperfect absorption correction for capillary samples or the superficial roughness of flat powder samples in the Bragg–Brentano technique could influence the final value of the overall thermal factor. If the Wilson plot provides a negative overall thermal factor, one should reconsider the entire process of diffraction-pattern decomposition.

5. The role of the weak reflections

To omit weak reflections from the set of integrated intensities is not critical for Patterson methods but may be ruinous for direct methods. The following may be observed:

(a) Omission of weak reflections in a Wilson-plot procedure causes the average intensity for each shell of reciprocal space to increase, so producing errors in the scaling factor (Rogers, Stanley & Wilson, 1955; Hall & Subramanian, 1982; Cascarano, Giacovazzo & Guagliardi, 1991). If the proportion of unobserved data increases with $\sin\theta/\lambda$ (as usual), the value of the overall thermal factor B will be smaller than the true one. As a consequence, the structure-factor distribution may appear acentric if the structure is centric, or centric if the structure is hypercentric or has some pseudo-translational symmetry.

(b) The Pawley (1981) technique for the full pattern decomposition frequently generates negative intensity estimates. Omitting the corresponding reflections from the Wilson-plot process causes the systematic errors described in (a); putting those intensities to zero generates an opposite type of error. In this case, the average intensity for each shell decreases and, in agreement with previous observations, the estimated B value will be too large and the structure-factor-amplitude distribution will appear centric for acentric structures or hypercentric for centric ones. Techniques such as those proposed by Jansen, Peschar & Schenk (1992) and Sivia & David (1994) should necessarily integrate the Pawley method.

(c) Weak reflections are of great importance for the phasing process. Several probabilistic formulas make explicit use of the information contained in the weak intensities: some examples are the P_{10} formula (Cascarano, Giacovazzo, Camalli, Spagna, Burla, Nunzi & Polidori, 1984), which estimates triplet invariants via a ten-node figure spanning all the reciprocal space, and the SAYTAN procedure (Debaerdemaker, Tate & Woolfson, 1988). For the reader not familiar with direct methods, we recall that the phase estimate of the triplet invariant

$$\Phi = \phi_{\mathbf{h}_1} + \phi_{\mathbf{h}_2} + \phi_{\mathbf{h}_3} \quad (\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = 0)$$

is obtained via the joint probability distribution

$$P(E_{\mathbf{h}_1}, E_{\mathbf{h}_2}, E_{\mathbf{h}_3}, E_{\mathbf{k}}, E_{\mathbf{h}_1+\mathbf{k}}, E_{\mathbf{h}_1-\mathbf{k}}, E_{\mathbf{h}_2+\mathbf{k}}, E_{\mathbf{h}_2-\mathbf{k}},$$
$$E_{\mathbf{h}_3+\mathbf{k}}, E_{\mathbf{h}_3-\mathbf{k}}),$$

where $\mathbf{k}$ is a free vector. The efficiency of the method (able to identify negative cosine triplets) relies on the simultaneous use of both strong and small $|E|$ values. If the small $|E|$'s are systematically omitted, the P_{10} formula cannot work: a similar conclusion holds for the SAYTAN procedure.

Weak reflections are involved in some important figures of merit aiming at recognizing the correct from the various false solutions provided by a multisolution process. The best known is the so-called psi-zero figure of merit (Cochran & Douglas, 1957) based on a special set of triplet invariants formed by two strong and one weak reflections. Of course, omission of weak reflections does not allow one to construct psi-zero triplets. Weak

reflections play a central role in all the formulas estimating negative quartet invariants (Schenk, 1973; Hauptman, 1975; Giacovazzo, 1975, 1976). Such invariants (statistically recognized because their cross magnitudes are small) are used both in an active way (Sheldrick, 1991; Cascarano, Giacovazzo, Moliterni & Polidori, 1994) and in a passive way as a figure of merit.

From the considerations made in points (a), (b), (c), it should be concluded that omitting weak reflections distorts the normalized amplitudes and therefore reduces the effectiveness of direct methods. The question is now: are weak reflections estimated sufficiently well to be safely used? Or are the errors in their estimates so large that their usefulness in a phasing procedure is negligible in practice? $SIRPOW.92$ applications suggest that the use of weak reflections is not only useful but often necessary, even if the problem of the reliability of their amplitude estimate is far from being optimally solved.

6. Amount of information in a powder diagram

If two or more peaks in a powder diffraction pattern heavily overlap, the integrated intensities of the reflections overlapping are highly correlated. As a consequence, as soon as the intensity clump is estimated the individual intensities cannot be considered as statistically independent. The question now is: is it possible to take into account the correlation among intensities in order to obtain an estimate of the statistically independent intensities contained in a powder diffraction pattern? This problem has several implications: (a) direct methods work well when the ratio (number of independent intensities)/(number of atoms to find) is sufficiently high (say larger than 7); (b) usual crystallographic least-squares refinements consider observations as statistically independent of each other: their efficiency is determined by the ratio (number of independent observations)/(number of parameters to refine); (c) how the maximum range of 2θ to be exploited in a direct phasing procedure should be defined. At high $\sin\theta/\lambda$ ranges, where the overlapping is extremely high and the contrast (diffraction spot intensity)/(background intensity) is extremely low, it may occur that the number of independent observations in the range is quite small. In such a case, the interval could be omitted from the calculations.

The above observations show that a qualitative estimate of the amount of information provided by a pattern allows one to guess the rate of success of a direct phasing process and about the reliability of the refinement process. An algorithm has recently been proposed by Altomare, Cascarano, Giacovazzo, Guagliardi, Moliterni, Burla & Polidori, 1995), which, via the systematic study of the peak overlap, is able to provide an estimate of the statistically independent observations. In Table 3 we give, for three test structures, the value of M (number of symmetry-independent reflections lying in

Table 3. *For the test structures BACO, SAPO and SBPO values of s^2_{max} [the maximum observed value of $(\sin\theta/\lambda)^2$], M (the number of symmetry-independent reflections lying in the measured range), M_{ind} (the number of statistically independent observations in the same range) and N_p (the number of structural parameters to vary in an isotropic least-squares procedure) are given*

	s^2_{max}	M	M_{ind}	N_p
BACO	0.26	278	109	22
SAPO	0.17	686	189	78
SBPO	0.30	1206	266	68

the measured 2θ range), M_{ind} (number of statistically independent observations belonging to the same range) and N_p (number of structural parameters to be refined in an isotropic least-squares procedure). The table clearly suggests that the ratio (number of independent observations)/(number of parameters to refine) is often very low, and this does not allow high accuracy in the determination of the structure parameters. Thus, supplementary information is needed for a sufficiently good refinement of a crystal structure. Usually, it comes from: (a) elimination of some of the parameters as independent variables by constraining them to be functions of others (Immirzi, 1980; Pawley, 1980); (b) introduction of restraints on some geometric or energy parameters (Baerlocher, 1993; Elsenhans, 1990; Izumi, 1989); (c) the integration of an X-ray diffraction pattern with a neutron pattern. Such a practice is suggested by fact that X-rays and neutrons are scattered by different mechanisms (Larson & Von Dreele, 1987).

7. The role of the preferred orientation

Texture in the powder specimen systematically distorts the intensity ratios. In the absence of any information on preferred orientation, decomposition programs will provide, as a measure of integrated Lorentz–polarization-corrected intensity, the biased value

$$|F'_k|^2 = |F_k|^2 O_k, \qquad (1)$$

where O_k is the preferred-orientation-correction factor. Since direct methods derive the phase values from the structure-factor amplitudes, systematic errors in their estimates will reduce the rate of success for the phasing process. To overcome this problem, one should be able to obtain intensities corresponding to a randomly oriented specimen. This may be done by measuring pole-density distributions (Bunge, Dahms & Brokmeier, 1989; Järvinen, 1993) but the procedure requires a multiaxis goniometer and additional experimental work. Performing such an additional experiment is rewarding in terms of direct-methods efficiency. However, if such work has not be carried out, an alternative procedure may be applied to the usual powder diffractometer data provided

the sample has cylindrical symmetry: a simple statistical analysis of the normalized structure factors may reveal the presence of preferred-orientation planes before starting the phasing process (Altomare, Cascarano, Giacovazzo & Guagliardi, 1994; Altomare, Cascarano, Giacovazzo, Guagliardi, Moliterni, Burla & Polidori, 1995; Peschar, Schenk & Capková, 1995). Once this has been done and the function O_k has been determined, then

$$|F_k|^2 = |F'_k|^2 O_k^{-1}$$

may be obtained. The normalized structure factor corresponding to $|F_k|^2$ should be closer to the true values than the $|E'_k|^2$'s: as a consequence, the phasing process should be more straightforward and, in the end, more accurate atomic parameters should be obtained.

8. About the radiation type

Positivity and atomicity of the electron density are traditionally considered as basic conditions for direct methods. The Cochran (1955) formula, which estimates the triplet phase invariants Φ, is based on the above postulates and may be written as follows:

$$P(\Phi | |E_{h_1}|, |E_{h_2}|, |E_{h_3}|) \simeq [2\pi I_0(G)]^{-1} \exp(G \cos \Phi),$$
$$(2)$$

where I_0 is the modified Bessel function of order zero and

$$G = 2|E_{h_1} E_{h_2} E_{h_3}| N_{eq}^{-1/2}.$$

N_{eq} is the so-called equivalent number of atoms in the unit cell and, for X-rays, is calculated via

$$N_{eq}^{-1/2} = \sum_{j=1}^{N} Z_j^3 \Big/ \left(\sum_{j=1}^{N} Z_j^2 \right)^{3/2}, \qquad (3)$$

where Z_j is the atomic number of the jth atom. Neutron powder data are often collected for crystal structure analysis. In this case, negative values of the scattering amplitude b are allowed and the positivity postulate is violated. It may be shown that (2) is still valid (Hauptman, 1976) provided (3) is replaced by

$$N_{eq}^{-1/2} = \sum_{j=1}^{N} b_j^3 \Big/ \left(\sum_{j=1}^{N} b_j^2 \right)^{3/2}. \qquad (4)$$

What are the consequences of the violation of positivity for direct methods? The question has been reviewed by Altomare, Giacovazzo, Guagliardi & Siliqi (1994). The conclusions may be summarized as follows:

(i) If negative amplitudes mix in (4) with positive ones, the absolute values of N_{eq} increases. This makes triplet relationships less reliable. If $\sum_{j=1}^{N} b_j^3 = 0$, then N_{eq} goes to infinity and the triplet estimate is impossible.

(ii) Quartet invariants depend on the even powers of the scattering amplitudes and are therefore not influenced by the signs of the various b. However, negative quartets

are not readily recognized when the positivity criterion is violated. This makes the figures of merit finding the correct solution among the various ones given by a multisolution approach less efficient.

(iii) The P_{10} formula (Cascarano, Giacovazzo, Camalli, Spagna, Burla, Nunzi & Polidori, 1984) depends on the lowest odd powers of the scattering amplitudes, and therefore suffers by the violation of the positivity criterion (as well as the Cochran relationship).

(iv) Crystal structure solution is in general easier for X-ray data than for neutron data. To explain this statement, we give two examples: (a) For a crystal structure with some heavy atoms, the value of (3) is markedly smaller than the value of (4). This makes triplet relationships more reliable when X-ray data are used and facilitates the immediate recognition of the heavy-atom positions. (b) If some H atoms are present in the crystal structure, they significantly contribute to N_{eq} while they are negligible when X-radiation is used. Again, the two values (3) and (4) may be remarkably different. Furthermore, H atoms incoherently contribute to the total scattering.

(v) The refinement process may be more straightforward for neutron data. Since the scattering amplitudes b are nearly constant with $\sin\theta/\lambda$, high $\sin\theta/\lambda$ data are more easily measured if neutron radiation is used. Furthermore, the dominating role of heavy atoms in the refinement stage, which is often a non-negligible obstacle to the positioning of the light atoms, is removed. In particular, H atoms can be carefully localized.

(vi) When negative scattering amplitudes are in the chemical formula, the location of negative peaks in the Fourier map is also necessary. In this case, after the location of the positive peaks, $\rho(r)$ should be automatically transformed into $-\rho(r)$ and negative peaks are found by the usual peak-search procedure.

9. About the size of the structures solvable by direct methods

Recent advances in experimental equipment and in the computational aspects have enlarged the size of the crystal structures that can be solved ab initio from powder data. Very recently (Morris, Harrison, Nicol, Wilkinson & Cheetham, 1992), it was noted that a disparity existed between the complexity of structures that could be solved ab initio from powder data and the complexity of the structures that in principle may be refined by the Rietveld profile method. Up to 1988, refinement of up to 34 atoms and 132 positional parameters was accomplished but the most complex unknown structure solved via powder data contained only 17 atoms in the asymmetric unit (McCusker, 1988). In 1992, the solution and refinement of $Ga_2(HPO_3)_3 \cdot 4H_2O$ with 29 atoms in the asymmetric unit and 117 structural parameters was reported (Morris, Harrison, Nicol, Wilkinson & Cheetham, 1992), for

which synchrotron X-ray and neutron powder diffraction data were used. The use of *SHELX86* (Sheldrick, 1985) determined 151 phases for E values greater than 1.2, which, used for the calculation of an E map, showed two large peaks where the two Ga atoms were accommodated. Subsequent Fourier synthesis revealed four more atoms, which were assigned as two P and two O atoms.

The situation has largely improved in the last two years. The structure of $La_3Ti_5Al_{15}O_{37}$, which has 60 atoms in the asymmetric unit (space group Cc), is the largest structure ever solved from powder data (Morris, Owen, Stalick & Cheetham, 1994). Direct methods localized a limited number of atoms: the structure could be completed by exploiting the supplemental information contained in a neutron powder pattern and structure refinement was accomplished by using suitable geometric restraints in the least-squares process.

The application of direct methods may not be straightforward even when the powder data of small structures are processed. Indeed, the moderate size of the unit cell and the frequent presence of heavy atoms may generate an undesirable (from the point of view of direct methods) effect: the pseudo-translational symmetry. This can greatly disturb the phasing approach and special procedures are often necessary to solve the crystal structures (Fan, Yao, Main & Woolfson, 1983; Böhme, 1982; Gramlich 1984; Cascarano, Giacovazzo & Luić, 1985, 1987, 1988; Cascarano, Giacovazzo, Luić & Vicković, 1989).

In Table 4, we give, for the test structures in Table 1, the mean fractional scattering power (MFSP) of the electron density in pseudo-translational symmetry (as estimated by *SIRPOW*.92). The percentage is often not negligible. A trivial application of direct methods easily determines the substructure, while the crystal structure completion presents additional difficulties. Luckily, the statistical analysis of the diffraction intensities can provide information about the nature of the pseudo-translational symmetry, and also an estimate of the MFSP. Thus, the user can apply special procedures for the complete crystal structure solution.

10. The preliminary structure refinement

In the most modern direct-methods programs for single-crystal data, the crystal structure is first solved by phasing a limited amount of structure factors, say up to the 500 largest intensities. Because of the unavoidable errors in the phasing procedure and owing to the effects of series truncation, the complete crystal structure is usually not evident from the first E map. Thus, an automatic refinement step follows the crystal-structure-solution process: phase determination is extended to a larger number of reflections and least-squares Fourier cycles are performed to recover automatically the full structure and refine parameters.

Table 4. *For the test structures in Table 1, the mean fractional scattering power (MFSP) of the electron density in pseudo-translational symmetry is shown*

Code	MFSP (%)
AGPZ	32
BACO	13
BAER	20
BENZ	–
CABA	59
CFBR3	13
CF3BR	21
CFCL	28
CFI	35
CIM	–
CROP	39
CROX	14
CUPZ	27
ERVO	19
FECOAS	33
GLYCIN	27
HOBA	42
HOFE	71
LAMO	13
LASI	21
LEV	23
MES	–
METYL	14
NBPO	34
NDBA	60
NIZR	52
PBS	52
PICO	13
RUCO	24
RUHG	25
SAPO	23
SBPO	42
SIGMA2	–
STRIA	58
SULPH	20
T83	39
TEAS	17
TENI	15
TIPORF	21
TRAN	32
YONO	16

SIRPOW.92 works in a similar way. The preliminary structure refinement is made by a technique such as that described by Will, Jansen & Schäfer (1982) and Jansen, Schäfer & Will (1988). The least-squares program minimizes

$$S = \sum_{i=1}^{n} w_i \left(I_{oi} - I_{ci}\right)^2,$$

where I_{oi} and I_{ci} represent the ith observed and calculated (Lp-corrected) intensities, whether they refer to a single reflection or to a set of reflections. In particular, if i denotes a set of reflections, then

$$I_{ci} = \sum_{\mathbf{k}} m(\mathbf{k})|F_{calc}(\mathbf{k})|^2.$$

The summation involves all the reflections belonging to the ith set [they are reflections that strongly overlap with

each other and slightly overlap (or not) with the others]. Such a technique requires the modification of the single-crystal least-squares approach. For example, the normal matrix corresponding to the minimization of S has, as generic element,

$$c_{lt} = \sum_{i=1}^{n} w_i \left[\sum_k m_k \partial |F_k|_c^2 / \partial x_l \right] \left[\sum_k m_c \partial |F_k|_c^2 / \partial x_t \right]$$

and, as generic column term,

$$b_t = \sum_{i=1}^{n} w_i \Delta I_i \left[\sum_k m_k \partial |F_k|_c^2 / \partial x_t \right].$$

Since the number of observations, that is the intensities of the reflection clusters, is often comparable with the number of parameters to be refined, the refinement process cannot be of the same quality as for single-crystal data. In particular, low residual values can be attained between observed and calculated I values but that does not mean that a fully satisfactory structure refinement has been accomplished. However, the values of the residuals suggest if the correct (even if unrefined) solution has been found.

In order to give an example, the performances of SIRPOW.92 on the test structures are shown in Table 5. It is evident from the table that SIRPOW.92 is often successful and that the percentage of atoms found is in general sufficient for allowing the structure completion by Rietveld refinement.

11. Conclusions

The great efficiency of direct methods (structures up to 300 atoms in the asymmetric unit can be solved) deteriorates when they are applied to powder data. The main limit is the uncertainty in the structure-factor amplitudes, which are estimated with limited accuracy by full-pattern-decomposition programs. Since errors are unavoidable, the risk of failure is remarkably higher than for single-crystal data. What are the perspectives for improving the state of the art? They should not be looked for in new direct-methods developments, but essentially in two fields:

(a) the decomposition of the diffraction pattern;

(b) the crystal structure refinement.

Significant improvements in (a) will generate higher efficiency in (b). Indeed, the recovery of a more complete structural model would be easier, the parameters under refinement would assume more accurate values and, very likely, the size of the structure accessible to direct methods would be enlarged.

Luckily, there is a reserve of power that has not been fully exploited by the current methods for pattern decomposition. Indeed, the Le Bail algorithm, as well as the constrained Pawley techniques, only uses a limited amount of prior information: i.e. the mere experimental data, the space-group symmetry and the non-negativity

Table 5. *Results after application of a default run of SIRPOW.92*

NAT is the number of atoms to be found, NATP the atoms found by SIRPOW.92, R_F is the final crystallographic residual R.

Code	NAT	NATP	$R_F \times 100$
AGPZ	6	5	14.35
BACO	6	6	11.69
BAER	16	10	–
BENZ	9	8	24.81
CABA	7	5	12.48
CFBR3	4	4	10.46
CF3BR	5	5	16.97
CFCL	3	3	8.64
CFI	4	4	18.25
CIM	17	14	28.03
CROP	8	7	–
CROX	15	13	11.04
CUPZ	6	3	14.24
ERVO	3	3	–
FECOAS	3	3	15.47
GLYCIN	9	9	22.16
HOBA	8	7	14.91
HOFE	4	4	16.54
LAMO	14	14	–
LASI	11	9	–
LEV	17	11	11.46
MES	12	10	23.17
METYL	5	5	18.41
NBPO	22	10	23.86
NDBA	7	5	17.71
NIZR	18	11	14.61
PBS	6	5	9.02
PICO	7	6	15.33
RUCO	5	4	21.57
RUHG	20	4	17.33
SAPO	21	14	15.55
SBPO	17	8	12.86
SIGMA2	17	13	15.62
STRIA	6	5	25.91
SULPH	5	5	14.31
T83	5	4	17.08
TEAS	3	3	12.48
TENI	3	3	5.77
TIPORF	17	7	23.35
TRAN	6	3	21.97
YONO	18	17	7.20

of the intensity values. No use is made of the positivity of the electron density, which is responsible for the X-ray scattering, and of the variety of information that becomes available during the crystal-structure-solution process. This includes: (a) pseudo-translational symmetry; (b) preferred orientation; (c) Patterson map; (d) partial structure. The crystallographic group in Bari has initiated the systematic integration of such information in the new version of EXTRA. We anticipate that the method will provide a dramatic improvement in the accuracy of structure-factor-amplitude estimates.

References

Altomare, A., Cascarano, G., Giacovazzo, C. & Guagliardi, A. (1994). *J. Appl. Cryst.* **27**, 1045–1050.

Altomare, A., Cascarano, G., Giacovazzo, C., Guagliardi, A., Burla, M. C., Polidori, G. & Camalli, M. (1994). *J. Appl. Cryst.* **27**, 435–436.

Altomare, A., Cascarano, G., Giacovazzo, C., Guagliardi, A., Moliterni, A. G., Burla, M. C. & Polidori, G. (1995). *J. Appl. Cryst.* **28**, 738–744.

Altomare, A., Giacovazzo, C., Guagliardi, A. & Siliqi, D. (1994). *Acta Cryst.* A**50**, 311–317.

Baerlocher, C. (1982). *The X-ray Rietveld System.* Institut für Kristallographie und Petrographie, Zürich, Switzerland.

Baerlocher, C. (1993). In *Crystallographic Computing 6*, edited by H. D. Flack, L. Párkányi & K. Simon, pp. 85–99. Oxford University Press.

Berar, J. F. (1990). IUCr Satellite Meeting on Powder Diffraction, Toulouse, France, July 1990.

Böhme, R. (1982). *Acta Cryst.* A**38**, 316–326.

Bricogne, G. (1991). *Acta Cryst.* A**47**, 803–829.

Bunge, H. J., Dahms, M. & Brokmeier, H. G. (1989). *Crystallogr. Rev.* **2**, 67–88.

Cascarano, G., Favia, L. & Giacovazzo, C. (1992). *J. Appl. Cryst.* **25**, 310–317.

Cascarano, G., Giacovazzo, C., Camalli, M., Spagna, R., Burla, M. C., Nunzi, A. & Polidori, G. (1984). *Acta Cryst.* A**40**, 278–283.

Cascarano, G., Giacovazzo, C. & Guagliardi, A. (1991). *Acta Cryst.* A**47**, 698–702.

Cascarano, G., Giacovazzo, C. & Guagliardi, A. (1992). *Z. Kristallogr.* **200**, 63–71.

Cascarano, G., Giacovazzo, C., Guagliardi, A. & Steadman, N. (1991). *Acta Cryst.* A**47**, 480–484.

Cascarano, G., Giacovazzo, C. & Luić, M. (1985). *Acta Cryst.* A**41**, 544–551.

Cascarano, G., Giacovazzo, C. & Luić, M. (1987). *Acta Cryst.* A**43**, 14–22.

Cascarano, G., Giacovazzo, C. & Luić, M. (1988). *Acta Cryst.* A**44**, 176–183, 183–188.

Cascarano, G., Giacovazzo, C., Luić, M. & Vicković, I. (1989). *Acta Cryst.* A**45**, 441–443.

Cascarano, G., Giacovazzo, C., Moliterni, A. G. & Polidori, G. (1994). *Acta Cryst.* A**50**, 22–27.

Cochran, W. (1955). *Acta Cryst.* **8**, 473–478.

Cochran, W. & Douglas, A. S. (1957). *Proc. R. Soc. London Ser. A*, **243**, 281.

David, W. I. F. (1987). *J. Appl. Cryst.* **20**, 316–319.

Debaerdemaeker, T., Tate, C. & Woolfson, M. M. (1988) *Acta Cryst.* A**44**, 353–357.

Elsenhans, O. (1990). *J. Appl. Cryst.* **23**, 73–76.

Estermann, M. A. & Gramlich, V. (1993). *J. Appl. Cryst.* **26**, 396–404.

Estermann, M. A., McCusker, L. B. & Baerlocher, C. (1992). *J. Appl. Cryst.* **25**, 539–543.

Fan, H., Yao, Y., Main, P. & Woolfson, M. M. (1983). *Acta Cryst.* A**39**, 566–569.

Fitch, A. N., Murray, A. D. & Jouanneaux, A. (1995). Private communication.

Giacovazzo, C. (1975). *Acta Cryst.* A**31**, 252–259.

Giacovazzo, C. (1976). *Acta Cryst.* A**32**, 91–99, 100–104.

Gilmore, C. J., Henderson, K. & Bricogne, G. (1991). *Acta Cryst.* A**47**, 830–841.

Gramlich, V. (1984). *Acta Cryst.* A**40**, 610–616.

Hall, S. R. & Subramanian, V. (1982). *Acta Cryst.* A**38**, 590–598, 598–608.

Hauptman, H. (1975). *Acta Cryst.* A**31**, 671–679, 680–687.

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

Immirzi, A. (1980). *Acta Cryst.* B**36**, 2378–2385.

Izumi, F. (1989). *Rigaku J.* **6**, No. 1, 10–19.

Jansen, E., Schäfer, W. & Will, G. (1988). *J. Appl. Cryst.* **21**, 228–239.

Jansen, J., Peschar, R. & Schenk, H. (1992). *J. Appl. Cryst.* **25**, 231–236.

Järvinen, M. (1993). *J. Appl. Cryst.* **26**, 525–531.

Larson, A. C. & Von Dreele, R. B. (1987). Report LA-UR-86-748. Los Alamos National Laboratory, NM, USA.

Le Bail, A., Duray, H. & Fourquet, J. L. (1988). *Math. Res. Bull.* **23**, 447–452.

Louër, D., Louër, M. & Touboul, M. (1992) *J. Appl. Cryst.* **25**, 617–623.

Lutterotti, L. & Scardi, P. (1990). *J. Appl. Cryst.* **23**, 246–252.

McCusker, L. B. (1988). *J. Appl. Cryst.* **21**, 305–310.

Morris, R. E., Harrison, W. T. A., Nicol, J. M. N., Wilkinson, A. P. & Cheetham, A. K. (1992). *Nature (London)*, **359**, 519–522.

Morris, R. E., Owen, J. J., Stalick, J. K. & Cheetham, A. K. (1994). *J. Solid State Chem.* **111**, 52–57.

Pawley, G. S. (1980). *J. Appl. Cryst.* **13**, 630–633.

Pawley, G. S. (1981). *J. Appl. Cryst.* **14**, 357–361.

Peschar, R., Schenk, H. & Capková, P. (1995). *J. Appl. Cryst.* **28**, 127–140.

Riello, P., Fagherazzi, G., Clemente, D. & Canton, P. (1995). *J. Appl. Cryst.* **28**, 115–120.

Rodriguez-Carvajal, J. (1990). Abstracts of the Satellite Meeting on Powder Diffraction of the XV Congress of the IUCr, Toulouse, France, p. 127.

Rogers, D., Stanley, E. & Wilson, A. J. C. (1955). *Acta Cryst.* **8**, 383–393.

Schenk, H. (1973). *Acta Cryst.* A**29**, 77–82, 480–481.

Sheldrick, G. M. (1985). *SHELX86. Crystallographic Computing 3*, edited by G. M. Sheldrick, C. Krüger & R. Goddard, pp. 175–189. Oxford University Press.

Sheldrick, G. M. (1991). In *Direct Methods in Solving Crystal Structures*, edited by H. Schenk. New York: Plenum.

Sivia, D. S. & David, W. I. F. (1994). *Acta Cryst.* A**50**, 703–714.

Suortti, P. & Jennings, L. D. (1977). *Acta Cryst.* A**33**, 1012–1027.

Toraya, H. (1986). *J. Appl. Cryst.* **19**, 440–447.

Will, G., Jansen, E. & Schäfer, W. (1982) *Neutron Scattering* (1981). *AIP Conf. Proc.* No. 89, pp. 205–208.