Multiple-wavelength anomalous dispersion techniques applied to powder data: a probabilistic method for finding the substructure *via* joint probability distribution functions

Angela Altomare,^a Benny Danilo Belviso,^a Maria Cristina Burla,^b Gaetano Campi,^a Corrado Cuocci,^a Carmelo Giacobazzo,^{a,c,*} Fabia Gozzo,^d Anna Moliterni,^a Giampiero Polidori^b and Rosanna Rizzi^a

^aIstituto di Cristallografia, Sede di Bari, Via G. Amendola 122/o, 70126 Bari, Italy, ^bDipartimento di Scienze della Terra – Piazza Università, 06100 Perugia, Italy, ^cDipartimento Geomineralogico, Università di Bari, Campus Universitario, Via Orabona 4, 70125 Bari, Italy, and ^dSwiss Light Source, Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland. Correspondence e-mail: carmelo.giacobazzo@ic.cnr.it

A new joint probability distribution function method is described to find the anomalous scatterer substructure from powder data. The method requires two wavelengths; the conclusive formulas provide estimates of the substructure structure factor moduli, from which the anomalous scatterer positions can be found by Patterson or direct methods. The theory has been preliminarily applied to two compounds, the first having Pt and the second having Fe as anomalous scatterer. Both substructures were correctly identified.

© 2009 International Union of Crystallography
Printed in Singapore – all rights reserved

1. Notation

The following notation has been used in this article.

N : number of atoms in the unit cell.

a : number of anomalous scatterers in the unit cell.

$na = N - a$: number of non-anomalous scatterers.

ε : statistical Wilson coefficient.

$f_j = f_j^o + \Delta f_j + if_j'' = f_j' + if_j''$: scattering factor of the j th atom; f' is the real and f'' the imaginary part; the thermal factor is included.

$\sum_{Np} = \sum_{j=1}^N (f_j'^2 + f_j''^2)$: the summation is calculated at the p th wavelength and is extended to all the atoms in the unit cell.

$\sum'_{Np} = \sum_{j=1}^N (f_j'^2)$: the summation is calculated at the p th wavelength and is extended to all the atoms in the unit cell.

$\sum_o = \sum_{j=1}^{na} (f_j^o)^2$: the summation is extended to all the non-anomalous scatterers in the unit cell.

$\sum_{oa} = \sum_{j=1}^a (f_j^o)^2$: the summation is extended to all the anomalous scatterers in the unit cell.

$F^+ = |F^+| \exp(i\varphi^+) = F_{\mathbf{h}} = \sum_{j=1}^N f_j \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j)$

$E^+ = F^+ / (\varepsilon \sum_N)^{1/2} = R \exp(i\varphi^+) = A^+ + iB^+$

$F^- = |F^-| \exp(i\varphi^-) = F_{-\mathbf{h}} = \sum_{j=1}^N f_j \exp(-2\pi i \mathbf{h} \cdot \mathbf{r}_j)$

$E^- = F^- / (\varepsilon \sum_N)^{1/2} = G \exp(i\varphi^-) = A^- + iB^-$

$F_p^+, F_p^-, E_p^+ = A_p^+ + iB_p^+, E_p^- = A_p^- + iB_p^-$: values for the p th wavelength.

$$F_{oa} = |F_{oa}| \exp(i\varphi_{oa}) = \sum_{j=1}^a f_j^o \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j)$$

$$E_{oa} = F_{oa} / (\varepsilon \sum_{oa})^{1/2} = R_{oa} \exp(i\varphi_{oa}) = A_{oa} + iB_{oa}$$

$$\Delta_{ano} = |F^+| - |F^-|$$

2. Introduction

Phase determination *via* multiple-wavelength anomalous dispersion (MAD) techniques is one of the most popular approaches for the solution of the phase problem. It is mostly applied when the amount of information available from the diffraction experiment is not sufficient to determine the structure. That occurs in practice when the complexity is high and/or when the data resolution is low. A classical MAD technique is essentially a three-step procedure:

(a) The estimation of the structure factor moduli of the anomalous scatterer substructure. Among the different approaches we quote the least-squares procedure suggested by Karle (1980), Hendrickson (1985) and Pähler *et al.* (1990), the Bayesian approach adopted by Terwilliger (1994), and the use of the joint probability distribution function suggested by Burla *et al.* (2002).

(b) The location of the anomalous scatterers *via* Patterson techniques (Terwilliger & Berendzen, 1999; Grosse-Kunstleve & Brunger, 1999; Burla *et al.*, 2007a) or direct methods (Miller *et al.*, 1994; Sheldrick, 1998; Smith *et al.*, 1998; Foadi *et al.*, 2000; Burla *et al.*, 2003).

(c) The protein phase estimation (Otwinowski, 1991; La Fortelle & Bricogne, 1997; Cowtan, 1994; Giacobazzo & Siliqi, 2004).

Most of the authors mentioned above applied anomalous dispersion to solve macromolecular structures *via* single-crystal data; little attention has been devoted to crystal structure solution *via* powder data. The reasons are basically twofold:

(i) MAD techniques cannot be applied to powder data without suitable modifications. Since Bijvoet pairs overlap in the powder diagram, $|F^+|^2$ and $|F^-|^2$ cannot be individually evaluated; only their sum (*i.e.* $|F^+|^2 + |F^-|^2$) is experimentally available, with a consequent loss of information. This feature makes the SAD method unsuitable for powders and reduces the efficiency of MAD.

(ii) Peak overlapping is frequently so severe that the full pattern decomposition (in terms of structure factor moduli) is often inefficient. Structure factor moduli are usually estimated with an average error of 40–50%, irrespective of whether the Le Bail (Le Bail *et al.*, 1988) or the Pawley (1981) technique is used. Average errors in single-crystal data of 5–10% on $|F^+|$ and on $|F^-|$ can lead SAD–MAD to fail.

The application of MAD to powder data is, however, not hopeless. Prandl (1990, 1994) adapted the classical Hendrickson procedure to the features of a powder diagram. He concluded that for the determination of the sign in a centric structure a single anomalous scatterer is sufficient; two data sets have to be collected, one at a wavelength close to the absorption edge of the anomalous scatterer and one far away from it. For acentric structures two different anomalous scatterers must be present. Therefore, three data sets have to be collected, two close to the absorption edges and the third far away from them. Prandl employed difference Patterson maps to locate the anomalous scatterers.

Gu *et al.* (2000) used simulated two-wavelength data for $C_{14}H_{20}O_3N_2 \cdot HBr$ to find the anomalous scatterer position (Br) and direct methods to complete the structure. Helliwell *et al.* (2005) successfully applied an algebraic approach for the *ab initio* structure solution of $Ni(SO_4) \cdot 6H_2O$; synchrotron-radiation data were collected at two wavelengths close to the K-edge for Ni, and at three wavelengths remote from the Ni absorption edge (1.3, 1.8 and 2.16 Å).

The probabilistic approaches exploiting anomalous dispersion are today largely applied to single-crystal data; they proved to be much more effective than algebraic techniques, which are nowadays considered obsolete. It may be expected that they are also more effective than algebraic techniques when applied to powder data. It is, however, impossible to transfer such probabilistic methods from single crystals to powders without heavy modifications. This task has been accomplished in this paper, where a two-wavelength probabilistic procedure is described which adapts to powder data the approach described by Burla *et al.* (2002, 2003). As for single-crystal data, probabilistic estimates are obtained for the structure factor moduli of the anomalous substructure, to which Patterson or direct methods may be applied.

While anomalous dispersion methods made a significant contribution to the advent of high-throughput crystallography in the macromolecular area, a similar situation is presently not realistic for powder crystallography. This is mainly a result of

the unavoidable peak overlapping in powder patterns. However, the spectacular improvements in the characteristics of synchrotron light, and the increasing ability to reduce the peak width, encourage a supplemental study of the anomalous dispersion methods. The approach described in this paper is an effort in this direction. The method has been successfully checked using two test structures; neither of them is challenging, and the tests have to be considered as preliminary.

3. The joint probability distribution $P(E_{oa}, \bar{E}_1, \bar{E}_2)$

Let us study the case in which a single species of anomalous scatterer is present. We assume that the positions of all the atoms in the asymmetric unit are the primitive random variables in our probabilistic approach; they are supposed to be uniformly distributed in the unit cell. Denoting the normalized structure factors at the p th wavelength by

$$E_p^+ = A_p^+ + iB_p^+, \quad E_p^- = A_p^- + iB_p^-, \quad p = 1, 2 \quad (1)$$

gives

$$A_p^+ = \left[\sum_{j=1}^{na} f_j^o \cos(2\pi \mathbf{h} \mathbf{r}_j) + \sum_{j=1}^a f_{jp}' \cos(2\pi \mathbf{h} \mathbf{r}_j) - \sum_{j=1}^a f_{jp}'' \sin(2\pi \mathbf{h} \mathbf{r}_j) \right] / (\varepsilon \sum_{Np})^{1/2}, \quad (2)$$

$$B_p^+ = \left[\sum_{j=1}^{na} f_j^o \sin(2\pi \mathbf{h} \mathbf{r}_j) + \sum_{j=1}^a f_{jp}' \sin(2\pi \mathbf{h} \mathbf{r}_j) + \sum_{j=1}^a f_{jp}'' \cos(2\pi \mathbf{h} \mathbf{r}_j) \right] / (\varepsilon \sum_{Np})^{1/2}, \quad (3)$$

$$A_p^- = \left[\sum_{j=1}^{na} f_j^o \cos(2\pi \mathbf{h} \mathbf{r}_j) + \sum_{j=1}^a f_{jp}' \cos(2\pi \mathbf{h} \mathbf{r}_j) + \sum_{j=1}^a f_{jp}'' \sin(2\pi \mathbf{h} \mathbf{r}_j) \right] / (\varepsilon \sum_{Np})^{1/2}, \quad (4)$$

$$B_p^- = \left[-\sum_{j=1}^{na} f_j^o \sin(2\pi \mathbf{h} \mathbf{r}_j) - \sum_{j=1}^a f_{jp}' \sin(2\pi \mathbf{h} \mathbf{r}_j) + \sum_{j=1}^a f_{jp}'' \cos(2\pi \mathbf{h} \mathbf{r}_j) \right] / (\varepsilon \sum_{Np})^{1/2}, \quad (5)$$

$$A_{oa} = \sum_{j=1}^a f_j^o \cos(2\pi \mathbf{h} \mathbf{r}_j) / (\varepsilon \sum_{oa})^{1/2}, \quad (6)$$

$$B_{oa} = \sum_{j=1}^a f_j^o \sin(2\pi \mathbf{h} \mathbf{r}_j) / (\varepsilon \sum_{oa})^{1/2}. \quad (7)$$

Giacovazzo & Siliqi (2002) derived, for single-crystal data applications, the joint probability distribution $P(E_{oa}, E_1^+, E_2^+, E_1^-, E_2^-)$, from which the conditional distribution $P(R_{oa} | R_1^+, R_2^+, R_1^-, R_2^-)$ was obtained, allowing the estimation of R_{oa} given the observed moduli $R_1^+, R_2^+, R_1^-, R_2^-$. Unfortunately it is impossible in powder diffraction to estimate Bijvoet pairs separately, and therefore the distribution $P(E_{oa}, E_1^+, E_2^+, E_1^-, E_2^-)$

E_2^-) is of little use. A more useful variable to bring into the probabilistic approach is

$$\bar{E}_p = (1/2)(E_p^+ + E_p^{-*}), \quad (8)$$

where E_p^{-*} is the complex conjugate of E_p^- . Indeed $|\bar{E}_p|$ may be experimentally estimated, provided (see Fig. 1)

$$|F''^+| = \left| \left[-\sum_{j=1}^a f''_{jp} \sin(2\pi \mathbf{h} \mathbf{r}_j) + i \sum_{j=1}^a f''_{jp} \cos(2\pi \mathbf{h} \mathbf{r}_j) \right] \right| \quad (9)$$

is negligible with respect to $|F^+|$ and $|F^{-*}|$. That occurs when $|E^+|$ and $|E^{-*}|$ are sufficiently large or, more generally, when f'' are nearly vanishing. In this case the phase values φ_p^+ and φ_p^{-*} differ by a few degrees, and the following approximation holds:

$$|\bar{E}_p| \simeq (1/2)(|E_p^+| + |E_p^{-*}|). \quad (10)$$

The coexistence of equations (8) and (10) allows us to derive, from the joint probability distribution function $P(E_{oa}, \bar{E}_1, \bar{E}_2)$, the conditional distribution $P(R_{oa} | \bar{R}_1, \bar{R}_2)$.

According to Giacovazzo & Siliqi (2002), it is useful to assign to $\bar{E}_p$ a complex error μ_p . It includes both the error on the measured diffraction intensities (at the angular steps of the experimental pattern) and the error arising from the full pattern decomposition procedure. It is defined as a complex quantity because its effects influence both the real and the imaginary components of $\bar{E}_p$. Then we can write

$$\bar{E}_p = (1/2)(E_p^+ + E_p^{-*}) + \mu_p \exp(i\theta_p) = \bar{A}_p + i\bar{B}_p, \quad (11)$$

where

$$\bar{A}_p = \left[\sum_{j=1}^{na} f_j^o \cos(2\pi \mathbf{h} \mathbf{r}_j) + \sum_{j=1}^a f'_{jp} \cos(2\pi \mathbf{h} \mathbf{r}_j) + |\mu_p| \cos \theta_p \right] / (\varepsilon \sum_{Np}')^{1/2}, \quad (12)$$

$$\bar{B}_p = \left[\sum_{j=1}^{na} f_j^o \sin(2\pi \mathbf{h} \mathbf{r}_j) + \sum_{j=1}^a f'_{jp} \sin(2\pi \mathbf{h} \mathbf{r}_j) + |\mu_p| \sin \theta_p \right] / (\varepsilon \sum_{Np}')^{1/2}. \quad (13)$$

We will suppose, in the absence of different prior information, that

Figure 1
The geometrical relations among F^+ , F^{-*} , $\bar{F}$, F''^+ and F''^{-*} .

$$\langle \mu_1 \rangle = \langle \mu_2 \rangle = 0 \text{ and } \langle \mu_1 \mu_2 \rangle = 0. \quad (14)$$

Let $u_a, u_1, u_2, v_a, v_1, v_2$ be carrying variables associated with $A_{oa}, \bar{A}_1, \bar{A}_2, B_{oa}, \bar{B}_1, \bar{B}_2$; then the characteristic function of $P(A_{oa}, \bar{A}_1, \bar{A}_2, B_{oa}, \bar{B}_1, \bar{B}_2)$ is

$$\begin{aligned} C(u_a, u_1, u_2, v_a, v_1, v_2) &\simeq \exp \left\{ -(1/4) [k_{11}(u_a^2 + v_a^2) \right. \\ &\quad + k_{22}(u_1^2 + v_1^2) + k_{33}(u_2^2 + v_2^2) + 2k_{12}(u_a u_1 + v_a v_1) \\ &\quad \left. + 2k_{13}(u_a u_2 + v_a v_2) + 2k_{23}(u_1 u_2 + v_1 v_2)] \right\} \\ &= \exp[-(1/4)(\bar{\mathbf{U}} \mathbf{K} \mathbf{U})], \end{aligned} \quad (15)$$

where

$$\bar{\mathbf{U}} = [u_a, u_1, u_2, v_a, v_1, v_2], \quad (16)$$

$$\mathbf{K} = \begin{bmatrix} 1 & k_{12} & k_{13} & 0 & 0 & 0 \\ k_{12} & k_{22} & k_{23} & 0 & 0 & 0 \\ k_{13} & k_{23} & k_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & k_{12} & k_{13} \\ 0 & 0 & 0 & k_{12} & k_{22} & k_{23} \\ 0 & 0 & 0 & k_{13} & k_{23} & k_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix}, \quad (17)$$

$\mathbf{Q}$ is a 3×3 matrix, and

$$k_{22} = k_{55} = e_1 = 1 + \sigma_1^2 \text{ with } \sigma_1^2 = \langle |\mu_1|^2 \rangle / (\varepsilon \sum_{N1}'), \quad (18)$$

$$k_{33} = k_{66} = e_2 = 1 + \sigma_2^2 \text{ with } \sigma_2^2 = \langle |\mu_2|^2 \rangle / (\varepsilon \sum_{N2}'), \quad (19)$$

$$k_{12} = k_{45} = S_9 / (\sum_{oa} \sum_{N1}')^{1/2}, \quad (20)$$

$$k_{13} = k_{46} = S_{11} / (\sum_{oa} \sum_{N2}')^{1/2}, \quad (21)$$

$$k_{23} = k_{56} = (\sum_o + S_5) / (\sum_{N1}' \sum_{N2}')^{1/2}, \quad (22)$$

$$S_5 = \sum_a f'_{j1} f'_{j2}, \quad S_9 = \sum_a f_j^o f'_{j1}, \quad S_{11} = \sum_a f_j^o f'_{j2}. \quad (23)$$

The joint probability distribution function $P(A_{oa}, \bar{A}_1, \bar{A}_2, B_{oa}, \bar{B}_1, \bar{B}_2)$ (hereafter abbreviated to P), is obtained by Fourier inversion of equation (3). We obtain

$$\begin{aligned} P &\simeq \pi^{-3} [\det \mathbf{K}]^{-1/2} \exp \left\{ -[\lambda_{11}(A_{oa}^2 + B_{oa}^2) + \lambda_{22}(\bar{A}_1^2 + \bar{B}_1^2) \right. \\ &\quad + \lambda_{33}(\bar{A}_2^2 + \bar{B}_2^2) + 2\lambda_{12}(A_{oa} \bar{A}_1 + B_{oa} \bar{B}_1) \\ &\quad \left. + 2\lambda_{13}(A_{oa} \bar{A}_2 + B_{oa} \bar{B}_2) + 2\lambda_{23}(\bar{A}_1 \bar{A}_2 + \bar{B}_1 \bar{B}_2)] \right\}. \end{aligned} \quad (24)$$

Defining

$$\begin{aligned} A_{oa} &= R_{oa} \cos \varphi_{oa}, & B_{oa} &= R_{oa} \sin \varphi_{oa}, \\ \bar{A}_j &= \bar{R}_j \cos \bar{\varphi}_j, & \bar{B}_j &= \bar{R}_j \sin \bar{\varphi}_j, \quad j = 1, 2, \end{aligned} \quad (25)$$

gives

$$\begin{aligned} P &\simeq \pi^{-3} R_{oa} \bar{R}_1 \bar{R}_2 [\det \mathbf{K}]^{-1/2} \exp \left\{ -[\lambda_{11} R_{oa}^2 + \lambda_{22} \bar{R}_1^2 + \lambda_{33} \bar{R}_2^2 \right. \\ &\quad + 2\lambda_{12} R_{oa} \bar{R}_1 \cos(\varphi_{oa} - \bar{\varphi}_1) + 2\lambda_{13} R_{oa} \bar{R}_2 \cos(\varphi_{oa} - \bar{\varphi}_2) \\ &\quad \left. + 2\lambda_{23} \bar{R}_1 \bar{R}_2 \cos(\bar{\varphi}_1 - \bar{\varphi}_2)] \right\}. \end{aligned} \quad (26)$$

The estimation of R_{oa} cannot be performed *via* exact calculations. The results obtained by Giacovazzo & Siliqi (2001a,b,c) suggest the usefulness of the approximation

$$\bar{\varphi}_1 \simeq \bar{\varphi}_2. \quad (27)$$

Introducing equations (27) into (26) and integrating over φ_1 gives

$$P(R_{oa}, \bar{R}_1, \bar{R}_2) \simeq 2\pi S R_{oa} \bar{R}_1 \bar{R}_2 \exp \left\{ -[\lambda_{11} R_{oa}^2 + \lambda_{22} \bar{R}_1^2 + \lambda_{33} \bar{R}_2^2] I_0(2R_{oa} X) \right\}, \quad (28)$$

where S is a suitable normalizing constant, I_0 is the modified Bessel function of order zero and

$$X = (\lambda_{12} \bar{R}_1 + \lambda_{13} \bar{R}_2). \quad (29)$$

Then [see Burla *et al.* (2002) for details]

$$P(R_{oa} | \bar{R}_1, \bar{R}_2) = 2\lambda_{11} R_{oa} \exp(-\lambda_{11} R_{oa}^2) \times \exp(-X^2/\lambda_{11}) I_0(2R_{oa} X), \quad (30)$$

from which

$$\langle R_{oa} | \bar{R}_1, \bar{R}_2 \rangle = 2^{-1} (\pi/\lambda_{11})^{1/2} {}_1F_1(-1/2; 1; -X^2/\lambda_{11}). \quad (31)$$

${}_1F_1$ is the confluent hypergeometric function; since ${}_1F_1(-1/2; 1; -z^2)$ is well approximated by the hyperbole $y = (1 + 4z^2/\pi)^{1/2}$ in the full range $(0, \infty)$, the expected value of R_{oa} may be calculated *via* the simpler expression

$$\langle R_{oa} | \bar{R}_1, \bar{R}_2 \rangle = (1/2)(\pi/\lambda_{11})^{1/2} [1 + (4/\pi)(X^2/\lambda_{11})]^{1/2}, \quad (32)$$

with standard deviation $\sigma_{R_{oa}}$ given by

$$\sigma_{R_{oa}} = [\langle R_{oa}^2 | \dots \rangle - \langle R_{oa} | \dots \rangle^2]^{1/2} = [(1 - \pi/4)\lambda_{11}^{-1}]^{1/2}, \quad (33)$$

and

$$\frac{\langle R_{oa} | \dots \rangle}{\sigma_{R_{oa}}} = \left(\frac{\pi/4 + X^2/\lambda_{11}}{1 - \pi/4} \right)^{1/2}. \quad (34)$$

4. Experimental

A full realization of the potential of the theory described in §3 requires extensive investigations involving compounds with different structural complexity, with different species of anomalous scatterers, and belonging both to centric and to acentric structures. All that is not in the scope of this paper. Some preliminary tests are, however, made. We used as test structures two compounds:

(1) Iron(II) phthalocyanine bis(pyridine), $C_{32}H_{16}N_8Fe(C_5H_5N)_2$, from now on denoted as IRON2. The powder was bought from Sigma–Aldrich (<http://www.sigma-aldrich.com>; declared purity over 95%). The structure was originally solved by Janczak & Kubiak (2003) using single-crystal data; it is therefore suitable to test our joint probability distribution function method. The published unit-cell parameters are $a = 9.576$, $b = 19.929$, $c = 9.179$ Å, $\beta = 111.7^\circ$, space group $P2_1/c$, $Z = 2$.

The Fe atom is on a special position (an inversion centre) and therefore does not contribute to half the reflections (*i.e.* those with $k + l = 2n + 1$). *Vice versa*, the reflections with $k + l = 2n + 1$ do not provide any information for the location of Fe.

Table 1

Scattering factors for the test structures IRON2 and PLAT4.

For each test structure we give the selected wavelengths, the literature values for $\Delta f'$ and f'' expected at the dip (for f'') and for the remote wavelengths, the experimental 2θ ranges (RAN), and the corresponding number of reflections (NREF). The last column gives the values of $RAT_a = \sum_a \Delta f'^2 / \sum_{j=1}^N (f_j^o)^2$ calculated at $2\theta_{max}$ for the dip wavelengths.

Code	λ (Å)	$\Delta f'$	f''	RAN (2θ)	NREF	RAT_a
IRON2	1.0729	0.0	0.0	1.0–40.9	464	
IRON2	1.7436	−8.0	0.6	1.0–60.9	332	0.09
PLAT4	0.9549	0.0	0.0	1.0–55.9	1826	
PLAT4	1.0729	−16.0	4.0	1.0–55.9	1289	0.10

(2) *trans*-Dichloridediacetateammine(1-adamantylamine) Pt^{IV} , $C_{14}H_{26}N_2O_4Cl_2Pt$, from now on denoted as PLAT4. Its structure was unknown, probably because single crystals of sufficient size are not available. Preliminary data were collected by laboratory diffractometer; indexing was performed by *N-TREOR* (Altomare *et al.*, 2000), providing unit-cell parameters of $a = 11.106$, $b = 13.982$, $c = 6.357$ Å, $\alpha = 102.28$, $\beta = 104.06$, $\gamma = 77.36^\circ$, allowed space groups $P1$ or $\bar{P}1$.

The MAD data for both the test structures were collected at the Swiss Light Source Materials Science beamline powder diffraction station (Gozzo *et al.*, 2004). The accurate positioning of the wavelength at or right before the scatterer absorption edge was performed by rocking the main axis of the monochromator around the nominal angle while collecting the transmitted signal from a thin pure absorber (Fe for IRON2 and Pt for PLAT4). The almost negligible backlash in the monochromator main rotation axis was minimized by running the motor always in the same direction and by looking at the intensity of the transmission curve. The diffraction patterns were collected using the high-resolution multicrystal analyser detector; a very accurate optical and detector setup was performed prior to the data acquisition to eliminate all residual aberrations coming from the beamline optics (Gozzo *et al.*, 2006). The 2θ independent peak shape was a pure Gaussian with an almost negligible Lorentzian component. Once the monochromator main axis angle was defined and fixed for data acquisition, the wavelength was very accurately determined by collecting and refining a 2θ diffraction pattern from silicon (NIST 640c). The remote wavelength diffraction patterns were collected for both compounds at wavelengths that were far enough from the absorption edge to make the anomalous corrections to the structure factor negligible, but not too far away from the near-edge wavelength to guarantee that the instrumental function had not substantially changed. The wavelengths and the corresponding literature values of Δf and f'' for the anomalous scatterers of the two test structures are given in Table 1.

5. Applications

The two-wavelength synchrotron data of each test structure were submitted to *EXPO2004* (Altomare *et al.*, 2004) for the identification of the unit cell and of the space group. The results previously obtained *via* home diffractometer data were

substantially confirmed: the new cell parameters were $a = 9.560$, $b = 19.937$, $c = 9.183$ Å, $\beta = 111.7^\circ$, space group $P2_1/c$, $Z = 2$ for IRON2, and $a = 11.100$, $b = 13.981$, $c = 6.352$ Å, $\alpha = 102.3$, $\beta = 104.0$, $\gamma = 77.5^\circ$, space group $P1$ or $P\bar{1}$ for PLAT4. The structure was actually solved both in $P1$ and in $P\bar{1}$; the analysis of the atomic positions obtained in $P1$ suggested that the true space group is $P\bar{1}$.

For each test structure and for each wavelength the following algorithm has been implemented in a modified version of *EXPO2004* and used in a default way:

(1) Each experimental pattern (independently from the others) was decomposed according to the Le Bail method and the integrated intensities for each reflection were extracted. The measured 2θ range [$\text{RAN} = (2\theta_{\min} - 2\theta_{\max})$] and the corresponding number of reflections (NREF) lying in the range are shown in Table 1.

(2) The reflection intensities were normalized according to the classical Wilson plot, to derive the $\bar{R}_1$, $\bar{R}_2$ values employed in the probabilistic approach described in §3.

(3) The correlation factor between $\bar{R}_1$ and $\bar{R}_2$ was calculated in intervals of $\sin^2 \theta / \lambda^2$ to allow the exclusion (in cases where the values are too small) of some intervals from the subsequent calculations (see Fig. 2). Such practice is currently used in MAD techniques for single-crystal protein data (Schneider & Sheldrick, 2002; Burla *et al.*, 2004), where the scattering power of the anomalous scatterers is a quite small percentage of the total unit-cell scattering. For both our test structures the anomalous scattering is a non-negligible percentage of the normal scattering, at least at high resolution. (See in Table 1 the ratio $\sum_a \Delta f^2 / \sum_{j=1}^N (f_j^0)^2$, calculated at $2\theta_{\max}$ for the dip wavelength, by considering thermal vibration to be equal for all the atoms. The calculated percentage should be larger if one takes into account the Pt or Fe atomic weights.) Accordingly, the low correlation value at high resolution for PLAT4 does not necessarily imply large errors in the full pattern decomposition process; it may be considered the consequence of the dominant anomalous scattering. In the present case, no interval was eliminated from the subsequent calculations for either of the test structures.

(4) The standard deviations σ_j , $j = 1, 2$, which take into account both measurement and peak decomposition errors, are not provided by the Le Bail method. We used, for each

Figure 2
The correlation between $\bar{R}_1$ and $\bar{R}_2$ versus 2θ for IRON2 (solid line) and PLAT4 (dashed line).

Table 2

Results obtained by the Patterson and tangent methods.

For each test structure the number of correct solutions (NSOL) is given versus the number of performed trials (NTRIALS) for each phasing method (PHAS). NFIG is the order of the first correct solution as ranked by the MABS figure of merit.

CODE	PHAS	NSOL/NTRIALS	NFIG
IRON2	TANG	8/32	1
IRON2	PATT	22/30	1
PLAT4	TANG	6/32	2
PLAT4	PATT	4/30	2

reflection, the weight $\sigma_j^2 = 1/w_j$, $j = 1, 2$, where w_j is the weight recently proposed by Altomare *et al.* (2006) to take into account the peak overlapping in the experimental pattern.

The weights w influence the size of the $\langle R_{oa} | \bar{R}_1, \bar{R}_2 \rangle$ moduli provided by equation (32). The algorithm automatically rescales the initial w values until the percentage of the $\langle R_{oa} | \bar{R}_1, \bar{R}_2 \rangle$ values larger than 2 is about 1.8% of the total, as suggested by the Wilson distribution.

(5) The reflections were ordered according to $\langle R_{oa} | \dots \rangle / \sigma_{R_{oa}}$ as given by equation (34), where $1/\sigma_{R_{oa}} = (w_1 + w_2)/2$. The best ranked reflections were submitted to the standard *EXPO2004* direct methods routine; the number of trials required (NTRIALS) was set at 32. For comparison, the Patterson approach for determining substructures from powder data, recently described by Burla *et al.* (2007b), was also applied: 30 trials were explored. The results obtained by the two methods are shown in Table 2: a relatively large number of trials (NSOL) provides the correct solutions for both the test structures (for PLAT4, the structure of which was originally unknown, we refer to the structural model obtained by this MAD approach). The trials were ranked by the MABS figure of merit (Germain *et al.*, 1970); other criteria based on negative triplet invariants (Cascarano *et al.*, 1984) or on negative quartet invariants (Hauptman, 1975; Giacovazzo, 1976) were ineffective. An *a posteriori* (given the two structures) analysis of the results shows that the distances between the calculated and true anomalous scatterer positions are 0.0 Å for IRON2 and 0.066 Å for PLAT4. Of particular interest is the location of the Fe position in IRON2: owing to its special position, for 153 reflections (those with $k + l = 2n + 1$), $R_{oa} = 0$ systematically; for 26 reflections with $h + l = 0$ or $k = 0$, $R_{oa} = 1$; for 153 reflections with $k + l = 2n$, $R_{oa} = 2^{1/2}$ systematically. Such values are, however, enough to fix the Fe position.

6. Conclusions

Recent probabilistic methods for finding the anomalous substructure from single-crystal data encompassed traditional algebraic methods. No probabilistic theory was available for powder data. This paper fills the gap: a probabilistic approach for finding the anomalous scatterer substructure from two-wavelength powder data is described. Formulas are provided for estimating the structure factor moduli of the anomalous substructure, to which Patterson or tangent methods may be

applied. The approach was applied to two test structures, one solved and one unsolved, by using a general phasing algorithm particularly designed for facing the unavoidable (for powder data) peak overlapping problem. The method succeeded in a straightforward manner in both cases. The applications are not challenging (in particular the Fe position may be determined by symmetry considerations) and very probably the two structures may be solved by other routine techniques, using and not using anomalous dispersion. The tests, however, show that (i) the theory is basically correct, (ii) although designed for acentric cases, it also works for the centric ones, and (iii) the full potential of the approach is still to be discovered.

References

- Altomare, A., Caliendo, R., Camalli, M., Cuocci, C., Giacovazzo, C., Moliterni, A. G. G. & Rizzi, R. (2004). *J. Appl. Cryst.* **37**, 1025–1028.
- Altomare, A., Cuocci, C., Giacovazzo, C., Moliterni, A. G. G. & Rizzi, R. (2006). *J. Appl. Cryst.* **39**, 558–562.
- Altomare, A., Giacovazzo, C., Guagliardi, A., Moliterni, A. G. G., Rizzi, R. & Werner, P.-E. (2000). *J. Appl. Cryst.* **33**, 1180–1186.
- Burla, M. C., Caliendo, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G. & Siliqi, D. (2007a). *J. Appl. Cryst.* **40**, 211–217.
- Burla, M. C., Caliendo, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G. & Siliqi, D. (2007b). *J. Appl. Cryst.* **40**, 834–840.
- Burla, M. C., Carrozzini, B., Cascarano, G. L., Giacovazzo, C., Moustiakimov, M., Polidori, G. & Siliqi, D. (2004). *Acta Cryst.* **D60**, 1683–1686.
- Burla, M. C., Carrozzini, B., Cascarano, G. L., Giacovazzo, C. & Polidori, G. (2003). *Acta Cryst.* **D59**, 662–669.
- Burla, M. C., Carrozzini, B., Cascarano, G. L., Giacovazzo, C., Polidori, G. & Siliqi, D. (2002). *Acta Cryst.* **D58**, 928–935.
- Cascarano, G., Giacovazzo, C., Camalli, M., Spagna, R., Burla, M. C., Nunzi, A. & Polidori, G. (1984). *Acta Cryst.* **A40**, 278–283.
- Cowtan, K. (1994). *Joint CCP4/ESF-EACBM Newsletter on Protein Crystallography*, No. 31, pp. 31–38.
- Foadi, J., Woolfson, M. M., Dodson, E. J., Wilson, K. S., Jia-xing, Y. & Chao-de, Z. (2000). *Acta Cryst.* **D56**, 1137–1147.
- Germain, G., Main, P. & Woolfson, M. M. (1970). *Acta Cryst.* **B26**, 274–285.
- Giacovazzo, C. (1976). *Acta Cryst.* **A32**, 958–966.
- Giacovazzo, C. & Siliqi, D. (2001a). *Acta Cryst.* **A57**, 40–46.
- Giacovazzo, C. & Siliqi, D. (2001b). *Acta Cryst.* **A57**, 414–419.
- Giacovazzo, C. & Siliqi, D. (2001c). *Acta Cryst.* **A57**, 700–707.
- Giacovazzo, C. & Siliqi, D. (2002). *Acta Cryst.* **A58**, 590–597.
- Giacovazzo, C. & Siliqi, D. (2004). *Acta Cryst.* **D60**, 73–82.
- Gozzo, F., De Caro, L., Giannini, C., Guagliardi, A., Schmitt, B. & Prodi, A. (2006). *J. Appl. Cryst.* **39**, 347–357.
- Gozzo, F., Schmitt, B., Bortolamedi, Th., Giannini, C., Gagliardi, A., Lange, M., Meister, D., Willmott, P. & Patterson, B. D. (2004). *J. Alloys Compd.* **362**, 206–217.
- Grosse-Kunstleve, R. W. & Brunger, A. T. (1999). *Acta Cryst.* **D55**, 1568–1577.
- Gu, Y. X., Liu, Y. D., Hao, Q. & Fan, H. F. (2000). *Acta Cryst.* **A56**, 592–595.
- Hauptman, H. (1975). *Acta Cryst.* **A31**, 671–679.
- Helliwell, J. R., Helliwell, M. & Jones, R. H. (2005). *Acta Cryst.* **A61**, 568–574.
- Hendrickson, W. A. (1985). *Trans. Am. Crystallogr. Assoc.* **21**, 11–21.
- Janczak, J. & Kubiak, R. (2003). *Inorg. Chim. Acta*, **342**, 64–76.
- Karle, J. (1980). *Int. J. Quantum Chem. Quantum Biol. Symp.* **7**, 357–367.
- La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
- Le Bail, A., Duroy, H. & Fourquet, L. (1988). *Mat. Res. Bull.* **23**, 447–452.
- Miller, R., Gallo, S. M., Khalak, H. G. & Weeks, C. M. (1994). *J. Appl. Cryst.* **27**, 613–621.
- Otwinowski, Z. (1991). Daresbury Study Weekend Proceedings. Daresbury Laboratory, Warrington, UK.
- Pähler, A., Smith, J. L. & Hendrickson, W. A. (1990). *Acta Cryst.* **A46**, 537–540.
- Pawley, G. S. (1981). *J. Appl. Cryst.* **14**, 357–361.
- Prandl, W. (1990). *Acta Cryst.* **A46**, 988–992.
- Prandl, W. (1994). *Acta Cryst.* **A50**, 52–55.
- Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* **D58**, 1772–1779.
- Sheldrick, G. M. (1998). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, pp. 131–141. Dordrecht: Kluwer Academic Publishers.
- Smith, G. D., Nagar, B., Rini, J. M., Hauptman, H. A. & Blessing, R. H. (1998). *Acta Cryst.* **D54**, 799–804.
- Terwilliger, T. C. (1994). *Acta Cryst.* **D50**, 11–16.
- Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* **D55**, 849–861.