For the cytochrome $c_{550}.PtCl_4^{2-}$ structure, the reported atomic coordinates were used to calculate normalized structure factors and form the sets of six TPSI for the 676 Friedel pairs with $|E| \geq 1.5$ in the resolution range $8 \geq d \geq 2$ Å at Cr $K\alpha$ and Cu $K\alpha$ wavelengths. A typical result for the six true and estimated TPSI for a strong $|E| \simeq 3$ reflection is illustrated in Fig. 4 and the essentially perfect behavior *on the average* over the 676 Friedel pairs is shown in Fig. 5. Over the 676 Friedel pairs, the percentage incidences of true TPSI with negative sign were 8.3, 8.1, 11.2, 8.0, 8.0 and 10.9% for the TPSI types ω_1 to ω_6, respectively. The scatter in the correlation between the true and estimated mixed-wavelength TPSI is illustrated in Figs. 6 and 7, which show that the range of true TPSI values is considerably wider than the range of the estimated values.

The incidence of true TPSI with negative sign and the scatter between the true and estimated TPSI magnitudes are, unsurprisingly, worse for the protein than for the small molecule. Still, the incidence of negative TPSI is only ~10% for the protein Friedel pairs with $|E| \geq 1.5$ and the errors in the estimated TPSI magnitudes are no worse for the mixed-wavelength estimates (4.3) and (4.4) than for the single-wavelength estimates (4.1) and (4.2). This indicates that the simple extension of the probabilistic TPSI-estimation theory from the one-wavelength case to the two-wavelength case, as expressed in the results (4.1)–(4.4), should be practically useful for structure determinations with two-wavelength anomalous-dispersion data.

We are grateful for the support of our work by grants nos. GM46733 and GM34073 from the National Institutes of Health.

References

Fanchon, E. & Hendrickson, W. A. (1990). *Acta Cryst.* A**46**, 809–820.

Fortier, S., Fraser, M. E. & Moore, N. J. (1986). *Acta Cryst.* A**42**, 149–156.

Guo, D. Y. (1990). *Acta Cryst.* A**46**, 942–944.

Guo, D. Y., Blessing, R. H. & Hauptman, H. A. (1991). *Acta Cryst.* A**47**, 340–345.

Guo, D. Y. & Hauptman, H. A. (1986). Am. Crystallogr. Assoc. Meeting, Hamilton, Ontario, Abstract L4. *ACA Abstr. Ser.* 2, **14**, 46.

Guo, D. Y. & Hauptman, H. A. (1989). *Chin. Sci. Bull.* **34**, 137–141.

Harker, D. (1956). *Acta Cryst.* **9**, 1–9.

Hauptman, H. A. (1982). *Acta Cryst.* A**38**, 632–641.

Hendrickson, W. A. (1985). *Trans. Am. Crystallogr. Assoc.* **21**, 11–21.

Hendrickson, W. A., Pahler, A., Smith, J. L., Satow, Y., Merritt, E. A. & Phizackerley, R. P. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 2190–2194.

Karle, J. (1980). *Int. J. Quantum Chem. Symp.* **7**, 356–367.

Shen, M., Ruble, J. R. & Hite, G. (1975). *Acta Cryst.* B**31**, 2706–2709.

Timkovich, R. & Dickerson, R. E. (1973). *J. Mol. Biol.* **79**, 39–56.

Timkovich, R. & Dickerson, R. E. (1976). *J. Biol. Chem.* **251**, 4033–4046.

Acta Cryst. (1994). A**50**, 311–317

Triplet and Quartet Relationships and the 'Positivity' Postulate

By A. Altomare, C. Giacovazzo, A. Guagliardi and D. Siliqi*

Istituto di Ricerca per lo Sviluppo di Metodologie Cristallografiche CNR, c/o Dipartimento Geomineralogico, Campus Universitario, 70124 *Bari, Italy*

(*Received* 11 *February* 1993; *accepted* 16 *September* 1993)

Abstract

Positivity and atomicity are traditionally considered as basic conditions for direct methods. While the role of atomicity is well understood, the theoretical and practical consequences for a direct-methods procedure generated by the violation of the positivity criterion are not quite clear. Triplet and quartet relationships are analyzed in connection with the positivity postulate.

Symbols and notation

f_j Scattering factor of the jth atom. It is assumed to be a real positive function of $\sin\theta/\lambda$ for X-ray scattering, coinciding with the scattering length b_j (a positive or negative real constant) for neutron scattering

$F_{\mathbf{h}}$ Structure factor with vectorial index $\mathbf{h} \equiv (h,k,l)$

---

* Permanent address: Laboratory of X-ray Diffraction, Department of Inorganic Chemistry, Faculty of Natural Sciences, Tirana University, Tirana, Albania.

Acta Crystallographica Section A
ISSN 0108-7673 ©1994

$\varphi_{\mathbf{h}}$ Phase of $F_{\mathbf{h}}$
$E_{\mathbf{h}}$ Normalized structure factor
$\varepsilon_{\mathbf{h}}$ $(|E^2_{\mathbf{h}}| - 1)$
N Number of atoms in the unit cell. For standard X-ray diffraction and for unequal-atom structures, N is replaced by $N_{\rm eq} \simeq \sigma_2^3/\sigma_3^2$ where $\sigma_i = \sum_{j=1}^N Z_j^i$; Z_j is the atomic number of the jth atom. N should be calculated for the primitive cell if the actual cell is centered
$\Phi_{\mathbf{h},\mathbf{k}} = (\varphi_{\mathbf{h}} - \varphi_{\mathbf{k}} - \varphi_{\mathbf{h}-\mathbf{k}})$
Triplet invariant phase
$G_{\mathbf{h},\mathbf{k}} = 2|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|N^{-1/2}$
Cochran's reliability parameter for a triplet invariant
$\Phi_{\mathbf{h},\mathbf{k},\mathbf{k}} = (\varphi_{\mathbf{h}} - \varphi_{\mathbf{k}} - \varphi_{\mathbf{l}} - \varphi_{\mathbf{h}-\mathbf{k}-\mathbf{l}})$
Quartet invariant phase
$G_{\mathbf{h},\mathbf{k},\mathbf{l}} = 2N^{-1}|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{l}}E_{\mathbf{h}-\mathbf{k}-\mathbf{l}}|[(1 + \varepsilon_{\mathbf{h}-\mathbf{k}} + \varepsilon_{\mathbf{h}-\mathbf{l}} + \varepsilon_{\mathbf{k}+\mathbf{l}}) \times (1+Q)^{-1}$
Giacovazzo's reliability parameter for a quartet invariant
$Q = [(\varepsilon_{\mathbf{h}}\varepsilon_{\mathbf{k}} + \varepsilon_{\mathbf{l}}\varepsilon_{\mathbf{h}-\mathbf{k}-\mathbf{l}})\varepsilon_{\mathbf{h}-\mathbf{k}} + (\varepsilon_{\mathbf{h}}\varepsilon_{\mathbf{l}} + \varepsilon_{\mathbf{k}}\varepsilon_{\mathbf{h}-\mathbf{k}-\mathbf{l}})\varepsilon_{\mathbf{h}-\mathbf{l}} + (\varepsilon_{\mathbf{h}}\varepsilon_{\mathbf{h}-\mathbf{k}-\mathbf{l}} + \varepsilon_{\mathbf{k}}\varepsilon_{\mathbf{l}})\varepsilon_{\mathbf{k}+\mathbf{l}}](2N)^{-1}$

Introduction

Recently, Navaza & Navaza (1992) presented a formulation of the phase problem that leads to exact algebraic relationships between Fourier coefficients starting from the atomicity condition. The hypothesis of positivity of the electron density is not used: this makes the theory particularly attractive when compared with direct methods, for which both positivity and atomicity are traditionally considered basic conditions for validity. The origin of this belief can be traced back to the early theory of inequalities where either the scattering factors (see Harker & Kasper, 1948; MacGillavry, 1950; Okaya & Nitta, 1952; Gillis, 1948) are assumed to be non-negative functions (the non-negativity of the electron density is just the physical counterpart of this property) or the positivity of the electron density is directly assumed (Karle & Hauptman, 1950). Early direct methods implicitly assumed positive scattering factors (the attention was mainly focused on standard X-ray diffraction). Later, it was perceived that positivity is not an essential ingredient of direct methods: there are several examples of crystal structures solved by direct methods (*via* data collected by neutron scattering) that involve atoms with negative scattering factor [for some comments in the field of powder data see Cascarano, Favia & Giacovazzo (1992)]. The major potential of direct methods with respect to traditional inequalities may also be inferred from a detail to which not much attention has been paid. Inequalities are usually expressed in terms of the *unitary structure factor*

$$U_{\mathbf{h}} = \sum_{j=1}^{N} n_j \exp(2\pi i\mathbf{h}\cdot\mathbf{r}_j)$$

with

$$n_j = f_j / \sum_{j=1}^{N} f_j \quad \text{and} \quad 0 \le |U_{\mathbf{h}}| \le 1.$$

$U_{\mathbf{h}}$ represents the fraction of the structure factor $F_{\mathbf{h}}$ with respect to its highest possible value (corresponding to the case when $\mathbf{h}\cdot\mathbf{r}_j$ is an integer for any value of j). This definition intrinsically relies on the positivity of the electron density. If some of the f_j's are negative, then $|U_{\mathbf{h}}|$ could be out of the interval (0,1): if $\sum_{j=1}^N f_j \simeq 0$ then $|U_{\mathbf{h}}|$ could become infinitely large. It is worthwhile noting that modern maximum-entropy techniques requiring the positivity of the electron density also work in term of U's.

Quite different are the properties of the *normalized structure factor* $E_{\mathbf{h}}$, in terms of which probabilistic phase relationships provided by direct methods are expressed:

$$E_{\mathbf{h}} = \sum_{j=1}^{N} \nu_j \exp(2\pi i\mathbf{h}\cdot\mathbf{r}_j),$$

where $\nu_j = f_j/(\sum_{j=1}^N f_j^2)^{1/2}$. The occurrence of negative scattering factors affects neither the definition of $E_{\mathbf{h}}$ nor its distribution (Wilson results hold even if the positivity condition is violated). Therefore, the E's are intrinsically more suitable for dealing with less-restrictive conditions on the properties of scattering matter.

The role of the positivity condition in direct methods is today not completely understood. In a paper devoted to the generalization of direct methods to the case that not all the atoms in the unit cell are equal, Hauptman (1976) explicitly treated the possible presence of negative scatterers. However, no description was made of the practical consequences for direct phasing procedures generated by the violation of the positivity condition. This is the primary aim of this paper.

Violation of the positivity condition: its effect on triplet and quartet relationships

Three cases are considered:

Case A. All the scattering factors are real non-negative functions of $\sin\theta/\lambda$ (as for standard X-ray or electron scattering).

Case B. All the scattering factors are real negative functions. Accordingly, ρ is supposed to be non-positive everywhere. This situation may seldom occur in neutron scattering. If isomorphous data are available, their direct processing may be performed according to either case A or case B (see later).

Case C. Some scattering factors are real positive functions, some are real negative functions (this may frequently occur in neutron scattering). Cases *A* and *B* are limiting situations of case *C*.

The case of complex scattering factors is not discussed because it finds a better description in a different framework (see Hauptman, 1982; Giacovazzo, 1983).

Case A

The basic formulas for this case are well known; they are quoted in order to allow comparison with the corresponding formulas for cases *B* and *C*.

If the electron density ρ satisfies the relationship

$$\rho \simeq \rho^2, \qquad (1)$$

then (Sayre, 1952)

$$F_{\mathbf{h}} \simeq \theta_{\mathbf{h}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}}, \qquad (2)$$

where $\theta_{\mathbf{h}}$ is a scaling constant of known value.

By application of the joint-probability-distribution method, the triplet phase distribution (Cochran, 1955)

$$P(\Phi_{\mathbf{h},\mathbf{k}}) \simeq [2\pi I_0(G_{\mathbf{h},\mathbf{k}})]^{-1} \exp(G_{\mathbf{h},\mathbf{k}} \cos \Phi_{\mathbf{h},\mathbf{k}}) \qquad (3)$$

and more generally the tangent formula (Karle & Hauptman, 1956)

$$\tan \varphi_{\mathbf{h}} = \sum_{\mathbf{h}} G_{\mathbf{h},\mathbf{k}} \sin(\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}}) / \sum_{\mathbf{k}} G_{\mathbf{h},\mathbf{k}} \cos(\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})$$
$$= T_{3,\mathbf{h}}/B_{3,\mathbf{h}} \qquad (4)$$

are derived. The reliability parameter of the phase estimation is

$$\alpha_{3,\mathbf{h}} \simeq (T_{3,\mathbf{h}}^2 + B_{3,\mathbf{h}}^2)^{1/2}. \qquad (5)$$

Equations (4) and (5) are the basis of traditional direct methods. It is worthwhile noting that the tangent formula tries to maximize $\sum_{\mathbf{h}} \alpha_{\mathbf{h}}$ or, equivalently, to satisfy

$$\sum_{\mathbf{h},\mathbf{k}} G_{\mathbf{h},\mathbf{k}} \cos(\varphi_{\mathbf{h}} - \varphi_{\mathbf{k}} - \varphi_{\mathbf{h}-\mathbf{k}}) = \max. \qquad (6)$$

If (1) holds, the relation

$$\rho \simeq \rho^3 \qquad (7)$$

is also fulfilled, which gives rise to the quartet relationship

$$F_{\mathbf{h}} \simeq q_h \sum_{\mathbf{k},\mathbf{l}} F_{\mathbf{k}} F_{\mathbf{l}} F_{\mathbf{h}-\mathbf{k}-\mathbf{l}}, \qquad (8)$$

where $q_{\mathbf{h}}$ is a suitable constant of known value. The application of the joint-probability-distribution method leads to the phase distribution [Giacovazzo (1976*a*, 1980); but see also Hauptman (1975) for a more complicated formula]

$$P(\Phi_{\mathbf{h},\mathbf{k},\mathbf{l}}) \simeq [2\pi I_0(G_{\mathbf{h},\mathbf{k},\mathbf{l}})]^{-1} \exp(G_{\mathbf{h},\mathbf{k},\mathbf{l}} \cos \Phi_{\mathbf{h},\mathbf{k},\mathbf{l}}) \qquad (9)$$

and, more generally, to the tangent formula

$$\tan \varphi_{\mathbf{h}} = \sum_{\mathbf{k},\mathbf{l}} G_{\mathbf{h},\mathbf{k},\mathbf{l}} \sin(\varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{h}-\mathbf{k}-\mathbf{l}})$$
$$\times \left[\sum_{\mathbf{h},\mathbf{l}} G_{\mathbf{h},\mathbf{k},\mathbf{l}} \cos(\varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{h}-\mathbf{k}-\mathbf{l}})\right]^{-1}$$
$$= T_{4,\mathbf{h}}/B_{4,\mathbf{h}}, \qquad (10)$$

with reliability parameter

$$\alpha_{4,\mathbf{h}} \simeq (T_{4,\mathbf{h}}^2 + B_{4,\mathbf{h}}^2)^{1/2}. \qquad (11)$$

A tangent formula based on (10) and (11) tries to maximize $\sum_{\mathbf{h}} \alpha_{4,\mathbf{h}}$ and therefore to satisfy

$$\sum_{\mathbf{h},\mathbf{k},\mathbf{l}} G_{\mathbf{h},\mathbf{k},\mathbf{l}} \cos(\varphi_{\mathbf{h}} - \varphi_{\mathbf{k}} - \varphi_{\mathbf{l}} - \varphi_{\mathbf{h}-\mathbf{k}-\mathbf{l}}) = \max. \qquad (12)$$

In a multisolution process, the correct solution is recognized by figures of merit (FOMs). Some FOMs, like the well known MABS, ABSFOM, R_α and NALF (Declercq, Germain & Woolfson, 1979; Karle & Karle, 1966; Cascarano, Giacovazzo & Viterbo, 1987) check if the distribution of the $\alpha_{3,\mathbf{h}}$'s fits the expected distribution. On the same principle are based the so-called psi(0) FOMs (Cochran & Douglas, 1957; Cascarano, Giacovazzo & Guagliardi, 1992) while the negative-quartet criterion (DeTitta, Edmonds, Langs & Hauptman, 1975; Cascarano, Giacovazzo & Viterbo, 1987) relies on the expectation that some quartet invariants are negative.

Case B

Let us replace f_j by $-f_j$. In this case, the structure factor is given by

$$\sum_{j=1}^{N} (-f_j) \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j) = -F_{\mathbf{h}} = F_{\mathbf{h}} \exp i\pi. \qquad (13)$$

Relation (1) is now replaced by

$$\rho \simeq -\rho^2 \qquad (14)$$

and

$$F_{\mathbf{h}} = -\theta_{\mathbf{h}} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}}. \qquad (15)$$

Comparison of (14) and (15) with (1) and (2), respectively, shows that the Sayre equation does not hold when ρ is non-positive definite. However, since (14) and (1) differ only by a sign, we will refer to (15) as a 'modified Sayre equation'.

By application of the joint-probability-distribution method, the triplet phase distribution

$$P(\Phi_{\mathbf{h},\mathbf{k}}) \simeq [2\pi I_0(-G_{\mathbf{h},\mathbf{k}})]^{-1} \exp(-G_{\mathbf{h},\mathbf{k}} \cos \Phi_{\mathbf{h},\mathbf{k}}) \qquad (16)$$

is obtained. More generally,

$$\begin{aligned}\tan\varphi_{\mathbf{h}} &= \sum_{\mathbf{k}} -G_{\mathbf{h,k}}\sin(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h-k}})\\ &\quad\times\left[\sum_{\mathbf{k}} -G_{\mathbf{h,k}}\cos(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h-k}})\right]^{-1}\\ &= \sum_{\mathbf{k}} G_{\mathbf{h,k}}\sin(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h-k}}+\pi)\\ &\quad\times\left[\sum_{\mathbf{k}} G_{\mathbf{h,k}}\cos(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h-k}}+\pi)\right]^{-1}\\ &= -T_{3,\mathbf{h}}/-B_{3,\mathbf{h}}, \qquad (17)\end{aligned}$$

with reliability parameter (unmodified with respect to case *A*)

$$\alpha_{3,\mathbf{h}} \simeq (T^2_{3,\mathbf{h}}+B^2_{3,\mathbf{h}})^{1/2}. \qquad (18)$$

The application of (17) tends again to maximize $\sum_{\mathbf{h}}\alpha_{3,\mathbf{h}}$ as for case *A*: this is now equivalent to satisfying the relation

$$\sum_{\mathbf{h,k}} G_{\mathbf{h,k}}\cos(\varphi_{\mathbf{h}}-\varphi_{\mathbf{k}}-\varphi_{\mathbf{h-k}}+\pi) = \max \qquad (19)$$

rather than (6). Since (18) and (5) coincide, a tangent routine that uses negative values of G (or, equivalently, assumes positive values of G and adds π to the triplet phase) is expected to be as efficient in finding the correct positions of the negative 'atoms' as the usual tangent routine for the location of the positive atoms. As a consequence, FOMs based on the distribution of the α moduli [psi(0) FOM included] are not sensitive to the sign of ρ. In other words, a structure with positive atomic scattering factors will be marked by the same FOMs as its negative image.

When we move from case *A* to case *B*, relation (1) is replaced by (14) but relation (7) still holds. Accordingly, the quartet relations (8)–(12) also hold. In short, the result is the following: while triplet phase relationships depend on the 'sign' of the scattering matter, quartet relationships and relative FOMs are insensitive to it. Therefore, the set of phases one obtains at the end of a direct procedure based only on quartet relationships can lead to the correct solution either *via* a collection of positive peaks or, equivalently, *via* a collection of negative peaks.

When isomorphous data are available (*e.g.* by addition of t heavy atoms to the protein structure), the differences $||F_{HP}|-|F_p||$ are often calculated in order to obtain normalized differences $|\Delta E|$. The same set $\{|\Delta E|\}$ may be interpreted as the reciprocal counterpart either of the differences $\Delta\rho=\rho_{HP}-\rho_P$ or of the difference $\rho_P-\rho_{HP}$. In the first case, $\Delta\rho$ is positive definite: thus, the use in (4) of positive values for the pseudo-Cochran parameter

$$G_{\mathbf{h,k}} = 2|\Delta E_{\mathbf{h}}\Delta E_{\mathbf{k}}\Delta E_{\mathbf{h-k}}|t^{-1/2}$$

may lead to identification of the heavy-atom positions *via* a Fourier map and a search routine that looks for the largest positive peaks. If the second interpretation is preferred, the same set $\{|\Delta E|\}$ may be processed *via* (16), *i.e. via* negative values of $G_{\mathbf{h,k}}$. Then, the search routine will explore the Fourier map looking for the largest negative peaks, which again will give the positions of the heavy atoms.

If quartet instead of triplet relationships are used for the phasing process of isomorphous data, the resulting Fourier maps will have to be examined in the following composite way: the structure is first searched, as usual, among the largest positive peaks and then, if the first search fails, the structure must also be searched among the largest negative peaks.

Case C

Now neither the Sayre equation (2) nor the modified Sayre equation (15) hold; however, the application of the joint-probability-distribution method leads to useful formulas. In particular, (3) is still valid but $G_{\mathbf{h,k}}$ must be calculated from the general expression

$$G_{\mathbf{h,k}} = 2C_{\mathbf{h,k}}|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h-k}}|, \qquad (20)$$

where

$$\begin{aligned}C_{\mathbf{h,k}} = \Sigma(\mathbf{h},\mathbf{k},\mathbf{h-k}) &= \sum_{j=1}^{N}\nu_j(\mathbf{h})\nu_j(\mathbf{k})\nu_j(\mathbf{h-k})\\ &\simeq \sigma_3\sigma_2^{-3/2}.\end{aligned}$$

If all the f_j's are positive (case *A*), then $C_{\mathbf{h,k}}\simeq\sigma_3\sigma_2^{-3/2}\simeq N^{-1/2}$ as in the standard formula. If all the f_j's are negative (case *B*), then $C_{\mathbf{h,k}}\simeq -N^{-1/2}$ in agreement with (16). If some f_j's are positive and some negative then $C_{\mathbf{h,k}}$ contains contributions of opposite sign. If their balance is close to zero, $G_{\mathbf{h,k}}$ is an inefficient reliability parameter. One could then say (in order to preserve the analogy with cases *A* and *B*) that the apparent value of N (*i.e.* $N=1/C^2_{\mathbf{h,k}}$) becomes infinitely large (accordingly, N is no longer the number of scatterers, as in cases *A* and *B*). In this case, the structure could hardly be solved *via* triplet invariants.

The equation $\rho\simeq\rho^3$ holds in case *C*; therefore, (8) is also valid. However, the estimation of a quartet phase *via* the joint-probability-distribution method requires some caution. The following result is obtained (see Giacovazzo, 1976*b*): $\Phi_{\mathbf{h,k,l}}$ is distributed according to the von Mises distribution (9) provided $G_{\mathbf{h,k,l}}$ is replaced by the general expression

$$\begin{aligned}G_{\mathbf{h,k,l}} = [2|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{l}}E_{\mathbf{h-k-l}}|/(1+Q)][&\Sigma(\mathbf{h},\mathbf{k},\mathbf{l},\mathbf{h-k-l})\\ &+\Sigma(\mathbf{h},\mathbf{k},\mathbf{h-k})\Sigma(\mathbf{l},\mathbf{h-k},\mathbf{h-k-l})\varepsilon_{\mathbf{h-k}}\\ &+\Sigma(\mathbf{h},\mathbf{l},\mathbf{h-l})\Sigma(\mathbf{k},\mathbf{h-l},\mathbf{h-k-l})\varepsilon_{\mathbf{h-l}}\\ &+\Sigma(\mathbf{h},\mathbf{k+l},\mathbf{h-k-l})\Sigma(\mathbf{k},\mathbf{l},\mathbf{k+l})\varepsilon_{\mathbf{k+l}}], \qquad (21)\end{aligned}$$

where

$$\Sigma(\mathbf{h},\mathbf{k},\mathbf{l},\mathbf{h}-\mathbf{k}-\mathbf{l}) = \sum_{j=1}^{N} \nu_j(\mathbf{h})\nu_j(\mathbf{k})\nu_j(\mathbf{l})\nu_j(\mathbf{h}-\mathbf{k}-\mathbf{l}),$$

$$\begin{aligned} Q = {} & \Sigma^2(\mathbf{h},\mathbf{k},\mathbf{h}-\mathbf{k})\varepsilon_{\mathbf{h}}\varepsilon_{\mathbf{k}}\varepsilon_{\mathbf{h}-\mathbf{k}} \\ & + \Sigma^2(\mathbf{l},\mathbf{h}-\mathbf{k},\mathbf{h}-\mathbf{k}-\mathbf{l})\varepsilon_{\mathbf{l}}\varepsilon_{\mathbf{h}-\mathbf{k}}\varepsilon_{\mathbf{h}-\mathbf{k}-\mathbf{l}} \\ & + \Sigma^2(\mathbf{h},\mathbf{l},\mathbf{h}-\mathbf{l})\varepsilon_{\mathbf{h}}\varepsilon_{\mathbf{l}}\varepsilon_{\mathbf{h}-\mathbf{l}} \\ & + \Sigma^2(\mathbf{k},\mathbf{h}-\mathbf{l},\mathbf{h}-\mathbf{k}-\mathbf{l})\varepsilon_{\mathbf{k}}\varepsilon_{\mathbf{h}-\mathbf{l}}\varepsilon_{\mathbf{h}-\mathbf{k}-\mathbf{l}} \\ & + \Sigma^2(\mathbf{h},\mathbf{k}+\mathbf{l},\mathbf{h}-\mathbf{k}-\mathbf{l})\varepsilon_{\mathbf{h}}\varepsilon_{\mathbf{k}+\mathbf{l}}\varepsilon_{\mathbf{h}-\mathbf{k}-\mathbf{l}} \\ & + \Sigma^2(\mathbf{k},\mathbf{l},\mathbf{k}+\mathbf{l})\varepsilon_{\mathbf{k}}\varepsilon_{\mathbf{l}}\varepsilon_{\mathbf{k}+\mathbf{l}}. \end{aligned}$$

The following approximations can be introduced:

$$\Sigma(\mathbf{h},\mathbf{k},\mathbf{l},\mathbf{h}-\mathbf{k}-\mathbf{l}) \simeq \sigma_4/\sigma_2^2,$$

$$\Sigma^2(\mathbf{h},\mathbf{k},\mathbf{h}-\mathbf{k}) \simeq \Sigma^2(\mathbf{h},\mathbf{l},\mathbf{h}-\mathbf{l}) \simeq \ldots \simeq (\sigma_3/\sigma_2^{3/2})^2.$$

With the assumption that $\sigma_4/\sigma_2^2 \simeq (\sigma_3/\sigma_2^{3/2})^2$, the classical quartet estimate is obtained, as described for cases *A* and *B*. If some f_j's are positive and some are negative, σ_4/σ_2^2 may be quite different from σ_3^2/σ_2^3. In particular, if $\sigma_3 \simeq 0$, the cross magnitudes do not influence the quartet estimate. This then depends only on the term

$$2(\sigma_4/\sigma_2^2)|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{l}}E_{\mathbf{h}-\mathbf{k}-\mathbf{l}}|$$

and the quartet is always expected to be positive. In particular, for situations in which $\sigma_3 \simeq 0$ while σ_4 is still large, the structure solution may be obtained more easily *via* the active use of quartet invariants rather than triplets.

The breakdown of the positivity condition and the P_{10} formula

A time-honored formula for the estimation of triplet invariants is the so-called P_{10} formula (Cascarano, Giacovazzo, Camalli, Spagna, Burla, Nunzi & Polidori, 1984): for each triplet $\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3}$ with $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = 0$, P_{10} exploits the information contained in the ten moduli

$$\{|E_{\mathbf{h}_1}|,|E_{\mathbf{h}_2}|,|E_{\mathbf{h}_3}|,|E_{\mathbf{k}}|,|E_{\mathbf{h}_1\pm\mathbf{k}}|,|E_{\mathbf{h}_2\pm\mathbf{k}}|,|E_{\mathbf{h}_3\pm\mathbf{k}}|\},$$

where $\mathbf{k}$ is a free vector. If P_{10} is used, the triplet phase is again distributed according to a von Mises distribution but the concentration parameter is now

$$G = 2|E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}|\sigma_3\sigma_2^{-3/2}[1 + A/(1+B)],$$

where *A* and *B* arise from the use of the special quintet invariants $\{\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3} + \varphi_{\mathbf{k}} - \varphi_{\mathbf{k}}\}$. Since *A* and *B* depend on the cube of $\sigma_3\sigma_2^{-3/2}$, the violation of the positivity condition has the same effect on the P_{10} formula as on the Cochran relationships.

Applications

In order to show the practical consequences for a direct-phasing process of the breakdown of the positivity condition, we have constructed the three fictitious structures shown in Fig. 1. The atomic coordinates correspond to non-H atoms of diethylene glycol ditosylate* $C_{18}H_{22}O_7S_2$ (Ferchaux, Villain & Navaza, 1990), space group *C*2/*c*, $a = 23.772$, $b = 5.472$, $c = 15.284$ Å, $\beta = 98.61°$. Three sets of structure factors were calculated using the numbers in the drawings as atomic scattering factors (from now on we refer to them as structures 1, 2 and 3, respectively). Tests were performed using a modified version of the program *SIR*88 (Burla, Camalli, Cascarano, Giacovazzo, Polidori, Spagna & Viterbo, 1989). In Table 1, for each test structure, the distributions of the triplet and quartet invariants as functions of the modulus of the concentration parameter *G* are shown. It should be noted:

(*a*) $\sigma_2^3/\sigma_3^2 \simeq \sigma_2^2/\sigma_4$ only for structure 1. They are remarkably different for structures 2 and 3.

(*b*) The distribution of the *G* values for triplets is sufficiently good for practical applications only for structures 1 and 2. For structure 3, the *G* values are between 0 and 0.4.

(*c*) The distribution of the *G* values for quartets is always good for practical applications (σ_2^2/σ_4 is a constant for the three structures). In accordance with (21), no quartet is estimated negative for structure 3.

(*d*) The reliability of the positive quartets is sufficiently good for all three structures; however, the breakdown of the positivity criterion makes unreliable the estimates of the negative quartets for structure 2 even when they are characterized by large |*G*| moduli. This behavior has been confirmed by other tests (not shown here for brevity) and is not well understood at the moment. It seems that, when the positivity criterion is violated, having cross magnitudes close to zero is not a sufficient condition for a quartet to be reliably estimated negative.

The behavior of the figures of merit (based on triplet and quartet relationships) may be deduced from Tables 2 and 3: both refer to the same case, $\rho \geq 0$ everywhere. In Table 2, the FOMs of the best two trial solutions are shown when only triplets are actively used in the phasing process. The trial with the highest CFOM is the correct solution. In Table 3, the FOMs of the best six trial solutions are shown when only quartets are actively used in the phasing process. It may now be expected that both ρ and $-\rho$ are among the set of trials and characterized by the same FOMs. Actually, pairs of solutions are marked by the same FOMs. A subsequent analysis of the Fourier maps led to the correct solution when the

---

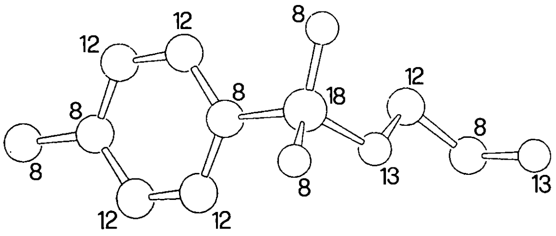
* IUPAC name 3-oxa-1,5-pentanediyl bis(*p*-toluenesulfonate).

peak-search routine analyzed the largest negative peaks of trial 1 ($-\rho$ structure) and the largest positive peaks of trial 2 (ρ structure).

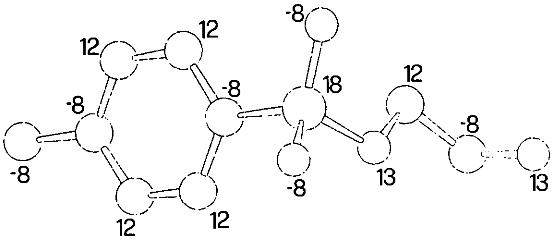
Concluding remarks

Triplet and quartet relationships have been discussed in connection with a possible violation of the positivity criterion. A similar analysis can be made for quintet or upper phase relationships. The behavior of the most common figures of merit has been described in order to clarify the practical aspects of a direct-phasing process applied to experimental data generated by scattering matter with positive and/or negative scattering factors.

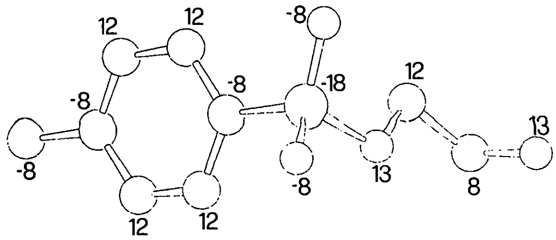
The authors thank Miss C. Chiarella for technical support.



Structure 1



Structure 2



Structure 3

Fig. 1. Fictitious structures with atomic positions corresponding to non-H atoms of diethylene glycol ditosylate $C_{18}H_{22}O_7S_2$ and arbitrary scatterers. The numbers in the drawings correspond to the atomic scattering factors used in the calculation of the diffraction intensities. The terminal atom with weight 13 is in a special position with occupation factor 0.5.

Table 1. *For the three isomorphous structures described in the text, the distributions of triplet and quartet invariants as functions of concentration parameter $|G|$ are given*

n is the number of invariants having concentration parameter larger than a given $|G|$, nw is the number of wrong estimates.

G	Triplets n (nw)	Positive-estimated quartets n (nw)	Negative-estimated quartets n (nw)
Test structure 1			
$\sigma_2^3/\sigma_3^2 = 43.8$, $\sigma_2^2/\sigma_4 = 41.0$			
0.6	3201 (194)	24779 (2326)	1577 (309)
1.0	1954 (41)	22728 (1686)	704 (74)
2.0	424 (0)	7592 (137)	91 (1)
3.8	58 (0)	3 (0)	4 (0)
5.0	24 (0)		
Test structure 2			
$\sigma_2^3/\sigma_3^2 = 88.1$, $\sigma_2^2/\sigma_4 = 41.0$			
0.6	1827 (273)	16584 (1984)	171 (114)
1.0	428 (35)	11011 (1002)	22 (19)
1.6	28 (0)	3763 (201)	2 (2)
2.0	4 (0)	1034 (23)	1 (1)
2.4		107 (0)	1 (1)
Test structure 3			
$\sigma_2^3/\sigma_3^2 = 2069$, $\sigma_2^2/\sigma_4 = 41.0$			
0.2	261 (89)	30000 (7647)	-
0.4	1 (0)	18448 (4192)	-
1.0		240 (17)	-
1.4		9 (1)	

Table 2. *Case with $\rho \geq 0$ everywhere, triplet invariants actively used*

FOMs for the two trial solutions with the largest combined figure of merit CFOM. MABS = $\sum_h \alpha / \sum \langle \alpha_h \rangle$, ALFCOMB is based on the α distribution for strong triplets, PSCOMB on the psi(0) triplets, CPHASE on negative estimated quartets.

Trial	MABS	ALFCOMB	PSCOMB	CPHASE	CFOM
1	1.082	0.994	0.619	1.0	0.997
2	0.972	0.422	0.200	0.999	0.684

Table 3. *Case with $\rho \geq 0$ everywhere, only quartet invariants actively used*

Trial	MABS	ALFCOMB	PSCOMB	CPHASE	CFOM
1	1.693	0.747	0.580	1.0	0.862
2	1.693	0.747	0.580	1.0	0.862
3	1.652	0.706	0.187	1.0	0.840
4	1.652	0.706	0.187	1.0	0.840
5	1.289	0.259	0.277	1.0	0.594
6	1.289	0.259	0.277	1.0	0.594

References

Burla, M. C., Camalli, M., Cascarano, G., Giacovazzo, C., Polidori, G., Spagna, R. & Viterbo, D. (1989). *J. Appl. Cryst.* **22**, 389–393.

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.* **A40**, 278–283.

Cascarano, G., Giacovazzo, C. & Guagliardi, A. (1992). *Acta Cryst.* A**48**, 859–865.
Cascarano, G., Giacovazzo, C. & Viterbo, D. (1987). *Acta Cryst.* A**43**, 22–29.
Cochran, W. (1955). *Acta Cryst.* **8**, 473–478.
Cochran, W. & Douglas, A. S. (1957). *Proc. R. Soc. London Ser. A*, **243**, 281.
Declercq, J.-P., Germain, G. & Woolfson, M. M. (1979). *Acta Cryst.* A**35**, 622–626.
DeTitta, G. T., Edmonds, J. W., Langs, D. A. & Hauptman, H. (1975). *Acta Cryst.* A**31**, 472–479.
Ferchaux, Y., Villain, F. & Navaza, A. H. (1990). *Acta Cryst.* C**46**, 346–348.
Giacovazzo, C. (1976*a*). *Acta Cryst.* A**32**, 91–99, 100–104.
Giacovazzo, C. (1976*b*). *Acta Cryst.* A**32**, 958–966.
Giacovazzo, C. (1980). In *Direct Methods in Crystallography.* London: Academic Press.
Giacovazzo, C. (1983). *Acta Cryst.* A**39**, 585–592.
Gillis, J. (1948). *Acta Cryst.* **1**, 76–80.
Harker, D. & Kasper, J. S. (1948). *Acta Cryst.* **1**, 70–75.
Hauptman, H. (1975). *Acta Cryst.* A**31**, 671–679, 680–687.
Hauptman, H. (1976). *Acta Cryst.* A**32**, 877–882.
Hauptman, H. (1982). *Acta Cryst.* A**38**, 632–641.
Karle, J. & Hauptman, H. (1950). *Acta Cryst.* **3**, 181–187.
Karle, J. & Hauptman, H. (1956). *Acta Cryst.* **9**, 635–651.
Karle, J. & Karle, I. (1966). *Acta Cryst.* **21**, 849–859.
MacGillavry, C. H. (1950). *Acta Cryst.* **3**, 214–217.
Navaza, A. & Navaza, J. (1992). *Acta Cryst.* A**48**, 695–700.
Okaya, J. & Nitta, I. (1952). *Acta Cryst.* **5**, 564–570.
Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.

Acta Cryst. (1994). A**50**, 317–325

(4;2)-Connected Three-Dimensional Nets Related to the Mixed-Coordinated Framework Structures $AlPO_4$-15, $AlPO_4$-CJ2 and $AlPO_4$-12

By K. J. Andries and J. V. Smith

The Consortium for Theoretical Frameworks, Department of the Geophysical Sciences, The University of Chicago, 5734 *South Ellis Avenue, Chicago, Illinois* 60637, *USA*

(*Received* 26 *March* 1993; *accepted* 27 *September* 1993)

Abstract

Application of the same pattern for linking neighbouring 4.8^2 two-dimensional nets as is found in the $AlPO_4$-CJ2 and $AlPO_4$-15 structures allows the enumeration of eleven (4;2)-connected three-dimensional nets with a maximum unit-cell repeat of ~ 10 Å. Net 551 is polytypic with net 398 (the tetrahedral analogue of $AlPO_4$-12). The (4;2)-connected nets related to $AlPO_4$-15 (net 400) and $AlPO_4$-CJ2 (net 725) have the highest symmetry of the present nets. Synthetic zeolite Linde J may have the latter framework topology. Nets 400 and 725 relax considerably upon lowering of the space group symmetry. Geometrical refinements indicate that the mixed-coordinated $AlPO_4$-15 is more feasible than its (4;2)-connected relative (same space group symmetry); the feasible net 398 and the topologically simpler net 551 have not been observed so far. Geometrically optimized atomic coordinates are given for the most feasible nets obtained here.

Introduction

Tetrahedrally coordinated frameworks occur in zeolitic materials used in ion exchange, catalysis and molecular sieving. Research is expanding into a wider range of scientific areas as new materials are synthesized [particularly ultra-pore materials, *e.g.* Kresge, Leonowicz, Roth, Vartuli & Beck (1992)]. The *Atlas of Zeolite Structure Types* (Meier & Olson, 1992) and Smith (1988) give references and structural information on most topologically distinct zeolite materials and their three-letter structure-type codes.

The systematic enumeration of (4;2)-connected three-dimensional (3D) nets is important for the characterization and classification of known framework structures as well as for the solution of as yet unknown phases. The notation describes four-connected nodes (*e.g.* silicon, aluminium, phosphorus) bridged by two-connected atoms (*e.g.* oxygen). Omission of the two-connected atoms leaves a fully four-connected net. Some exploratory contributions in the field of theoretical enumeration have come from Wells (1977), Smith (1977, 1978), Alberti (1979), Smith & Bennett (1984), Smith & Dytrych (1986), Hawthorne & Smith (1988), Kokotailo, Fyfe, Gies & Cox (1989), Andries (1990) and O'Keeffe (1992). It is expected that, with the aid of modern computational techniques, more general approaches can be undertaken towards the systematic enumeration of four-connected 3D nets. Four techniques for generating 3D net topologies are: (i) simulated annealing (Deem & Newsam, 1989); (ii) a combinatorial method (Treacy, Rao & Rivin, 1992); (iii) application of stacking operators to two-dimensional