

or equal to  $1.5 \text{ \AA}^{-1}$  were all 2.0, which meant that all the corresponding values of  $F_{\text{tab}}(s)$  were zero. It was only possible to obtain a fit to the sum of four Gaussians (an eight coefficient fit) and hence  $a_5$  and  $b_5$  are listed as zero. The relative errors for  $s$  less than or equal to  $0.8 \text{ \AA}^{-1}$  are all less than 1% while the remainder are rather large or could not be calculated because of division by the zero  $F_{\text{tab}}(s)$ . However, because the actual values of  $F_{\text{tab}}(s)$  are very small for  $s$  greater than  $0.8 \text{ \AA}^{-1}$ , the fit is in fact quite acceptable.

For the other elements the mean percentage differences are all less than 0.3% while the worst percentage differences are all less than 1% except for Li, where the problems were similar to those encountered with He, and K. The rather large value of 1.25% at  $s = 1.5 \text{ \AA}^{-1}$  obtained in the latter case is not readily explicable but if the calculated and fitted values of  $Z - I_{\text{inc}}(s)$  are plotted it is obvious that the overall fit is very good. The quality of the overall fit for the other elements is as good as or better than that for K.

The computations were performed using the CDC 6600 and 7600 computers at the University of London Computer Centre.

### References

- CROMER, D. T. (1969). *J. Chem. Phys.* **50**, 4857–4859.  
 CROMER, D. T. & MANN, J. B. (1967). *J. Chem. Phys.* **47**, 1892–1893.  
 CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104–109.  
 HAJDU, F. (1972). *Acta Cryst.* **A28**, 250–252.  
*International Tables for X-ray Crystallography* (1962). Vol. III. Birmingham: Kynoch Press.  
 PENNINGTON, R. H. (1970). *Introductory Computer Methods and Numerical Analysis*, 2nd ed., pp. 417–427. New York: Macmillan.  
 SMITH, G. H. & BURGE, R. E. (1962). *Acta Cryst.* **15**, 182–186.  
 TAVARD, C., NICOLAS, D. & ROUAULT, M. (1967). *J. Chem. Phys.* **64**, 540–554.

*Acta Cryst.* (1975). **A31**, 602

## Strengthening of the $\Sigma_1$ Relationship: a Quartet Method

BY C. GIACOVAZZO

*Istituto di Mineralogia e Petrografia, Università degli Studi, Bari, Italy*

(Received 11 April 1975; accepted 14 April 1975)

A probabilistic theory of special quartets is described which is able to fix the sign of reflexions with even indices. The method proves suitable for removing a large percentage of the  $\Sigma_1$  formula failures.

### Statement of the problem

Direct methods are today the most powerful crystallographic tool for solving equal-atom structures. Although various types of phase-determining formulae have been proposed, the most effective is the sign relationship

$$\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{-\mathbf{h}-\mathbf{k}} \approx 0.$$

It has long been recognized, however, that this formula may not lead to a physically correct solution without the introduction of auxiliary phase information. Several procedures, therefore, have been proposed (*i.e.* symbolic addition procedures or multisolution methods) which are able to give routine solutions even to complicated structures. As the choice of starting reflexions constitutes a very critical element influencing success or failure of a direct solution of the phase problem, much effort has been devoted to improving the selection of the starting set (Germain, Main & Woolfson, 1970).

It is a matter of experience that the phase-determining process is more effective if one has a large starting set: unfortunately the size of the set is

limited by the number of phase combinations it is practical to explore. Thus it is useful to be able to enlarge the starting set by knowing in advance a number of phases.

In order to achieve this, several additional formulae have been exploited by different authors. In this paper we shall devote our attention to some auxiliary formulae which are able to fix the phase of the  $E_{2\mathbf{h}}$  reflexions. We shall discuss their limits and shall try to improve their effectiveness. We treat here only the centrosymmetric case: formulae useful in non-centrosymmetric cases will be described in a subsequent paper.

### The mathematical approach

A variety of conditional probability distributions will be described. We denote by  $P(E_1, E_2, \dots, E_n)$  the joint probability distribution function of  $n$  normalized structure factors. The characteristic function of the distribution (Giacovazzo, 1975a) is

$$C(u_1, u_2, \dots, u_n) = \exp \left[ -\frac{1}{2}(u_1^2 + u_2^2 + \dots + u_n^2) \right] \\ \times \left\{ 1 + \frac{S_3}{t^{3/2}} + \frac{S_4}{t^2} + \frac{S_3^2}{2t^3} + \dots \right\}, \quad (1)$$

where  $u_j, j=1, \dots, n$  are carrying variables associated with  $E_j, t=N/2$ , and

$$S_v = t \sum_{r+s+\dots+w=v} \frac{\lambda_{rs\dots w}}{r!s!\dots w!} (iu_1)^r (iu_2)^s \dots (iu_n)^w.$$

$\lambda_{rs\dots w}$  are the standardized cumulants of the distribution. The probability distribution can be written down by calculating the Fourier transform of (1). Hermite polynomials  $H_v$  will be involved, defined by

$$H_v(x) = (-1)^v \exp(\frac{1}{2}x^2) \frac{d^v}{dx^v} \exp(-\frac{1}{2}x^2).$$

**The distribution  $P(E_h, E_{2h})$**

This distribution has been studied by several authors (Hauptman & Karle, 1953; Cochran & Woolfson, 1955; Klug, 1958). The probabilistic result is

$$P_+(E_{2h}) = \frac{1}{2} + \frac{1}{2} \tanh |E_{2h}| \frac{(E_h^2 - 1)}{2\sqrt{N}}. \tag{2}$$

The physical counterpart of (2) in real space is the Patterson function (Cochran & Woolfson, 1954): the determination of the sign of  $E_{2h}$  in space group  $P\bar{1}$  depends on identifying all peaks except those in  $\pm 2r_j$ . Some authors (e.g. Karle, 1970) tend to use  $\sum_1$  relationships at the end of a phase determination process in order to decide the value of a symbol, others allow its optional use in the early stages (Germain, Main & Woolfson, 1970).

**The distribution  $P(E_{2h}, E_k, E_{h+k})$**

This distribution (Hauptman & Karle, 1953; Cochran & Woolfson, 1955) leads to

$$P_+(E_{2h}) = \frac{1}{2} + \frac{1}{2} \tanh \frac{1}{4N^{3/2}} |E_{2h}| \times \frac{(E_k^2 - 1)(E_{h+k}^2 - 1)}{1 - [H_4(E_k) + H_4(E_{h+k})]/8N}. \tag{3}$$

Owing to its high order ( $N^{3/2}$ ), (3) cannot reach high probability levels. Its big advantage, however, consists in the ability to vary  $k$  over reciprocal space (see next distribution).

**The distribution  $P(E_{2h}, E_{k_1}, E_{h+k_1}, E_{k_2}, E_{h+k_2}, \dots)$**

This distribution leads to

$$P_+(E_{2h}) \simeq \frac{1}{2} + \frac{1}{2} \tanh |E_{2h}| \cdot \frac{1}{4N^{2/3}} \times \frac{\sum_k (E_k^2 - 1)(E_{h+k}^2 - 1)}{1 - \frac{1}{8N} \left[ \sum_k H_4(E_k) + H_4(E_{h+k}) \right]^* + \frac{1}{N} \left( \sum_{k_1 \neq k_2} E_{k_1} E_{k_2} E_{h+k_1} E_{h+k_2} \right)^*}. \tag{4}$$

In this expression we have marked by an asterisk those terms which are not given by previous authors, but which are relevant for appreciating the limits of (4).

Because

$$\int_0^\infty H_4(x) \exp(-x^2/2) dx = 0,$$

the contribution of the first term marked by an asterisk in (4) vanishes when  $k$  varies over all reciprocal space. The second term, on the contrary, is strongly positive, since it is a summation of quartets (Giacovazzo, 1975b).

Owing to the large value of the denominator, (4) cannot reach high levels of probability and is unsuitable for sign determination.

**The distribution  $P(E_{2h}, E_h, E_k, E_{h+k})$**

The study of this distribution leads to

$$P_+(E_{2h}) \simeq \frac{1}{2} + \frac{1}{2} \tanh |E_{2h}| \frac{G}{Q}, \tag{5}$$

where

$$G = \frac{1}{2\sqrt{N}} (E_h^2 - 1) + \frac{1}{N} E_h E_k E_{h+k} + \frac{1}{2N} H_3(E_h) E_k E_{h+k} + \frac{1}{N^{3/2}} \left\{ -\frac{1}{3} H_4(E_h) + \frac{1}{4} (E_k^2 - 1)(E_{h+k}^2 - 1) + (E_h^2 - 1)(E_k^2 - 1)(E_{h+k}^2 - 1) - \frac{1}{16} [H_6(E_h) + (E_h^2 - 1)H_4(E_k) + (E_h^2 - 1)H_4(E_{h+k}) - 4H_4(E_h)(E_k^2 - 1)(E_{h+k}^2 - 1)] \right\}, \tag{6}$$

$$Q = 1 + \frac{1}{\sqrt{N}} E_h E_k E_{h+k} - \frac{1}{8N} [H_4(E_h) + H_4(E_k) + H_4(E_{h+k})] + \frac{1}{4N} H_4(E_h) + \frac{1}{2N} (E_h^2 - 1)(E_k^2 - 1)(E_{h+k}^2 - 1) + \frac{1}{N^{3/2}} \left\{ -\frac{1}{8} E_h E_k E_{h+k} [H_4(E_h) + H_4(E_k) + H_4(E_{h+k})] + \frac{1}{6} E_h E_k E_{h+k} (E_h^2 - 3)(E_k^2 - 3)(E_{h+k}^2 - 3) + H_3(E_h) E_k E_{h+k} \right\}. \tag{7}$$

Most of the terms in (6) and (7) have been neglected by previous authors. Most attention, in fact, has been

devoted to the asymptotic formula,

$$S(E_{2h}) = S[2(E_h^2 - 1) - N\langle(E_k^2 - 1)(E_{h+k}^2 - 1)\rangle], \quad (8)$$

which may be derived from the squared Patterson function (Cochran, 1954; Hauptman & Karle, 1957; von Eller, 1973). Relations (5) and (8) are unsuitable for sign determination. To large values of (6), in fact, correspond large values of (7), so that values only marginally different from 0.50 may be expected from (5).

#### The distribution $P(E_{2h}, E_h, E_k, E_{h-k}, E_{2h-k}, E_{h+k})$

We will describe this distribution in more detail because the final formulae will prove suitable for phase determination. We note that  $P(E_h, E_{2h})$ ,  $P(E_{2h}, E_k, E_{h+k})$ ,  $P(E_{2h}, E_h, E_k, E_{h+k})$  are particular cases of that treated in this paragraph. We denote for brevity

$$\begin{aligned} E_1 &= E_{2h}, & E_2 &= E_h, & E_3 &= E_k, \\ E_4 &= E_{h-k}, & E_5 &= E_{2h-k}, & E_6 &= E_{h+k}. \end{aligned}$$

In accordance with (1) we derive

$$\begin{aligned} S_3/t^{3/2} &= \frac{1}{\sqrt{N}} \{ \frac{1}{2}(iu_1)(iu_2)^2 + (iu_1)(iu_3)(iu_5) \\ &+ (iu_1)(iu_4)(iu_6) + (iu_2)(iu_3)(iu_4) \\ &+ (iu_2)(iu_3)(iu_6) + (iu_2)(iu_4)(iu_5) \}; \\ S_4/t^2 &= -\frac{1}{8N} [(iu_1)^4 + \dots + (iu_6)^4] \\ &+ \frac{1}{N} [(iu_1)(iu_2)(iu_3)(iu_4) \\ &+ (iu_1)(iu_2)(iu_3)(iu_6) + (iu_1)(iu_2)(iu_4)(iu_5) \\ &+ (iu_1)(iu_2)(iu_5)(iu_6) + (iu_3)(iu_4)(iu_5)(iu_6) \\ &+ \frac{1}{2}(iu_2)^2(iu_3)(iu_5) + \frac{1}{2}(iu_2)^2(iu_4)(iu_6) \\ &+ \frac{1}{2}(iu_3)^2(iu_4)(iu_6)]. \end{aligned}$$

The probability distribution is then

$$\begin{aligned} P(E_1, E_2, E_3, E_4, E_5, E_6) &= \frac{1}{(2\pi)^3} \exp -\frac{1}{2}(E_1^2 + \dots + E_6^2) \\ &\times \left\{ 1 + \frac{1}{\sqrt{N}} [\frac{1}{2}E_1(E_2^2 - 1) + E_1E_3E_5 \right. \\ &+ E_1E_4E_6 + E_2E_3E_4 + E_2E_3E_6 + E_2E_4E_5] \\ &- \frac{1}{8N} [H_4(E_1) + \dots + H_4(E_6)] \\ &+ \frac{1}{N} [E_1E_2E_3E_4 + E_1E_2E_3E_6 + E_1E_2E_4E_5 \\ &+ E_1E_2E_5E_6 + E_3E_4E_5E_6] \\ &+ \frac{1}{2N} [(E_2^2 - 1)E_4E_6 + (E_3^2 - 1)E_4E_6 \\ &+ (E_2^2 - 1)E_3E_5] + \frac{1}{2N} [\frac{1}{4}(E_1^2 - 1)H_4(E_2) \end{aligned}$$

$$\begin{aligned} &+ (E_1^2 - 1)(E_3^2 - 1)(E_5^2 - 1) \\ &+ (E_1^2 - 1)(E_4^2 - 1)(E_6^2 - 1) \\ &+ (E_2^2 - 1)(E_3^2 - 1)(E_4^2 - 1) \\ &+ (E_2^2 - 1)(E_3^2 - 1)(E_6^2 - 1) \\ &+ (E_2^2 - 1)(E_4^2 - 1)(E_5^2 - 1) + (E_1^2 - 1)(E_2^2 - 1)E_3E_5 \\ &+ (E_1^2 - 1)(E_2^2 - 1)E_4E_6 + E_1H_3(E_2)E_3E_4 \\ &+ E_1H_3(E_2)E_3E_6 + E_1H_3(E_2)E_4E_5 \\ &+ 2(E_1^2 - 1)E_3E_4E_5E_6 + 2E_1E_2(E_3^2 - 1)E_4E_5 \\ &+ 2E_1E_2(E_3^2 - 1)E_5E_6 + 2E_1E_2E_3E_4(E_5^2 - 1) \\ &+ 2E_1E_2E_3(E_4^2 - 1)E_6 + 2E_1E_2E_3E_4(E_6^2 - 1) \\ &+ 2E_1E_2(E_4^2 - 1)E_5E_6 + 2(E_2^2 - 1)(E_3^2 - 1)E_4E_6 \\ &+ 2(E_2^2 - 1)E_3(E_4^2 - 1)E_5 + 2(E_2^2 - 1)E_3E_4E_5E_6 \}. \quad (9) \end{aligned}$$

From the marginal distribution  $P(E_{2h}|E_h, E_k, E_{h-k}, E_{2h-k}, E_{h+k})$  one derives

$$P_+(E_1) \simeq \frac{1}{2} + \frac{1}{2} \tanh |E_1| \frac{\langle E_1 \rangle}{\sigma_1^2}, \quad (10)$$

where

$$\begin{aligned} \langle E_1 \rangle &= \frac{1}{\sqrt{N}} [\frac{1}{2}(E_2^2 - 1) + E_3E_5 + E_4E_6] \\ &+ \frac{1}{2N} [E_2E_3E_4(E_2^2 + 2E_5^2 + 2E_6^2 - 5) \\ &+ E_2E_3E_6(E_2^2 + 2E_4^2 - 3) + E_2E_4E_5(E_2^2 + 2E_3^2 - 3) \\ &+ E_2E_5E_6(2E_3^2 + 2E_4^2 - 2)], \\ \sigma_1^2 &\simeq 1 + \frac{1}{\sqrt{N}} [E_2E_3E_4 + E_2E_3E_6 + E_2E_4E_5] \\ &- \frac{1}{4N} (E_2^2 - 1)^2 - \frac{1}{N} E_3^2E_5^2 - \frac{1}{N} E_4^2E_6^2 \\ &+ \frac{1}{8N} H_4(E_2) \\ &- \frac{1}{8N} [H_4(E_3) + H_4(E_4) + H_4(E_5) + H_4(E_6)] \\ &+ \frac{1}{N} E_3E_4E_5E_6 \cdot E_2^2 \\ &+ \frac{1}{2N} [2(E_3^2 - 1)(E_5^2 - 1) + 2(E_4^2 - 1)(E_6^2 - 1) \\ &+ (E_2^2 - 1)(E_3^2 - 1)(E_4^2 - 1) \\ &+ (E_2^2 - 1)(E_3^2 - 1)(E_6^2 - 1) \\ &+ (E_2^2 - 1)(E_4^2 - 1)(E_5^2 - 1) \\ &+ 2E_3E_5(E_2^2 - 1)(E_4^2 - 1) \\ &+ 2E_4E_6(E_2^2 - 1)(E_3^2 - 1)]. \end{aligned}$$

As we see, the value of (10) critically depends on the unknown signs of the semi-invariants  $E_k E_{2h-k}$ ,  $E_{h-k} E_{h+k}$ ,  $E_h E_{2h-k} E_{h+k}$  as well as on the signs of the invariants  $E_1 E_k E_{h-k}$ ,  $E_h E_k E_{h+k}$ ,  $E_h E_{2h-k} E_{h-k}$ . The signs of the invariants could be considered positive when the  $E$ 's involved are large enough, whereas the same cannot be said for semi-invariants. The signs of

$E_{h-k}E_{h+k}$  (and of  $E_kE_{2h-k}$  with a trivial change of notation), nevertheless, could be deduced (Giacovazzo, 1974) from

$$P_+(E_{h+k}E_{h-k}) \simeq \frac{1}{2} + \frac{1}{2} \tanh \frac{1}{2N} |E_{h+k}E_{h-k}| \times (2E_h^2E_k^2 - E_h^2 - E_k^2), \quad (11)$$

which is able to assign negative values to  $E_{h-k}E_{h+k}$ . Unfortunately the probability levels involved in (11) are in general not so high as to warrant their use for defining signs which play a role in the first stages of the sign determination process. We will follow, therefore, a different approach. As  $E_1, E_2, E_3, E_4$  constitute a quartet, we derive

$$P_+(E_{2h}E_hE_kE_{h-k}) \simeq \frac{1}{2} + \frac{1}{2} \tanh \frac{1}{N} |E_{2h}E_hE_kE_{h-k}| \times \frac{E_{2h-k}^2 + E_{h+k}^2 - 1}{1 + 4[E_{2h-k}^2 + E_{h+k}^2 - 1]/N}. \quad (12)$$

(12) tells us that the product  $E_{2h}E_hE_kE_{h-k}$  is probably positive when  $E_{2h-k}^2 + E_{h+k}^2 - 1 > 0$ , probably negative when  $E_{2h-k}^2 + E_{h+k}^2 - 1 < 0$ . The character of positivity or negativity is strengthened by large values of  $|E_{2h}E_hE_kE_{h-k}|$ . (12) may be usefully compared with the formula valid for general quartets (Giacovazzo, 1975b)

$$P_+(E_hE_kE_lE_{h+k+l}) \simeq \frac{1}{2} + \frac{1}{2} \tanh \frac{1}{N} |E_hE_kE_lE_{h+k+l}| \times \frac{E_{h+k}^2 + E_{h+l}^2 + E_{k+l}^2 - 2}{1 + 4[E_{h+k}^2 + E_{h+l}^2 + E_{k+l}^2 - 3]/N}. \quad (13)$$

As we see, when  $l = -2h$ , the second of the cross vectors  $h+k, h+l, k+l$  coincides with a vector which is part of the quartet: in this case it is not relevant for defining the positivity or negativity of  $E_hE_kE_lE_{h+k+l}$ . A first conclusion is that (12) may improve the reliability of the quartet theory because it allows us to distinguish between special and general quartets. (12) is correct up to order  $N^{-1}$ . In a recent paper (Giacovazzo, 1975c) the inspection of the distribution

$$P(E_h, E_k, E_l, E_{h+k+l}, E_{h+k}, E_{h+l}, E_{k+l})$$

has been performed up to order  $N^{-2}$ . Using that result, we rewrite (12) in the more complete form

$$P_+(E_{2h}, E_h, E_k, E_{h+k}) \simeq \frac{1}{2} + \frac{1}{2} \tanh \frac{1}{N} |E_{2h}E_hE_kE_{h-k}| \times \frac{E_{2h-k}^2 + E_{h+k}^2 - 1 + R}{1 + 4[E_{2h-k}^2 + E_{h+k}^2 - 1]/N + Q}, \quad (14)$$

where

$$R \simeq \frac{1}{N} \{ (E_{2h-k}^2 - 1)(E_{h+k}^2 - 1) - [H_6(E_{2h-k}) + H_6(E_{h+k}) + H_4(E_{2h-k})(E_{h+k}^2 + 8) + H_4(E_{h+k})(E_{2h-k}^2 + 8)]/8 \},$$

$$Q \simeq \frac{1}{18N^2} [H_6(E_{2h-k}) + H_6(E_{h+k})] + \frac{9}{N^2} (E_{2h-k}^2 + E_{h+k}^2 - 1) + \frac{1}{128N^2} [H_8(E_{2h-k}) + H_8(E_{h+k}) + 2H_4(E_{2h-k})H_4(E_{h+k})] + \frac{52}{N^2} (E_{2h-k}^2 - 1)(E_{h+k}^2 - 1) - \frac{4}{3N^2} [H_4(E_{2h-k}) + H_4(E_{h+k})].$$

### Calculations

In this section we explore the usefulness of (12) and (14) by comparing their reliability with that of triplets. As the probability that a triplet has a positive sign is given by

$$P_+(E_hE_kE_{h+k}) \simeq \frac{1}{2} + \frac{1}{2} \tanh \frac{1}{N} |E_hE_kE_{h+k}|, \quad (15)$$

the comparison may be made by calculating the failure percentages as a function of the argument of the hyperbolic tangent in (12), (13), (14) and (15). In Tables 1 and 2 we give for two model structures the number of relations and the percentage of correct ones above the corresponding value of the argument. In accordance with the aim of this paper we include in the tables only those triplets which are part of the quartets considered. For the sake of simplicity,  $2h, h, k, h-k$  reflexions with  $|E| < 0.8$  are excluded from the calculations.

Table 1. *Number (nr) and percentage of correct relations for triplets and quartets according to (15), (12), (14) and (13) in a model structure with  $N=40$*

Reflexions  $2h, h, k, h-k$  with  $|E| < 0.8$  are excluded from the calculations.

Tanh argu-ments	Triplets (15)		Quartets (12)		Quartets (14)		Quartets (13)	
	nr	%	nr	%	nr	%	nr	%
0.15	1492	72	356	73	362	73	407	68
0.30	788	80	161	86	167	85	207	73
0.50	360	91	83	87	86	88	107	86
0.70	168	95	55	91	56	95	75	92
0.90	98	97	39	95	37	97	48	92
1.10	61	100	28	100	25	100	40	95
1.40	31	100	21	100	21	100	31	94
2.00	12	100	15	100	14	100	22	91
2.50	8	100	10	100	8	100	17	94
3.00	1	100	6	100	1	100	10	90

From the tables we may conclude that the reliability of (12) and (14) is comparable with that of the triplets: on the contrary the general formula (13) is not suitable for special quartets. A last observation is that the estimate of the terms of order  $1/N^2$  in (14) tends to lower the higher probability values as measured by (12).

Table 2. Number (*nr*) and percentage of correct relations for triplets and quartets according to (15), (12), (14) and (13) in a model structure with  $N=60$

Reflexions  $2h, h, k, h-k$  with  $|E| < 0.8$  are excluded from the calculations.

Tanh arguments	Triplets (15)		Quartets (12)		Quartets (14)		Quartets (13)	
	nr	%	nr	%	nr	%	nr	%
0.15	5256	67	734	66	741	66	1043	62
0.30	2413	74	106	73	209	71	354	63
0.50	781	83	60	82	53	81	116	72
0.70	245	91	23	83	24	79	52	75
0.90	111	98	10	90	9	89	28	82
1.10	42	100	5	80	4	75	16	75
1.40	13	100	2	100	2	100	7	71
2.00	1	100	1	100	1	100	3	67
2.50	1	100					3	67
3.00								

The sign of  $E_{2h}$

A distinctive feature of the quartet  $E_{2h}, E_h, E_k, E_{h-k}$  is that it contains the triplet  $E_k, E_h, E_{h-k}$ . Under the assumption that (15) is large enough, (12) fixes the sign of  $E_{2h}$  when  $P_+(E_{2h}E_hE_kE_{h-k})$  or  $1 - P_+(E_{2h}E_hE_kE_{h-k})$  is sufficiently large. It should be noted that the sign of  $E_{2h}$  is positive when both  $E_{2h}E_hE_kE_{h-k}$  and  $E_hE_kE_{h-k}$  are positive or both are negative. If we consider that the sign of the triplet does not depend on that of the quartet, we may write

$$P_+(E_{2h}) = P_Q P_T + (1 - P_Q)(1 - P_T), \quad (16)$$

where  $P_Q$  and  $P_T$  are given respectively by (12) and (15). It may be verified that  $P_+(E_{2h})$  does not approach unity too closely unless  $P_Q$  and  $P_T$  are both close to unity; or,  $P_+(E_{2h})$  does not approach zero too closely unless  $P_Q$  is close to zero and  $P_T$  close to unity. In other words, the use of (16) involves reliability levels lower than those involved separately in the corresponding triplet and quartet relationships. One could conclude that this type of sign relationship is not advisable in the first stages of sign determination without the cooperation of further favourable circumstances.

A great advantage of the theory, in fact, is that, for a given crystal structure, a large number of triplet and quartet relationships can occur so as to make certain the determination of the sign of  $E_{2h}$ . The  $k$  index, indeed, can vary over reciprocal space so giving rise to a large number of quartets which have  $h$  as a fixed index.

Let us denote by  $P_{k_j}^+$  the probability of a positive sign for  $E_{2h}$ , as calculated by (16) when  $k=k_j$ , and let

$$P_{k_j}^- = 1 - P_{k_j}^+.$$

A measure of the overall probability that  $E_{2h}$  has a positive sign is given by (Woolfson, 1961)

$$P_+(E_{2h}) = \left( 1 + \frac{\prod_j P_{k_j}^-}{\prod_j P_{k_j}^+} \right)^{-1}. \quad (17)$$

Experimental

We have tested (17) on a real structure (Kalyani & Vijayan, 1969) and by the same two model structures as in Tables 1 and 2. As usual in the real structure we have replaced  $N$  by  $\sigma_3^{-2}\sigma_2^3$  in all the probabilistic formulae (Karle & Karle, 1966). We have first tried to involve for each  $E_{2h}$  reflexion the largest number of quartets possible. This procedure proved unsuitable:

(1) because of the number of failures in triplet as well as in quartet relationships.

(2) as, in these conditions, (17) is too sensitive eventually to overestimate or underestimate the positivity of the quartets compared with their negativity. This danger is real (Giacovazzo, 1975*b*; Schenk, 1975), and is due to the fact that we have represented the probability distributions as asymptotic series and the actual values of the probability we obtain are correct to the degree of approximation we choose.

It seemed then preferable to fix a threshold value for the magnitudes of the quartet vectors (not for that of the cross vectors, of course) so as to warrant high probability levels both for quartets and, even more, for triplets. Threshold values about 1.30-1.40 seemed advisable.

Experimental tests showed that (12) overestimates the positivity of the quartets: (14) improves this behaviour but not in a conclusive way. We have used, therefore, an empirical formula which, saving computing time, lowers the weight of the positive quartets. (12) and (14) have been replaced then by

$$P_+(E_{2h}E_hE_kE_{h-k}) \simeq \frac{1}{2} + \frac{1}{2} \tanh \frac{1}{N} |E_{2h}E_hE_kE_{h-k}| \times \frac{E_{2h-k}^2 + E_{h+k}^2 - 1}{1 + 2 \tanh [(E_{2h-k}^2 + E_{h+k}^2 - 1)/3]}. \quad (18)$$

Table 3. List of pairs  $E_{2h}, E_h$  which form quartets  $E_{2h}, E_h, E_k, E_{h-k}$  whose  $|E|$  magnitudes are larger than 1.35 in a real structure (Kalyani & Vijayan, 1969) with  $\sigma_3^{-2}\sigma_2^3 \simeq 20$

2h	$E_{2h}$	h	$E_h$	PSIG	PQ	PC	NQ
2 -4 0	1.494	1 -2 0	3.158	0.956	1.0	1.0	19
2 -2 -6	1.822	1 -1 -3	1.685	0.683	1.0	1.0	15
2 4 6	-2.007	1 2 3	1.879	0.762	0.403	0.683	6
4 0 -2	1.807	2 0 -1	-1.851	0.732	1.0	1.0	14
2 -2 -6	1.822	1 -1 -3	1.685	0.683	1.0	1.0	15

(18) proved satisfactory: the results are shown in Tables 3, 4 and 5. In these tables for each pair  $E_{2h}$ ,  $E_k$  are given the probability values as calculated by (2), and by (17) through (16) and (18) (respectively PSIG and PQ in the tables).

In Tables 3 and 4 five pairs  $E_{2h}$ ,  $E_h$  are tested. The  $\sum_1$  relationship fails in two cases, unfortunately marked by the highest values of  $|E_{2h}|$ : on the other hand, the values of PQ correctly suggest for them negative values.

In Table 5 the negativity of  $E_{2,-2,4}$  and  $E_{4,0,8}$  is strongly indicated, although the  $\sum_1$  formula assigns to them comparatively large probability values. Three negative signs, nevertheless, are not denounced by the quartet method. So it seemed advisable to introduce the combined figure

$$PC = \left[ 1 + \frac{(1 - \text{PSIG})(1 - \text{PQ})}{\text{PSIG} \cdot \text{PQ}} \right]^{-1}$$

the  $E_{2h}$  structure factors with the highest values of PC will have large probabilities of being positive.

Quartets in which  $k=h/2$  must be removed from the procedure: the product  $E_h E_k E_{h-k}$  in fact reduces to  $E_h E_{h/2}^2$ , whose sign is *a priori* unknown. The use of these special quartets, nevertheless, may be advisable when the sign of  $E_h$  has been determined by a high PC value. We have marked by an asterisk in Table 5 a vector  $2h$  whose PQ value has been fixed by this device. For the sake of clarity we give in Table 6 the list of the quartets which contribute to fixing the sign

of  $E_{4,-4,8}$ , and their respective values  $P_+(E_{2h}E_h, E_k E_{h-k})$  as given by (18). For the first of these quartets  $k=h/2$ ; the value of  $P_+(E_{2h})$  in (16) may be calculated replacing  $P_T$  by 0.152, as derived from Table 5. Let us note two further applications of the theory:

(1) the use of (18) for finding the correct solution out of a set of  $\sum_2$  solutions, particularly in symmorph space groups. A good criterion would be

$$SQC = \sum_{h,k} E_{2h} E_h E_k E_{h-k} (E_{2h-k}^2 + E_{h+k}^2 - 1).$$

(2) The removal of probably failing triplets from the  $\sum_2$  listing. Let us suppose, for example, that  $E_{2,-2,4} = -1.417$  has been included in the starting set. From Table 7 we see that 14 of the 38 quartets which contribute to fixing the value of  $E_{2,-2,4}$  suggest, through (18) and (16), a positive sign for  $E_{2h}$ . As the failures may be attributed to quartet as well as to triplet relationships, it seems advisable to remove these triplets from the  $\sum_2$  listing. Two triplets, marked by an asterisk in Table 7, prove to be fallacious.

### Conclusions

The  $\sum_1$  formula often generates phases whose reliability is mathematically very high. Unfortunately the frequency of the failures advises against the introduction of these phases in the starting set.

The quartet method seems able to reduce the number of failures and to warrant reliability levels such as are

Table 4. List of pairs  $E_{2h}, E_h$  which form quartets  $E_{2h}, E_h, E_k, E_{h-k}$  whose  $|E|$  magnitudes are larger than 1.35 in a model structure with  $N=40$

$2h$	$E_{2h}$	$h$	$E_h$	PSIG	PQ	PC	NQ
6 2 4	1.799	3 1 2	3.446	0.957	1.0	1.0	7
2 -2 4	2.241	1 -1 2	-2.760	0.913	1.0	1.0	23
4 4 0	2.480	2 2 0	-1.757	0.694	1.0	1.0	12
2 -6 2	-2.505	1 -3 1	1.655	0.666	0.053	0.100	4
6 -6 0	1.594	3 -3 0	-1.584	0.594	0.690	0.770	3

Table 5. List of pairs  $E_{2h}, E_h$  which form quartets  $E_{2h}, E_h, E_k, E_{h-k}$  whose  $|E|$  magnitudes are larger than 1.35 in a model structure with  $N=60$

$2h$	$E_{2h}$	$h$	$E_h$	PSIG	PQ	PC	NQ
2 -2 4	-1.417	1 -1 2	2.016	0.637	0.093	0.152	38
0 6 -8	1.369	0 3 -4	1.935	0.619	0.721	0.808	15
8 0 -2	1.673	4 0 -1	-1.856	0.629	0.971	0.983	22
4 0 -8	-1.390	2 0 -4	-1.848	0.607	0.434	0.542	17
4 6 6	1.751	2 3 3	1.688	0.603	0.971	0.981	13
4 8 0	1.496	2 4 0	-1.654	0.583	0.536	0.618	10
6 -6 -2	1.753	3 -3 -1	-1.610	0.589	0.448	0.538	16
6 -8 8	1.848	3 -4 4	-1.595	0.591	0.663	0.740	1
8 0 -6	1.712	4 0 -3	-1.562	0.579	0.409	0.488	11
6 8 6	-1.711	3 4 3	1.491	0.567	0.596	0.658	3
*4 -4 8	1.528	2 -2 4	-1.417	0.550	0.594	0.641	7
8 0 4	-1.697	4 0 2	1.406	0.553	0.688	0.732	9
4 8 4	3.010	2 4 2	1.402	0.593	0.696	0.769	11
2 4 -4	-1.384	1 2 -2	1.363	0.538	0.686	0.718	28
6 -2 -2	1.916	3 -1 -1	1.356	0.730	0.350	0.593	21
6 -2 2	1.366	3 -1 1	1.354	0.537	0.768	0.793	35

required in the early stages of procedures for phase assignment. Two additional applications – the recognition of the correct solution out of a set of  $\Sigma_2$  solutions and the removal of some probably fallacious triplets from the  $\Sigma_2$  listing – are also suggested.

In subsequent papers we shall describe the use of the quartet method in non-centrosymmetric groups and in centrosymmetric ones with symmetry higher than  $P\bar{1}$ .

The author is indebted to Mr G. Gastellaneta (C.S.A.T.A., Bari) who performed the calculations summarized in the tables.

## References

- COCHRAN, W. (1954). *Acta Cryst.* **7**, 581–583.  
 COCHRAN, W. & WOOLFSON, M. M. (1954). *Acta Cryst.* **7**, 450–451.  
 COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1–12.  
 ELLER, G. VON (1973). *Acta Cryst.* **A29**, 63–67.  
 GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1970). *Acta Cryst.* **B26**, 274–285.  
 GIACOVAZZO, C. (1974). *Acta Cryst.* **A30**, 481–484.  
 GIACOVAZZO, C. (1975a). *I Metodi Diretti nella Soluzione delle Strutture Cristalline*. Milan: Tamburini.

Table 6. List of the quartets which contribute to fixing the sign of  $E_{4,-4,8}$ 

$E_{2h}$	$h$	$E_{2h}$	$k$	$E_k$	$h-k$	$E_{h-k}$	$2h-k$	$E_{2h-k}$	$h+k$	$E_{h+k}$	(18)
1.528	2 -2	4 -1.417	1 -1	2.016	1 -1	2 2.016	3 -3	6 -0.030	3 -3	6 -0.030	0.306
			1 4	3 1.846	1 -6	1 -3.058	3 -8	5 0.094	3 2	7 -0.935	0.487
			4 2	3 -1.616	2 4	-1 2.447	0 6	-5 0.258	6 0	7 0.444	0.400
			2 1	0 -1.587	0 3	-4 1.935	2 -5	8 0.187	4 -1	4 1.268	0.525
			4 5	-4 1.500	6 3	0 -2.180	8 1	4 0.531	2 7	-8 -2.190	0.587
			4 4	0 1.404	6 2	4 -1.702	8 0	8 0.097	2 6	-4 -0.025	0.384
			2 4	2 1.402	0 6	-2 1.685	2 -8	6 -0.525	4 2	6 -1.245	0.523

Table 7. List of the triplets which contribute to fixing the sign of  $E_{2,-2,4}$ 

$2h$	$E_{2h}$	$h$	$E_h$	$k$	$E_k$	$h-k$	$E_{h-k}$	$2h-k$	$E_{2h-k}$	$h+k$	$E_{h+k}$	(16)
2 -2	4 -1.417	1 -1	2 2.016	1 -6	1 -3.058	2 -7	3 -3.417	3 -8	5 0.094	0 5	1 1.043	0.523
			2 4	4 -2.067	1 5	2 -2.119	0 6	0 -0.377	3 3	6 -1.423	0.569	
			6 -2	-2 1.916	5 -1	-4 3.598	4 0	-6 0.915	7 -3	0 0.169	0.476	
			3 -4	3 1.912	4 -5	5 3.598	5 -6	7 0.917	2 -3	1 0.167	0.476	
			3 -5	-2 -1.827	4 -6	0 -2.125	5 -7	2 -0.004	2 -4	-4 -0.605	0.401	
			6 -1	3 -1.824	5 0	1 -2.121	4 1	-1 -0.002	7 -2	5 -0.606	0.401	
			2 7	3 -1.793	3 6	5 -2.435	4 5	7 -0.572	1 8	1 2.505	0.690	
			6 -7	4 -1.776	5 -6	2 -1.858	4 -5	0 0.018	7 -8	6 -0.274	0.327	
			3 1	-3 -1.771	4 0	-1 -1.856	5 -1	1 0.018	2 2	-5 -0.273	0.327	
			2 6	1 -1.695	3 5	3 -1.820	4 4	5 1.028	1 7	-1 -0.321	0.511	
			2 4	0 -1.654	3 3	2 -3.067	4 2	4 -0.576	1 5	-2 1.033	0.538	
			3 -4	5 1.606	2 -3	3 1.687	1 -2	1 -0.666	4 -5	7 -0.463	0.472	
			6 -2	-4 1.603	7 -3	-2 1.686	8 -4	0 -0.665	5 -1	-6 -0.461	0.471	
			1 7	2 1.594	2 6	4 2.461	3 5	6 0.187	0 8	0 0.516	0.382	
			1 -2	4 1.560	2 -3	6 1.962	3 -4	8 0.603	0 1	-2 -0.480	0.460	
			3 2	5 1.560	2 3	3 1.688	1 4	1 0.883	4 1	7 -1.129	0.539	
			5 1	-3 1.552	4 2	-5 2.050	3 3	-7 -0.122	6 0	-1 -1.109	0.516	
			7 2	2 -1.472	6 3	0 -2.180	5 4	-2 -0.582	8 1	4 0.531	0.461	
			2 3	6 -1.464	1 4	4 -1.497	0 5	2 0.370	3 2	8 0.335	0.424	
			*1 4	2 1.461	0 5	0 -1.630	1 -6	2 -0.728	2 3	4 -0.788	0.508	
			5 2	-6 -1.451	6 1	-4 -1.786	7 0	-2 -0.428	4 3	-8 0.744	0.480	
			3 1	4 -1.446	2 2	2 -2.547	1 3	0 -0.977	4 0	6 1.332	0.574	
			2 -3	-2 -1.424	1 -2	-4 -1.509	0 1	6 -1.408	3 -4	0 0.331	0.533	
			3 3	6 -1.423	2 4	4 -2.067	1 5	2 -2.119	4 2	8 -0.006	0.591	
			5 -6	-1 1.408	4 -5	-3 1.714	3 -4	-5 -0.001	6 -7	1 -0.663	0.449	
			0 1	6 -1.408	1 -2	-4 -1.509	2 -3	-2 -1.424	1 0	8 1.099	0.550	
			4 0	2 1.406	5 -1	4 1.711	6 -2	6 -0.002	3 1	0 -0.664	0.450	
			4 -4	1 -1.405	3 -3	-1 -1.610	2 -2	-3 0.184	5 -5	3 0.309	0.394	
			5 1	0 1.404	4 2	-2 1.635	3 3	-4 -0.046	6 0	2 -0.856	0.482	
			5 -2	0 -1.404	6 -3	2 -1.609	7 -4	4 0.186	4 -1	-2 0.310	0.394	
			4 -1	6 -1.389	3 0	4 -1.679	2 1	2 0.160	5 -2	8 0.681	0.457	
			2 2	-3 -1.388	1 3	-5 -2.294	0 4	-7 0.086	3 1	-1 -0.035	0.304	
			5 -5	-5 -1.387	6 -6	-3 -1.680	7 -7	-1 0.156	4 -4	-7 0.681	0.457	
			2 4	-4 -1.384	3 3	-2 -2.140	4 2	0 -0.478	1 5	-6 -0.096	0.395	
			3 -4	-1 1.369	4 -5	1 1.520	5 -6	3 -0.813	2 -3	-3 0.341	0.487	
			6 -2	2 1.366	5 -1	0 1.517	4 0	-2 -0.813	7 -3	4 0.344	0.487	
			*1 2	-2 1.363	2 1	0 -1.587	3 0	2 0.936	0 3	-4 1.935	0.569	
			1 2	-2 1.363	0 3	-4 1.935	1 -4	6 0.891	2 1	0 -1.587	0.563	

- GIACOVAZZO, C. (1975*b*). *Acta Cryst.* **A31**, 252–259.  
 GIACOVAZZO, C. (1975*c*). In preparation.  
 HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal*. A.C.A. Monograph No. 3. Pittsburg: Polycrystal Book Service.  
 HAUPTMAN, H. & KARLE, J. (1957). *Acta Cryst.* **10**, 267–270.  
 KALYANI, V. & VIJAYAN, M. (1969). *Acta Cryst.* **B25**, 252–260.  
 KARLE, I. L. (1970). *General Procedure for Phase Determination: Discussion*. In *Crystallographic Computing*. Munksgaard: Copenhagen.  
 KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.  
 KLUG, A. (1958). *Acta Cryst.* **11**, 513–543.  
 SCHENK, H. (1975). *Acta Cryst.* **A31**, 259–263.  
 WOOLFSON, M. M. (1961). *Direct Methods in Crystallography*. Oxford: Clarendon Press.

*Acta Cryst.* (1975). **A31**, 609

## Anisotropies Optiques de Quelques Molécules de Stéroïdes $C_{21}$

PAR R. SOBRY ET G. VAN DEN BOSSCHE

*Laboratoire de Cristallographie et de Physique de l'Etat solide, Institut de Physique, Université de Liège au Sart Tilman, B-4000 Liège, Belgique*

(Reçu le 6 février 1975, accepté le 10 mars 1975)

The molecular refractivities of some corticosteroids  $C_{21}$  are calculated from the crystal refractivities. The latter are obtained from the measured principal refractive indices. The orientations of the principal molecular axes with regard to the steroid skeleton are given and the magnetic and optical anisotropies compared.

Dans un article précédent, nous avons étudié les propriétés diamagnétiques de quelques corticostéroïdes (Van den Bossche & Sobry, 1975). Préalablement à la mesure magnétique, nous avons dû entreprendre, pour chacun des cristaux, une étude goniométrique de façon à pouvoir repérer les axes cristallins par rapport au faciès et ainsi pouvoir réaliser les différentes suspensions lors des mesures d'anisotropie. Aussi en avons-nous profité pour déterminer les caractéristiques principales des ellipsoïdes des indices de réfraction et ce pour les différentes longueurs d'onde de la lampe à vapeur de mercure.

A partir de ces indices, il nous est possible de calculer les réfractivités cristallines des composés correspondants par la formule de Lorenz et Lorentz:

$$R_i = \frac{n_i^2 - 1}{n_i^2 + 2} \cdot \frac{M}{\rho} = \frac{4\pi}{3} N\alpha_i \quad (1)$$

où  $R_i$ ,  $n_i$  et  $\alpha_i$  sont respectivement la réfractivité molaire, l'indice de réfraction et la polarisabilité électrique dans la direction principale  $i$ ,  $M$  la masse moléculaire,  $\rho$  la masse volumique et  $N$  le nombre d'Avogadro.

Toutes les réfractivités sont exprimées en unités c.g.s. c'est-à-dire en  $\text{cm}^3/\text{mole}$ . Le Tableau 1 reprend les valeurs calculées à partir de (1) pour les principales longueurs d'onde de la lampe à vapeur de mercure. Dans le cas des composés monocliniques, pour conserver le parallélisme de notations avec les susceptibilités magnétiques, la réfractivité suivant  $b$  est appelée  $R_3$  et la plus grande dans le plan (010),  $R_1$ ; cette dernière est repérée par l'angle  $\varphi$  qu'elle fait avec l'axe cristallographique  $a$ , cet angle étant défini positivement dans l'angle obtus  $\beta$  compris entre les axes  $a$  et  $c$ .

Tableau 1. Réfractivités cristallines

$\lambda$ (Å)	6234	5780	5461	4916	4358	4047
Progesterone						
$R_a$		101,69	102,27	102,97	103,78	104,35
$R_b$		86,05	86,31	86,69	87,01	87,59
$R_c$		84,75	85,01	85,40	85,66	86,31
17 $\alpha$ -Hydroxyprogesterone						
$R_a$	87,10	87,49	87,88	88,26	88,65	89,55
$R_b$	98,76	99,37	99,85	100,57	101,04	102,58
$R_c$	89,55	90,06	90,38	90,83	91,33	92,09
11-Désoxycorticostérone $\varphi = 38^\circ$						
$R_1$	103,33	104,16	104,63	105,33	106,03	107,64
$R_3$	82,63	82,90	83,16	83,70	83,97	84,63
$R_2$	87,66	88,05	88,31	88,83	89,35	90,12
11-Désoxycortisol $\varphi = 42^\circ 07'$						
$R_1$	102,24	102,84	103,32	103,91	104,51	105,81
$R_3$	91,25	91,64	91,90	92,28	92,80	93,82
$R_2$	84,54	84,94	85,21	85,74	86,14	87,07
Corticostérone $\varphi = 34^\circ 25'$						
$R_1$	105,81	106,52	107,11	107,70	108,40	109,91
$R_3$	85,50	85,86	86,10	86,46	86,80	87,46
$R_2$	92,15	92,67	92,93	93,44	93,83	94,72

Parmi les premières tentatives d'interprétation des réfractivités, citons celles de Bragg (1924*a, b*) sur les carbonates et ensuite de Brasseur (1932; Brasseur & de Rassenfosse, 1941), de Hendricks & Deming (1935) sur les oxalates, et plus récemment de Bothorel (1959) sur une série de molécules polyphényliques.