Computer Analysis of Molecular Geometry.
V. Symmetry Aspects of Factor Analysis

By Peter Murray-Rust

Department of Chemistry, University of Stirling, Stirling FK9 4LA, Scotland

Abstract

The application of group theory to factor analysis of molecular geometry is discussed and illustrated with an examination of 211 distorted phosphate groups.

Introduction

In previous papers [Murray-Rust & Motherwell, 1978*a* (MM*a*), 1978*b* (MM*b*); Murray-Rust & Bland, 1978 (MB)] we have described the application of factor analysis (particularly the principal-components method) to molecular geometry. Specifically we showed how the multivariate distribution of torsion angles describing a nucleoside fragment could be analysed by this method. The procedure described in these papers is, however, only appropriate for molecular fragments which cannot show any symmetry or, put another way, can be given a unique labelling scheme for the atoms. In this paper the method is extended to cover configurations which are symmetrical or which show small distortions from a symmetrical structure (and therefore have an ambiguity in the atomic labelling).

The multivariate distribution

A multivariate distribution of m parameters $\mathbf{p}$ can be expressed by the equation:

$$P(\mathbf{p}) \propto \exp\left[-\tfrac{1}{2}(\mathbf{p} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{A}^{-1}(\mathbf{p} - \boldsymbol{\mu})\right]. \qquad (1)$$

Here P is the probability density at the point with coordinates $\mathbf{p}$, $\boldsymbol{\mu}$ are the means of the parameters $\mathbf{p}$, and $\mathbf{A}$ is the dispersion matrix (the matrix of variances and covariances). In our previous treatment we found it convenient to transform $\mathbf{p}$ values to $\mathbf{z}$ values (which have zero mean and unit variance) by a simple shift of origin and change of scale:

$$\mathbf{z} = (\mathbf{p} - \boldsymbol{\mu})/\boldsymbol{\sigma} \quad [i.e.\ z_i = (\mu_i - p_i)/\sigma_i]. \qquad (2)$$

The expression for the probability density then becomes:

$$P(z) \propto \exp\left[-\tfrac{1}{2}\mathbf{z}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{z}\right] \qquad (3)$$

where $\mathbf{R}$ is the correlation matrix, with unities in the leading diagonal and Pearson correlation coefficients off the diagonal. If there are no missing observations in the data, $\mathbf{R}$ is Gramian (or semi-positive definite), *i.e.* all the eigenvalues are real and non-negative. The eigenvectors of $\mathbf{R}$, weighted by the eignevalues, are the factors or principal components of the distribution:

$$\mathbf{F} = \boldsymbol{\lambda}^{1/2}\, \mathbf{E}. \qquad (4)$$

For a given probability distribution density P_i we can write:

$$\mathbf{z}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{z} = 2 \ln P_i + \text{constant}, \qquad (5)$$

which is a standard quadratic form and represents the equation of a hyperellipsoid. For different P_i we can therefore draw hyperellipsoidal contours concentric about the mean of the distribution.

In the treatment that follows, we shall assume that all distributions have the form of (1), although the use of z values and (3) may show simpler forms. In the general case where the centre of the distribution ($\boldsymbol{\mu}$) represents a configuration of atoms with no symmetry (as for the nucleosides) the derivation of factors is straightforward and can be carried out with standard algorithms (*e.g. SPSS*). Where the mean $\boldsymbol{\mu}$ represents a configuration with higher symmetry than the configuration represented by any other point, group-theoretical considerations must be introduced into the factor analysis. These considerations are very similar to those involved in the analysis of molecular vibrations of symmetrical molecules and we shall assume familiarity with the application of group-theoretical methods as described in standard works (*e.g.* Wilson, Decius & Cross, 1955). The effect of symmetry on factor analysis is introduced by a two-dimensional (bivariate) example.

A bivariate distribution with symmetry

Consider a linear triatomic *XYX* fragment in different crystal or molecular environments (Fig. 1*a*). Each case is completely described by two parameters, r_1 and r_2, which in general will not be equal. Unless there has

been some system in labelling the molecule we cannot say *a priori* in any particular fragment whether r_1 or r_2 is the longer. If a large number of cases were taken we would expect r_1 and r_2 to have identical distributions since there is no way of distinguishing between them. Although there is no symmetry relating r_1 and r_2 in any *particular* fragment, it is clear that in general there is symmetry relating the *distributions* of r_1 and r_2 and this is exemplified by Fig. 1(*b*). In this figure the distribution is shown for a hypothetical series of XYX molecules which are labelled at random. Although not *precisely* determined by symmetry the plot has an almost exact diagonal mirror line about $r_1 = r_2$. If the distribution is unimodal and conforms to (1), the mean will lie on the diagonal with $r_1 = r_2 = r_0$ (say). Moreover, the dispersion matrix **A** must have elements such that the axes of the distribution lie at approximately 45° to the parameters r_i.
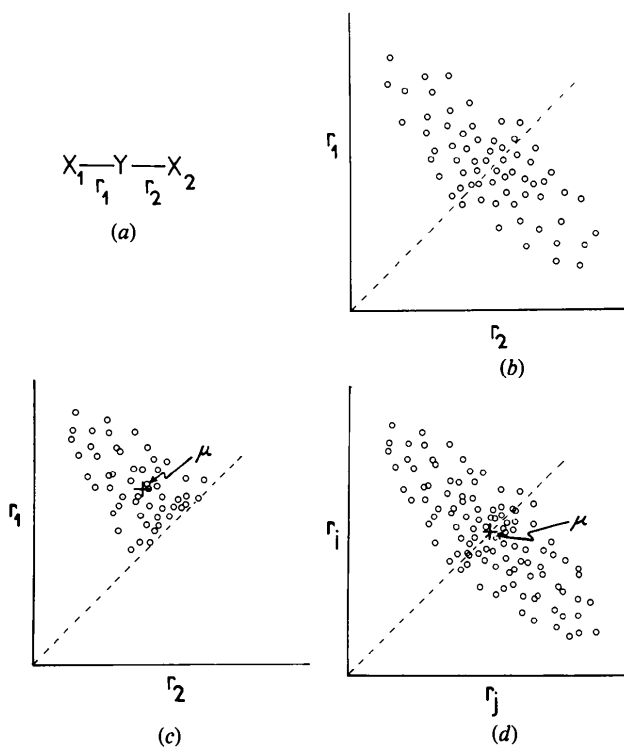
When this problem of labelling arises there are three possible ways of proceeding. The first is simply to accept whatever (random) labelling is already on the molecule, or to label it at random. In doing this we can expect near, but not exact, symmetry in the results of statistical analysis. In more complicated systems it may be difficult to see whether this symmetry is coincidental or arises from the intrinsic ways of labelling the molecule. A further disadvantage is that two different workers using different labelling may produce slightly different answers from the same data. The second possibility is that of using the observed geometry of each fragment to calculate a labelling scheme; thus for the XYX system we might always call the longest $X-Y$ bond r_1. The result of this is to produce a diagram such as Fig. 1(*c*) where all the points lie in one asymmetric unit. If the distortion is normal and few points border the edge of another asymmetric unit, this approach is the most appropriate, but where the distribution is similar to Fig. 1(*c*) there are several disadvantages. The symmetry of the system is not explicitly considered and the distribution is most unlikely to be normal. The mean of the distribution will not correspond to a symmetrical configuration and will give a false idea of the most favoured geometry. Moreover, in more complicated systems it may be impossible to produce a unique labelling scheme. In a tetrahedral molecule it might be reasonable to label the atoms such that $r_1 \le r_2 \le r_3 \le r_4$ or that a_{12} was the largest angle (*etc.*). In any particular molecule it is possible that both of these conditions cannot be met and in general a satisfactory unique labelling scheme is not available.

The third approach involves recognizing the labelling problem explicitly. For the example given, Fig. 1(*b*) or (*c*) can be reflected in the diagonal mirror line to give Fig. 1(*d*), which has twice the number of points. Only if the resulting distribution is unimodal is this method applicable. Obviously the mean of the distribution must lie on $r_i = r_j$* and the axes of the distribution must be at 45° to the parameter axes. This application of symmetry can be described in two isomorphic ways. We can think of a probability distribution or scattergram constructed from the original data and *subsequently* acted on by the operations of an n-dimensional point group (usually n will be 2 or 3 but it could be higher). Alternatively we can apply the operations of a permutation group (in this case S_2) to the data immediately and construct a scattergram or probability distribution from the transformed data. We shall show the formalism of the permutational approach by considering the XYX example.

Fig. 1. (*a*) Parameters describing a linear $X-Y-X$ fragment. (*b*) Hypothetical scattergram of r_1 and r_2 for a linear XYX fragment whose atoms have been randomly labelled 1 and 2 for each individual point. (*c*) Hypothetical scattergram of r_1 and r_2 for the same data as (*b*) but where the atoms have been relabelled so that $r_1 \ge r_2$. Note that the distribution can be reflected about the line $r_1 = r_2$ and fills one asymmetric unit. Note also that the mean of the distribution shown is not on the diagonal line $r_1 = r_2$ and that normality is not well obeyed for the points plotted. (*d*) Hypothetical scattergram of r_i with r_j using the same raw data as (*b*) and (*c*) where the atom labels have been permuted ($i = 1$, $j = 2$; $i = 2, j = 1$). There are twice as many points plotted as in (*b*) and (*c*) and there is an exact line of symmetry along the diagonal. This diagram can be produced by reflecting either (*b*) or (*c*) about this line.

The dispersion matrix, **A**, for a normal bivariate distribution with m data for parameters p_1 and p_2 (where p_1 and p_2 are not potentially interchangeable, i.e. it is meaningless to permute their labels) has the form:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{6}$$

where

$$a_{11} = \sum_k^m (p_{1k} - \mu_1)^2,$$

$$a_{12} = a_{21} = \sum_k^m (p_{1k} - \mu_1)(p_{2k} - \mu_2)$$

and

$$a_{22} = \sum_k^m (p_{2k} - \mu_2)^2. \tag{7}$$

The covariances a_{12} and a_{21} are always equal but in general the variances a_{11} and a_{22} are not. The eigenvectors of this matrix describe axes whose angles to the coordinate axes p_1 and p_2 are not determined by symmetry. If we apply the operations of the permutation group $\mathbf{S}_2$ ($i = 1, j = 2$; $i = 2, j = 1$) to the labels, the number of terms contributing to the variances and covariances is doubled. The new values are:†

$$a_{ii}^* = \sum_k^{2m} (p_{ik} - \mu_i)^2, \tag{8}$$

$$a_{ij}^* = \sum_k^{2m} (p_{ik} - \mu_i)(p_{jk} - \mu_i) \tag{9}$$

and

$$\mu_i^* = \frac{1}{2m} \sum_k^{2m} (p_{ik}). \tag{10}$$

We can then show by standard methods that the diagonalization of the new covariance matrix, **A*** is:

$$\begin{bmatrix} a_{ii}^* & a_{ij}^* \\ a_{ij}^* & a_{ii}^* \end{bmatrix}$$

$$= \begin{bmatrix} 2^{-1/2} & -2^{-1/2} \\ 2^{-1/2} & 2^{-1/2} \end{bmatrix} \begin{bmatrix} a_{ii}^* + a_{ij}^* & 0 \\ 0 & a_{ii}^* + a_{ij}^* \end{bmatrix} \begin{bmatrix} 2^{-1/2} & 2^{-1/2} \\ -2^{-1/2} & 2^{-1/2} \end{bmatrix} \tag{11}$$

or

$$\mathbf{A}^* = \mathbf{E}^{*T} \boldsymbol{\lambda}^* \mathbf{E}^* = (\mathbf{E}^{*T} \boldsymbol{\lambda}^{*1/2})(\boldsymbol{\lambda}^{*1/2} \mathbf{E}^*) = \mathbf{F}^{*T} \mathbf{F}^*. \tag{12}$$

---

† All statistical quantities (including matrices) derived *after* the operation of a permutation group on the labelling will be marked with asterisks. (The transpose of a matrix **M** will be denoted by **M**^{T}, its inverse by **M**$^{-1}$.)

The eigenvectors now represent a pair of lines at 45° to the parameter axes p_i and p_j. The factors (**F**) are the major and minor axes of the elliptical distribution.

[When z scores are used a false similarity to (11) and (12) can occur in the bivariate case. Even where the permutation operations do not apply the variances of both parameters are equal since they are normalized to unity. The eigenvectors of the correlation matrix must therefore be at 45° to z_1 and z_2. However, when the factors are retransformed by (2) the scales are affected differently and the false symmetry disappears from the distribution and scattergrams. In multivariate distributions of z scores there is normally no false symmetry.]

The general problem is therefore to determine how the operation of permuting atom labels before computing variances and covariances affects the symmetry of the eigenvector matrix **E*** and thus the orientation of the principal components. Fortunately the mathematical treatment is isomorphic to the application of symmetry relationships to normal vibrations of molecules.

Symmetry and normal vibrations

At the level of the harmonic approximation the potential energy (V) of a molecule (relative to the equilibrium configuration) can be represented by the general quadratic form:

$$2V = \mathbf{x}^T \mathbf{B} \mathbf{x}. \tag{13}$$

Here **B** is a symmetric matrix of force constants,† and **x** is a vector representing the distortion of the molecule from its equilibrium geometry. Two-coordinate systems are generally used for **x**: internal coordinates, **d**, corresponding to changes in bond lengths and angles, etc.; and symmetry coordinates, **s**, which are orthogonal linear combinations of the **d** transforming as the irreducible representations (IR's) of G, the point group corresponding to the symmetry of the equilibrium configuration. An account of the use of symmetry coordinates in describing the geometry and energy of molecules distorted from a reference symmetry G is given in Murray-Rust, Bürgi & Dunitz (1978a,b, 1979) (MBDa,b,c).

In the general valence force field the potential energy is represented by:

$$2V = \mathbf{d}^T \mathbf{K} \mathbf{d} \tag{14}$$

where the matrix of force constants, **K**, contains interaction terms k_{ij} for every pair of parameters p_i and p_j, and for symmetrical molecules many of the k_{ij} will necessarily be identical. Thus in a tetrahedron the

---

† The symbols used may be different from those in standard textbooks since otherwise there is confusion between factors and force constants, etc.

interaction constants $k(r_1,\alpha_{12})$ and $k(r_1,\alpha_{13})$ will be the same and in fact there are twelve such equalities of this type; see Appendix. Where symmetry is present **K** can be factorized to give an alternative expression for the energy

$$2V = \mathbf{s}^T \mathbf{K}' \mathbf{s} \qquad (15)$$

where **K**' has a blocked form, the blocks corresponding to IR's of G. (14) and (15) describe the same potential-energy function (with the same hyperellipsoidal contours) but referred to two different sets of orthogonal axes. These are related by a rotation matrix **U**:

$$\mathbf{s} = \mathbf{U}\mathbf{d} \qquad (16)$$

where the elements u_{ij} are partly, or sometimes completely, defined by the symmetry G.

In general, therefore, the matrix of force constants **K**' can be diagonalized to give an equation of the form of (12), from which the eigenvectors can be uniquely determined. There is, however, a problem where the group G has high symmetry and some of the IR's are degenerate. Some of the eigenvalues are equal and there is some freedom in how the eigenvectors can be written. This can be seen in the different representations of normal coordinates used by different authors for the same normal mode. The problem is discussed in McDowell (1965) and at length in MBDa, MBDc. For an n-tuply degenerate IR there are $n - 1$ rotational degrees of freedom in choosing the representation of the eigenvectors, and for some particular orientations of the vectors, the normal coordinate has higher molecular symmetry than in general. These higher symmetries are called cokernel symmetries for the particular IR. Thus the bond vibration of a regular tetrahedron transforming as T_2 is described by some authors as a lengthening of two bonds and a shortening of the others (cokernel symmetry C_{2v}) and by others as a lengthening of one bond and a shortening of the other three (cokernel symmetry C_{3v}). These two descriptions simply correspond to a different choice of basis vectors and several other choices are also possible. In the present context we need simply note that the eigenvectors for a degenerate IR can be chosen in different ways and an example of this is given in the Appendix.

Symmetry and factor analysis

Factor analysis of static distortions of a molecule from a symmetrical mean configuration follows in an isomorphic manner. The probability distribution can be represented either by valence parameters [$cf.$ (14)]:

$$-2 \ln P(\mathbf{d}) = \mathbf{d}^T(\mathbf{A}^{*-1})\mathbf{d} \qquad (17)$$

or by symmetry coordinates:

$$-2 \ln P(\mathbf{s}) = \mathbf{s}^T(\mathbf{C}^{-1})\mathbf{s} \qquad (18)$$

where **C** has a blocked form exactly analogous to **K**' in (15). Its elements are the covariances between symmetry coordinates **s**. In some cases the explicit use of (18) is useful and is exemplified by the description of a distorted tetrahedron in MBDb. Since (17) and (18) differ only in the choice of axes, eigenvectors and eigenvalues of **A*** will also show the symmetry required by G and (17) is often easier to use in practice. The symmetry of G is manifested in **A*** by label permutations which exactly parallel the equality of force constants in **K** (14).

Factor analysis can be carried out with either **A*** or **C** or, more conveniently, the corresponding correlation matrices. If no representations are degenerate there will be a unique representation for each factor. Factors will occur in several blocks (if several IR's are involved) and the particular IR for each factor can easily be worked out by referring to character tables. The same principles apply when degenerate representations are involved, where factors occur in pairs (E) or triplets (T). The total variance due to a degenerate factor is found by summing the two or three equal eigenvalues. The form of the eigenvectors is indeterminate and if determined automatically will depend on the idiosyncrasy of the computer and its algorithm. By suitable choice of a rotation matrix the factors can always be rotated to give eigenvectors corresponding to cokernel symmetries (see Appendix).

Even without the problems of cokernel symmetry, factor axes can now be rotated to simplify the coefficients in each factor (see MB and MMb) but this requires that rotation occurs only $within$ blocks. The procedure cannot be carried out on **A*** and so **C** must be used instead, one block at a time. An example is given in the Appendix but the warning against carefree use of factor rotation (MB) is even more relevant to symmetrical matrices.

The relationship of probability density and energy

It has been suggested (Bürgi, Dunitz & Shefter, 1973; Bürgi, 1975; Dunitz, 1975; Murray-Rust, Bürgi & Dunitz, 1975) that the probability distribution of geometrical parameters can give information about the potential-energy surface of a molecular fragment. Most relevantly, in MBDb it was shown that the symmetry properties of the probability distribution of bond lengths and angles of distorted tetrahedra were experimentally the same as those expected for the energy surface of a tetrahedral molecule. Here we shall merely note the very close mathematical relationship between probability and energy.

The harmonic approximation to the potential energy of an isolated molecule is given by (14). At a temperature T an assemblage of molecules will classi-

cally have the potential energy distributed according to the Boltzmann distribution:

$$P(V) = A \exp(-V/kT),\qquad(19)$$

where $P(V)$ is the probability of finding the molecule with energy V and A is a normalizing constant. Substituting (14) in (19) we get

$$P(V) = A \exp(-\tfrac{1}{2}\mathbf{d}^T \mathbf{Kd}/kT)\qquad(20)$$

or

$$\ln P(V) = \ln A - \mathbf{d}^T \mathbf{Kd}/2kT.\qquad(21)$$

This has exactly the form of the normal distribution (1) and if the classical approximation holds we should expect the geometries of gaseous molecules vibrating harmonically to be distributed as (20) at any instant in time. It is less clear whether the imposition of crystal packing forces in a random manner on a molecular fragment will necessarily give a normal distribution although there is evidence (MBDb) that this can happen.

APPENDIX

An example

The distortions of tetrahedral molecules provide an excellent illustration of the factor analysis of a highly symmetrical distribution. The use of symmetry coordinates has been presented at length (MBDa, MBDb) and here we shall show how the same raw data used by MBD for distorted PO_4 groups can be factor-analysed. The two approaches are complementary in that factor analysis shows clearly what proportion of the variance is attributable to any symmetry species but it does not detect deviations from normal distribution, either multimodality or non-linearity.

The point group T_d is isomorphic with the permutation group S_4 (of order 24) which must be applied to the data before calculating the dispersion matrix $\mathbf{A}^*$. Thus instead of 211 data being used, 24×211 are used, corresponding to the permutations $(i = 1, j = 2, k = 3, l = 4)$, $(i = 2, j = 1, k = 3, l = 4)$, ... etc., where i, j, k, l are the labels of the four O atoms. The formulae for calculation of the variances and covariances are given in Table 1. (In practice the permutations were applied to the bond lengths and angles and a set of 5064 data was analysed by the normal method.) The corresponding correlation matrix, $\mathbf{R}^*$, is easily calculated and given in Table 2 for the bond lengths and angles. It has, of course, the same symmetry as $\mathbf{A}^*$.

The eigenvalues and eigenvectors of $\mathbf{R}^*$ are given in Table 2 exactly as they came from the computer (*SPSS* program running on an ICL 4130). One factor is identically zero, corresponding to the redundant A_1

Table 1. *Dispersion matrix* $\mathbf{A}^*$ *for the geometrical parameters (lengths and angles) of a tetrahedral* MX_4 *molecule*

The lengths are labelled r_1 to r_4 and the angles α_{12} to α_{34}. There are seven distinct matrix elements for the variances and covariances. For the bond-length distortions we write:

$$\Delta r_i = r_i - r_0$$

where r_0 is the mean bond length and for the angles we have:

$$\Delta\alpha_{ij} = \alpha_{ij} - 109.47°.$$

Then for n observed tetrahedra:

$$A = 6 \sum_i^n \Delta r_i^2, \quad B = 4 \sum_i \sum_{j>i}^n \Delta\alpha_{ij}^2,$$

$$C = 4 \sum_i \sum_{j>i}^n \Delta r_i \Delta r_j, \quad D = 2 \sum_i \sum_{j\neq i}^n \Delta r_i \Delta\alpha_{ij},$$

$$E = 2 \sum_i \sum_j \sum_k^n \Delta r_i \Delta\alpha_{jk}, \quad F = 2 \sum_i \sum_j \sum_k^n \Delta\alpha_{ij} \Delta\alpha_{ik},$$

$$G = 8 \sum_i \sum_{j\neq i}^n \Delta\alpha_{ij} \Delta\alpha_{kl}$$

represent the variance and covariance terms where the first summation is taken over the unique (*i.e.* unpermuted) data set. The dispersion matrix is symmetrical and here the lower triangle is shown:

	r_1	r_2	r_3	r_4	α_{12}	α_{13}	α_{14}	α_{23}	α_{24}	α_{34}
r_1	A									
r_2	C	A								
r_3	C	C	A							
r_4	C	C	C	A						
α_{12}	D	D	B	B	E					
α_{13}	D	E	D	E	F	B				
α_{14}	D	E	E	D	F	F	B			
α_{23}	E	D	D	E	F	F	G	B		
α_{24}	E	D	E	D	F	G	F	F	B	
α_{34}	E	E	D	D	G	F	F	F	F	B

angle coordinate describing the non-independence of the six angles. The other nine factors can be seen to be made up of two different triply-degenerate factors (1–3, 7–9), one doubly-degenerate factor (4–5) and one non-degenerate factor (6). These are identified by standard methods as:

$$2T_2 + E + A_1.$$

The second T_2 factor (7–9) accounting for only 4% of the total variance is not experimentally significant and corresponds to errors of various sorts. The A_1 factor (6) is also very small and shows that the average bond length changes very little (see MBDb). By symmetry it cannot account for any angular variance. The E factor similarly cannot account for any variance in bond lengths but describes 17% of the angular variation. The eigenvectors have (coincidentally) come out in a simple

Table 2. *Factor analysis of 211 distorted phosphate* (PO_4) *groups; data taken from Baur's* (1974) *compilation*

The following quantities were available for each case: bond lengths (r_i), bond angles (a_{ij}), and the average e.s.d. for bonds and angles (author's estimates), $\sigma(r)$ and $\sigma(a)$.

(a) The following statistical quantities were derived:

	Mean	s.d.	Variance	Mean e.s.d.	(Mean e.s.d.)²
r_i	1·537 Å	0·042 Å	0·0018 Å²	0·007 Å	0·00005 Å²
a_{ij}	109·35°	4·12°	17·0 (°)²	0·37°	0·14 (°)²

(Note that the estimated variance only accounts for about 2–3% of the observed variance.)

(b) The elements of the correlation matrix R^* (formed after permuting all lengths and angles by the symmetry operations of S_4). The symmetry of the matrix is given in Table 1 and the elements were found to be:

$$A = B = 1; C = 0·27; D = -0·47; E = 0·46; F = -0·09; G = -0·65.$$

(c) Eigenvalues and eigenvectors (for the first six eigenvalues only) of R^*. The columns are as follows: the serial number of the eigenvalue; the IR of T_d appropriate to the eigenvector; the eigenvalue (λ); the percentage of total variance (PTV) explained by each factor; the cumulative percentage of total variance (CTV); the cumulative percentage of bond variance (CBV); the cumulative percentage of angle variance (CAV); the coefficients ($\times 100$) of the bonds in each factor (r_i); the coefficients ($\times 100$) of the angles in each factor (a_{ij}).

Factor	IR	λ	PTV	CTV	CBV	CAV	r_1	r_2	r_3	r_4	a_{12}	a_{13}	a_{14}	a_{23}	a_{24}	a_{34}
1	T_2	2·78	27·8				32	−84	−9	61	43	−19	−76	76	19	−43
2	T_2	2·78	27·8				7	−43	87	−50	29	−76	36	−36	76	−29
3	T_2	2·78	27·8	83·6	89·8	79·6	−89	0	37	52	72	42	30	−30	−42	−72
4	E	0·52	5·2				0	0	0	0	40	−30	−10	−10	−30	40
5	E	0·52	5·2	94·2	89·8	97·0	0	0	0	0	10	30	−40	−40	30	10
6	A_1	0·19	1·9	96·1	94·4	97·0	−21	−21	−21	−21	0	0	0	0	0	0
7	T_2	0·13	1·3													
8	T_2	0·13	1·3													
9	T_2	0·13	1·3	100·0	100·0	100										
10	A_1	0·00	0	100	100	100										

(d) Rotation of the first three factors (T_2) according to the varimax criterion

Rotation matrix (elements $\times 100$)

$$\begin{bmatrix} 85 & 21 & 48 \\ -40 & 85 & 33 \\ -34 & -47 & 81 \end{bmatrix}$$

Rotated factors ($\times 100$)

	r_1	r_2	r_3	r_4	a_{12}	a_{13}	a_{14}	a_{23}	a_{24}	a_{34}
	55	−55	−55	55	0	0	−89	89	0	0
	55	−55	55	−55	0	−89	0	0	89	0
	−55	−55	55	55	89	0	0	0	0	−89

Note that the factors now all show cokernel C_{2v} symmetry.

(e) Rotation of factors 4–5 (E) to show cokernel symmetry

Rotation matrix

$$\begin{bmatrix} 2^{-1/2} & -2^{-1/2} \\ 2^{-1/2} & 2^{-1/2} \end{bmatrix}$$

Rotated factors ($\times 100$)

a_{12}	a_{13}	a_{14}	a_{23}	a_{24}	a_{34}
35	0	−35	−35	0	35
−21	42	−21	−21	42	−21

Note that the second of the rotated factors shows the cokernel D_{2d} symmetry, but that the first only shows D_2 (kernel) symmetry.

form; they are both at 45° to a symmetry line in the deformation space. Rotation by 45° gives more symmetrical formulations for these two vectors, the second one of which now corresponds to a distortion with D_{2d} symmetry, the cokernel of $E(T_d)$.

The T_2 factor is the most interesting in that it involves both bond and angle deformation and gives rise to the correlations found by MBD*b* between the two T_2 symmetry coordinates, S_3 and S_4. The eigenvectors initially showed no symmetry but after

rotation (in this case carried out by the Kaiser varimax procedure) the three eigenvectors all corresponded to distortions with C_{2v} cokernel symmetry. The relation between bond and angle deformations is easily seen from the appropriate components of the factors. Thus as r_1 and r_4 increase by 0·55 standard deviations (bond), r_2 and r_3 decrease by the same amount, α_{14} decreases by 0·89 standard deviations (angle) and α_{23} increases by the same amount. Multiplying these quantities by the observed standard deviations (Table 2a) we find that for a change of 1° in α_{14} there is a corresponding decrease in bond length r_1 of 0·0063 Å, very similar to the results of MBDb from correlation of symmetry coordinates.

References

BAUR, W. H. (1974). *Acta Cryst.* B30, 1195–1215.

BÜRGI, H.-B. (1975). *Angew. Chem. Ind. Ed. Engl.* 14, 460–473.

BÜRGI, H.-B., DUNITZ, J. D. & SHEFTER, E. (1973). *J. Am. Chem. Soc.* 95, 5066–5067.

DUNITZ, J. D. (1975). *Proc. R. Soc. London Ser. B,* 272, 99–108.

MCDOWELL, R. S. (1965). *J. Mol. Spectrosc.* 17, 365–367.

MURRAY-RUST, P. & BLAND, R. (1978). *Acta Cryst.* B34, 2527–2533.

MURRAY-RUST, P., BÜRGI, H.-B. & DUNITZ, J. D. (1975). *J. Am. Chem. Soc.* 97, 921–922.

MURRAY-RUST, P., BÜRGI, H.-B. & DUNITZ, J. D. (1978a). *Acta Cryst.* B34, 1787–1793.

MURRAY-RUST, P., BÜRGI, H.-B. & DUNITZ, J. D. (1978b). *Acta Cryst.* B34, 1793–1803.

MURRAY-RUST, P., BÜRGI, H.-B. & DUNITZ, J. D. (1979). *Acta Cryst.* A35, 703–713.

MURRAY-RUST, P. & MOTHERWELL, W. D. S. (1978a). *Acta Cryst.* B34, 2518–2526.

MURRAY-RUST, P. & MOTHERWELL, W. D. S. (1978b). *Acta Cryst.* B34, 2534–2546.

WILSON, E. B., DECIUS, J. C. & CROSS, P. C. (1955). *Molecular Vibrations.* New York: McGraw-Hill.

Neutron Diffraction Study of the Crystallographic and Magnetic Structures of Potassium Tribromoferrate(II)

BY E. GUREWITZ

Nuclear Research Centre–Negev, POB 9001, Beer Sheva 84190, Israel

AND H. SHAKED

Nuclear Research Centre–Negev, and Ben-Gurion University of the Negev, POB 653, Beer Sheva 84120, Israel

Abstract

A neutron diffraction study of a powder sample of $KFeBr_3$ was carried out at various temperatures. (Weighted R factors are 0·076, 0·075 for 26, 31 intensities measured at room temperature and liquid-helium temperature, respectively.) This compound was found to be isostructural with $KFeCl_3$ and belongs to the orthorhombic space group *Pnma* with four molecules per unit cell. It is paramagnetic at room temperature and undergoes a transition to a magnetically ordered state at $T_N \sim 9 \cdot 5$ K. The magnetic structure as determined from diffraction patterns at 4·2 K consists of antiferromagnetically coupled ferromagnetic chains parallel to **b**. The antiferromagnetic axis is along **b** and the magnetic moment per Fe^{2+} ion is $3 \cdot 7 \pm 0 \cdot 2$ BM (1 BM $\equiv 9 \cdot 27 \times 10^{-24}$ J T^{-1}). The temperature dependence of the magnetic reflections shows some residual coherent reflections above T_N. This is interpreted in terms of strong one-dimensional intra-chain correlations.

I. Introduction

Most of the ABX_3 compounds, where A is an alkaline metal, B a transition metal and X a halogen or O, have crystallographic structures which are derived from either the ideal cubic perovskite or the hexagonal perovskite structures. However, some ABX_3 compounds have different structures. For example, the structure of $KCdCl_3$ (Wyckoff, 1964) cannot be obtained from the cubic or the hexagonal perovskite-like structures by a series of continuous distortions.

The compounds $KFeCl_3$ and $KFeBr_3$ are isostructural with $KCdCl_3$ (Gurewitz, Makovsky &