No theoretical justification has been made, on the other hand, for the use of the pseudo-normalized structure factors $E'$ proposed by Karle & Karle (1966).

References


The Phase Problem and its Implications in the Least-squares Refinement of Crystal Structures

BY A.D. RAE

Department of Physical Chemistry, The University of New South Wales, Kensington, N.S.W. 2033, Australia

(Received 16 April 1974; accepted 6 May 1974)

It is found that covariance problems encountered in pseudosymmetric crystal structures are caused by an incorrect use of the least-squares refinement procedure. Rederivation of the least-squares equations for the situation in which the residual need not have the phase angle $\sigma$ or $\sigma + \pi$, where $\sigma$ is the phase angle associated with a trial structure, reveals that the minimization of the component at $\pi/2$ to $\sigma$ has been omitted from the least-squares equations. Inclusion of the extra terms associated with this minimization reveals that it should now be possible to refine a centrosymmetric crystal in an non-centrosymmetric space group. It is also shown that the use of weights derived from counting statistics alone is incorrect and, with a correct weighting scheme, $\sum w^2(n-m)$ should reduce to one in a single cycle. The weighting scheme is re-evaluated for further refinement cycles.

Introduction

For the least-squares refinement of single-crystal structures from X-ray diffraction data it has been customary (International Tables for X-ray Crystallography, 1959) to minimize the sum of the $n$ weighted squared residuals $S_1 = \sum w_n(F_{0\alpha h} - |F_{\alpha h}|)^2$, where $w_n^{-1} = \text{var}(|F_{0\alpha h}|)$, by solving the equations

$$\sum_j \sum_n w_n \left( \frac{\partial A}{\partial u_j} \right)_{\alpha h} \left( \frac{\partial A}{\partial u_j} \right)_{\alpha h} [u_j - (u_j)_0] = -\sum_n w_n A_{\alpha h} \left( \frac{\partial A}{\partial u_j} \right)_{\alpha h} ; \quad i = 1 \text{ to } m,$$

where $A = |F_{\alpha}| - |F_{\alpha}|$, the difference between the measured and calculated amplitudes. Throughout this paper the subscript $h$ implies the $h$th observable and the subscript 0 implies evaluation with parameters $(u_j)_0$ of a trial structure. Now for $A = |F_{\alpha}| - |F_{\alpha}|$ and tan $\alpha = (B_{\alpha}/A_{\alpha})_0$ we use $(\partial A/\partial u_j)_0 = -(\partial |F_{\alpha}|/\partial u_j)_0 = -[\cos \alpha_0(\partial A_{\alpha}/\partial u_j)_0 + \sin \alpha_0(\partial B_{\alpha}/\partial u_j)_0]$. The application of this procedure reveals two apparent faults. Firstly the assumption that $w_n^{-1} = \text{var}(|F_{0\alpha h}|)$ because $(F_{\alpha})_0$ is without error does not produce the expected result that $\sum w_n A^2 = n-m$ in a single cycle. Secondly, in pseudosymmetric structures the apparent variances of parameters $u_j$ are usually in excess of calculated variances (Rae, 1973). This suggests that both the weighting scheme and the actual least-squares equations are at fault. Investigation shows that this hypothesis is indeed true and that the situation may be remedied.

Theory

The least-squares equations

Account should be taken of the fact that we are dealing with quantities $F = A + iB$ that do not have a fixed phase. If $(F_{\alpha})_0$ is an initial estimate of the phased quantity $F_{\alpha}$ for which only the magnitude $|F_{\alpha}|$ has been experimentally determined, then it is found that a different set of least-squares equations are generated by
THE PHASE PROBLEM AND ITS IMPLICATIONS

considering the minimization of \( \sum w_h \Delta_h \) where \( \Delta_h \) is the phased quantity \( (F_o)_{oh} - (F_c)_{oh} \) rather than the unphased quantity \( |(F_o)_{oh} - |(F_c)_{oh}| \). These equations are seen to minimize \( S_1 + S_2 \) rather than \( S_1 \) where

\[
S_1 = \sum w_h (\Delta_o \sin \alpha_0 + (B_o - B_c) \cos \alpha_0)^2,
\]

\[
S_2 = \sum w_h (\Delta_o \cos \alpha_0 - (B_o - B_c) \sin \alpha_0)^2.
\]

and

\[
\tan \alpha_0 = (B_c/A_o)_{oh}.
\]

If \( (F_o)_{oh} \) and \( (F_c)_{oh} \) are assumed to have the same phase \( \alpha_0 \), then

\[
B_i = \sum (\cos \alpha_{oh} \partial A_c/\partial u_i)_{oh} + \sin \alpha_{oh} \partial B_c/\partial u_i)_{oh} w_h ((F_{oh} - |F_{coh}|).
\]

The weighting scheme

To evaluate the variance–covariance matrix of real parameters \( \hat{u}_i \) we need to evaluate \( M_{ij} = \sum (A^{-1})_{ik} \bar{B}_{ki} \) without error only if the estimate of the true value \( F_h \) given by

\[
F_h = (F_c)_{oh} + \sum a_{hi} \hat{u}_j - (u_j)_{oh}
\]

is used to evaluate \( \bar{B}_{ki} \), for then we may say

\[
(F_{oh}) - F_h = \hat{A}_h + \sum a_{hi} \hat{u}_j - (u_j)_{oh}
\]

so that

\[
\bar{B}_{ki} = \sum (a_{hi} w_h \Delta_h + a_{hi} w_h \Delta_c)^2
\]

and

\[
M_{ij} = \sum (A^{-1})_{ik} \langle \bar{B}_k \bar{B}_h \rangle (A^{-1})_{kj}.
\]

However we do not know \( \hat{u}_j \) and the estimation of \( M_{ij} \) as \( (A^{-1})_{ik} \sigma^2 \) implies that \( \langle \bar{B}_k \bar{B}_h \rangle = A_{ih} \sigma^2 \). If \( \sigma^2 \) is chosen to be \( \sum w_h \Delta_h^2 A_h/(n-m) \) then we imply that the quantities \( w^2_h [(F_{oh}) - F_h] \) all belong to the same distribution with variance \( \sigma^2 \), but imply nothing about the form of this distribution and so significance levels can only be obtained by experimental determination of the actual \( w^2_h [(F_{oh}) - F_h] \) values. However if weights of \( \langle (F_{oh}) - F_h \rangle \) are estimated and used and \( \sum w_h \Delta_h^2 \Delta_c = \sum w_h \Delta_h^2 \Delta_c + 2 \sum B_i (\hat{u}_i - (u_i)_{oh}) \) does not equal \( n-m \) then we can say that the estimates of \( \langle (F_{oh}) - F_h \rangle \) are wrong.

The rigorous derivation of the correct least-squares equations

A derivation of the correct least-squares equations to solve can be obtained by minimizing the variance
in a function \( f - f_0 = \sum_j d_j[u_j - (u_\ell)] \) where \( d_j = (\partial f/\partial u_\ell)_0 \).

If \( f \) is the true value of \( f \) using the true values \( u_\ell \), and \( \hat{f} \) is the value of \( f \) using parameters \( \hat{u}_\ell \) obtained from least-squares equations involving residuals, \( D_h = (F_\ell)_{oh} - F_h \), then

\[
\hat{f} - f = \sum_j d_j(\hat{u}_j - u_\ell) = \sum_h C_h D_h = \sum_{hj} C_{h\ell} a_{h\ell}(\hat{u}_j - u_\ell)
\]

where

\[
a_{h\ell} = \left( \frac{\partial F_\ell}{\partial u_\ell} \right)_{oh}.
\]

\((F_\ell)_{oh} \) is the observed estimate of the true quantity \( F_\ell \) for which calculated estimates are made using \( \hat{F}_\ell = (F_\ell)_{oh} + \sum a_{h\ell}(\hat{u}_j - u_\ell) \). \( F_h \) is the estimate of \( F_\ell \) using \( F_h = (F_\ell)_{oh} + \sum a_{h\ell}(\hat{u}_j - u_\ell) \) since only then can we say \( \sum C_{h\ell}(F_\ell)_{oh} - F_h \) exactly. The \( C_{h\ell} \) coefficients are not unique if \( n \), the number of \( (F_\ell)_{oh} \) values, exceeds \( m \), the number of \( u_j \) parameters.

The variance of \( \hat{f} \) is \( \text{var}(\hat{f}) = \langle (\hat{f} - f)^2 \rangle = \langle \sum_{hj} (D_{h\ell}^* D_{h\ell} C_{h\ell} C_{h\ell}) \rangle \). It is customary (Hamilton, 1964) at this stage to continue assuming that the problem is a real-number problem, but this assumes that \( D_h \) can only have a phase of \( \alpha \) or \( \alpha + \pi \), so that only the value of \( D_h \) in the phase direction of \( (F_\ell)_{oh} \) need be considered. However it is beneficial to assume \( D_h \) can have any phase and we shall continue using components parallel (subscript 1) and at \( \pi/2 \) (subscript 2) to an arbitrary direction.

To find the values of \( C_{h\ell} \) we minimize

\[
\text{var}(\hat{f}) - \sum_j (\lambda_j \sum_{h\ell} C_{h\ell} a_{h\ell} + \lambda_{j*} \sum_{h\ell} C_{h\ell}^* a_{h\ell}^*) + \sum_j (\lambda_j d_j + \lambda_{j*} d_j^*)
\]

Since \( d_j = \sum_{h\ell} C_{h\ell} a_{h\ell} \) we are still only minimizing \( \text{var}(\hat{f}) \), but the Lagrange multipliers \( \lambda_j \) enable the determination of the \( C_{h\ell} \) values under the constraint \( d_j = \sum_{h\ell} C_{h\ell} a_{h\ell} \).

Great simplification is achieved if we choose a scale for our calculations of \( (F_h)_{oh} \) either by adding a constant to all \( (F_\ell)_{oh} \) or by multiplying all \( (F_\ell)_{oh} \) by a constant depending on the type of problem, so that \( \langle \sum_{h1 h2} (D_{h1}^* D_{h2} C_{h1} C_{h2}) \rangle = 0 \) making \( \text{var}(\hat{f}) = \langle \sum_{h} |D_h|^2 \rangle C_{h\ell}^2 \). This will eliminate any correlation between observations and avoid the possibility of having to minimize sums of the form \( \sum_{h1 h2} D_{h1}^* D_{h2} A_{h1} A_{h2} \). For example the \( n \) real numbers \( x_i \) can be put on such a scale to minimize the variance of \( \sum (x_i - \bar{x})^2 \) by adding \( k - \bar{x} \) where \( \bar{x} = \sum_i x_i/n \).

Then \( \sum_{i \neq j} (x_i + k - \bar{x}) (x_j + k - \bar{x}) = n(n-1)k^2 - \sum_i (x_i - \bar{x})^2 = 0 \) if \( k = \pm \sqrt{\sum_i (x_i - \bar{x})^2/n(n-1)} \), i.e. the real numbers \( x_i \) are put on scale by making \( \sum_i x_i = \sum_i \bar{x} \). Any dependence of a value for an observation upon any other observation should be included in the model for calculating \( (F_\ell)_{oh} \) by including parameters \( u_j \) whose coefficients \( a_{h\ell} \) depend on \( (F_\ell)_{oh} \) values.

A method of notation will be adopted henceforth in which the appropriate subscripts are included in order in brackets after the terms to which the subscripts apply. For example the real part of \( \lambda_j C_{h\ell} a_{h\ell} = (\lambda_j)(C_{h\ell})(a_{h\ell})_1 - (\lambda_j)_2 (C_{h\ell})_2 (a_{h\ell})_2 - (\lambda_j)_3 (C_{h\ell})_3 (a_{h\ell})_3 \) will be written as \( \lambda_j C_{h\ell} a_{h\ell} (111-222-212) \).

From \( \frac{\partial \text{var}(\hat{f})}{\partial C_{h1}} = 0 \) we obtain

\[
\langle |D_h|^2 \rangle (C_{h1}) = \sum_j \lambda_j a_{h\ell} (11 - 22).
\]

From \( \frac{\partial \text{var}(\hat{f})}{\partial C_{h2}} = 0 \) we obtain

\[
\langle |D_h|^2 \rangle (C_{h2}) = -\sum_j \lambda_j a_{h\ell} (12 + 21).
\]

If it is assumed that

\[
\sum_j a_{hi} a_{h\ell}/\langle |D_h|^2 \rangle (12 - 21) = 0
\]

for any \( i, j \) it follows that

\[
\sum_h C_{h\ell} (11 - 22) = \sum A_{ij}(\lambda_j)_1
\]

and that

\[
-\sum_h C_{h\ell} (12 + 21) = \sum A_{ij}(\lambda_j)_2
\]

where

\[
A_{ij} = A_{ji} = \sum_h |D_h|^2 - \sum_h (a_{h\ell} a_{h\ell} + a_{hi} a_{hi})
\]

is independent of the phase chosen for evaluating components. \( A_{ij} \) and \( (A_{ij})_2 \) are real and consequently we do not associate component subscripts 1 and 2 with these quantities.

We now obtain

\[
(\lambda_j)_1 = \sum_{j h} C_{h\ell} a_{h\ell} (A^{-1})_{ij} (11 - 22)
\]

\[
(\lambda_j)_2 = -\sum_{j h} C_{h\ell} a_{h\ell} (A^{-1})_{ij} (12 + 21)
\]

and

\[
\langle |D_h|^2 \rangle (C_{h1}) = \sum_{j h} C_{h\ell} a_{h\ell} f(A^{-1})_{ij} a_{h\ell} (111 - 222 + 112 + 221 + 121 - 112 + 222)
\]

It follows that

\[
\sum_j (d_j)_j(\hat{u}_j - u_\ell) = \sum_h (C_h D_h)_1 = \sum_{j h} (d_j)_j (A^{-1})_{ij} \bar{B}_i
\]

and

\[
\sum_j (d_j)_j(\hat{u}_j - u_\ell) = \sum_h (C_h D_h)_2 = -\sum_{j h} C_h D_h (12 + 21)
\]

\[
= \sum_{j h} (d_j)_j (A^{-1})_{ij} \bar{B}_i
\]

\[
= \sum_{j h} (d_j)_j (A^{-1})_{ij} \bar{B}_i
\]
where
\[ (d_j)_1 = \sum_n (C_n a_{h,j} a_{h,j}) = \sum_n C_n a_{h,j} (11 - 22), \]
\[ (d_j)_2 = -\sum_n C_n a_{h,j} (12 + 21) \]
and
\[ B_i = \sum_h |D_h|^2 i a_{h,i} D_h (11 + 22) = \frac{1}{2} \sum_h |D_h|^2 i - (a^*_n D_h + a_n D^*_h). \]

The equations \( \sum_j d_j (\tilde{u}_j - \bar{u}_j) = \sum_j d_j (A^{-1})_j \tilde{B}_i \) are satisfied if \( \sum_j A_{ji} (\tilde{u}_j - \bar{u}_j) = \tilde{B}_i \) for \( i = 1 \) to \( m \) irrespective of the values of \( d_j \).

It can also be seen at this stage that the minimum value \( \text{var} (\tilde{f}) \) takes up is
\[ \sum_h [C_n a_{h,j} (11 - 22) (A^{-1})_j a_{n,h} C_h (11 - 22) + C_n a_{h,j} (12 + 21) (A^{-1})_j a_{n,h} C_h (12 + 21)] = \sum_{ij} d_j (A^{-1})_j d_i (11 + 22) = \sum_{ij} d_j^* (A^{-1})_j d_i. \]
Since \( \tilde{f} - f = \sum_j d_j (\tilde{u}_j - \bar{u}_j) \) and
\[ \text{var} (\tilde{f}) = \sum_{ij} d_j^* (\tilde{u}_j - \bar{u}_j) (\tilde{u}_i - \bar{u}_i) \]
it is seen that \( \text{cov} (\tilde{u}_j, \tilde{u}_i) = (A^{-1})_{ji} \) provided the parameters \( \bar{u}_j \) are obtained from the least-squares equations \( \sum_j A_{ji} (\tilde{u}_j - \bar{u}_j) = \tilde{B}_i \) where
\[ B_i = \frac{1}{2} \sum_h [a^*_n w_h (F_0 h - F_0 h) + a_n w_h (F_0 h - (F_0 h))^*], \]
\[ A_{ij} = \frac{1}{2} \sum_h [a^*_n w_h a_{h,j} + a_n w_h a_{j,h}] \] and \( w_{h^{-1}} = |D_h|^2 \).

To do this it is important to preserve the linearity of the residuals and this implies that \( D_h = (F_0 h - F_h) \) where \( F_h = (F_0 h + \sum a_{h,j} [u_j - (u_j)_0]) \) and is not necessarily the true value \( F_h \) for two reasons. Firstly the coefficients \( a_{h,j} \) are evaluated for the parameters \( (u_j)_0 \), not the parameters \( u_j \) and any non-linearity of the residuals causes a variation of the coefficient \( a_{h,j} \) with choice of \( (u_j)_0 \). Secondly it has been assumed that the model for calculation is correct with all relevant parameters \( u_j \) included. It is therefore wrong to say \( |D_h|^2 = \text{var} (F_0 h) \) as this assumes that there is no error in the phase of \( F_0 h \) and that there is no error in the model \( F_h = (F_0 h + \sum a_{h,j} [u_j - (u_j)_0]) \) for calculating \( F_h \) from the true parameters \( u_j \). It is equally wrong to say that any error not accounted for can be included by saying \( |D_h|^2 = \sigma^2 \text{var} (F_0 h) \) where \( \sigma^2 \) is a constant.

Now \( D_h = (F_0 h - F_h) + [F_h - F_h] - [F_0 h - F_h] \) where \( F_0 h - F_h = \tilde{A}_h = (F_0 h - (F_0 h) + \sum a_{h,j} [u_j - (u_j)_0]) \) and \( F_h - F_h = \sum a_{h,j} (\tilde{u}_j - \bar{u}_j) \). We readily see that if \( \tilde{F}_h - F_h = 0 \) and \( w_{h^{-1}} = |D_h|^2 \), then
\[ \sum_h w_h |D_h|^2 = n = \sum_h w_h |\tilde{A}_h|^2 + \sum_{h,ij} w_h a^*_n a_{h,j} M_{ij}. \]
Now the variance-covariance matrix \( M_{ij} = (A^{-1})_{ij} \) so that
\[ n = \sum h w_h |\tilde{A}_h|^2 + \sum_{ij} A_{ij} (A^{-1})_{ij} \]
giving \( \sum h w_h |\tilde{A}_h|^2 = n - m \) for an ideal weighting system.

Applied to the problem of obtaining the estimate \( \tilde{x} \) of the true value \( x \) from a number of observations \( x_h \) we see that if we have \( n \) equally weighted observations a consistent result is obtained for \( \tilde{x} = \sum x_h/n \), and
\[ w_{h^{-1}} = \sum (x_h - \tilde{x})^2/n + \sum (x_h - \tilde{x})^2/(n - 1) \]
where
\[ \sum (x_h - \tilde{x})^2/n \] is the estimate of \( (x_h - \tilde{x})^2 \)
and
\[ \sum (x_h - \tilde{x})^2/(n - 1) = \text{var} (\tilde{x}). \]
We see that \( w_{h^{-1}} = \sum (x_h - \tilde{x})^2/(n - 1) \) is the estimate of \( (x_h - \tilde{x})^2 \) and since \( a_{h,1} = 1 \) we also see that \( A_{1,1} = \sum w_h a_{h,1}^2 \) gives
\[ (A^{-1})_{1,1} = \sum (x_h - \tilde{x})^2/(n - 1) = \text{var} (\tilde{x}) = \langle (\tilde{x} - \bar{x})^2 \rangle \]
and that
\[ \sum_h w_h (x_h - \tilde{x})^2 = n - 1. \]

The estimation of \( w_{h^{-1}} = |D_h|^2 \)

We have seen that \( D_h = \tilde{A}_h + \sum a_{h,j} [u_j - (u_j)_0] = (F_0 h - F_h) \) where \( \tilde{A}_h = (F_0 h) - (F_0 h) - \sum a_{h,j} [u_j - (u_j)_0] \). We assume \( \langle u_j - (u_j)_0 \rangle = 0 \) which gives \( \langle D_h \rangle = \langle \tilde{A}_h \rangle + \sum a_{h,j} M_{ij} \) where \( M_{ij} \) is an estimate of the variance-covariance matrix. Now the approximate value of \( \langle |D_h|^2 \rangle \) is \( (n - m)/n \), so the accuracy of the estimate of \( M_{ij} \) will not matter too much if \( (n - m)/m \gg 1 \). An actual value of \( \tilde{A}_h \) can be obtained by iteration from the initial values \( (u_j)_0 \). However if the actual value of \( |D_h|^2 \) is used to evaluate \( \langle |D_h|^2 \rangle \) the refinement will iterate to the point where \( m \) values \( F_h \) have perfect agreement and only those other values of \( F_h \) which also perfectly agree with \( (F_0 h) \) will have a non-zero weight.

It is best to evaluate \( \langle |D_h|^2 \rangle \) by saying \( D_h = (F_0 h - F_h) = E_1 + E_2 + E_3 \) where \( E_1 = (F_0 h - (F_0 h) F_h) / |F_h| \), \( E_2 = (F_0 h - F_h) / |F_h| F_h \) and \( E_3 = F_h - F_h \) so that \( \langle |D_h|^2 \rangle = \langle |E_1|^2 + |E_2|^2 + |E_3|^2 \rangle \), since each of the terms \( E_1, E_2 \) and \( E_3 \) may be reasonably assumed to show no covariance. \( \langle |E_1|^2 \rangle \) is simply \( \text{var} (F_0 h) \), \( \langle |E_2|^2 \rangle \) is the variance associated with the choice of phase of \( (F_0 h) \) and \( E_3 \) is the systematic error associated with the model \( F_h = (F_0 h) + \sum a_{h,j} [u_j - (u_j)_0] \) to estimate \( \tilde{A}_h \). \( E_3 \) will have a contribution due to the variation of the coefficients \( a_{h,j} \) with choice of \( (u_j)_0 \) but it is also possible to have a contribution to \( E_3 \) due to the inaccuracy of the algebraic form for calculating \( (F_0 h) \). In X-ray crystal
structure analyses the omission of atoms, absorption and extinction corrections and the use of isotropic temperature factors and rigid groups of atoms all fall into this category.

The estimation of error in the phase of the observation

We can attribute a variance–covariance matrix to the components of \( (F_c)_{on} = (A \cdot)_{on} + i(B \cdot)_{oh} \) associated with parameters \((u_j)_{0}\) since \((F \cdot)_{on}\) changes to \(F_h\) when \((u_j)_{0}\) changes to \(f_{ij}\) in the model

\[
F_n = (F_c)_{on} + \sum_j a_{nj}[f_{ij} - (u_j)_{0}].
\]

Since \(a_{nj} = (a_{nj})_{1} + i(a_{nj})_{2}\) we can say

\[
\text{var} (A \cdot)_{on} = \sum_{ij} (a_{nj})_{1} a_{nj} (M_{ij})_{0},
\]

\[
\text{var} (B \cdot)_{oh} = \sum_{ij} (a_{nj})_{2} a_{nj} (M_{ij})_{0}
\]

and

\[
\text{cov} (A \cdot)_{on}(B \cdot)_{oh} = \sum_{ij} (a_{nj})_{1} a_{nj} (M_{ij})_{0}
\]

where \((M_{ij})_{0}\) can be determined from an initial setting up of the least-squares equations \(\sum_j [\tilde{u}_j - (u_j)_{0}] = B_i\) with weights \(w_{oh}\). Then

\[
(M_{ij})_{0} = [\tilde{u}_i - (u_i)_{0}] [\tilde{u}_j - (u_j)_{0}] + \text{cov} (\tilde{u}_i \tilde{u}_j)
\]

where \text{cov} \((\tilde{u}_i \tilde{u}_j)\) is

\[
(A^{-1})_{ij}\{(\sum_{h}w_{oh}\tilde{A}_{oh}) + 2\sum_{k}B_{ik}[\tilde{u}_i - (u_i)_{0}]/(n - m).
\]

The inverse of the variance in the phase direction \(\alpha\) is given by

\[
\tan 2\beta = 2 \text{cov} (A \cdot)_{on}(B \cdot)_{oh}/(\text{var} (A \cdot)_{on} - \text{var} (B \cdot)_{oh});
\]

\[
\sigma_{\alpha}^2 = \cos^2 \beta \text{var} (A \cdot)_{on} + \sin^2 \beta \text{var} (B \cdot)_{oh}
\]

\[
+ 2\beta \text{cov} (A \cdot)_{on}(B \cdot)_{oh}
\]

and

\[
\sigma_{\beta}^2 = \sin^2 \beta \text{var} (A \cdot)_{on} + \cos^2 \beta \text{var} (B \cdot)_{oh}
\]

\[
- 2\beta \text{cov} (A \cdot)_{on}(B \cdot)_{oh}.
\]

This gives

\[
\sigma^2 = \frac{\text{var} (A \cdot)_{on} \text{var} (B \cdot)_{oh} - [\text{cov} (A \cdot)_{on}(B \cdot)_{oh}]^2}{\cos^2 \alpha \text{var} (B \cdot)_{oh} + \sin^2 \alpha \text{var} (A \cdot)_{on} - 2\alpha \text{cov} (A \cdot)_{on}(B \cdot)_{oh}}.
\]

If \text{var} \((A \cdot)_{oh}\) \text{var} \((B \cdot)_{oh}\) = \(\text{cov} (A \cdot)_{oh}(B \cdot)_{oh}\) then \(\sigma_{\alpha}^2 = 0 \) and \(\sigma_{\beta}^2 = \sigma_{\alpha}^2\) when \(\alpha = \beta + \pi\) and zero elsewhere. This situation is well known in X-ray crystallography, being the case for all reflexions in crystals having a centre of symmetry and for special reflexions in many other space groups, e.g. \(hkl\) data of space group \(P2\).

Now

\[
E_2 = (F_o)_{oh} - F_h|F_{oh}|/|F_h| = |F_{oh}|(\exp i\alpha_h - \exp i\bar{\alpha}_h)
\]

where the assumed value of the phase angle of \((F_o)_{oh}\) is \(\alpha_{oh}\) while the true value of \(\alpha_h\) is \(\bar{\alpha}_h\), and this makes \(\langle E_2^2 \rangle = |F_{oh}|^2[2 - 2\langle \cos (\alpha_{oh} - \bar{\alpha}_h) \rangle].\) Obviously if \((F_o)_{oh} = 0, \langle \cos (\alpha_{oh} - \bar{\alpha}_h) \rangle = 0\) and \(\langle E_2^2 \rangle = 2|F_{oh}|^2\) while if \(|F_{oh}|^2 = 0, |E_2|^2 = 0\) and we see that this term discriminates in favour of observations for which \(|F_{oh}|_h < |(F_o)_{oh}|.\)

To evaluate \(\langle \cos (\alpha_{oh} - \bar{\alpha}_h) \rangle, \bar{\alpha}_h\) is unknown and some probability function is necessary. If we assume that quantities \(|(F_o)_{oh} - (F_{oh})_o|/\sigma\) all belong to the same normal distribution, \(\sigma^2\) being \text{var} \((F_c)_{oh}\) in the direction of \((F_o)_{oh} - (F_{oh})_o,\) then we can allow \(\bar{\alpha}_h\) to range over \(0 - 2\pi\) and say

\[
\langle \cos (\alpha_{oh} - \bar{\alpha}_h) \rangle = \int_0^{2\pi} \cos (\alpha_{oh} - \bar{\alpha}_h) \exp - \frac{X^2}{2} \frac{d\alpha}{\sqrt{2\pi}}
\]

where

\[
X^2 = \frac{A^2 \text{var} (B \cdot)_{oh} + B^2 \text{var} (A \cdot)_{on} - 2AB \text{cov} (A \cdot)_{on}(B \cdot)_{oh}}{\text{var} (A \cdot)_{on} \text{var} (B \cdot)_{oh} - [\text{cov} (A \cdot)_{on}(B \cdot)_{oh}]^2},
\]

\(A = |F_{oh}| \cos \bar{\alpha}_h - (A \cdot)_{oh}\) and \(B = |F_{oh}| \sin \bar{\alpha}_h - (B \cdot)_{oh}.\) In the case when \(\sigma^2 = 0, \cos (\alpha_{oh} - \bar{\alpha}_h)\) can only take values of \pm 1 and \(\langle E_2^2 \rangle = 4|F_{oh}|^2[1 + \exp (2|F_{oh}|/|F_{oh}|)]\).

The estimation of systematic error

An amount \(\langle E_3^2 \rangle = \langle |F_o - F_h|^2 \rangle\) has to be estimated and included in the weight so that \(\sum_h w_h \tilde{A}_h^2 = n - m\) after refinement. It is advantageous to distribute this error in a more meaningful way than simply to say that it is a constant. One obvious contribution in an X-ray crystallographic application is \(\langle f^2 \rangle\) the mean-squared scattering power of any omitted atoms. The most meaningful distribution of error will minimize

\[
\sum_h \left( \frac{\tilde{A}_h^2}{\langle |E_1|^2 + |E_2|^2 + |E_3|^2 \rangle_h - \frac{n - m}{n}} \right)^2
\]

and expressing \(E_3\) as a function of three or four variables is a justifiable attempt to locate the cause of systematic error.

Application to X-ray crystallography

The refinement of non-centrosymmetric crystal structures

We see that it has been customary to refine a wrong set of least-squares equations since the component of \((F_o)_{oh} - F_h\) at \(\pi/2\) to the phase angle \(\alpha_{oh}\) of \((F_o)_{oh}\) has been ignored. It is important that we try to explain the difficulties that are encountered in the refinement of pseudo-centrosymmetric crystal structures. We can
most simply do this in a two-dimensional example in which \((\partial A_c/\partial u_i)_{00} = (\partial A_c/\partial u_i)_{00} = (\partial B_i/\partial u_i)_{00} = - (\partial B_i/\partial u_i)_{00}\).

The least-squares equations thus obtained have
cov \((u_1 + u_2) (u_1 - u_2) = 0\) and are

\[
\begin{bmatrix}
A_1 + A_2 & A_1 - A_2 \\
A_1 + A_2 & A_1 + A_2
\end{bmatrix}
\begin{bmatrix}
u_1 - (u_1) \\
u_2 - (u_2)
\end{bmatrix}
= \begin{bmatrix} B_1 + B_2 \\ B_1 - B_2 \end{bmatrix}
\]

where \(B_1 = \sum_h w_h (\partial A_c/\partial u_i)_{00} (F_o - F_i) \cos \alpha_{0h}\) and \(B_2 = \sum_h w_h (\partial B_c/\partial u_i)_{00} (F_o - F_i) \sin \alpha_{0h}\) but the values of \(A_1\) and \(A_2\) depend on whether or not the correct equations are used. Using the correct equations \(A_1 = \sum_h w_h (\partial A_c/\partial u_i)_{00}^2\) and \(A_2 = \sum_h w_h (\partial B_c/\partial u_i)_{00}^2\), while using the incorrect equations \(A_1' = \sum_h w_h \cos^2 \alpha_{0h} (\partial A_c/\partial u_i)_{00}^2\) and \(A_2' = \sum_h w_h \sin^2 \alpha_{0h} \cos \alpha_{0h} (\partial B_c/\partial u_i)_{00}^2\) assuming \(\sum_h w_h \cos \alpha_{0h} \sin \alpha_{0h} \times (\partial A_c/\partial u_i)_{00} (\partial B_c/\partial u_i)_{00} = 0\). We can thus say \(A_1' = \langle \cos^2 \alpha \rangle A_1\) and \(A_2' = \langle \sin^2 \alpha \rangle A_2\) though \(\langle \cos^2 \alpha \rangle + \langle \sin^2 \alpha \rangle = 1\) only if \(A_1' + A_2' = 1\). If certain data can only have phase \(\alpha_0\) or \(\alpha_0 + \pi\) then \(\langle \cos^2 \alpha \rangle + \langle \sin^2 \alpha \rangle\) can be greater than one. The solutions of the least-squares equations are

\[
\Delta u_1 + \Delta u_2 = B_1/A_1, \quad \Delta u_1 - \Delta u_2 = B_2/A_2
\]

if the correct equations are used. If the incorrect equations are used these shifts are \(B_1/A'_1\) and \(B_2/A'_2\) respectively. Thus when the incorrect equations are used the shifts are overemphasized by \(1/\langle \cos^2 \alpha \rangle\) and \(1/\langle \sin^2 \alpha \rangle\). The corresponding variances are estimated as being \(\langle \cos^2 \alpha \rangle\) and \(\langle \sin^2 \alpha \rangle\) greater than before also.

However these variances are underestimated as we can see using the true parameters \(u_1\), \(u_2\) as our starting values of \((u_1)_0\), \((u_2)_0\).

The variances are defined as \(\langle (u_1 + u_2 - \bar{u}_1 - \bar{u}_2)^2 \rangle\) and \(\langle (u_1 - u_2 - \bar{u}_1 + \bar{u}_2)^2 \rangle\) and are thus \(1/\langle \cos^2 \alpha \rangle^2\) and \(1/\langle \sin^2 \alpha \rangle^2\) greater for the incorrect equations than for the correct equations. The correct least-squares equations to use were derived for an expansion about the true values \(\bar{u}_1\), \(\bar{u}_2\) using coefficient \(a_{ij}\) evaluated for \((u_1)_0\).

The uncertainty of the phase of \((F_2)_{00} - F_h\) is not the same as the uncertainty of the phase of \(\bar{F}_h - F_h\) and this allows refinement of the incorrect least-squares equations despite the fact that no restraint is imposed on the component of \((F_2)_{00} - F_h\) at \(\pi/2\) to \(\alpha_{0h}\). It should be noted that in the example we have used \(A_1\) approximates \(A_2\) and the covariance of parameters \(u_1\) and \(u_2\), which has bedevilled least-squares refinement with the incorrect equations, is largely removed. As a consequence it should now be possible with the correct equations to refine a centrosymmetric structure in a non-centrosymmetric space group since atoms at \(x, y, z\) and \(\bar{x}, \bar{y}, \bar{z}\) will no longer have a correlation coefficient of 1. Thus in a space group such as \(C2/m\) it will be possible to say more certainly that this is indeed the space group and not \(C2\) or \(Cm\).

The refinement of a structure showing pseudo translational symmetry

In a crystal showing pseudo translational symmetry only a small fraction of the data (e.g. \(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\) etc.) will have high average intensity and the ordered nature of the variation from translational symmetry is best determined by the weak data. For example only data with \(h = 2n + 1\) can distinguish between fractional coordinates \(x + 1/2\) and \(x + 1/2\). When data are weighted according to counting statistics only, it is commonly found that \(\langle 1 \rangle\), the value of \(\langle w_h | D_h |^2 \rangle\) for the \(h = 2n\) data is greater than \(\langle 2 \rangle\), the value of \(\langle w_h | D_h |^2 \rangle\) for the \(h \neq 2n\) data because of an overestimate of the ability of high-intensity data to refine the structure.

It is of interest to investigate the consequence of this situation. Let us consider a two-parameter problem for fractional coordinates \(x_1\) and \(x_2\) separated by about \(\frac{1}{2}\) and \(\text{cov} \{(x_1 + x_2) (x_1 - x_2)\} = 0\). For \(h = 2n\) we obtain contributions to the least-squares equations of

\[
\begin{bmatrix}
A_1 (1 - \delta_1) \\
A_1 (1 - \delta_2)
\end{bmatrix} \quad \begin{bmatrix} \Delta x_1 \end{bmatrix} = \begin{bmatrix} B_1 + b_1 \\ B_1 - b_1 \end{bmatrix}
\]

where \(\Delta x_1 = \hat{x}_1 - (x_1)_0\). On their own these data give \(\Delta x_1 + \Delta x_2 = 2B_2/A_1 (2 - \delta_1)\) and \(\Delta x_1 - \Delta x_2 = 2b_1/A_1 \delta_1\) where \(\text{var} \{(\Delta x_1 + \Delta x_2) = 2(1)/A_1 (2 - \delta_1)\}\) and \(\text{var} \{(\Delta x_1 - \Delta x_2) = 2(1)/A_1 \delta_1\}\) for \(h \neq 2n\) data we obtain contributions

\[
\begin{bmatrix}
A_2 \\
A_2
\end{bmatrix} \quad \begin{bmatrix} \Delta x_1 \end{bmatrix} = \begin{bmatrix} B_2 + b_2 \\ -B_2 - b_2 \end{bmatrix}
\]

On their own these data give \(\Delta x_1 + \Delta x_2 = 2b_2/A_2 \delta_2\) and \(\Delta x_1 - \Delta x_2 = 2B_2/A_2 (2 - \delta_2)\) where

\[
\text{var} \{(\Delta x_1 + \Delta x_2) = 2(2)/A_2 \delta_2\}
\]

and \(\text{var} \{(\Delta x_1 - \Delta x_2) = 2(2)/A_2 (2 - \delta_2)\}\).

If we combine the results of the \(h = 2n\) and the \(h \neq 2n\) data to obtain the minimum variances for the non-covarying parameters \(x_1 + x_2\) and \(x_1 - x_2\) then we should weight the data as the inverse of the variances.

Then

\[
\begin{align*}
\Delta x_1 + \Delta x_2 &= (2B_2 (2) + 2b_2 (1)) / (A_1 (2 - \delta_1) (2) + A_2 \delta_2 (1)) \\
\Delta x_1 - \Delta x_2 &= (2B_2 (1) + 2b_2 (2)) / (A_1 (2 - \delta_2) (1) + A_2 \delta_2 (2))
\end{align*}
\]

where \(\text{var} \{(x_1 + x_2) = 2(1) / A_1 (2 - \delta_1) (1)\}\)

and \(\text{var} \{(x_1 - x_2) = 2(1) / A_2 \delta_2 (1)\}\).

Combining the two sets of data, ignoring the fact that \(\langle 1 \rangle \neq \langle 2 \rangle\), we obtain

\[
\begin{bmatrix}
A_1 + A_2 \\
A_1 (1 - \delta_1) - A_2 (1 - \delta_2)
\end{bmatrix} \quad \begin{bmatrix} \Delta x_1 \end{bmatrix} = \begin{bmatrix} B_1 + B_2 + b_1 + b_2 \\ B_1 - B_2 - b_1 + b_2 \end{bmatrix}
\]
giving
\[ \Delta x_1 + \Delta x_2 = \frac{2B_1 + 2b_2}{A_1(2-\delta_1) + A_2\delta_2}, \]
\[ \Delta x_1 - \Delta x_2 = \frac{2B_2 + 2b_1}{A_1\delta_1 + A_2(2-\delta_2)}, \]
\[ \text{var}(\Delta x_1 + \Delta x_2) = \frac{\langle 1 \rangle + \langle 2 \rangle}{A_1(2-\delta_1) + A_2\delta_2} \]
and
\[ \text{var}(\Delta x_1 - \Delta x_2) = \frac{\langle 1 \rangle + \langle 2 \rangle}{A_1\delta_1 + A_2(2-\delta_2)}. \]

These answers are only the same as before if \( \langle 1 \rangle = \langle 2 \rangle \).

The differences in the two answers are given as
\[ \langle (1) - (2) \rangle \]
\[ \frac{[2B_1/A_1(2-\delta_1) - 2b_2/A_1\delta_2]A_1(2-\delta_1)A_2\delta_2}{[A_1(2-\delta_1) + A_2\delta_2][A_1(2-\delta_1)(2) + A_2\delta_2(1)]} \]
for \( \Delta x_1 + \Delta x_2 \)
and
\[ \langle (1) - (2) \rangle \]
\[ \frac{[2b_1/A_1\delta_1 - 2B_2/A_2(2-\delta_2)]A_1\delta_1A_2(2-\delta_2)}{[A_1\delta_1 + A_2(2-\delta_2)][A_1\delta_1(2) + A_2(2-\delta_2)(1)]} \]
for \( \Delta x_1 - \Delta x_2 \).

Now the true variance from the combined data set is the variance from the minimum-variance combination of the data sets plus the expectation value of the square of the difference in the answers using \( \text{var}[2B_1/A_1(2-\delta_1) - 2b_2/A_1\delta_2] = 2\langle 1 \rangle/A_1(2-\delta_1) \) etc. and assuming that the covariance of the various contributions \( B_1, B_2, b_1, b_2 \) are zero. Thus
\[ \text{var}(x_1 + x_2) = \frac{2\langle 1 \rangle \langle 2 \rangle}{A_1(2-\delta_1) + A_2\delta_2} \]
\[ \times \left[ 1 + \frac{A_1(2-\delta_1)A_2\delta_2}{[A_1(2-\delta_1) + A_2\delta_2]^2} \frac{\langle 1 \rangle - \langle 2 \rangle^2}{\langle 1 \rangle \langle 2 \rangle} \right] \]
and
\[ \text{var}(x_1 - x_2) = \frac{2\langle 1 \rangle \langle 2 \rangle}{A_1\delta_1A_2(2-\delta_2) + A_1(2-\delta_2)} \]
\[ \times \left[ 1 + \frac{A_1\delta_1A_2(2-\delta_2)}{[A_1\delta_1 + A_2(2-\delta_2)]^2} \frac{\langle 1 \rangle - \langle 2 \rangle^2}{\langle 1 \rangle \langle 2 \rangle} \right]. \]

When \( \langle 1 \rangle > \langle 2 \rangle \) the incorrectly estimated variances are too high for \( \Delta x_1 - \Delta x_2 \) and too low for \( \Delta x_1 + \Delta x_2 \). It should be noted that the removal of the \( x + \frac{1}{2} \) translational symmetry element halves the number of symmetry elements in the supercell and the implications of non-centrosymmetric refinements discussed earlier are commonly an additional feature of these refinements.

The crystal structure of \( K_2ZnCl_4 \) has been investigated (Dix, 1972) and has pseudo translational symmetry elements of \( x + \frac{1}{2}, x + \frac{3}{2} \) and also a pseudo mirror plane, crystallizing in the space group \( Pna2_1 \). Using unit weights with discrimination to exclude 10% of the data with low \( |F_{clh}| < |F_{olh}| \) and fractional coordinate shifts proved to be the best initial method of refinement with the conventional but incorrect least-squares equations. Final refinement using weights from counting statistics gave a final value for \( R = 0.064 \). At this stage weighting by counting statistics gave
\[ \langle w_n|D_n|^2 \rangle_n = 3 \langle w_n|D_n|^2 \rangle_{n \neq n} = 6.5 \text{ not } 1.0. \]

There are twelve different Zn–Cl bonds in this structure and a standard deviation evaluated from the set of twelve bond lengths gave a value of 0.025 Å compared with the average estimate of 0.0072 Å. We consider the value of 0.025 Å to be more representative of the true variance using this incorrect approach. As a result of these considerations further refinement of this structure is in progress.

**Polar space groups**

In the past when a full-matrix refinement procedure has been used it has been found necessary to impose some restraint on atom shifts for polar space groups. This need is directly attributable to the fact that the component of \( (F_{oh})_c - (F_{oh})_b \) at \( \pi/2 \) to the phase direction of \( (F_{oh})_b \) has been completely ignored. Translation of atoms in the crystal implies multiplication by a phase factor, or in other words a change in phase angle. If the component of \( (F_{oh})_b - (F_{oh})_b \) at \( \pi/2 \) to \( \alpha_{oh} \) is minimized then so is the amount of translation of atoms in the polar direction.

**Approximations to the error in the phase of \( (F_{oh})_h \) using X-ray data**

We saw earlier that
\[ \langle |E_2|^2 \rangle = |F_{oh}|^2 [2 - 2 \cos(\alpha_{oh} - \bar{\alpha}_h)]. \]

For an acentric crystal we assume that at a particular value of \( \sin\theta/\lambda, \sigma_1^2 = \sigma_2^2 \) is a constant \( \sigma^2 \) for general reflexions and \( \sigma_1^2 = 2\sigma_2^2 \) when \( \sigma_2 = 0 \). For general reflexions \( \cos(\alpha_{oh} - \bar{\alpha}_h) \) is evaluated as
\[ \int_{0}^{2\pi} \cos(\alpha_{oh} - \bar{\alpha}_h) \exp\left(-\frac{X^2}{2}\right) \, d\bar{\alpha}_h, \]
where \( X^2 \) is now
\[ \int_{0}^{2\pi} \cos(\alpha_{oh} - \bar{\alpha}_h) \exp\left(-\frac{X^2}{2}\right) \, d\bar{\alpha}_h. \]

\[ \langle \cos(\alpha_{oh} - \bar{\alpha}_h) \rangle = \int_{0}^{2\pi} \cos(\alpha_{oh}) \exp((|F_{oh}|(|F_{oh}| \cos \alpha/\sigma)^2) \, d\alpha. \]

\( \sigma^2 \) may be estimated as \( \langle |(F_{oh})_b - (F_{oh})_b|^2 \rangle \) at angle \( \theta \) and to a good approximation will vary as \( \sum_i f_i^2 \sin^2 \theta/\lambda^2 \)
where \( f_i \) is the scattering factor of the \( i \)th atom corresponding to the value of \( 2 \sin \theta/\lambda \). Values of \( \langle \cos(\alpha_{oh} - \bar{\alpha}_h) \rangle \) for various values of \( |F_{oh}|(|F_{oh}|/\sigma)^2 \) can
Conclusion

It has been demonstrated that the major cause of covariance problems in the least-squares refinement of crystal structures has been caused by the use of incorrect least-squares equations and the insistence on weighting according to counting statistics. With the approach outlined in this paper it is possible to refine a centrosymmetric structure in a non-centrosymmetric space group. In such an application, if one starts from exactly centrosymmetric coordinates, refinement proceeds to a false minimum since the structure will stay centrosymmetric. However, if one is trying to test the validity of a structural parameter defining a rigid group of atoms, for example the angle which a plane of atoms makes with a symmetry axis, then an initial perturbation involving this parameter may be tested.

The problem of false minima is a necessary part of refining a crystal structure with a fixed weighting scheme (Rae, 1974). It is unrealistic to rely on the redefinition of the phase of \( (F_0)_h \) to enable refinement to proceed. The contribution to the weighting scheme of \( \langle |E_i|^2 \rangle \) and \( \langle |E_i|^2 \rangle \) far outweighs the counting statistic contribution \( \langle |E_i|^2 \rangle \) in all but the final refinement cycle since \( \sum_{h} w_h |d_{ii}|^2 > 1 \) with \( w_h^{-1} = \langle |E_i|^2 \rangle \).

The point is made that the refinement is only as good as the weighting scheme. Because data with \( |(F_0)_h| \) and \( |(F_0)_h| \) are weighted preferentially, scale constants should be refined on their own in a separate least-squares cycle in the initial refinement stages with \( \langle |F_h|^2 \rangle = \sum f_i^2 \) where \( f_i \) is the scattering factor for the \( i \)th atom in the crystal. The better the weighting scheme the more rapid and correct the convergence. The only extra computing time involved is in the actual multiplication of derivatives to form the matrix.

When refinement is complete \( w_h^{1/2} \alpha_h \) values will enable a probability distribution to be evaluated. By multiplying \( w_h^{1/2} \alpha_h \) by the probability associated with the value of \( w_h^{1/2} \alpha_h \) it should be possible to improve the refinement since the variance associated with the square of a normal distribution is half the variance associated with the normal distribution function itself.

References


On the Libration of 9, 10-Anthraquinone at Five Temperatures

BY URI SHMUELI* AND PETER A. KROON
Laboratorium voor Structuurchimie, Rijksuniversiteit te Groningen, Zernikelaan, Paddepoel, Groningen,
The Netherlands

(Received 25 March 1974; accepted 29 April 1974)

Potential energy profiles corresponding to libration of 9,10-anthraquinone about its molecular axes were constructed for the five structures derived from data which were collected at \(-170, -112, -72, -12.5\) and \(20.5^\circC\) [Lonsdale, Milledge & El Sayed (1966). Acta Cryst. 20, 1–13]. These profiles were represented by fourth-degree least-squares polynomials, whereafter r.m.s. libration amplitudes and rigid rotator frequencies of 9,10-anthraquinone were evaluated in the quadratic approximation. The temperature dependence of the calculated quantities is in most cases close to that of the observed ones thus reproducing, by comparison with observed Raman frequencies, the pseudoharmonic behaviour of 9,10-anthraquinone. Calculated r.m.s. libration amplitudes are only qualitatively comparable to the experimental ones and appear to be somewhat too low. The present representation of energy profiles makes it possible to estimate conveniently the contribution of anharmonicity to the profile shape.

Introduction

The availability of semiempirical potential functions enables one to construct approximate potential energy

* On leave from the Department of Chemistry, Tel-Aviv University, Ramat-Aviv, Israel.