

Significance Tests on the Crystallographic R Factor*

BY WALTER C. HAMILTON

Chemistry Department, Brookhaven National Laboratory, Upton, New York, U.S.A.

(Received 13 April 1964)

The theory of linear hypothesis tests is discussed in relationship to its use in crystallographic problems. In particular, formulas are developed which allow significance tests on the R -factor ratio $\mathcal{R} = R_1/R_0$, where R_0 and R_1 are respectively the generalized weighted R factors for a structure resulting from an unrestrained least-squares refinement and a refinement with restraints on some of the parameters. These tests allow one to decide whether the addition of parameters or the imposition of fixed relationships between parameters results in a *significant* improvement or a *significant* worsening of the agreement between observed and calculated structure factors. Tables for determining significance levels of \mathcal{R} are presented, and several applications from the crystallographic literature and the author's experience are discussed.

1. Introduction

In the refinement of a crystal structure, one assumes a model with variable parameters, the values of which are chosen so as to achieve the best agreement between the observed and calculated structure amplitudes. Frequently, one wishes to choose between two models which differ essentially in the number of parameters used to describe the structure. For example, a model with anisotropic thermal parameters on each atom has in general five more parameters per atom than a model where each atom is assumed to vibrate isotropically. A model in an acentric space group may have about twice as many positional parameters as does a model in a centrosymmetric space group. The model with the fewer restraints, that is, with the greater number of parameters, can usually be made to fit the data better than can the more restrained model†; the crystallographer is thus often tempted to add more and more parameters to his model to obtain better fits to his data. It is the purpose of this paper to introduce a convenient method which will allow the crystallographer to decide whether the increase in the number of parameters leads to a significant improvement in the agreement between observed and calculated structure amplitudes.

We would like in some way to be able to test the validity of statements such as those that follow:

(a) The use of anisotropic rather than isotropic temperature factors in the description of the structure resulted in a *significant* improvement in the agreement between observed and calculated structure factors.

(b) We can reject as being incompatible with the data a model in which all the distances in a benzene ring are equal rather than alternating.

(c) The two molecules in the asymmetric unit are coplanar.

(d) A model with free rotation of the methyl group gives significantly better agreement with the data than one which places half-hydrogen atoms with normal temperature factors in six equivalent positions.

(e) Addition of hydrogen atom contributions to the calculated structure factors results in a significantly lower R factor.

(f) In this acentric structure, we can conclude that the absolute configuration is A rather than B , because the R factors for the best refinements in the two cases are 0.100 and 0.110 respectively.

The probability of the truth of these statements may be difficult to judge by simply examining the final parameter values and estimated standard deviations obtained from a least-squares refinement. We recommend here a far simpler approach which can be used to answer a variety of questions; we find the approximate statistical significance levels for an R -factor ratio refined by $\mathcal{R} = R_1/R_0$, where R_1 is the generalized R factor for a model with restrictions on the parameter values, and R_0 is the generalized R factor for a model without these restrictions.

In §2 we introduce the generalized R factor. In §3 we present briefly the theory of linear hypothesis tests and show how it may be related to tests of the R -factor ratio. Following a brief discussion of non-linearity in §4 we present in §5 several applications of the proposed test. The Appendix contains computed values of the significance levels of the R -factor ratio.

We will not attempt in this paper to prove all the pertinent results. The reader is referred to some recent texts (Graybill, 1962; Hamilton, 1964; Plackett, 1960) for a full discussion. The use of the R -factor ratio as a test of hypotheses concerning crystal structures

* Research performed under the auspices of the U.S. Atomic Energy Commission.

† The model with the greater number of parameters can *always* be made to fit the data at least as well as the model with the fewer parameters, provided that the parameters in the latter are a subset of those in the former.

was briefly discussed by the author a few years ago (Hamilton, 1961).

2. The generalized R factor

Given a set of observed structure amplitudes $|F_i|_o$, a corresponding set of values $|F_i|_c$ calculated for a model, and a set of weights $w_i = 1/\sigma_i^2$, where σ_i^2 is the variance of $|F_i|_o$, the conventional crystallographic agreement factor may be defined by

$$R' = \frac{\sum_i w_i^{\frac{1}{2}} ||F_i|_o - |F_i|_c|}{\sum_i w_i^{\frac{1}{2}} |F_i|_o} \quad (1)$$

In computing R' , the weights w_i are usually chosen to be equal for all reflections.

R' as defined in (1) is essentially a normalized *mean deviation*. A more convenient quantity from a statistical point of view is

$$R'' = \left[\frac{\sum_i w_i (|F_i|_o - |F_i|_c)^2}{\sum_i w_i |F_i|_o^2} \right]^{\frac{1}{2}}. \quad (2)$$

R'' is essentially a normalized *standard deviation* and is the 'weighted R factor' often calculated in crystallographic least-squares computer programs.

If there are correlations between the observations such that the variance-covariance matrix $\mathbf{M}^{(F)}$ for the observations has elements

$$m_{ij} = \text{covariance} (|F_i|, |F_j|), \quad (3)$$

the corresponding weight matrix \mathbf{W} is defined by

$$\mathbf{W} = \sigma^2 \mathbf{M}^{-1} \quad (4)$$

where σ^2 is the (perhaps unknown) variance of an observation of unit weight. We may define a generalized weighted R factor (Hamilton & Ibers, 1963a) by

$$\bar{R}^G = \left[\frac{\sum_i \sum_j w_{ij} (|F_i|_o - |F_i|_c) (|F_j|_o - |F_j|_c)}{\sum_i \sum_j w_{ij} |F_i|_o |F_j|_o} \right]^{\frac{1}{2}}. \quad (5)$$

\bar{R}^G is identical with R'' defined in (2) if the weight matrix is diagonal with $w_{ii} = w_i$ and $w_{ij} = 0$ for $i \neq j$.*

3. Least-squares and linear hypothesis tests

A brief summary of the least-squares method in matrix notation may be useful. We denote by $\mathbf{A}_{n,m}$ a matrix with n rows and m columns. The transpose of \mathbf{A} is denoted by \mathbf{A}' and the inverse of a square matrix $\mathbf{B}_{n,n}$ by \mathbf{B}^{-1} .

Let a set of observations $\mathbf{F}_{n,1}$ be linearly related to a set of parameters $\mathbf{X}_{m,1}$ ($n > m$) by the observational equations

$$\mathbf{F} = \mathbf{A}_{n,m} \mathbf{X} + \boldsymbol{\varepsilon}_{n,1} \quad (6)$$

where \mathbf{A} is the design matrix and $\boldsymbol{\varepsilon}$ is a random vector with zero mean, representing the errors in the experimental observations. We denote by $\mathbf{M}_{n,n}^{(F)}$ the variance-covariance matrix for the observations:

$$\mathbf{M}_{n,n}^{(F)} = \mathcal{E} \{ \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \} \quad (7)$$

where the notation $\mathcal{E}(\mathbf{P})$ means the matrix whose elements are the expected values of the elements of \mathbf{P} . If a weight matrix is defined as in (4)

$$\mathbf{W} = \sigma^2 [\mathbf{M}^{(F)}]^{-1} \quad (4)$$

the least-squares estimates $\hat{\mathbf{X}}$ of the parameters \mathbf{X} are given by

$$\hat{\mathbf{X}} = \mathbf{B}^{-1} \mathbf{A}' \mathbf{W} \mathbf{F} \quad (8)$$

where \mathbf{B} , the matrix of the normal equations, is defined by

$$\mathbf{B} = \mathbf{A}' \mathbf{W} \mathbf{A}. \quad (9)$$

The variance-covariance matrix for the parameters may be shown to be

$$\mathbf{M}^{(X)} = \sigma^2 \mathbf{B}^{-1}. \quad (10)$$

An unbiased estimate of σ^2 is given by

$$\hat{\sigma}^2 = \frac{(\mathbf{F} - \mathbf{A} \hat{\mathbf{X}})' \mathbf{W} (\mathbf{F} - \mathbf{A} \hat{\mathbf{X}})}{n - m} \quad (11a)$$

$$\equiv \frac{\mathbf{V}' \mathbf{W} \mathbf{V}}{n - m} \equiv \frac{G_0}{n - m} \quad (11b)$$

where (11b) is a definition of the matrix of residuals \mathbf{V} and the weighted sum of squares G_0 . An unbiased estimate of $\mathbf{M}^{(X)}$ is thus given by

$$\hat{\mathbf{M}}^{(X)} = \mathbf{B}^{-1} \frac{\mathbf{V}' \mathbf{W} \mathbf{V}}{n - m}. \quad (12)$$

In crystallographic and other non-linear problems to which the least-squares method is applied, an iterative process is used in which the observational equations (6) retain their meaning if the following definitions of \mathbf{F} , \mathbf{A} and \mathbf{X} in (6) are made.

$$\begin{aligned} \mathbf{F} &= \{|F_i|_o - |F_i|_c\} \equiv \{\Delta F_i\} \\ \mathbf{A} &= \{a_{ij}\} \equiv \{\partial |F_i| / \partial x_j\} \\ \mathbf{X} &= \{\Delta x_j\} \end{aligned} \quad (13)$$

where $\{\Delta x_j\}$ are the corrections to be applied to the parameters before the next iteration.

We note that equation (5) may be written

$$R^G = \left[\frac{(|\mathbf{F}|_o - |\mathbf{F}|_c)' \mathbf{W} (|\mathbf{F}|_o - |\mathbf{F}|_c)}{|\mathbf{F}'_o \mathbf{W} \mathbf{F}|_o} \right]^{\frac{1}{2}} \quad (14)$$

with the result that equation (11) may be written

$$\hat{\sigma}^2 = \frac{(R^G)^2}{n - m} \mathbf{F}'_o \mathbf{W} \mathbf{F}_o, \quad (15)$$

* In definitions (1), (2) and (5), we may wish to use F^2 rather than $|F|$, particularly if the least-squares refinements have been carried out on F^2 .

a relationship that we will use in developing a test for R -factor ratios.

A linear hypothesis is a hypothesis which specifies fixed linear relationships between the parameters of a physical statistical model. It may be expressed by the set of linear equations

$$\mathbf{Q}_{b,m}\mathbf{X}_{m,1} = \mathbf{Z}_{b,1} \quad (16)$$

The hypothesis that (16) is true is a b -dimensional linear hypothesis if \mathbf{Q} has rank b . An example of a linear hypothesis for a five-parameter problem is

$$\begin{aligned} x_1 &= x_2 \\ x_3 &= x_4 + 0.50 \\ x_5 &= 0 \end{aligned}$$

i.e.

$$\mathbf{Q} \equiv \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (17)$$

$$\mathbf{Z} \equiv \begin{pmatrix} 0 \\ 0.50 \\ 0 \end{pmatrix}$$

In crystallographic problems, some hypotheses of interest might be strictly linear in the refined parameters, others might be linearized by means of a Taylor's series expansion. An example of the former is a hypothesis which specifies restrictions on the thermal parameters, *e.g.*

$$\begin{aligned} \beta_{11} &= \beta_{22} = \beta_{33} \\ \beta_{12} &= \beta_{13} = \beta_{23} = 0, \end{aligned}$$

when the β_{ij} are the refined parameters. An example of the latter might be a hypothesis that two bond lengths are equal when the refined parameters are the atomic coordinates.

If the problem is strictly linear, the least-squares estimate $\bar{\bar{\mathbf{X}}}$ of \mathbf{X} under the condition that the hypothesis (16) is true is

$$\bar{\bar{\mathbf{X}}} = \hat{\mathbf{X}} + \mathbf{B}^{-1}\mathbf{Q}'(\mathbf{Q}\mathbf{B}^{-1}\mathbf{Q}')^{-1}(\mathbf{Z} - \mathbf{Q}\hat{\mathbf{X}}) \quad (18)$$

The weighted sum of squares of residuals is

$$G_Q \equiv (\mathbf{F} - \mathbf{A}\bar{\bar{\mathbf{X}}})'\mathbf{W}(\mathbf{F} - \mathbf{A}\bar{\bar{\mathbf{X}}}) \quad (19a)$$

$$= G_0 + (\bar{\bar{\mathbf{X}}} - \hat{\mathbf{X}})'\mathbf{B}(\bar{\bar{\mathbf{X}}} - \hat{\mathbf{X}}) \quad (19b)$$

If the hypothesis is true, it may be shown that the quantity

$$F \equiv \frac{G_Q - G_0}{G_0} \cdot \frac{(n-m)}{b} \quad (20)$$

is distributed as $F_{b,n-m}$, the usual analysis-of-variance ratio for b and $n-m$ degrees of freedom. We reject the hypothesis at the significance level α (often 0.05) if F defined by (20) exceeds the value of $F_{b,n-m,\alpha}$ for which

$$\text{Prob}[F > F_{b,n-m,\alpha}] = \alpha.$$

We will thus subject ourselves to the risk of rejecting a true hypothesis $100\alpha\%$ of the time.

From (15), we see that (20) may be written

$$F = \frac{R_Q^2 - R_0^2}{R_0^2} \cdot \frac{(n-m)}{b} \quad (21)$$

if R_Q is the generalized R factor calculated for the best parameter values consistent with the linear restrictions on the parameters, and R_0 is that calculated for the best model without the restrictions. We may thus test the significance of improvement obtained when parameters are added to a least-squares refinement by calculating the R factors and using them in an F test as suggested in (21). In practice, it may be more convenient to examine the ratio of R factors:

$$\mathcal{R} = R_Q/R_0 \quad (22)$$

Combining (22) and (21), we obtain

$$\mathcal{R} = \left[\frac{b}{n-m} \cdot F + 1 \right]^{\frac{1}{2}} \quad (23)$$

Since \mathcal{R} is a monotonic function of F , if

$$F > F_{b,n-m,\alpha}$$

then

$$\mathcal{R} > \mathcal{R}_{b,n-m,\alpha}$$

where the significance points of \mathcal{R} are obtained as

$$\mathcal{R}_{b,n-m,\alpha} = \left[\frac{b}{n-m} \cdot F_{b,n-m,\alpha} + 1 \right]^{\frac{1}{2}} \quad (24)$$

A tabulation of $\mathcal{R}_{b,n-m,\alpha}$ will be found in Table 1. Shorter versions of these tables have been presented elsewhere (Hamilton, 1964).

We have frequently found for typical X-ray diffraction data that R (equation (1)) and R^G (equation (5)) are, at the end of a refinement, comparable. Even when they are not, the ratio \mathcal{R} is very likely to be relatively insensitive to how the R factors are computed. Hence, for quick work, the use of the conventional R factors in calculating \mathcal{R} may be adequate.*

4. Non-linearity

The preceding analysis is exact in the case of linear observational equations (1) and linear hypotheses (16). Two *separate* assumptions of linearity are (i) that

$$\Delta\mathbf{F} = \{\partial F_i / \partial x_j\} \Delta\mathbf{X} \quad (25)$$

and (ii) that the hypotheses can be cast into linear form. These two conditions of linearity are almost never strictly satisfied in crystallographic problems. The best that one can hope is that the regions of approximate linearity extend over the entire region

* Some least-squares programs give the sum of squares of deviations as part of the output. In such a case, one could of course simply use tables of F .

of the multiparameter space where there is an appreciable probability density. Fortunately, experience has shown this to be often the case in crystallographic problems.

The usual Student's t test made on individual parameters resulting from a least-squares refinement is equivalent to the use of (19b) rather than (19a) to calculate G_Q . This presupposes the linearity expressed by (25) entirely aside from whether the hypothesis itself is linear. If there are severe non-linearities here, comparison of the R factors (essentially calculation of G_Q by (19a)) does much to alleviate the uncertainties caused by these non-linearities. This is part of the basis for the recommendation of the present method: it is the agreement of the observed and calculated structure factors that is important in assessing the reasonableness of any parameter set. If an alternative set of parameters is to be tested, it is the real effect of the alternative parameters on the data rather than that extrapolated from the derivatives of the point of best fit which should be taken into account.

The method suggested here, although better than any approach which demands linearity of the observational equations, is not exact. It still demands linearity of the hypotheses, and further, because of the non-linearities of the observational equations, the number of degrees of freedom associated with the hypothesis is not well defined. This latter difficulty is probably not serious, and the effect of the former is readily investigated. In any case, a test of \mathcal{R} will always be at least as meaningful, and often more meaningful, than the usual arguments which are made on the basis of standard deviations resulting from the least-squares refinement of crystal structures.

Methods of estimating the effects of non-linearity on confidence regions in non-linear estimation have been discussed by Beale (1960), who introduces an average non-linearity parameter, the value of which allows one to assess whether non-linearity is likely to demand a *serious* modification of the usual hypothesis tests. The calculation of this or a related parameter would form a useful addition to computer programs designed to carry out non-linear least-squares refinements.

5. Some examples

In this section, we will attempt to illustrate by concrete examples the value of the \mathcal{R} test. (Where we quote an ' R factor', we mean either R'' or R^G as defined in (2) and (5) above.)

Example 1

A structure refinement of 30 parameters, based on 90 observations, has resulted in an R factor of 0.100. Another investigator (A) has found a different set of values for the 30 parameters. The R factor calculation using his parameter values and our data is 0.123.

Can we conclude that his model is inconsistent with our data? The hypothesis to be tested may be formulated

H_0 : The thirty parameters given by A correctly describe the structure.

The dimension of the hypothesis is $b=30$. The estimate of σ (essentially the goodness of the data) is based on $n-m=90-30$ degrees of freedom. We form the R -factor ratio

$$\mathcal{R} = 0.123/0.100 = 1.23$$

and test as follows: From Table 1, we find that $\mathcal{R}_{30, 60, 0.05} = 1.351$. We cannot reject the hypothesis at the 5% level of significance. An examination of Table 1 also reveals that $\mathcal{R}_{30, 60, 0.50} = 1.222$, indicating that the hypothesis cannot be rejected at the 50% level, *i.e.* we would be wrong half the time if we rejected similar hypotheses at this level. We conclude that our data do not indicate the incorrectness of investigator A 's model.*

Example 2

In an X-ray study of the crystal structure of XeF_4 (Hamilton & Ibers, 1963b), least-squares refinements were carried out for three cases:

- All atoms have isotropic temperature factors.
- The Xe atom only has an anisotropic temperature factor.
- All atoms have anisotropic temperature factors.

The numbers of parameters refined in the three cases were $m_a=15$, $m_b=20$, and $m_c=30$. There were 262 reflections. The R factors achieved were $R_a=0.114$, $R_b=0.100$, and $R_c=0.097$.

Let us test the hypothesis

H_0 : All atoms vibrate isotropically.

We compare the R factors for the restrained model (a), corresponding to the hypothesis and the unrestrained model (c). The appropriate R -factor ratio is

$$\mathcal{R} = R_a/R_c = 0.114/0.097 = 1.175.$$

The dimension of the hypothesis is $m_c - m_a = 15$ and the number of degrees of freedom for the refinement is $n - m = 262 - 30 = 232$. A pertinent tabular value of \mathcal{R} , obtained by interpolation from Table 1, is

$$\mathcal{R}_{15, 232, 0.005} = 1.072.$$

We can reject the hypothesis at the 0.005 level, *i.e.* we are quite sure that there is anisotropic motion. Let us however test the hypothesis

H_0 : The fluorine atoms have isotropic temperature factors, but the Xe does not.

* Assuming that both refinements are complete, we would be justified in taking a weighted mean of the two sets of results as the best current estimate of the true parameter set.

We compute

$$\mathcal{R} = R_b/R_c = 0.100/0.097 = 1.03$$

and test as \mathcal{R} with 10 and 232 degrees of freedom:

$$\begin{aligned}\mathcal{R}_{10, 232, 0.05} &= 1.040 \\ \mathcal{R}_{10, 232, 0.10} &= 1.035 \\ \mathcal{R}_{10, 232, 0.25} &= 1.027.\end{aligned}$$

The probability of error if the hypothesis is rejected lies between 10% and 25%. Usual practice would not demand the rejection of the hypothesis at this level.

We conclude that the results of our experiment do not preclude a model where only the Xe is allowed to vibrate anisotropically.

Example 3

In a study of the crystal structure of DCrO_2 (see Hamilton & Ibers, 1963a), a least-squares refinement of three parameters was carried out with 19 intensities from a powder pattern. The resulting value of the generalized R factor was 0.083, and the value of the single parameter determining the deuterium atom position was estimated to be $x_D = 0.4814$ with an estimated standard deviation of $\sigma_D = 0.0015$. Let us test the hypothesis

$$H_0: x_D = 0.5000.$$

A refinement of the other parameters with x_D fixed at 0.5000 resulted in a value for R of 0.124. We test the ratio

$$\mathcal{R} = 0.124/0.083 = 1.494.$$

The dimension of the hypothesis is 1. Since

$$\mathcal{R}_{1, 16, 0.005} = 1.289,$$

we can reject at the 0.005 level* the hypothesis that $x_D = 0.5$.

Let us now test the hypothesis

$$H_0: x_D = 0.486.$$

The calculated R factor is 0.090, and

$$\mathcal{R} = 0.090/0.083 = 1.084.$$

Since

$$\mathcal{R}_{1, 16, 0.10} = 1.091,$$

we cannot reject H_0 at the 0.10 significance level. If we had applied the usual Student's t test however, we would have calculated

$$t = (0.4860 - 0.4814)/0.0015 = 3.0667,$$

to be compared with $t_{15, 0.01} = 2.947$. Hence, the t test would have rejected the hypothesis at the 0.01 level. Fig. 1 illustrates the non-linearities which cause the two tests to be inconsistent; the test of \mathcal{R} is to be preferred.

* Clearly, the hypothesis could be rejected at a much lower level, but the smallest value of α in Table I was used.

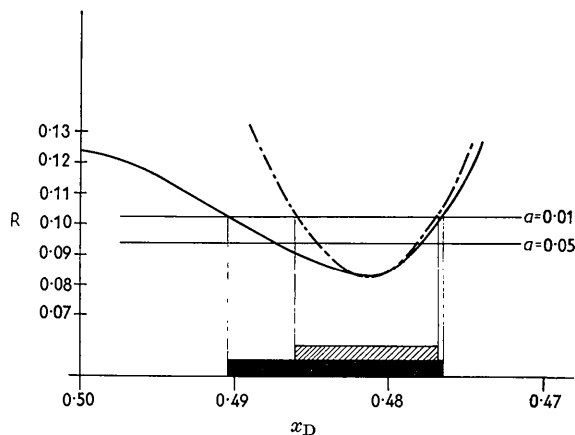


Fig. 1. The heavy curve is the R factor for DCrO_2 as a function of x_D . The dash-dot curve is the R factor predicted on the assumption that the structure factors are linear in the parameter changes. The horizontal lines are the R factors at the 0.01 and 0.05 significance levels. The solid black bar is thus the 99% confidence interval for x_D based on a test of \mathcal{R} , and the cross-hatched bar is the 99% confidence interval for x_D based on a Student's t test. The former is seen to be considerably longer.

Example 4

A refinement of 60 parameters in an acentric structure was carried out with 540 reflections. The calculated structure factors correctly included the effects of anomalous scattering and hence were sensitive to the absolute configuration. The generalized R factor for the better of the two absolute configurations was 0.140; that for the alternative absolute configuration 0.141. The statement

H_0 : The second absolute configuration is correct

is a one-dimensional hypothesis. In this particular problem, it involved changing the coordinate z to $-z$ for all atoms, but the choice of the sign of z for one atom determines it for all the rest. The value of the R -factor ratio is

$$\mathcal{R} = 0.1410/0.1400 = 1.007.$$

From Table 1, we find that

$$\mathcal{R}_{1, 480, 0.01} = 1.007.$$

Hence, we may reject at approximately the 0.01 level the hypothesis that the absolute configuration is the second of the two tried.

This method of absolute configuration determination without the necessity of measuring intensities for (hkl) and $(\bar{h}\bar{k}\bar{l})$ pairs was first suggested by Ibers & Hamilton (1964). One must assume, in this case as well as in others, that there are no systematic errors present in the data which would favor one configuration over the other. All statistical tests demand the assumption of random errors in the data, and systematic errors can lead one to make gross errors in the application of hypothesis tests. If large

systematic errors are feared, it is well to carry out the hypothesis tests conservatively, perhaps by using a smaller value of α than if the errors were known to be random.*

Most crystallographers will greet with an air of skepticism any conclusion based on an R factor change from 0.141 to 0.140. The author believes that the method applied here is entirely valid, provided that one bears in mind the comments above regarding systematic errors. Since this paper was first prepared, another example of the same type has been treated, resulting in a highly significant preference for the correct configuration over the incorrect one. The conventional R factors were 0.097 and 0.095, and the weighted R factors were 0.127 and 0.125 for the two configurations. The refinement was carried out for 612 observations and 201 parameters; the R factor ratio is significant at a level of less than $\frac{1}{2}\%$. The absolute configuration preferred agrees with that determined by 'classical' methods and by a method due to Kartha (1964) which gives considerably more weight to the reflections which show the greatest effects of anomalous scattering. The validity of the present approach, which the author by no means recommends as being the best for the determination of absolute configuration, can only be assessed by its application to many sets of data; the author is optimistic about the results of such an assessment.

Example 5

In a recent refinement of the structure of the methyl ester of *o*-nitrobenzenesulfenic acid, Hamilton & LaPlaca (1964) carried out an isotropic least-squares refinement of 49 parameters with 576 reflections. The resulting generalized R factor, based on 527° of freedom, was 0.173. The C-C bonds in the benzene ring had the following lengths: 1.44, 1.39, 1.36, 1.46, 1.34, 1.41, each with an estimated marginal standard deviation of 0.02 Å. It was desired to test the hypothesis:

H_0 : The bond lengths are equal to the average value of 1.40 Å and the ring is planar.

Using the group refinement program described by La Placa & Ibers (1963), a least-squares refinement was carried out, freely varying the parameters of the nitro group and the -S-O-CH₃ function but refining only the position of the centroid and the orientation of the benzene ring, keeping the bond lengths fixed at 1.40 Å and the bond angles at 120°. The resulting R factor was 0.181. The value of \mathcal{R} is

$$\mathcal{R} = 0.181/0.173 = 1.046$$

and the number of degrees of freedom of the hypothesis is $6 \times 4 - 7 = 17$. (For purposes of this example, both

* We include among systematic errors the use of an incorrect model for the calculated structure factors.

refinements were carried out with isotropic temperature factors, one for each atom in the unrestrained case and one for the group in the restrained case. Hydrogen atoms were not included in either refinement.) Interpolation from Table 1 gives us a value

$$\mathcal{R}_{17, 527, 0.005} = 1.034,$$

and the hypothesis can therefore be strongly rejected.

Example 6

Cruickshank, Lynton & Barclay (1962) refined the structure of thortveitite in three space groups with the following results:

Space group	Parameters	R^G
<i>Cm</i>	58	$\sqrt{0.0283} = 0.168$
<i>C2</i>	54	$\sqrt{0.0279} = 0.167$
<i>C2/m</i>	33	$\sqrt{0.0325} = 0.180$

The refinement with 252 pieces of data was carried out on F^2 . It was desired to test the hypothesis that the space group is *C2/m*, rather than one of the two acentric space groups. Let us take the refinement in *Cm* as the unrestricted case. Restriction to *C2/m* is a linear hypothesis of dimension $58 - 33 = 25$. We may test the ratio $0.180/0.168 = 1.071$ as $\mathcal{R}_{25, 194}$. Interpolation from Table 1 results in

$$\mathcal{R}_{25, 194, 0.25} = 1.073.$$

Hence we cannot reject at the 25% significance level the hypothesis that *C2/m* is the correct space group. This is the conclusion reached by Cruickshank *et al.*, who preferred the centrosymmetric space group for stereochemical reasons. One can make a similar comparison of *C2/m* and *C2* with a similar result. The fact that *C2* with 54 parameters resulted in a slightly lower value for R than that in *Cm* with 58 parameters is perhaps to be understood on the basis that the structure in *C2* cannot be obtained simply by putting restrictions on the parameters in *Cm*, but by relaxing some restrictions that are present as well. Non-linearities are doubtless of some importance also.

APPENDIX

The values in Table 1 were calculated by means of equation (23), using the tables of significance points of F tabulated by Merrington & Thompson (1946). Interpolation procedures for values of b and $(n-m)$ not found in the tables are based on the fact that interpolation in F may be carried out on the reciprocals of the degrees of freedom. For interpolation on $(n-m)$, the denominator degrees of freedom, we have to a sufficient degree of accuracy, where $N_0 < N_1 < N_2$:

$$\mathcal{R}_{b, N_1, \alpha} \simeq \frac{1}{N_1(N_2 - N_0)} \times [(N_1 - N_0)N_2\mathcal{R}_{b, N_2, \alpha} + (N_2 - N_1)N_0\mathcal{R}_{b, N_0, \alpha}] \quad (\text{Ia})$$

SIGNIFICANCE TESTS ON THE CRYSTALLOGRAPHIC R FACTOR

Table 1. Significance points of the R-factor ratio

TABLES FOR TESTING THE R FACTOR AT SIGNIFICANCE LEVEL 0.500

Table with columns for significance level (0.500) and rows for R-factor ratios (1 to INF). Values range from 1.414 to 1.000.

TABLES FOR TESTING THE R FACTOR AT SIGNIFICANCE LEVEL 0.250

Table with columns for significance level (0.250) and rows for R-factor ratios (1 to INF). Values range from 2.613 to 1.000.

TABLES FOR TESTING THE R FACTOR AT SIGNIFICANCE LEVEL 0.100

Table with columns for significance level (0.100) and rows for R-factor ratios (1 to INF). Values range from 6.392 to 1.000.

Table 1 (cont.)

TABLES FOR TESTING THE R FACTOR AT SIGNIFICANCE LEVEL 0.05C

Table with columns for R (1-120) and N-P (1-120). Values range from 1.000 to 99.999. Includes a note: A VALUE OF 99.999 FOR THE R FACTOR RATIO INDICATES THAT THE TRUE VALUE IS 100.00 OR GREATER.

A VALUE OF 99.999 FOR THE R FACTOR RATIO INDICATES THAT THE TRUE VALUE IS 100.00 OR GREATER

TABLES FOR TESTING THE R FACTOR AT SIGNIFICANCE LEVEL 0.025

Table with columns for R (1-120) and N-P (1-120). Values range from 1.000 to 99.999. Includes a note: A VALUE OF 99.999 FOR THE R FACTOR RATIO INDICATES THAT THE TRUE VALUE IS 100.00 OR GREATER.

A VALUE OF 99.999 FOR THE R FACTOR RATIO INDICATES THAT THE TRUE VALUE IS 100.00 OR GREATER

TABLES FOR TESTING THE R FACTOR AT SIGNIFICANCE LEVEL 0.01C

Table with columns for R (1-120) and N-P (1-120). Values range from 1.000 to 99.999. Includes a note: A VALUE OF 95.999 FOR THE R FACTOR RATIO INDICATES THAT THE TRUE VALUE IS 100.00 OR GREATER.

A VALUE OF 95.999 FOR THE R FACTOR RATIO INDICATES THAT THE TRUE VALUE IS 100.00 OR GREATER

Table 1 (cont.)

TABLES FOR TESTING THE R FACTOR AT SIGNIFICANCE LEVEL 0.005

N-P	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999	99.999
2	10.012	14.142	17.333	15.987	22.344	24.474	26.434	28.257	29.971	31.591	34.605	38.688	44.671	48.934	54.709	63.170	77.365	99.999
3	4.418	5.848	6.922	7.912	8.755	9.522	10.231	10.894	11.517	12.109	13.212	14.711	16.917	18.493	20.632	23.772	29.051	40.995
4	2.972	3.761	4.361	4.915	5.392	5.828	6.232	6.611	6.969	7.309	7.944	8.811	10.091	11.008	12.255	14.090	17.180	24.188
5	2.357	2.895	3.304	3.667	3.992	4.291	4.569	4.831	5.078	5.314	5.755	6.359	7.253	7.896	8.771	10.062	12.240	17.192
6	2.026	2.418	2.731	3.003	3.249	3.475	3.686	3.884	4.073	4.252	4.590	5.053	5.741	6.237	6.913	7.912	9.403	13.455
7	1.822	2.132	2.380	2.597	2.793	2.974	3.144	3.304	3.456	3.602	3.875	4.251	4.812	5.216	5.770	6.589	7.978	11.150
8	1.684	1.935	2.144	2.324	2.488	2.635	2.781	2.915	3.042	3.164	3.394	3.712	4.186	4.529	4.999	5.695	6.880	9.590
9	1.585	1.802	1.976	2.130	2.269	2.399	2.521	2.636	2.746	2.851	3.050	3.325	3.736	4.035	4.444	5.052	6.088	8.466
10	1.511	1.695	1.850	1.984	2.106	2.220	2.326	2.427	2.524	2.617	2.792	3.034	3.398	3.663	4.026	4.568	5.491	7.616
11	1.453	1.615	1.753	1.871	1.980	2.080	2.175	2.265	2.352	2.434	2.591	2.808	3.135	3.373	3.700	4.189	5.024	6.951
12	1.407	1.555	1.675	1.781	1.879	1.965	2.055	2.136	2.214	2.289	2.430	2.627	2.924	3.141	3.439	3.885	4.649	6.415
13	1.369	1.503	1.612	1.708	1.796	1.876	1.957	2.031	2.102	2.170	2.299	2.479	2.751	2.950	3.225	3.635	4.341	5.974
14	1.338	1.460	1.559	1.647	1.728	1.804	1.875	1.943	2.008	2.071	2.190	2.356	2.607	2.791	3.046	3.427	4.082	5.604
15	1.311	1.424	1.515	1.596	1.671	1.740	1.806	1.869	1.929	1.987	2.098	2.252	2.485	2.657	2.894	3.250	3.863	5.289
16	1.289	1.392	1.477	1.552	1.621	1.686	1.747	1.806	1.862	1.916	2.019	2.162	2.381	2.541	2.763	3.097	3.674	5.018
17	1.269	1.366	1.444	1.514	1.579	1.635	1.696	1.751	1.803	1.854	1.950	2.085	2.290	2.441	2.650	2.965	3.509	4.781
18	1.252	1.342	1.416	1.481	1.542	1.598	1.652	1.703	1.752	1.800	1.890	2.017	2.211	2.353	2.550	2.849	3.365	4.573
19	1.237	1.322	1.391	1.452	1.509	1.562	1.613	1.661	1.707	1.752	1.838	1.957	2.140	2.275	2.463	2.746	3.237	4.388
20	1.224	1.303	1.368	1.426	1.480	1.530	1.578	1.624	1.667	1.710	1.791	1.904	2.078	2.206	2.384	2.654	3.122	4.223
21	1.212	1.287	1.349	1.403	1.454	1.502	1.547	1.590	1.632	1.672	1.749	1.857	2.022	2.144	2.314	2.572	3.019	4.074
22	1.201	1.272	1.331	1.383	1.431	1.476	1.519	1.560	1.600	1.638	1.711	1.814	1.972	2.088	2.251	2.497	2.926	3.940
23	1.191	1.255	1.315	1.364	1.410	1.453	1.494	1.533	1.571	1.607	1.677	1.775	1.926	2.038	2.193	2.430	2.842	3.818
24	1.182	1.247	1.300	1.347	1.391	1.432	1.471	1.508	1.545	1.579	1.646	1.740	1.885	1.992	2.141	2.368	2.765	3.706
25	1.174	1.236	1.287	1.332	1.374	1.413	1.450	1.486	1.521	1.554	1.618	1.708	1.847	1.950	2.093	2.312	2.695	3.603
26	1.167	1.226	1.274	1.318	1.358	1.395	1.431	1.465	1.499	1.531	1.592	1.679	1.812	1.911	2.050	2.260	2.630	3.508
27	1.160	1.217	1.263	1.305	1.343	1.375	1.414	1.447	1.478	1.509	1.568	1.652	1.780	1.875	2.009	2.213	2.570	3.421
28	1.154	1.208	1.253	1.293	1.330	1.364	1.397	1.429	1.460	1.489	1.546	1.627	1.750	1.843	1.972	2.168	2.514	3.340
29	1.148	1.200	1.243	1.282	1.317	1.351	1.382	1.413	1.443	1.471	1.526	1.604	1.723	1.812	1.937	2.127	2.463	3.264
30	1.143	1.193	1.234	1.271	1.306	1.338	1.369	1.399	1.427	1.454	1.507	1.582	1.698	1.784	1.905	2.089	2.415	3.194
40	1.105	1.142	1.172	1.195	1.224	1.248	1.270	1.292	1.313	1.334	1.373	1.429	1.516	1.582	1.674	1.815	2.068	2.682
60	1.068	1.092	1.112	1.130	1.146	1.162	1.176	1.191	1.205	1.218	1.244	1.282	1.340	1.384	1.447	1.545	1.721	2.161
120	1.034	1.045	1.055	1.062	1.071	1.075	1.086	1.093	1.100	1.107	1.120	1.139	1.168	1.191	1.223	1.274	1.369	1.614
INF	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

A VALUE OF 99.999 FOR THE R FACTOR RATIO INDICATES THAT THE TRUE VALUE IS 100.00 OR GREATER

In particular, for interpolation between $N_0=120$ and $N_2=\infty$, we have

$$\mathcal{R}_{b, N_1, \alpha} \approx 1 + \frac{120}{N_1} (\mathcal{R}_{b, 120, \alpha} - 1). \quad (1b)$$

For interpolation on b , we have for $b_0 < b_1 < b_2$:

$$\mathcal{R}_{b_1, N, \alpha}^2 \approx \frac{1}{b_2 - b_0} [(b_1 - b_0)\mathcal{R}_{b_2, N, \alpha}^2 + (b_2 - b_1)\mathcal{R}_{b_0, N, \alpha}^2].$$

Although the selection of degrees of freedom for the tabulation is perhaps not the best possible for crystallographic purposes, it was felt that the tabulation should be presented for the values found in the F' tables and allow the reader to make the interpolations for himself.

The author would like to express his appreciation to J. A. Ibers and Sam J. La Placa for some stimulating conversations, to Dr Gopinath Kartha for supplying part of the data used in Example 4 and to Professor D. W. J. Cruickshank for some helpful criticism.

References

BEALE, E. M. L. (1960). *J. Roy. Stat. Soc. B*, **22**, 41.

CRUICKSHANK, D. W. J., LYNTON, H. & BARCLAY, G. A. (1962). *Acta Cryst.* **15**, 491.
 GRAYBILL, F. A. (1962). *An Introduction to Linear Statistical Models*. Vol. 1, New York: McGraw-Hill.
 HAMILTON, W. C. (1961). Abstracts, American Crystallographic Association Meeting, Boulder, Colorado.
 HAMILTON, W. C. (1964). *Statistics in Physical Science; Estimation, Hypothesis Testing, and Least Squares*. New York: The Ronald Press.
 HAMILTON, W. C. & IBERS, J. A. (1963a). *Acta Cryst.* **16**, 1209.
 HAMILTON, W. C. & IBERS, J. A. (1963b). *Noble Gas Compounds*, H. H. HYMAN, Ed. p. 195. Chicago: Univ. of Chicago Press.
 HAMILTON, W. C. & LA PLACA, S. J. (1964). *J. Amer. Chem. Soc.* **86**, 2289.
 IBERS, J. A. & HAMILTON, W. C. (1964). *Acta Cryst.* **17**, 781.
 KARTHA, G. (1964). Private communication.
 LA PLACA, S. J. & IBERS, J. A. (1963). *J. Amer. Chem. Soc.* **85**, 3501.
 MERRINGTON, M. & THOMPSON, C. M. (1946). *Biometrika*, **33**, 73.
 PLACKETT, R. L. (1960). *Regression Analysis*. Oxford Univ. Press.